

Algebra & Number Theory

Volume 19

2025

No. 4



Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR

Antoine Chambert-Loir
Université Paris-Diderot
France

EDITORIAL BOARD CHAIR

David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Jason P. Bell	University of Waterloo, Canada	Philippe Michel	École Polytechnique Fédérale de Lausanne
Bhargav Bhatt	University of Michigan, USA	Martin Olsson	University of California, Berkeley, USA
Frank Calegari	University of Chicago, USA	Irena Peeva	Cornell University, USA
J.-L. Colliot-Thélène	CNRS, Université Paris-Saclay, France	Jonathan Pila	University of Oxford, UK
Brian D. Conrad	Stanford University, USA	Anand Pillay	University of Notre Dame, USA
Samit Dasgupta	Duke University, USA	Bjorn Poonen	Massachusetts Institute of Technology, USA
Hélène Esnault	Freie Universität Berlin, Germany	Victor Reiner	University of Minnesota, USA
Gavril Farkas	Humboldt Universität zu Berlin, Germany	Peter Sarnak	Princeton University, USA
Sergey Fomin	University of Michigan, USA	Michael Singer	North Carolina State University, USA
Edward Frenkel	University of California, Berkeley, USA	Vasudevan Srinivas	SUNY Buffalo, USA
Wee Teck Gan	National University of Singapore	Shunsuke Takagi	University of Tokyo, Japan
Andrew Granville	Université de Montréal, Canada	Pham Huu Tiep	Rutgers University, USA
Ben J. Green	University of Oxford, UK	Ravi Vakil	Stanford University, USA
Christopher Hacon	University of Utah, USA	Akshay Venkatesh	Institute for Advanced Study, USA
Roger Heath-Brown	Oxford University, UK	Melanie Matchett Wood	Harvard University, USA
János Kollár	Princeton University, USA	Shou-Wu Zhang	Princeton University, USA
Michael J. Larsen	Indiana University Bloomington, USA		

PRODUCTION

production@msp.org
Silvio Levy, Scientific Editor


See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2025 is US \$565/year for the electronic version, and \$820/year (+\$70, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW[®] from MSP.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2025 Mathematical Sciences Publishers

Odd moments in the distribution of primes

Vivian Kuperberg

Montgomery and Soundararajan showed that the distribution of $\psi(x+H) - \psi(x)$, for $0 \leq x \leq N$, is approximately normal with mean $\sim H$ and variance $\sim H \log(N/H)$, when $N^\delta \leq H \leq N^{1-\delta}$. Their work depends on showing that sums $R_k(h)$ of k -term singular series are $\mu_k(-h \log h + Ah)^{k/2} + O_k(h^{k/2-1/(7k)+\varepsilon})$, where A is a constant and μ_k are the Gaussian moment constants. We study lower-order terms in the size of these moments. We conjecture that when k is odd, $R_k(h) \asymp h^{(k-1)/2}(\log h)^{(k+1)/2}$. We prove an upper bound with the correct power of h when $k = 3$, and prove analogous upper bounds in the function field setting when $k = 3$ and $k = 5$. We provide further evidence for this conjecture in the form of numerical computations.

1. Introduction	617
2. Three-term integer sums: proof of Theorem 1.2	621
3. Function field analogs: proof of Theorem 1.3	643
4. The fifth moment of reduced residues in the function field setting	651
5. Numerical evidence for odd moments	662
6. Toy models and open problems	664
References	666

1. Introduction

What is the distribution of primes in short intervals? Cramér [1936] modeled the indicator function of the sequence of primes by independent random variables X_n , for $n \geq 3$, which are 1 (“ n is prime”) with probability $1/\log n$, and 0 (“ n is composite”) with probability $1 - 1/\log n$. Cramér’s model predicts that the distribution of $\psi(n+h) - \psi(n)$, a weighted count of the number of primes in an interval of size h starting at n , follows a Poisson distribution when n varies in $[1, N]$ and when $h \asymp \log N$. Gallagher [1976] proved that this follows from a quantitative version of the Hardy–Littlewood prime k -tuple conjecture: namely, that if $\mathcal{D} = \{d_1, d_2, \dots, d_k\}$ is a set of k distinct integers, then

$$\sum_{n \leq N} \prod_{i=1}^k \Lambda(n + d_i) = (\mathfrak{S}(\mathcal{D}) + o(1))N,$$

The author is supported by NSF GRFP grant DGE-1656518 and would like to thank Kannan Soundararajan for many helpful comments and discussions, as well as Régis de la Bretèche, Alexandra Florea, Andrew Granville, Zeev Rudnick, Yuval Wigderson, and the anonymous referee for helpful feedback.

MSC2020: 11N05, 11N13.

Keywords: sums of singular series, distribution of primes.

where $\mathfrak{S}(\mathcal{D})$ is the singular series, a constant dependent on \mathcal{D} given by

$$\mathfrak{S}(\mathcal{D}) = \prod_p \left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{\nu_p(\mathcal{D})}{p}\right),$$

where $\nu_p(\mathcal{D})$ denotes the number of distinct residue classes modulo p among the elements of \mathcal{D} . The singular series is also given by the formula

$$\mathfrak{S}(\mathcal{D}) = \sum_{\substack{q_1, \dots, q_k \\ 1 \leq q_i < \infty}} \left(\prod_{i=1}^k \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^k \frac{a_i d_i}{q_i}\right). \tag{1}$$

The Hardy–Littlewood prime k -tuple conjectures give us a better lens through which to understand the distribution of primes: by understanding sums of singular series. For example, Gallagher used the estimate that

$$\sum_{\mathcal{D} \subset [1, h]} \mathfrak{S}(\mathcal{D}) \sim \sum_{\mathcal{D} \subset [1, h]} 1$$

to prove that the Hardy–Littlewood conjectures imply Poisson behavior in intervals of logarithmic length. Our concern is the distribution of primes in somewhat longer intervals, namely, those of size H , where $H = o(N)$ and $H/\log N \rightarrow \infty$ as $N \rightarrow \infty$. In this setting, the Cramér model would predict that the distribution of $\psi(n + H) - \psi(n)$ for $n \leq N$ is approximately normal, with mean $\sim H$ and variance $\sim H \log N$. However, the Hardy–Littlewood prime k -tuple conjecture gives a different answer in this case. Montgomery and Soundararajan [2004] provided evidence based on the Hardy–Littlewood prime k -tuple conjectures that the distribution ought to be approximately normal with variance $\sim H \log(N/H)$. They consider the K -th moment $M_K(N; H)$ of the distribution of primes in an interval of size H , given by

$$M_K(N; H) = \sum_{n=1}^N (\psi(n + H) - \psi(n) - H)^K.$$

They conjecture that these moments should be given by the Gaussian moments

$$M_K(N; H) = (\mu_K + o(1))N \left(H \log \frac{N}{H}\right)^{K/2},$$

where $\mu_K = 1 \cdot 3 \cdots (K - 1)$ if K is even and 0 if K is odd, uniformly for $(\log N)^{1+\delta} \leq H \leq N^{1-\delta}$. Their technique relies on more refined estimates of sums of the singular series constants $\mathfrak{S}(\mathcal{D})$. Instead of the von Mangoldt function $\Lambda(n)$, they consider sums of $\Lambda_0(n) = \Lambda(n) - 1$, where the main term has been subtracted from the beginning. The corresponding form of the Hardy–Littlewood conjecture states that

$$\sum_{n \leq N} \prod_{i=1}^k \Lambda_0(n + d_i) = (\mathfrak{S}_0(\mathcal{D}) + o(1))N$$

as $N \rightarrow \infty$, where $\mathfrak{S}_0(\mathcal{D})$ is given by

$$\mathfrak{S}_0(\mathcal{D}) = \sum_{\mathcal{J} \subseteq \mathcal{D}} (-1)^{|\mathcal{D} \setminus \mathcal{J}|} \mathfrak{S}(\mathcal{J}),$$

and in turn

$$\mathfrak{S}(\mathcal{D}) = \sum_{\mathcal{J} \subseteq \mathcal{D}} \mathfrak{S}_0(\mathcal{J}).$$

We can combine this with (1) to see that

$$\mathfrak{S}_0(\mathcal{D}) = \sum_{\substack{q_1, \dots, q_k \\ 1 < q_i < \infty}} \left(\prod_{i=1}^k \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^k \frac{a_i d_i}{q_i} \right). \tag{2}$$

Montgomery and Soundararajan considered the sum

$$R_k(h) := \sum_{\substack{d_1, \dots, d_k \\ 1 \leq d_i \leq h \\ d_i \text{ distinct}}} \mathfrak{S}_0(\mathcal{D}), \tag{3}$$

showing that, for any nonnegative integer k , for any $h > 1$, and for any $\varepsilon > 0$,

$$R_k(h) = \mu_k(-h \log h + Ah)^{k/2} + O_k(h^{k/2-1/(7k)+\varepsilon}), \tag{4}$$

where $A = 2 - \gamma - \log 2\pi$. Their estimate on $R_k(h)$ implies their bound on the moments. For more on the distribution of primes in short intervals, see for example [Chan 2006; Granville and Lumley 2023; Montgomery and Soundararajan 2004].

For all k , the optimal error term in (4) is expected to be smaller. In the case of the variance, this was studied in [Montgomery and Soundararajan 2002]. In this paper, we restrict our attention to the cases when k is odd. We conjecture the following, which was mentioned in [Lemke Oliver and Soundararajan 2016].

Conjecture 1.1. *Let $k \geq 3$ be an odd integer, and let $h > 1$. With $R_k(h)$ defined as above,*

$$R_k(h) \asymp h^{(k-1)/2} (\log h)^{(k+1)/2}.$$

The conjectured power of $\log h$ here comes from numerical evidence, which we present in Section 5. For k odd, we do not know, even heuristically, which terms contribute to the main term in $R_k(h)$; for this reason, we do not know what the constant should be in front of the asymptotic in Conjecture 1.1. Nevertheless, our goal in this paper is to provide evidence for Conjecture 1.1. When $k = 3$, we can show an upper bound with the correct power of h .

Theorem 1.2. *For $h \geq 4$ and R_3 defined in (3),*

$$R_3(h) \ll h(\log h)^5.$$

Another source of evidence for Conjecture 1.1 is the analog of this problem in the function field setting, which is also studied in [Keating and Rudnick 2014]. As we discuss in Section 3, we can consider analogous questions over $\mathbb{F}[T]$, where \mathbb{F} is a finite field, instead of over \mathbb{Z} . To state the analog, we first revisit the

techniques of Montgomery and Soundararajan in the integer case. Upon expanding (3) using (2), we get

$$R_k(h) = \sum_{\substack{d_1, \dots, d_k \\ 1 \leq d_i \leq h \\ d_i \text{ distinct}}} \sum_{\substack{q_1, \dots, q_k \\ 1 < q_i < \infty}} \left(\prod_{i=1}^k \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^k \frac{a_i d_i}{q_i}\right) = \sum_{\substack{q_1, \dots, q_k \\ 1 < q_i < \infty}} \left(\prod_{i=1}^k \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} \prod_{i=1}^k E\left(\frac{a_i}{q_i}\right),$$

where $E(\alpha) = \sum_{m=1}^h e(m\alpha)$. The sums $E(\alpha)$ approximately detect when $\|\alpha\| \leq 1/h$.

This expression for $R_k(h)$ is closely related to a quantity studied by in [Montgomery and Vaughan 1986]. They considered the related problem of the k -th moment of reduced residues modulo a fixed q , given by

$$m_k(q; h) = \sum_{n=1}^q \left(\sum_{\substack{1 \leq m \leq h \\ (m+n, q) = 1}} 1 - h \frac{\phi(q)}{q} \right)^k.$$

The moment m_k satisfies $m_k(q; h) = q(\phi(q)/q)^k V_k(q; h)$, where $V_k(q; h)$ is the “singular series sum”,

$$V_k(q; h) = \sum_{\substack{d_1, \dots, d_k \\ 1 \leq d_i \leq h}} \sum_{\substack{q_1, \dots, q_k \\ 1 < q_i | q}} \left(\prod_{i=1}^k \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^k \frac{a_i d_i}{q_i}\right),$$

which differs from $R_k(h)$ only in that the q_i are now constrained to divide a fixed q . In this paper, as well as in the work of Montgomery and Soundararajan, estimating $V_k(q; h)$ when q is a product of primes $p \leq h^{k+1}$ is a key step towards estimating $R_k(h)$. Similarly, understanding $m_k(q; h)$ is closely related to understanding $R_k(h)$. For example, Conjecture 1.1 predicts that $R_k(h) \asymp h^{(k-1)/2} (\log h)^{(k+1)/2}$ when k is odd; this conjecture is closely related to the prediction that when q is a product of primes $p \leq h^A$ for a fixed power A , and when k is odd, then we should have $m_k(q; h) \asymp q(h/(\log h))^{(k-1)/2}$. Montgomery and Vaughan [1986] predicted that $m_k(q; h) \ll q(h/(\log h))^{(k-1)/2}$ in this setting. In the function field setting, we study an analog of the moments $m_k(q; h)$.

Let \mathbb{F}_q be a finite field with q elements, and let Q be a fixed monic polynomial in $\mathbb{F}_q[t]$. Note that Q in the function field case serves the same role as q in the integer case, since q now represents the size of the field. The moment $m_k(Q; h)$, an analog of the k -th moment of reduced residues in short intervals which is defined precisely in (15), is the k -th moment of the distribution of polynomials that are relatively prime to Q lying in intervals of size q^h in the function field $\mathbb{F}_q[t]$. In this case an “interval” of size q^h centered at a polynomial $G(t)$ consists of all polynomials $F(t)$ such that $F(t) \equiv G(t) \pmod{t^h}$. We can adapt the methods of Montgomery–Vaughan to prove a bound on $m_k(Q; h)$ that has the same shape as the bounds of Montgomery–Vaughan and Montgomery–Soundararajan.

Theorem 1.3. *For any fixed $k \geq 3$ and for $Q \in \mathbb{F}_q[t]$ square-free, for $h \geq 2$,*

$$m_k(Q; h) \ll \begin{cases} |Q|(q^h)^{k/2} \left(\frac{\phi(Q)}{|Q|}\right)^{k/2} \left(1 + (q^h)^{-1/(k-2)} \left(\frac{\phi(Q)}{|Q|}\right)^{-2^k+k/2}\right) & \text{if } k \text{ is even,} \\ |Q|((q^h)^{k/2-1/2} + (q^h)^{k/2-1/(k-2)}) \left(\frac{\phi(Q)}{|Q|}\right)^{-2^k+k/2} & \text{if } k \text{ is odd.} \end{cases}$$

The function field exponential sums are cleaner than their integer analogs, making this proof more streamlined than the proof of Montgomery–Vaughan. As a result, the bound is tighter; in fact, for $k = 3$, Theorem 1.3 already yields a bound where the exponent of q^h is 1. This is of the same shape as Theorem 1.2, where the exponent of h is 1.

Using a more involved argument we can achieve a bound on the fifth moment of reduced residues in short intervals.

Theorem 1.4. *Let $h \geq 2$ and let*

$$Q = \prod_{\substack{P \text{ irred.} \\ |P| \leq q^{6h}}} P.$$

For all $\varepsilon > 0$,

$$m_5(Q; h) \ll |Q|q^{2h+\varepsilon}.$$

As discussed above, Conjecture 1.1 would predict in the integer case that for k odd and $q = \prod_{p \leq h^\Lambda} p$, we have $m_k(q; h) \asymp q(h/(\log h))^{(k-1)/2}$. In the function field case, we have a polynomial $Q(t)$ in place of the modulus q , and an interval of size q^h instead of one of size h , so the analog of Conjecture 1.1 would predict that $m_5(Q; h) \asymp |Q|q^{2h}(\log q^h)^{-2}$. In particular, Theorem 1.4 matches the exponent of q^h in this prediction. Our techniques do not quite succeed in proving such a bound for any higher odd moments, as we note in Section 4. However, we do get as a corollary the following bound on sums of singular series in function fields. The sum $R_k(q^h)$ of singular series in function fields is defined very analogously to the sum $R_k(h)$ in the integer setting; a precise definition is given in (19).

Corollary 1.5. *Let $h \geq 2$ and let*

$$Q = \prod_{\substack{P \text{ irred.} \\ |P| \leq q^{6h}}} P.$$

Then

$$R_3(q^h) \ll V_3(Q; h) + q^h \left(\frac{|Q|}{\phi(Q)} \right)^2 \ll q^h \left(\frac{|Q|}{\phi(Q)} \right)^8,$$

and, for all $\varepsilon > 0$,

$$R_5(q^h) \ll V_5(Q; h) + \left(\frac{|Q|}{\phi(Q)} \right)^{21/2} q^{2h} \ll q^{(2+\varepsilon)h}.$$

This paper is organized as follows. In Section 2 we prove Theorem 1.2. In Section 3, we discuss the analogous problem in $\mathbb{F}_q[T]$, and adapt the framework of Montgomery and Vaughan to the function field setting to prove Theorem 1.3. In Section 4 we prove Theorem 1.4. Finally, in Section 5 we provide numerical evidence for Conjecture 1.1, and in Section 6 we discuss toy problems, further directions of inquiry, and possible applications of these questions.

2. Three-term integer sums: proof of Theorem 1.2

Our goal is to bound

$$R_3(h) = \sum_{\substack{d_1, d_2, d_3 \\ 1 \leq d_i \leq h \\ d_i \text{ distinct}}} \mathfrak{S}_0(\mathcal{D}).$$

Expanding $\mathfrak{S}_0(\mathcal{D})$ as an exponential sum yields

$$R_3(h) = \sum_{\substack{d_1, d_2, d_3 \\ 1 \leq d_i \leq h \\ d_i \text{ distinct}}} \sum_{\substack{q_1, q_2, q_3 \\ 1 < q_i < \infty}} \left(\prod_{i=1}^3 \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, a_2, a_3 \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^3 \frac{a_i d_i}{q_i}\right).$$

Our argument will follow the same thread as that of [Montgomery and Soundararajan 2004], which in turn relies on the analysis of [Montgomery and Vaughan 1986] of the distribution of reduced residues. To that end, we consider $V_3(q; h)$, which is approximately the third centered moment of the number of reduced residues mod q in an interval of length h . Precisely, $V_3(q; h)$ is given by

$$V_3(q; h) = \sum_{\substack{d_1, d_2, d_3 \\ 1 \leq d_i \leq h}} \sum_{\substack{q_1, q_2, q_3 \\ 1 < q_i \mid q}} \left(\prod_{i=1}^3 \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, a_2, a_3 \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^3 \frac{a_i d_i}{q_i}\right). \tag{5}$$

This is very similar to the above expression for $R_3(h)$; the two differences are that the outer sum in $R_3(h)$ is taken over *distinct* d_i 's, whereas the outer sum for $V_3(q; h)$ is not, and that the summands q_i range over all integers for $R_3(h)$, but are restricted to factors of q for $V_3(q; h)$.

Theorem 2.1. *Let $h \geq 4$ and let q be the product of primes $p \leq h^4$. Then*

$$V_3(q; h) \ll h(\log h)^5.$$

We use Theorem 2.1 to establish Theorem 1.2. In order to derive Theorem 1.2, it suffices to show that terms arising from transforming $V_3(q; h)$ into $R_3(h)$ do not contribute more than $O(h(\log h)^5)$; in fact they contribute on the order of $h(\log h)^2$, which is the conjectured asymptotic size of $R_3(h)$. We begin with this derivation of Theorem 1.2 from Theorem 2.1.

To account for terms where d_1, d_2, d_3 are not necessarily distinct, we make the following definition.

Definition 2.2. Let $k \geq 2$, and let $\mathcal{D} = \{d_1, \dots, d_k\}$ be a k -tuple of not necessarily distinct integers, and fix q a square-free integer. Then the *singular series at \mathcal{D} with respect to q* is given by

$$\mathfrak{S}(\mathcal{D}; q) := \sum_{q_1, \dots, q_k \mid q} \left(\prod_{i=1}^k \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^k \frac{a_i d_i}{q_i}\right).$$

Just as for $\mathfrak{S}(\mathcal{D})$, one can subtract off the main term of $\mathfrak{S}(\mathcal{D}; q)$ to define

$$\mathfrak{S}_0(\mathcal{D}; q) := \sum_{\mathcal{J} \subset \mathcal{D}} (-1)^{|\mathcal{D} \setminus \mathcal{J}|} \mathfrak{S}(\mathcal{J}; q).$$

Combining this with the definition for $\mathfrak{S}(\mathcal{D}; q)$ yields the formula

$$\mathfrak{S}_0(\mathcal{D}; q) = \sum_{1 < q_1, \dots, q_k \mid q} \left(\prod_{i=1}^k \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^k \frac{a_i d_i}{q_i}\right). \tag{6}$$

If the d_i are not all distinct, this expression converges for any fixed q but not in the $q \rightarrow \infty$ limit. The singular series at \mathcal{D} with respect to q is equal to a finite Euler product.

Lemma 2.3. *Let $k \geq 2$, and let $\mathcal{D} = \{d_1, \dots, d_k\}$ be a k -tuple of not necessarily distinct integers, and fix q a square-free integer. Then*

$$\mathfrak{S}(\mathcal{D}; q) = \prod_{p|q} \left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{v_p(\mathcal{D})}{p}\right),$$

where $v_p(\mathcal{D})$ is the number of distinct residue classes mod p occupied by elements of \mathcal{D} .

This lemma is proven in [Montgomery and Soundararajan 2004, Lemma 3]; it is stated there for sets with distinct elements, but their proof holds in this setting as well. They note first that $\mathfrak{S}(\mathcal{D}; q)$ is multiplicative in q , so that it suffices to check the lemma for primes p . For a given prime p , they express the condition that $\sum_{i=1}^k a_i/q_i \in \mathbb{Z}$ in terms of additive characters mod p , and then rearrange to get the result.

Consider the following expression for \mathfrak{S}_0 , which is [Montgomery and Soundararajan 2004, equation (45)]. For all $y \geq h$,

$$\mathfrak{S}_0(\mathcal{D}) = \sum_{\substack{q_1, q_2, q_3 \\ q_i > 1 \\ p|q_i \Rightarrow p \leq y}} \prod_{i=1}^3 \frac{\mu(q_i)}{\phi(q_i)} A(q_1, q_2, q_3; \mathcal{D}) + O\left(\frac{(\log y)}{y}\right),$$

where

$$A(q_1, q_2, q_3; \mathcal{D}) = \sum_{\substack{a_1, a_2, a_3 \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i/q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^3 \frac{d_i a_i}{q_i}\right).$$

Apply this to $R_3(h)$ with $y = h^4$ and $q = \prod_{p \leq y} p$ to get

$$R_3(h) = \sum_{\substack{q_1, q_2, q_3 \\ q_i > 1 \\ q_i | q}} \prod_{i=1}^3 \frac{\mu(q_i)}{\phi(q_i)} S(q_1, q_2, q_3; h) + O(1),$$

where

$$S(q_1, q_2, q_3; h) := \sum_{\substack{d_1, d_2, d_3 \\ 1 \leq d_i \leq h \\ d_i \text{ distinct}}} A(q_1, q_2, q_3; \{d_1, d_2, d_3\}) = \sum_{\substack{d_1, d_2, d_3 \\ 1 \leq d_i \leq h \\ d_i \text{ distinct}}} \sum_{\substack{a_1, a_2, a_3 \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i/q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^3 \frac{d_i a_i}{q_i}\right).$$

If the condition that the d_i should be distinct were omitted, then the main term in $R_3(h)$ would be exactly $V_3(q; h)$. So, it suffices to remove this condition.

Put $\delta_{i,j} = 1$ if $d_i = d_j$ and 0 otherwise, so that

$$\prod_{1 \leq i < j \leq 3} (1 - \delta_{i,j}) = \begin{cases} 1 & \text{if the } d_i \text{ are all pairwise distinct,} \\ 0 & \text{otherwise,} \end{cases}$$

and

$$S(q_1, q_2, q_3; h) = \sum_{\substack{d_1, d_2, d_3 \\ 1 \leq d_i \leq h}} \left(\prod_{1 \leq i < j \leq 3} (1 - \delta_{i,j}) \right) \sum_{\substack{a_1, a_2, a_3 \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^3 \frac{d_i a_i}{q_i}\right).$$

Expanding the product over the $\delta_{i,j}$ yields

$$1 - \delta_{1,2} - \delta_{1,3} - \delta_{2,3} + \delta_{1,2}\delta_{2,3} + \delta_{1,3}\delta_{1,2} + \delta_{2,3}\delta_{1,3} - \delta_{1,2}\delta_{2,3}\delta_{1,3}.$$

Note that the last four terms each require precisely that $d_1 = d_2 = d_3$ in order to be nonzero; each of these can be written as $\delta_{1,2,3}$, so that their sum is $2\delta_{1,2,3}$. The following lemma addresses the contribution of these last four terms.

Lemma 2.4. *Let $h \geq 4$ be an integer. Then*

$$2 \sum_{d \leq h} \sum_{\substack{q_1, q_2, q_3 \\ q_i > 1 \\ q_i \nmid q}} \prod_{i=1}^3 \frac{\mu(q_i)}{\phi(q_i)} \sum_{\substack{a_1, a_2, a_3 \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(\sum_{i=1}^3 \frac{d a_i}{q_i}\right) = 2h \left(\frac{q}{\phi(q)}\right)^2 - 6h \frac{q}{\phi(q)} + 4h.$$

Proof. Note that the left-hand expression is precisely $2 \sum_{d \leq h} \mathfrak{S}_0(\{d, d, d\}; q)$. Expanding \mathfrak{S}_0 and applying Lemma 2.3 yields

$$\begin{aligned} 2 \sum_{d \leq h} \mathfrak{S}_0(\{d, d, d\}; q) &= 2 \sum_{d \leq h} (\mathfrak{S}(\{d, d, d\}; q) - 3\mathfrak{S}(\{d, d\}; q) + 3\mathfrak{S}(\{d\}; q) - 1) \\ &= 2 \sum_{d \leq h} \left(\prod_{p|q} \left(1 - \frac{1}{p}\right)^{-2} - 3 \prod_{p|q} \left(1 - \frac{1}{p}\right)^{-1} + 2 \right) \\ &= 2h \frac{q^2}{\phi(q)^2} - 6h \frac{q}{\phi(q)} + 4h, \end{aligned}$$

as desired. □

Now consider the contribution to $R_3(h)$ from the terms $-\delta_{1,2}$, $-\delta_{1,3}$, and $-\delta_{2,3}$. Via relabeling, it suffices to only consider the term with $-\delta_{1,2}$, which is nonzero when $d_1 = d_2$ and otherwise 0.

Lemma 2.5. *Let $h \geq 4$ be an integer. Then*

$$\sum_{d, d_3 \leq h} \sum_{\substack{q_1, q_2, q_3 \\ q_i > 1 \\ q_i \nmid q}} \prod_{i=1}^3 \frac{\mu(q_i)}{\phi(q_i)} \sum_{\substack{a_1, a_2, a_3 \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} e\left(d \left(\frac{a_1}{q_1} + \frac{a_2}{q_2}\right)\right) e\left(\frac{d_3 a_3}{q_3}\right) = \left(\frac{q}{\phi(q)} - 2\right) \left(h \frac{q}{\phi(q)} - h \log h + Bh + O(h^{1/2+\epsilon})\right).$$

Proof. As in the previous lemma, we note that the left-hand side is $\sum_{d, d_3 \leq h} \mathfrak{S}_0(\{d, d, d_3\}; q)$. We again expand and apply Lemma 2.3, to get

$$\begin{aligned} \sum_{d, d_3 \leq h} \mathfrak{S}_0(\{d, d, d_3\}; q) &= \sum_{d, d_3 \leq h} (\mathfrak{S}(\{d, d, d_3\}; q) - 2\mathfrak{S}(\{d, d_3\}; q) - \mathfrak{S}(\{d, d\}; q) + 2) \\ &= \left(\frac{q}{\phi(q)} - 2\right) \left(\sum_{d, d_3 \leq h} \mathfrak{S}(\{d, d_3\}; q) - h^2\right). \end{aligned}$$

By [Montgomery and Soundararajan 2004, Lemma 4],

$$\sum_{d, d_3 \leq h} \mathfrak{S}(\{d, d_3\}; q) = \sum_{q_1 | q} \frac{\mu(q_1)^2}{\phi(q_1)^2} \sum_{\substack{1 \leq a \leq q_1 \\ (a, q_1) = 1}} \left| E\left(\frac{a}{q_1}\right) \right|^2 = h \frac{q}{\phi(q)} + h^2 - h \log h + Bh + O(h^{1/2+\varepsilon}),$$

with $B = 1 - \gamma - \log 2\pi$. Thus our expression becomes

$$= \left(\frac{q}{\phi(q)} - 2\right) \left(h \frac{q}{\phi(q)} - h \log h + Bh + O(h^{1/2+\varepsilon})\right),$$

as desired. □

Combining these computations yields

$$\begin{aligned} R_3(h) &= V_3(q; h) + 2h \left(\frac{q}{\phi(q)}\right)^2 - 6h \frac{q}{\phi(q)} + 4h - 3 \left(\frac{q}{\phi(q)} - 2\right) \left(h \frac{q}{\phi(q)} - h \log h + Bh + O(h^{1/2+\varepsilon})\right) \\ &= V_3(q; h) - h \left(\frac{q}{\phi(q)}\right)^2 + 3h \log h \frac{q}{\phi(q)} - 3Bh \frac{q}{\phi(q)} - 6h \log h + 6Bh + 4h + O\left(h^{1/2+\varepsilon} \frac{q}{\phi(q)}\right). \end{aligned}$$

By Theorem 2.1, $V_3(q; h) \ll h(\log h)^5$, so $R_3(h) \ll h(\log h)^5$, which completes the proof of Theorem 1.2.

2.1. Preparing for the proof of Theorem 2.1. The rest of this section will be devoted to the proof of Theorem 2.1; here we begin by fixing some notation and proving several preparatory lemmas. Specifically, Lemmas 2.8, 2.9, 2.10, and 2.11 are general results on adding integer reciprocals along hyperplanes. Lemmas 2.12, 2.13, and 2.14 rely on these general results to prove bounds on specific sums that will appear in the proof of Theorem 2.1.

We begin with a reparametrization of variables into a system of common divisors. Let (q_1, q_2, q_3) be a triple in the sum in (5) defining $V_3(q; h)$. The contribution of the (q_1, q_2, q_3) -term to $V_3(q; h)$ is zero unless there are nontrivial solutions to

$$\frac{a_1}{q_1} + \frac{a_2}{q_2} + \frac{a_3}{q_3} \in \mathbb{Z},$$

or equivalently

$$a_1 q_2 q_3 + a_2 q_1 q_3 + a_3 q_1 q_2 \equiv 0 \pmod{q_1 q_2 q_3},$$

where $(a_i, q_i) = 1$ for all i . This implies that $q_1 \mid q_2 q_3$ (and likewise $q_2 \mid q_1 q_3$ and $q_3 \mid q_1 q_2$), since reducing mod q_1 shows that $a_1 q_2 q_3 \equiv 0 \pmod{q_1}$, and by assumption $(a_1, q_1) = 1$. Since q is square-free, so are q_1, q_2 , and q_3 , so we can reparametrize as follows. Let $g = \gcd(q_1, q_2, q_3)$ be the product of all primes dividing all three q_i 's. Define $x = \gcd(q_2/g, q_3/g)$, $y = \gcd(q_1/g, q_3/g)$, and $z = \gcd(q_1/g, q_2/g)$. Then $q_1 = gyz$, $q_2 = gxz$, and $q_3 = gxy$, with g, x, y, z pairwise coprime and square-free. This reparametrization is the same as writing the system of *relative greatest common divisors* for q_1, q_2 , and q_3 ; see for example [Elsholtz and Planitzer 2020] for more details.

Then

$$V_3(q; h) = \sum_{\substack{g, x, y, z | q \\ gxy, gxz, gyz > 1}} \frac{\mu(g)\mu(gxyz)^2}{\phi(g)\phi(gxyz)^2} \sum_{\substack{a_1, a_2, a_3 \\ 0 \leq a_1 < gyz, \dots \\ (a_1, gyz) = \dots = 1 \\ a_1/(gyz) + a_2/(gxz) + a_3/(gxy) \in \mathbb{Z}}} E\left(\frac{a_1}{gyz}\right) E\left(\frac{a_2}{gxz}\right) E\left(\frac{a_3}{gxy}\right).$$

We start by taking absolute values, using the bound that, for all $0 \leq \alpha < 1$, $|E(\alpha)| \leq F(\alpha)$, where

$$F(\alpha) := \min\{h, \|\alpha\|^{-1}\}, \quad (7)$$

so that

$$V_3(q; h) \leq \sum_{\substack{g, x, y, z | q \\ gxy, gxz, gyz > 1}} \frac{\mu(gxyz)^2}{\phi(g)\phi(gxyz)^2} \sum_{\substack{a_1, a_2, a_3 \\ 0 \leq a_1 < gyz, \dots \\ (a_1, gyz) = \dots = 1 \\ \sum a_i / gyz \in \mathbb{Z}}} F\left(\frac{a_1}{gyz}\right) F\left(\frac{a_2}{gxz}\right) F\left(\frac{a_3}{gxy}\right). \quad (8)$$

We now split the sum $V_3(q; h)$ into three different sums, addressed separately. Let T_1 consist of all terms g, x, y, z in (8) with $gx \geq h$. Let T_2 consist of all terms g, x, y, z in (8) with $gx < h$, $gy < h$, and $gz < h$, and $\|a_2/q_2\|, \|a_3/q_3\| > 1/h$. Finally, let T_3 consist of all terms g, x, y, z in (8) with $gx < h$, $gy < h$, and $gz < h$, as well as the constraints that $\|a_1/(gyz)\| \leq 2/h$, $\|a_2/(gxz)\| \leq 2/h$, and $\|a_3/(gxy)\| \leq 2/h$.

We claim that, after permuting the names of the variables as necessary, each term $g, x, y, z, a_1, a_2, a_3$ is contained in sums for T_1, T_2 , or T_3 . Terms where any of gx, gy , or gz are $\geq h$ are included in a copy of T_1 . For remaining terms we have $gx < h, gy < h$, and $gz < h$. If two of the three fractions a_i/q_i satisfy $\|a_i/q_i\| \leq 1/h$ (say $i = 1, 2$), then the third one must satisfy $\|a_3/q_3\| \leq 2/h$ because $a_1/q_1 + a_2/q_2 + a_3/q_3 \in \mathbb{Z}$; therefore, these terms are included up to permutation of indices in T_3 . The remaining terms must be included, up to permuting the indices, in T_2 . This implies in particular that

$$V_3(q; h) \ll T_1 + T_2 + T_3.$$

We will show in Lemmas 2.15, 2.16, and 2.17 respectively that $T_1 \ll h(\log h)^5$, $T_2 \ll h(\log h)^4(\log \log h)^2$, and $T_3 \ll h(\log h)^4(\log \log h)^2$, which completes the proof of Theorem 2.1.

In what follows, it will be helpful for us to approximate fractions a/q by a nearby multiple of $1/h$; to do so, we make the following definition.

Definition 2.6. Fix $h \geq 4$. Let $q > 1$ and let $1 \leq a < q$ with $(a, q) = 1$. If $q > h$, the h -approximate numerator $n(a, q)$ is defined as

$$n(a, q) = \left\lceil h \left\| \frac{a}{q} \right\| \right\rceil = \begin{cases} \left\lceil \frac{ha}{q} \right\rceil & \text{if } \frac{a}{q} \leq \frac{1}{2}, \\ h - \left\lfloor \frac{ha}{q} \right\rfloor & \text{if } \frac{a}{q} > \frac{1}{2}. \end{cases}$$

Meanwhile, if $q \leq h$, the h -approximate numerator $n(a, q)$ is defined to be a itself.

For example, if $q > h$ and $1/h < a/q \leq 2/h$, say, then the h -approximate numerator $n(a, q)$ is equal to 2, so that $n(a, q)/(2h) \leq a/q \leq n(a, q)/h$. The definition is arranged so that $n(a, q)$ is never zero

when $(a, q) = 1$; if $0 < a/q \leq 1/h$, then $n(a, q) = 1$. The key consequence of this definition is the following property.

Claim 2.7. *Let $h \geq 4$. For $F(\alpha)$ defined in (7), we have*

$$F\left(\frac{a}{q}\right) \leq 2 \left\| \frac{n(a, q)}{\min\{q, h\}} \right\|^{-1}. \tag{9}$$

Proof. If $q \leq h$, then (9) states that $\|a/q\|^{-1} \leq 2\|a/q\|^{-1}$, which is true.

For $q > h$, we restrict to considering the case when $a/q \in (0, \frac{1}{2}]$, so that $\|a/q\| = a/q$; the case when $a/q \in (\frac{1}{2}, 1)$ is analogous. Assume first that $0 < a/q \leq 1/h$. Then $F(a/q) = h$ and $n(a, q) = 1$, so that (9) states that $h \leq 2h$, which is true. Finally assume that $1/h < a/q$. By definition, $n(a, q) = \lceil ha/q \rceil = ha/q + e$, where $0 \leq e < 1$. For any such e ,

$$\left\| \frac{a}{q} + \frac{e}{h} \right\| \leq \left\| \frac{a}{q} \right\| + \frac{1}{h} \leq 2 \frac{a}{q}.$$

Thus

$$\left(\frac{a}{q}\right)^{-1} \leq 2 \left\| \frac{a}{q} + \frac{e}{h} \right\|^{-1} = 2 \left\| \frac{\lceil ha/q \rceil}{h} \right\|^{-1},$$

which is precisely (9) in this case. □

We write $\tilde{q} := \min\{q, h\}$, so that $F(a/q) \leq 2\|n(a, q)/\tilde{q}\|^{-1}$. For any fraction a/q , we then have that $a/q \approx n(a, q)/\tilde{q}$ in the sense that $|a/q - n(a, q)/\tilde{q}| < 1/h$, since if $q \leq h$ then $a/q = n(a, q)/\tilde{q}$, and if $q > h$ then this follows from the definition of $n(a, q)$.

We are now ready to proceed with several lemmas concerning sums of fractions, sums over quantities $\|a/q\|^{-1}$, and sums of $F(\alpha)$. The following four lemmas are general results on adding integer reciprocals of points lying close to certain hyperplanes. Loosely speaking, these lemmas will appear in our argument in the following way. For each of T_1, T_2 , and T_3 , we will have to evaluate a sum of the form

$$\sum_{\substack{a_1, a_2, a_3 \\ a_1/q_1 + a_2/q_2 + a_3/q_3 \in \mathbb{Z}}} F\left(\frac{a_1}{q_1}\right) F\left(\frac{a_2}{q_2}\right) F\left(\frac{a_3}{q_3}\right),$$

where in practice there will be further constraints on the terms a_i and q_i . After applying (9) and the observation that $a/q \approx n(a, q)/\tilde{q}$, and dealing with a little casework on the sign of $n(a_i, q_i)$, we arrive at a sum that is roughly of the form

$$8 \prod_{i=1}^3 \min\{q_i, h\} \sum_{\substack{a_1, a_2, a_3 \\ \|n(a_1, q_1)/\tilde{q}_1 + n(a_2, q_2)/\tilde{q}_2 + n(a_3, q_3)/\tilde{q}_3\| \approx 0}} \frac{1}{n(a_1, q_1)n(a_2, q_2)n(a_3, q_3)}.$$

In particular, in order to analyze T_1, T_2, T_3 , we will have to understand sums of reciprocals of lattice points. Understanding the precise sums requires some amount of casework, largely coming from the cases $q_i < h$ versus $q_i \geq h$ and the cases $a_i/q_i \leq \frac{1}{2}$ versus $a_i/q_i > \frac{1}{2}$. This casework is accomplished by the Lemmas 2.8, 2.10, and 2.11.

Lemma 2.8. Let $v_2 \geq v_1$ and $\alpha_1 \geq 1$ be real numbers, and let $h \in \mathbb{N}$ with $h \geq 4$. Then

$$\sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ 1 \leq n_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ -\alpha_1 n_1 + n_2 + n_3 \in [v_1, v_2]}} \frac{1}{\alpha_1 n_1 n_2 n_3} \ll \begin{cases} (v_2 - v_1 + 1) \frac{\log h}{\alpha_1} \left(\frac{2 - v_1}{\alpha_1} + 1 \right) & \text{if } v_1 < 0, \\ (v_2 - v_1 + 1) \frac{\log h}{\alpha_1^2} & \text{if } v_1 \geq 0, \end{cases}$$

where n_1, n_2 , and n_3 range over integers.

Proof. Since $n_2 + n_3 \geq v_1 + \alpha_1 n_1$ and $n_2 + n_3 \geq 2$,

$$\frac{1}{\alpha_1 n_1 n_2 n_3} = \frac{1}{\alpha_1 n_1 (n_2 + n_3)} \left(\frac{1}{n_2} + \frac{1}{n_3} \right) \leq \frac{1}{\alpha_1 n_1 \max\{2, v_1 + \alpha_1 n_1\}} \left(\frac{1}{n_2} + \frac{1}{n_3} \right).$$

The sum is then bounded by

$$\begin{aligned} \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ 1 \leq n_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ -\alpha_1 n_1 + n_2 + n_3 \in [v_1, v_2]}} \frac{1}{\alpha_1 n_1 n_2 n_3} &\leq \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \frac{1}{\alpha_1 n_1 \max\{v_1 + \alpha_1 n_1, 2\}} \sum_{\substack{1 \leq n_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ -\alpha_1 n_1 + n_2 + n_3 \in [v_1, v_2]}} \frac{1}{n_2} + \frac{1}{n_3} \\ &= \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \frac{1}{\alpha_1 n_1 \max\{v_1 + \alpha_1 n_1, 2\}} \sum_{\substack{1 \leq n_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ -\alpha_1 n_1 + n_2 + n_3 \in [v_1, v_2]}} \frac{2}{n_2}, \end{aligned}$$

where equality follows because the roles of n_2 and n_3 are symmetric. For fixed values of n_1 and n_2 , the integer n_3 must satisfy $1 \leq n_3 \leq h/2$ and $n_3 \in [v_1 + \alpha_1 n_1 - n_2, v_2 + \alpha_1 n_1 - n_2]$; the number of valid choices of n_3 is $\ll v_2 - v_1 + O(1)$. Thus the sum is

$$\begin{aligned} &\ll (v_2 - v_1 + 1) \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \frac{1}{\alpha_1 n_1 \max\{v_1 + \alpha_1 n_1, 2\}} \sum_{1 \leq n_2 \leq h/2} \frac{1}{n_2} \\ &\ll (v_2 - v_1 + 1) \log h \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \frac{1}{\alpha_1 n_1 \max\{v_1 + \alpha_1 n_1, 2\}}. \end{aligned}$$

If $v_1 \geq 0$, then $v_1/\alpha_1 + n_1 \geq 1$ and the sum is

$$\begin{aligned} &\ll (v_2 - v_1 + 1) \log h \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \frac{1}{\alpha_1 n_1 (v_1 + \alpha_1 n_1)} \\ &\ll (v_2 - v_1 + 1) \frac{\log h}{\alpha_1^2} \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \frac{1}{n_1 (v_1/\alpha_1 + n_1)} \ll (v_2 - v_1 + 1) \frac{\log h}{\alpha_1^2}, \end{aligned}$$

since the sum over n_1 is bounded by $\sum_{n=1}^{\infty} 1/n^2$, and thus by a constant. This completes the proof for this case.

On the other hand, if $\nu_1 < 0$, then the sum is

$$\begin{aligned} &\ll (\nu_2 - \nu_1 + 1) \log h \left(\sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ n_1 < (2-\nu_1)/\alpha_1 + 1}} \frac{1}{\alpha_1 n_1} + \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ \nu_1 + \alpha_1 n_1 \geq 2 + \alpha_1}} \frac{1}{\alpha_1 n_1 (\nu_1 + \alpha_1 n_1)} \right) \\ &\ll (\nu_2 - \nu_1 + 1) \log h \left(\frac{1}{\alpha_1} \left(\frac{2 - \nu_1}{\alpha_1} + 1 \right) + \frac{1}{\alpha_1^2} \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ \nu_1/\alpha_1 + n_1 \geq 2/\alpha_1 + 1}} \frac{1}{n_1 (\nu_1/\alpha_1 + n_1)} \right). \end{aligned}$$

The final sum is bounded by $\sum_{n=1}^{\infty} 1/n^2$, and thus by a constant. □

Lemma 2.9. *Let $\nu_2 \geq \nu_1 \geq 3$ and $\alpha_1 \geq 1$ be real numbers, and let $h \in \mathbb{N}$ with $h \geq 4$. Then*

$$\sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ 1 \leq n_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ \alpha_1 n_1 + n_2 + n_3 \in [\nu_1, \nu_2]}} \frac{1}{\alpha_1 n_1 n_2 n_3} \ll \frac{(\nu_2 - \nu_1 + 1)}{\nu_1} \log \min\{\nu_2, h\} \left(\nu_2 - \nu_1 + 1 + \frac{1}{\alpha_1} \log \min\{\nu_1, h\} \right),$$

where n_1, n_2 , and n_3 range over integers.

Proof. The first part of this proof follows along identical lines to those of Lemma 2.8, but with α_1 having opposite signs. By following the first part of the argument of Lemma 2.8, we get that the sum we want to bound is

$$\begin{aligned} &\ll (\nu_2 - \nu_1 + 1) \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ n_1 \leq (\nu_2 - 2)/\alpha_1}} \frac{1}{\alpha_1 n_1 \max\{\nu_1 - \alpha_1 n_1, 2\}} \sum_{\substack{1 \leq n_2 \leq h/2 \\ n_2 \leq \nu_2 - \alpha_1 n_1}} \frac{1}{n_2} \\ &\ll (\nu_2 - \nu_1 + 1) \log \min\{\nu_2, h\} \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ n_1 \leq (\nu_2 - 2)/\alpha_1}} \frac{1}{\alpha_1 n_1 \max\{\nu_1 - \alpha_1 n_1, 2\}}. \end{aligned}$$

If $\max\{\nu_1 - \alpha_1 n_1, 2\} = 2$, then $\nu_1 - 2 < \alpha_1 n_1 \leq \nu_2 - 2$. The number of such terms is $\ll \nu_2 - \nu_1$, and for these terms the summand is $1/(2\alpha_1 n_1) \ll 1/\nu_1$, so these terms provide an overall contribution of size $\ll (\nu_2 - \nu_1 + 1) \log \min\{\nu_2, h\} (\nu_2 - \nu_1)/\nu_1$. For the remaining terms, $\alpha_1 n_1 \leq \nu_1 - 2$.

We rewrite

$$\frac{1}{\alpha_1 n_1 (\nu_1 - \alpha_1 n_1)} = \frac{1}{\nu_1 \alpha_1 n_1} + \frac{1}{\nu_1 (\nu_1 - \alpha_1 n_1)},$$

so that for the remaining terms we have

$$\begin{aligned} \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ n_1 \leq (\nu_1 - 2)/\alpha_1}} \frac{1}{\alpha_1 n_1 \max\{\nu_1 - \alpha_1 n_1, 2\}} &= \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ n_1 \leq (\nu_1 - 2)/\alpha_1}} \left(\frac{1}{\nu_1 \alpha_1 n_1} + \frac{1}{\nu_1 (\nu_1 - \alpha_1 n_1)} \right) \\ &\ll \frac{1}{\nu_1 \alpha_1} \log \min\{\nu_1, h\} + \frac{1}{\nu_1} \left(1 + \frac{1}{\alpha_1} \log \min\{\nu_1, h\} \right). \quad \square \end{aligned}$$

Lemma 2.10. *Let $\alpha_1 \geq 1$ and $v_2 \geq v_1$ be (possibly negative) real numbers, and let $h \in \mathbb{N}$ with $h \geq 4$. Then*

$$\sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ 1 \leq n_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ \alpha_1 n_1 - n_2 + n_3 \in [v_1, v_2]}} \frac{1}{\alpha_1 n_1 n_2 n_3} \ll (v_2 - v_1 + 1) \left(\frac{\log h}{\alpha_1} + 1 \right) \frac{\log \max\{v_1, \alpha_1 + 1\} + 1}{\max\{v_1, \alpha_1\} + 1},$$

where n_1, n_2 , and n_3 range over integers.

Proof. Since $\alpha_1 n_1 + n_3 \geq v_1 + n_2$ and $\alpha_1 n_1 + n_3 \geq \alpha_1 + 1$, we have

$$\frac{1}{\alpha_1 n_1 n_2 n_3} = \frac{1}{n_2 (\alpha_1 n_1 + n_3)} \left(\frac{1}{\alpha_1 n_1} + \frac{1}{n_3} \right) \leq \frac{1}{n_2 \max\{v_1 + n_2, \alpha_1 + 1\}} \left(\frac{1}{\alpha_1 n_1} + \frac{1}{n_3} \right).$$

The sum is then bounded by

$$\sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ 1 \leq n_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ \alpha_1 n_1 - n_2 + n_3 \in [v_1, v_2]}} \frac{1}{\alpha_1 n_1 n_2 n_3} \leq \sum_{1 \leq n_2 \leq h/2} \frac{1}{n_2 \max\{v_1 + n_2, \alpha_1 + 1\}} \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ 1 \leq n_3 \leq h/2 \\ \alpha_1 n_1 - n_2 + n_3 \in [v_1, v_2]}} \left(\frac{1}{\alpha_1 n_1} + \frac{1}{n_3} \right).$$

For fixed values of n_1 and n_2 , the integer n_3 must satisfy $n_3 \in [v_1 - \alpha_1 n_1 + n_2, v_2 - \alpha_1 n_1 + n_2]$ and $1 \leq n_3 \leq h/2$; the number of valid choices of n_3 is $\ll v_2 - v_1 + O(1)$. Thus

$$\begin{aligned} \sum_{1 \leq n_2 \leq h/2} \frac{1}{n_2 \max\{v_1 + n_2, \alpha_1 + 1\}} \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ 1 \leq n_3 \leq h/2 \\ \alpha_1 n_1 - n_2 + n_3 \in [v_1, v_2]}} \frac{1}{\alpha_1 n_1} &\ll \frac{(v_2 - v_1 + 1)}{\alpha_1} \log h \sum_{1 \leq n_2 \leq h/2} \frac{1}{n_2 \max\{v_1 + n_2, \alpha_1 + 1\}} \\ &\ll (v_2 - v_1 + 1) \frac{\log h \log \max\{v_1, \alpha_1 + 1\} + 1}{\alpha_1 \max\{v_1, \alpha_1\} + 1}. \end{aligned}$$

It remains to evaluate the $1/n_3$ -term in the sum. Since $n_3 \geq v_1 - \alpha_1 n_1 + n_2$, we have

$$\begin{aligned} \sum_{1 \leq n_2 \leq h/2} \frac{1}{n_2 \max\{v_1 + n_2, \alpha_1 + 1\}} \sum_{\substack{1 \leq n_1 \leq h/(2\alpha_1) \\ 1 \leq n_3 \leq h/2 \\ \alpha_1 n_1 - n_2 + n_3 \in [v_1, v_2]}} \frac{1}{n_3} \\ &\ll \sum_{1 \leq n_2 \leq h/2} \frac{1}{n_2 \max\{v_1 + n_2, \alpha_1 + 1\}} \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \frac{v_2 - v_1 + 1}{\lceil v_1 - \alpha_1 n_1 + n_2 \rceil} \\ &\ll \sum_{1 \leq n_2 \leq h/2} \frac{v_2 - v_1 + 1}{n_2 \max\{v_1 + n_2, \alpha_1 + 1\}} \left(\frac{\log h}{\alpha_1} + 1 \right) \\ &\ll (v_2 - v_1 + 1) \left(\frac{\log h}{\alpha_1} + 1 \right) \frac{\log \max\{v_1, \alpha_1 + 1\} + 1}{\max\{v_1, \alpha_1\} + 1}. \quad \square \end{aligned}$$

If $\alpha_1 = 1$, we have the following stronger bound.

Lemma 2.11. *There exist absolute constants C and D such that, for all integers $v \geq 3$ and $h \geq 4$,*

$$\sum_{\substack{1 \leq n_1 \leq v-2 \\ 1 \leq n_2 \leq v-2 \\ 1 \leq n_3 \leq v-2 \\ n_1+n_2+n_3=v}} \frac{1}{n_1 n_2 n_3} \leq C \quad \text{and} \quad \sum_{\substack{1 \leq n_1 \leq h \\ 1 \leq n_2 \leq v+h \\ 1 \leq n_3 \leq v+h \\ n_2+n_3=v+n_1}} \frac{1}{n_1 n_2 n_3} \leq D,$$

where the sum ranges over integer values of n_1, n_2, n_3 .

Proof. For real numbers $x, x' \geq 1$ with $|x - x'| \leq 1$, we have $|1/x - 1/x'| \leq 2/x$. Thus

$$\begin{aligned} \sum_{\substack{1 \leq n_1 \leq h/2 \\ 1 \leq n_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ n_1+n_2+n_3=v}} \frac{1}{n_1 n_2 n_3} &\leq 8 \int_1^{v-2} \int_1^{v-x_1-1} \frac{1}{x_1 x_2 (v-x_1-x_2)} dx_2 dx_1 \\ &= 8 \int_1^{v-2} \frac{2 \ln(v-x_1-1)}{x_1(v-x_1)} dx_1 \\ &\leq 16 \ln v \int_1^{v-2} \frac{1}{x_1(v-x_1)} dx_1 \\ &= 16 \ln v \frac{2 \ln(v-1)}{v} = 32 \frac{(\ln v)(\ln(v-1))}{v}. \end{aligned}$$

The function $(\ln v)(\ln(v-1))/v$ has a global maximum M ; setting $C = 16M$ completes the proof of the first claim.

For the second claim, we similarly have

$$\begin{aligned} \sum_{\substack{1 \leq n_1 \leq h \\ 1 \leq n_2 \leq v+h \\ 1 \leq n_3 \leq v+h \\ n_2+n_3=v+n_1}} \frac{1}{n_1 n_2 n_3} &\leq 8 \int_1^h \int_1^{v+x_1-1} \frac{1}{x_1 x_2 (v+x_1-x_2)} dx_2 dx_1 \\ &= 16 \int_1^h \frac{\ln(v+x_1-1)}{x_1(v+x_1)} dx_1 \\ &\leq 16D_1 + 16 \int_{10}^h \frac{\ln(x_1-1)}{x_1^2} dx_1 \end{aligned}$$

for some constant D_1 , since $\ln(x-1)/x$ is decreasing for $x \geq 10$. The integral converges to a constant as $h \rightarrow \infty$, so setting $D = 16D_1 + 16 \int_{10}^{\infty} \ln(x-1)/x^2 dx$ completes the proof. \square

The next two lemmas concern triple sums over $\|a/q\|^{-1}$, which arise because of their role in the definition of $F(\alpha)$ and make use of the previous four lemmas.

Lemma 2.12. *Fix an integer $h \geq 4$. Then*

$$\sum_{\substack{1 \leq n_i \leq h-1 \\ \|\sum_i n_i/h\| \leq 3/h}} \left\| \frac{n_1}{h} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} \ll h^3,$$

where n_1, n_2 , and n_3 range over integers.

Proof. We will split into cases based on whether $n_i \leq h/2$ or $n_i > h/2$, i.e., based on the value of $\|n_i/h\|$.

Assume first that $1 \leq n_i \leq h/2$ for all $i = 1, 2, 3$. Then $\|n_i/h\| = n_i/h$, so we have

$$\sum_{\substack{1 \leq n_i \leq h/2 \\ \|\sum_i n_i/h\| \leq 3/h}} \left\| \frac{n_1}{h} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} = h^3 \sum_{\substack{1 \leq n_i \leq h/2 \\ \|\sum_i n_i/h\| \leq 3/h}} \frac{1}{n_1 n_2 n_3}.$$

To satisfy $\|\sum_i n_i/h\| \leq 3/h$, we must have $n_1 + n_2 + n_3 \in \{3\} \cup [h-3, h+3] \cup [2h-3, 2h+3] \cup \{3h-3\}$. There are finitely many possible integer values for $n_1 + n_2 + n_3$; for each one, by Lemma 2.11, the sum over $1/(n_1 n_2 n_3)$ is bounded by an absolute constant. Thus the lemma holds in this case.

Now consider terms where $h/2 < n_i \leq h-1$ for all i . For each i , define $m_i = h - n_i$, so that $1 \leq m_i \leq h/2$. Then $\|n_i/h\| = m_i/h$, and

$$\left\| \sum_i \frac{n_i}{h} \right\| = \left\| 3h - \sum_i \frac{m_i}{h} \right\| = \left\| \sum_i \frac{m_i}{h} \right\|.$$

Then

$$\sum_{\substack{h/2 < n_i \leq h-1 \\ \|\sum_i n_i/h\| \leq 3/h}} \left\| \frac{n_1}{h} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} \ll \sum_{\substack{1 \leq m_i \leq h/2 \\ \|\sum_i m_i/h\| \leq 3/h}} \left\| \frac{m_1}{h} \right\|^{-1} \left\| \frac{m_2}{h} \right\|^{-1} \left\| \frac{m_3}{h} \right\|^{-1},$$

which is precisely the previous case, since $1 \leq m_i \leq h/2$ for all i . Thus this case is also $\ll h^3$.

Finally consider terms where, for some i , $n_i \in [1, h/2]$, whereas for others $n_i \in (h/2, h-1]$. As in the previous paragraph, we can always flip all three n_i 's with $h - n_i$. Moreover, the roles of n_1, n_2 , and n_3 are entirely symmetric. Thus it suffices to bound those terms where $n_2, n_3 \in [1, h/2]$ and $n_1 \in (h/2, h-1]$. Set $m_1 = h - n_1$. Then

$$\sum_{\substack{h/2 < n_1 \leq h-1 \\ 1 \leq n_2, n_3 \leq h/2 \\ \|\sum_i n_i/h\| \leq 3/h}} \left\| \frac{n_1}{h} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} = h^3 \sum_{\substack{1 \leq m_1 \leq h/2 \\ 1 \leq n_2, n_3 \leq h/2 \\ \|-m_1/h + n_2/h + n_3/h\| \leq 3/h}} \frac{1}{m_1 n_2 n_3}.$$

Just as before, there are finitely many possible integer values for $-m_1 + n_2 + n_3$ satisfying the constraint that $\|\sum_i n_i/h\| \leq 3/h$. For each value v , by Lemma 2.11, the sum

$$\sum_{\substack{1 \leq m_1 \leq h/2 \\ 1 \leq n_2, n_3 \leq h/2 \\ -m_1 + n_2 + n_3 = v}} \frac{1}{m_1 n_2 n_3}$$

is bounded by a constant, which completes the proof. □

Lemma 2.13. *Let $h \geq 4$ and $1 \leq q_1 < h$ be integers. Then*

$$\sum_{\substack{1 \leq n_1 \leq q_1 - 1 \\ 1 \leq n_2, n_3 \leq h - 1 \\ \|n_1/q_1 + n_2/h + n_3/h\| \leq 3/h}} \left\| \frac{n_1}{q_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} \ll h^2 q_1 (\log h),$$

where n_1, n_2 , and n_3 range over integers.

Proof. We will split into cases based on whether each of n_1/q_1 , n_2/h , and n_3/h lie in $(0, \frac{1}{2}]$ or $(\frac{1}{2}, 1)$; for each case, we will show that the bound holds. Assume first that all three of n_1/q_1 , n_2/h , and n_3/h lie in $(0, \frac{1}{2}]$. Note that

$$\frac{n_1}{q_1} + \frac{n_2}{h} + \frac{n_3}{h} \geq \frac{1}{q_1} + \frac{2}{h} > \frac{3}{h},$$

so the constraint that $\|n_1/q_1 + n_2/h + n_3/h\| \leq 3/h$ is equivalent to the constraint that

$$\begin{aligned} \frac{n_1}{q_1} + \frac{n_2}{h} + \frac{n_3}{h} &\in \left[1 - \frac{3}{h}, 1 + \frac{3}{h}\right] \cup \left[2 - \frac{3}{h}, 2 + \frac{3}{h}\right] \cup \left[3 - \frac{3}{h}, 3\right] \\ &\iff \frac{h}{\tilde{q}_1} n_1 + n_2 + n_3 \in [h - 3, h + 3] \cup [2h - 3, 2h + 3] \cup [3h - 3, 3h]. \end{aligned}$$

These are finitely many intervals, each of bounded size. Thus these terms are given by

$$\sum_{\substack{1 \leq n_1 \leq q_1/2 \\ 1 \leq n_2, n_3 \leq h/2 \\ \|n_1/q_1 + n_2/h + n_3/h\| \leq 3/h}} \frac{q_1 h^2}{n_1 n_2 n_3} = \sum_{\substack{[v_1, v_2] \in \{[h-3, h+3], \\ [2h-3, 2h+3], [3h-3, 3h]\}}} \sum_{\substack{1 \leq n_1 \leq q_1/2 \\ 1 \leq n_2, n_3 \leq h/2 \\ (h/q_1)n_1 + n_2 + n_3 \in [v_1, v_2]}} \frac{h^3}{(h/q_1)n_1 n_2 n_3}.$$

We apply Lemma 2.9, with $\alpha_1 = h/q_1$ and $[v_1, v_2] = [h - 3, h + 3]$, $[2h - 3, 2h + 3]$, or $[3h - 3, 3h]$, respectively. By Lemma 2.9, each of these three terms is

$$\ll h^3 \frac{1}{h} \log h \left(1 + \frac{\log h}{\alpha_1}\right) \ll h^2 \log h \left(1 + \frac{q_1 \log h}{h}\right),$$

which is $\ll h^2 q_1 \log h$, as desired.

Now assume that all three of n_1/q_1 , n_2/h , and n_3/h lie in $(\frac{1}{2}, 1)$. Define $m_1 = q_1 - n_1$, $m_2 = h - n_2$, and $m_3 = h - n_3$, so that

$$\sum_{\substack{q_1/2 < n_1 \leq q_1 - 1 \\ h/2 < n_2, n_3 \leq h - 1 \\ \|n_1/q_1 + n_2/h + n_3/h\| \leq 3/h}} \left\| \frac{n_1}{q_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} = \sum_{\substack{1 \leq m_1 \leq q_1/2 \\ 1 \leq m_2, m_3 \leq h/2 \\ \|m_1/q_1 + m_2/h + m_3/h\| \leq 3/h}} \frac{h^3}{(h/q_1)m_1 m_2 m_3}.$$

This is identical to the previous case, which we have already shown to be $\ll h^2 q_1 \log h$.

We now tackle the cases where not all fractions lie in the same half of $(0, 1)$. Assume that $n_1/q_1 \in (\frac{1}{2}, 1)$ but $n_2/h, n_3/h \in (0, \frac{1}{2}]$. Define $m_1 = q_1 - n_1$, so that

$$\sum_{\substack{q_1/2 < n_1 \leq q_1 - 1 \\ 1 \leq n_2, n_3 \leq h/2 \\ \|n_1/q_1 + n_2/h + n_3/h\| \leq 3/h}} \left\| \frac{n_1}{q_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} = \sum_{\substack{1 \leq m_1 \leq q_1/2 \\ 1 \leq n_2, n_3 \leq h/2 \\ \|-m_1/q_1 + n_2/h + n_3/h\| \leq 3/h}} \frac{h^3}{(h/q_1)m_1 n_2 n_3}.$$

The constraint that

$$\left\| -\frac{m_1}{q_1} + \frac{n_2}{h} + \frac{n_3}{h} \right\| \leq \frac{3}{h}$$

is equivalent to the constraint that $-h/q_1 m_1 + n_2 + n_3$ lies in one of the intervals $[-3, 3]$ or $[h - 3, h + 3]$. Applying Lemma 2.8 to the sum over m_1, n_2, n_3 , with $\alpha_1 = h/q_1$ and $[v_1, v_2]$ equal to each of these intervals respectively, we get that

$$\sum_{\substack{q_1/2 < n_1 \leq q_1 - 1 \\ 1 \leq n_2, n_3 \leq h/2 \\ \|n_1/q_1 + n_2/h + n_3/h\| \leq 3/h}} \left\| \frac{n_1}{q_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} \ll h^3 \frac{\log h}{(h/q_1)} \left(1 + \frac{1}{(h/q_1)} \right) \ll h^2 q_1 \log h.$$

If $n_1/q_1 \in (0, \frac{1}{2}]$ but $n_2/h, n_3/h \in (\frac{1}{2}, 1)$, then we can once again replace n_1 by $m_1 = q_1 - n_1$, n_2 by $m_2 = h - n_2$, and n_3 by $m_3 = h - n_3$ to revert to the previous case.

Finally assume that $n_1/q_1 \in (0, \frac{1}{2}]$, $n_2/h \in (\frac{1}{2}, 1)$, and $n_3/h \in (0, \frac{1}{2}]$. The roles of n_2 and n_3 are symmetric, and we can always replace all three n_i 's by the corresponding m_i -value, so this is the only remaining case.

Define $m_2 = h - n_2$, so that

$$\sum_{\substack{1 \leq n_1 \leq q_1/2 \\ h/2 < n_2 \leq h-1 \\ 1 \leq n_3 \leq h/2 \\ \|n_1/q_1 + n_2/h + n_3/h\| \leq 3/h}} \left\| \frac{n_1}{q_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} = \sum_{\substack{1 \leq n_1 \leq q_1/2 \\ 1 \leq m_2 \leq h/2 \\ 1 \leq n_3 \leq h/2 \\ \|n_1/q_1 - m_2/h + n_3/h\| \leq 3/h}} \frac{h^3}{(h/q_1)n_1 m_2 n_3}.$$

The constraint that $\|n_1/q_1 - m_2/h + n_3/h\| \leq 3h$ is equivalent to the constraint that $-(h/q_1)n_1 - m_2 + m_3$ lies in one of the intervals $[-3, 3]$ or $[h - 3, h + 3]$. Applying Lemma 2.10 to the sum over n_1, m_2, n_3 with $\alpha_1 = h/q_1$ and $[v_1, v_2]$ equal to each of these intervals respectively, we get that

$$\sum_{\substack{1 \leq n_1 \leq q_1/2 \\ h/2 < n_2 \leq h-1 \\ 1 \leq n_3 \leq h/2 \\ \|n_1/q_1 + n_2/h + n_3/h\| \leq 3/h}} \left\| \frac{n_1}{q_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} \ll h^3 \left(\frac{q_1 \log h}{h} + 1 \right) \frac{q_1 \log(h/q_1 + 1)}{h}.$$

Since $\log x/x$ is uniformly bounded for $x \geq 1$, we have $q_1/h \log(h/q_1) \ll 1$, so these terms are also $\ll h^2 q_1 \log h$, which completes the proof. \square

Finally, the following lemma directly bounds a sum over triple products of $F(a_i/q_i)$.

Lemma 2.14. *Let $h \in \mathbb{N}$ with $h \geq 4$ and let $d_1 \geq 1$ and $d_2 \geq 2$ be positive integers with $d_1 | d_2$ and $d_2 < h$. Then*

$$\sum_{\substack{1 \leq n_1 < d_1 \\ 1 \leq n_2 < d_2}} F\left(\frac{n_1}{d_1}\right) F\left(\frac{n_2}{d_2}\right) F\left(\frac{n_1}{d_1} - \frac{n_2}{d_2}\right) \ll h d_1^2 + d_1^2 d_2 \log d_2,$$

where n_1 and n_2 range over integers.

Proof. Write $f := d_2/d_1$. Then

$$\frac{n_1}{d_1} - \frac{n_2}{d_2} = \frac{f n_1 - n_2}{d_2}.$$

Since $d_2 < h$,

$$F\left(\frac{fn_1 - n_2}{d_2}\right) = \left\| \frac{fn_1 - n_2}{d_2} \right\|^{-1}$$

unless $fn_1 - n_2 = 0$. Moreover, in the range where $1 \leq n_1 < d_1$ and $1 \leq n_2 < d_2$,

$$F\left(\frac{n_1}{d_1}\right) = \left\| \frac{n_1}{d_1} \right\|^{-1} \quad \text{and} \quad F\left(\frac{n_2}{d_2}\right) = \left\| \frac{n_2}{d_2} \right\|^{-1}.$$

Thus

$$\begin{aligned} \sum_{\substack{1 \leq n_1 < d_1 \\ 1 \leq n_2 < d_2}} F\left(\frac{n_1}{d_1}\right) F\left(\frac{n_2}{d_2}\right) F\left(\frac{n_1}{d_1} - \frac{n_2}{d_2}\right) \\ = \sum_{\substack{1 \leq n_1 < d_1 \\ 1 \leq n_2 < d_2 \\ fn_1 = n_2}} h \left\| \frac{n_1}{d_1} \right\|^{-1} \left\| \frac{n_2}{d_2} \right\|^{-1} + \sum_{\substack{1 \leq n_1 < d_1 \\ 1 \leq n_2 < d_2 \\ fn_1 \neq n_2}} \left\| \frac{n_1}{d_1} \right\|^{-1} \left\| \frac{n_2}{d_2} \right\|^{-1} \left\| \frac{fn_1 - n_2}{d_2} \right\|^{-1}. \end{aligned}$$

The first sum is bounded by

$$\sum_{\substack{1 \leq n_1 < d_1 \\ 1 \leq n_2 < d_2 \\ fn_1 = n_2}} h \left\| \frac{n_1}{d_1} \right\|^{-1} \left\| \frac{n_2}{d_2} \right\|^{-1} = h \sum_{1 \leq n_1 < d_1} \left\| \frac{n_1}{d_1} \right\|^{-2} \leq 2hd_1^2 \sum_{1 \leq n_1 \leq d_1/2} \frac{1}{n_1^2} \ll hd_1^2.$$

It remains to bound the second sum. As in the proofs of Lemmas 2.12 and 2.13, we will split into cases based on whether n_1/d_1 and n_2/d_2 are in $(0, \frac{1}{2}]$ or $(\frac{1}{2}, 1)$.

Assume first that both n_1/d_1 and n_2/d_2 are in $(0, \frac{1}{2}]$, or that both n_1/d_1 and n_2/d_2 are in $(\frac{1}{2}, 1)$. In the latter case, we can substitute $m_1 = d_1 - n_1$ and $m_2 = d_2 - n_2$ to revert precisely to the former case, so it suffices to assume that both n_1/d_1 and n_2/d_2 are in $(0, \frac{1}{2}]$. Then

$$\sum_{\substack{1 \leq n_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ fn_1 \neq n_2}} \frac{d_1 d_2}{n_1 n_2} \left\| \frac{fn_1 - n_2}{d_2} \right\|^{-1} = \sum_{\substack{1 \leq n_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ fn_1 > n_2}} \frac{d_1 d_2^2}{n_1 n_2 (fn_1 - n_2)} + \sum_{\substack{1 \leq n_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ fn_1 < n_2}} \frac{d_1 d_2^2}{n_1 n_2 (n_2 - fn_1)}.$$

By applying Lemma 2.8 with $\alpha_1 = f$ and $\nu_1 = \nu_2 = 0$, the first sum is bounded by $\ll d_2^3 \log d_2 / f^2 = d_1^2 d_2 \log d_2$. For the second sum, we can achieve a bound that is somewhat stronger than the bound furnished by Lemma 2.10 in this special case. Specifically we have, writing $n_3 = n_2 - fn_1$,

$$\begin{aligned} d_2^3 \sum_{\substack{1 \leq n_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ 1 \leq n_3 \leq d_2/2 \\ fn_1 - n_2 + n_3 = 0}} \frac{1}{fn_1 n_2 n_3} &= d_2^3 \sum_{1 \leq n_1 \leq d_1/2} \frac{1}{fn_1} \sum_{\substack{fn_1 \leq n_2 \leq d_2/2 \\ 1 \leq n_3 \leq d_2/2 \\ fn_1 + n_3 = n_2}} \frac{1}{n_2 - n_3} \left(\frac{1}{n_3} - \frac{1}{n_2} \right) \\ &= d_2^3 \sum_{1 \leq n_1 \leq d_1/2} \frac{1}{(fn_1)^2} \sum_{1 \leq n_3 \leq d_2/2 - fn_1} \left(\frac{1}{n_3} - \frac{1}{n_3 + fn_1} \right) \\ &\ll d_2^3 \sum_{1 \leq n_1 \leq d_1/2} \frac{1}{(fn_1)^2} \log d_2 \ll d_2^3 \frac{\log d_2}{f^2} = d_1^2 d_2 \log d_2. \end{aligned}$$

Thus in this case, the second sum is $\ll d_1^2 d_2 \log d_2$.

Now assume that $n_1/d_1 \in (\frac{1}{2}, 1)$ but $n_2/d_2 \in (0, \frac{1}{2}]$; by swapping both n_i 's with $m_i = d_i - n_i$, this is the same as the case that $n_1/d_1 \in (0, \frac{1}{2}]$ but $n_2/d_2 \in (\frac{1}{2}, 1)$, so it is our only remaining case.

On substituting $m_1 = d_1 - n_1$, the sum in this case becomes

$$\sum_{\substack{1 \leq m_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ fm_1+n_2 < d_2}} \frac{d_1 d_2}{m_1 n_2} \left\| \frac{fm_1 + n_2}{d_2} \right\|^{-1} = \sum_{\substack{1 \leq m_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ fm_1+n_2 \leq d_2/2}} \frac{d_1 d_2^2}{m_1 n_2 (fm_1 + n_2)} + \sum_{\substack{1 \leq m_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ d_2/2 < fm_1+n_2 < d_2}} \frac{d_1 d_2^2}{m_1 n_2 (d_2 - n_2 - fm_1)}.$$

The first sum is

$$\leq d_1 d_2^2 \sum_{\substack{1 \leq m_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ fm_1+n_2 < d_2}} \frac{1}{fm_1^2 n_2} \ll \frac{d_1 d_2^2}{f} \log d_2 = d_1^2 d_2 \log d_2.$$

As for the second sum, setting $n_3 = d_2 - n_2 - fm_1$, we can bound it by applying Lemma 2.9, where $\alpha_1 = f$ and $\nu_1 = \nu_2 = d_2$, to get that

$$d_2^3 \sum_{\substack{1 \leq m_1 \leq d_1/2 \\ 1 \leq n_2 \leq d_2/2 \\ 1 \leq n_3 \leq d_2/2 \\ fm_1+n_2+n_3=d_2}} \frac{1}{fm_1 n_2 n_3} \ll d_2^3 \frac{1}{d_2} \log d_2 \left(1 + \frac{\log d_2}{f}\right) \ll d_2^2 \log d_2 + d_1 d_2 \log d_2,$$

both of which are $\ll d_1^2 d_2 \log d_2$. □

2.2. Bounding T_1 : terms with $gx \geq h$. Define

$$T_1 = \sum_{\substack{g,x,y,z | q \\ gx \geq h}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{\substack{a_1, a_2, a_3 \\ (a_1, gyz) = \dots = 1 \\ a_1/gyz + \dots \in \mathbb{Z}}} F\left(\frac{a_1}{gyz}\right) F\left(\frac{a_2}{gxz}\right) F\left(\frac{a_3}{gxy}\right). \tag{10}$$

For these terms, the rough argument that “the probability that each of a_2/q_2 and a_3/q_3 are sufficiently small is about $1/h$, making the size of the sum $h^{1+\varepsilon}$ instead of $h^{3+\varepsilon}$ ” can be made precise, although some of the counting arguments are rather involved, and rely on the lemmas of the previous section. Nevertheless, we will use this basic idea to prove the following bound.

Lemma 2.15. *Let $h \geq 4$, let q be the product of primes $p \leq h^4$, and define T_1 by (10). Then*

$$T_1 \ll h(\log h)^5.$$

Proof. Recall that $q_1 = gyz$, $q_2 = gxz$, and $q_3 = gxy$. Since $gx \geq h$, gxy and gxz (i.e., q_2 and q_3) must also both be $\geq h$. Recall the notation that $\tilde{q}_i = \min\{q_i, h\}$, so that $\tilde{q}_2 = \tilde{q}_3 = h$.

Since $a_1/q_1 + a_2/q_2 + a_3/q_3 \in \mathbb{Z}$, the sum

$$\frac{n(a_1, q_1)}{\tilde{q}_1} + \frac{n(a_2, q_2)}{\tilde{q}_2} + \frac{n(a_3, q_3)}{\tilde{q}_3}$$

satisfies

$$\left\| \frac{n(a_1, q_1)}{\tilde{q}_1} + \frac{n(a_2, q_2)}{\tilde{q}_2} + \frac{n(a_3, q_3)}{\tilde{q}_3} \right\| \leq \left\| \frac{a_1}{q_1} + \frac{a_2}{q_2} + \frac{a_3}{q_3} \right\| + \sum_{i=1}^3 \left\| \frac{n(a_i, q_i)}{\tilde{q}_i} - \frac{a_i}{q_i} \right\| \leq \frac{3}{h},$$

since $|a/q - n(a, q)/\tilde{q}| < 1/h$ always. We can then bound the sum by replacing the fractions a_i/q_i by their h -approximations $n(a_i, q_i)/q_i$. Precisely, we have

$$\begin{aligned} T_1 &= \sum_{\substack{g,x,y,z|q \\ gx \geq h}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{\substack{a_1, a_2, a_3 \\ (a_i, q_i)=1 \\ \sum_i a_i/q_i \in \mathbb{Z}}} F\left(\frac{a_1}{q_1}\right) F\left(\frac{a_2}{q_2}\right) F\left(\frac{a_3}{q_3}\right) \\ &\ll \sum_{\substack{g,x,y,z|q \\ gx \geq h}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{\substack{a_1, a_2, a_3 \\ (a_i, q_i)=1 \\ \sum_i a_i/q_i \in \mathbb{Z}}} \left\| \frac{n(a_1, q_1)}{\tilde{q}_1} \right\|^{-1} \left\| \frac{n(a_2, q_2)}{\tilde{q}_2} \right\|^{-1} \left\| \frac{n(a_3, q_3)}{\tilde{q}_3} \right\|^{-1} \\ &\ll \sum_{\substack{g,x,y,z|q \\ gx \geq h}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{\substack{1 \leq n_1, n_2, n_3 \leq \tilde{q}_i - 1 \\ \|\sum_i n_i/\tilde{q}_i\| \leq 3/h}} \left\| \frac{n_1}{\tilde{q}_1} \right\|^{-1} \left\| \frac{n_2}{\tilde{q}_2} \right\|^{-1} \left\| \frac{n_3}{\tilde{q}_3} \right\|^{-1} \sum_{\substack{a_1, a_2, a_3 \\ (a_i, q_i)=1 \\ \sum_i a_i/q_i \in \mathbb{Z} \\ n(a_i, q_i)=n_i}} 1. \end{aligned}$$

The inside sum is the number of triplets a_1, a_2, a_3 with $n(a_i, q_i) = n_i$ for all i , $(a_i, q_i) = 1$, and $\sum_i a_i/q_i \in \mathbb{Z}$. The constraint that $n(a_i, q_i) = n_i$ implies that each a_i lies in an interval of length $\ll q_i/h + 1$; that is, for $q_i \geq h$, $(q_i/h)n_i \leq a_i \leq (q_i/h)(n_i + 1)$.

The constraint that $\sum_i a_i/q_i \in \mathbb{Z}$, after multiplying out denominators, is equivalent to the constraint

$$a_1x + a_2y + a_3z \equiv 0 \pmod{gxyz}. \tag{11}$$

Once the q_i 's (or equivalently g, x, y , and z) are fixed, there are $\ll q_1/h + 1$ choices of a_1 such that $n(a_1, q_1) = n_1$. Once a_1 is fixed, a_2 is determined mod z by (11). Since $1 \leq a_2 \leq gxz$, fixing a_2 is equivalent to choosing a congruence class mod gx for a_2 ; there are $\ll gx/h + 1$ choices of this congruence class such that a_2 lies within the interval where $n(a_2, q_2) = n_2$. Since $gx \geq h$ by assumption, $gx/h + 1 \ll gx/h$. Once a_1 and a_2 have been fixed, a_3 is entirely determined by (11). Thus the total number of triplets a_1, a_2, a_3 satisfying all constraints is $\ll (q_1/h + 1)(gx/h)$.

Thus T_1 is bounded by

$$T_1 \ll \sum_{\substack{g,x,y,z|q \\ gx \geq h}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \left(\frac{q_1}{h} + 1\right) \frac{gx}{h} \sum_{\substack{1 \leq n_i \leq \tilde{q}_i - 1 \\ \|\sum_i n_i/\tilde{q}_i\| \leq 3/h}} \left\| \frac{n_1}{\tilde{q}_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1}.$$

Consider first those terms where $\tilde{q}_1 = h$. Thus $q_1/h \gg 1$, and by Lemma 2.12, the inside sum is $\ll h^3$. This implies that the terms with $\tilde{q}_1 = h$ are bounded by

$$\begin{aligned} &\ll \sum_{\substack{g,x,y,z|q \\ gx \geq h}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \frac{q_1}{h} \frac{gx}{h} h^3 \\ &\ll h \sum_{\substack{g,x,y,z|q \\ gx \geq h}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} g^2 xyz \quad (\text{since } q_1 = gyz). \end{aligned}$$

Recalling that q is the product of all primes $p \leq h^4$, this sum is

$$\ll h \prod_{p \leq h^4} \left(1 + \frac{p^2}{(p-1)^3} + \frac{3p}{(p-1)^2} \right) \ll h(\log h)^4.$$

The remaining terms are those where $\tilde{q}_1 = q_1 < h$. By applying Lemma 2.13 to the inside sum, the terms with $\tilde{q}_1 = q_1 < h$ are bounded by

$$\begin{aligned} &\ll \sum_{\substack{g,x,y,z|q \\ gx \geq h}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \frac{gx}{h} (h^2 q_1 \log h) \\ &\ll h \log h \sum_{\substack{g,x,y,z|q \\ gx \geq h}} \frac{\mu(gxyz)^2 g^2 xyz}{\phi(g)^3 \phi(xyz)^2} \quad (\text{since } q_1 = gyz) \\ &\ll h(\log h) \prod_{p \leq h^4} \left(1 + \frac{p^2}{(p-1)^3} + \frac{3p}{(p-1)^2} \right) \ll h(\log h)^5. \end{aligned}$$

Thus $T_1 \ll h(\log h)^4 + h(\log h)^5 \ll h(\log h)^5$, as desired. □

2.3. Bounding T_2 : terms with gx, gy, gz small and a_2, a_3 large. We now consider T_2 , which is the sum of terms in (8) where $gx, gy,$ and gz are all $< h$ and $\|a_2/gxz\| \geq 1/h$, and $\|a_3/(gxy)\| \geq 1/h$. That is, define

$$T_2 := \sum_{\substack{g,x,y,z|q \\ x,y,z < h/g}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{\substack{a_1, a_2, a_3 \\ (a_1, gyz) = \dots = 1 \\ a_1/gyz + \dots \in \mathbb{Z} \\ \|a_2/gxz\| \geq 1/h \\ \|a_3/gxy\| \geq 1/h}} F\left(\frac{a_1}{gyz}\right) F\left(\frac{a_2}{gxz}\right) F\left(\frac{a_3}{gxy}\right). \tag{12}$$

The strategy for bounding T_2 is very different from that used to bound T_1 . Intuitively, since the fractions $a_2/(gxz)$ and $a_3/(gxy)$ are far from an integer, we are now considering terms where the values of $F(a_2/(gxz))$ and $F(a_3/(gxy))$ are relatively small, except perhaps at the boundary where $a_2/(gxz)$ and $a_3/(gxy)$ are very close to $1/h$. Since the denominators are loosely constrained to be small, there cannot be too many points on this boundary. We will prove a precise bound in the following lemma.

Lemma 2.16. *Let $h \geq 4$, let q be the product of primes $p \leq h^4$, and let T_2 be defined as in (12). Then*

$$T_2 \ll h(\log h)^4 (\log \log h)^2.$$

Proof. We begin by reparametrizing the sum in (12) over a_1, a_2, a_3 . For fixed g, x, y, z and fixed a_1, a_2, a_3 satisfying the constraints of the sums in (12), we will fix parameters a, b, c as follows. By the Chinese remainder theorem, and since $g, x,$ and y are pairwise relatively prime, there exist unique values $1 \leq a \leq x$ and $1 \leq b \leq gy$ such that

$$\frac{a_3}{gxy} \equiv \frac{a}{x} - \frac{b}{gy} \pmod{1}.$$

Similarly, there exist unique values $1 \leq a' \leq x$ and $1 \leq c \leq gz$ such that

$$\frac{a_2}{gxz} \equiv \frac{c}{gz} - \frac{a'}{x} \pmod{1}.$$

Since $a_1/(gyz) + a_2/(gxz) + a_3/(gxy) \in \mathbb{Z}$, we have

$$gyz \left(\frac{a_2}{gxz} + \frac{a_3}{gxy} \right) \in \mathbb{Z} \implies gyz \left(\frac{a}{x} - \frac{b}{gy} + \frac{c}{gz} - \frac{a'}{x} \right) \in \mathbb{Z} \implies gyz \frac{(a - a')}{x} \in \mathbb{Z}.$$

Since $(gyz, x) = 1$, this implies $x|(a - a')$; thus $a = a'$. Finally, $a_1/(gyz) + a_2/(gxz) + a_3/(gxy) \in \mathbb{Z}$ implies that

$$\frac{a_1}{gyz} \equiv -\frac{a_2}{gxz} - \frac{a_3}{gxy} \equiv \frac{b}{gy} - \frac{c}{gz} \pmod{1},$$

so that the triple a_1, a_2, a_3 uniquely determines (and is uniquely determined by) a triple a, b, c with $1 \leq a \leq x$, $1 \leq b \leq gy$, and $1 \leq c \leq gz$ such that

$$\frac{a_1}{gyz} \equiv \frac{b}{gy} - \frac{c}{gz} \pmod{1}, \quad \frac{a_2}{gxz} \equiv \frac{c}{gz} - \frac{a}{x} \pmod{1}, \quad \text{and} \quad \frac{a_3}{gxy} \equiv \frac{a}{x} - \frac{b}{gy} \pmod{1}.$$

Upon moving the sums over y and z in (12) inside, we get

$$T_2 = \sum_{\substack{g,x|q \\ x < h/g}} \frac{\mu(gx)^2}{\phi(g)^3 \phi(x)^2} \sum_{\substack{a \\ (a,x)=1}} S_2(g, x, a),$$

where $S_2(g, x, a)$ denotes the sum

$$S_2(g, x, a) = \sum_{\substack{y,z|q \\ y,z < h/g}} \frac{\mu(gxyz)^2}{\phi(yz)^2} \sum_{\substack{b,c \\ (b,gy)=(c,gz)=1 \\ \|c/gz - a/x\| \geq 1/h \\ \|a/x - b/gy\| \geq 1/h}} F\left(\frac{a}{x} - \frac{b}{gy}\right) F\left(\frac{b}{gy} - \frac{c}{gz}\right) F\left(\frac{c}{gz} - \frac{a}{x}\right). \tag{13}$$

Since $gy < h$ and $gz < h$, the product yz is less than h^2 , so that

$$\frac{yz}{\phi(yz)} \ll \log \log(h^2) \ll \log \log h.$$

Thus we can replace the expression $1/\phi(yz)^2$ in (13) with $(\log \log h)^2/(y^2 z^2)$.

Let ℓ and m be such that $2^\ell < y \leq 2^{\ell+1}$ and $2^m < z \leq 2^{m+1}$, and further define n_ℓ and n_m to be variables ranging from 1 to $g2^\ell$ and 1 to $g2^m$ respectively.

If

$$\frac{n_\ell}{g2^{\ell+1}} \leq \left\| \frac{a}{x} - \frac{b}{gy} \right\| \leq \frac{n_\ell + 1}{g2^{\ell+1}},$$

then

$$F\left(\frac{a}{x} - \frac{b}{gy}\right) \ll F\left(\frac{n_\ell}{g2^{\ell+1}}\right);$$

crucially, this upper bound depends only on ℓ and n_ℓ , and does not depend on b or y . Similarly, if

$$\frac{n_m}{g2^{m+1}} \leq \left\| \frac{c}{gz} - \frac{a}{x} \right\| \leq \frac{n_m + 1}{g2^{m+1}},$$

then

$$F\left(\frac{c}{gz} - \frac{a}{x}\right) \ll F\left(\frac{n_m}{g2^{m+1}}\right).$$

Because of the assumption that $\|a/x - b/(gy)\| \geq 1/h$, the constraint

$$\frac{n_\ell}{g2^{\ell+1}} \leq \left\| \frac{a}{x} - \frac{b}{gy} \right\| \leq \frac{n_\ell + 1}{g2^{\ell+1}}$$

is satisfied for some n_ℓ with $1 \leq n_\ell \leq g2^\ell$; in particular, the case that $n_\ell = 0$ is ruled out. Similarly, the case that $n_m = 0$ is ruled out by our assumptions on $c/(gz) - a/x$.

Thus

$$S_2(g, x, a) \ll (\log \log h)^2 \sum_{\ell, m=1}^{\log_2(h/g)} \sum_{n_\ell=1}^{g2^{\ell+1}-1} \sum_{n_m=1}^{g2^{m+1}-1} \frac{1}{2^{2\ell+2m}} \times F\left(\frac{n_\ell}{g2^{\ell+1}}\right) F\left(\frac{n_m}{g2^{m+1}}\right) F\left(\frac{n_\ell 2^m - n_m 2^\ell}{g2^{\ell+m+1}}\right) \sum_{\substack{2^\ell < y \leq 2^{\ell+1} \\ 2^m < z \leq 2^{m+1} \\ n_\ell \leq g2^{\ell+1} \|a/x - b/(gy)\| \leq n_\ell + 1 \\ n_m \leq g2^{m+1} \|c/(gz) - a/x\| \leq n_m + 1}} 1.$$

Define

$$C_{\ell, n_\ell} = \#\left\{ b, y : \frac{b}{y} \in \left(\frac{ga - n_\ell + 1}{x - y}, \frac{ga - n_\ell}{x - y} \right) \cup \left(\frac{ga + n_\ell}{x + y}, \frac{ga + n_\ell + 1}{x + y} \right), 1 \leq b < g2^{\ell+1}, 2^\ell < y \leq 2^{\ell+1} \right\},$$

and define C_{m, n_m} in the same way, so that the inside sum of $S_2(g, x, a)$ is $C_{\ell, n_\ell} C_{m, n_m}$. The minimum spacing of two distinct points b_1/y_1 and b_2/y_2 with denominators $y_i \leq 2^{\ell+1}$ is $O(2^{-2\ell})$, so

$$C_{\ell, n_\ell} \ll \frac{2^{2\ell}}{2^\ell} \ll 2^\ell,$$

and similarly $C_{m, n_m} \ll 2^m$. This implies that

$$S_2(g, x, a) \ll (\log \log h)^2 \sum_{\ell, m=1}^{\log_2(h/g)} \frac{2^{\ell+m}}{2^{2\ell+2m}} \sum_{n_\ell=1}^{g2^{\ell+1}} \sum_{n_m=1}^{g2^{m+1}} F\left(\frac{n_\ell}{g2^{\ell+1}}\right) F\left(\frac{n_m}{g2^{m+1}}\right) F\left(\frac{n_\ell 2^m - n_m 2^\ell}{g2^{\ell+m+1}}\right).$$

By the symmetry of ℓ and m , we can restrict the sum to the terms where $\ell \leq m$. Applying Lemma 2.14 to the sums over n_ℓ, n_m with $d_1 = g2^\ell$ and $d_2 = g2^m$ gives

$$S_2(g, x, a) \ll (\log \log h)^2 \sum_{\substack{\ell, m=1 \\ \ell \leq m}}^{\log_2(h/g)} \frac{1}{2^{\ell+m}} (hg^2 2^{2\ell} + g^3 2^{2\ell+m} m) \\ \ll h(\log \log h)^2 g^2 \sum_{\substack{\ell, m=1 \\ \ell \leq m}}^{\log_2(h/g)} \frac{1}{2^{m-\ell}} + (\log \log h)^2 g^3 \sum_{\substack{\ell, m=1 \\ \ell \leq m}}^{\log_2(h/g)} m 2^\ell \ll h(\log \log h)^2 g^2 \left(\log \frac{h}{g}\right)^2,$$

and thus

$$\begin{aligned}
 T_2 &\ll h(\log \log h)^2 \sum_{\substack{g,x|q \\ x < h/g}} \frac{\mu(gx)^2}{\phi(g)^3 \phi(x)^2} \sum_{\substack{a \\ (a,x)=1}} g^2 \left(\log \frac{h}{g} \right)^2 \\
 &\ll h(\log h)^2 (\log \log h)^2 \sum_{\substack{g,x|q \\ x < h/g}} \frac{\mu(gx)^2 g^2}{\phi(g)^3 \phi(x)} \\
 &\ll h(\log h)^2 (\log \log h)^2 \prod_{p \leq h^4} \left(1 + \frac{p^2}{(p-1)^3} + \frac{1}{p-1} \right) \left(\text{since } q = \prod_{p \leq h^4} p \right) \\
 &\ll h(\log h)^4 (\log \log h)^2. \quad \square
 \end{aligned}$$

2.4. Bounding T_3 : terms with gx, gy, gz small and each a_i small. All that remains is to analyze the sum T_3 , which consists of the terms in (8) where $gx, gy,$ and $gz < h$, and, for each i , $\|a_i/q_i\| \leq 2/h$. Precisely, we define

$$T_3 := \sum_{\substack{g,x,y,z|q \\ x,y,z < h/g}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{\substack{a_1,a_2,a_3 \\ (a_1,xyz)=\dots=1 \\ a_1/gyz+\dots \in \mathbb{Z} \\ \|a_1/gyz\| < 2/h \\ \|a_2/gxz\| < 2/h \\ \|a_3/gxy\| < 2/h}} F\left(\frac{a_1}{gyz}\right) F\left(\frac{a_2}{gxz}\right) F\left(\frac{a_3}{gxy}\right). \quad (14)$$

Intuitively, there are simply not many triples of fractions a_i/q_i where the denominators are not too big, each fraction is close to an integer, and the sum of all three is in \mathbb{Z} . We will make this precise in the following lemma bounding T_3 , where the key savings come from bounding the number of satisfactory triples.

Lemma 2.17. *Let $h \geq 4$, let q be the product of all primes $p \leq h^4$ and define T_3 by (14). Then*

$$T_3 \ll h(\log h)^4 (\log \log h)^2.$$

Proof. Since $\|a_3/(gxy)\| < 2/h$, we must have $(1/gxy) < 2/h$, so if $y < \sqrt{h/(2g)}$, then $x > \sqrt{h/(2g)}$. By the same logic with a_1 and a_2 , at most one of x, y, z can be $< \sqrt{h/(2g)}$. By relabeling if necessary, we get that

$$T_3 \ll \sum_{\substack{g,x,y,z|q \\ x,y,z < h/g \\ y,z \geq \sqrt{h/(2g)}}} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{\substack{a_1,a_2,a_3 \\ (a_1,xyz)=\dots=1 \\ a_1/gyz+\dots \in \mathbb{Z} \\ \|a_1/gyz\| < 2/h \\ \|a_2/gxz\| < 2/h \\ \|a_3/gxy\| < 2/h}} F\left(\frac{a_1}{gyz}\right) F\left(\frac{a_2}{gxz}\right) F\left(\frac{a_3}{gxy}\right).$$

As in the proof of Lemma 2.16, there are unique values a, b, c with

$$\frac{a_1}{gyz} \equiv \frac{b}{gy} - \frac{c}{gz} \pmod{1}, \quad \frac{a_2}{gxz} \equiv \frac{c}{gz} - \frac{a}{gx} \pmod{1}, \quad \text{and} \quad \frac{a_3}{gxy} \equiv \frac{a}{gx} - \frac{b}{gy} \pmod{1},$$

and we can reparametrize T_3 in terms of sums over a, b, c instead of a_1, a_2, a_3 . Doing so, and moving the sums over $b, y, c,$ and z inside, we get that

$$T_3 \ll h^3 \sum_{\substack{g,x|q \\ gx \leq h}} \frac{\mu(gx)^2}{\phi(g)^3 \phi(x)^2} \sum_{\substack{a \leq x \\ (a,x)=1}} S_3(g, x, a),$$

where

$$S_3(g, x, a) := \sum_{\substack{\sqrt{h/(2g)} \leq y \leq h/(2g) \\ \sqrt{h/(2g)} \leq z \leq h/(2g)}} \frac{\mu(yz)^2}{\phi(yz)^2} \# \left\{ b, c : \frac{b}{gy}, \frac{c}{gz} \in \left(\frac{a}{x} - \frac{2}{h}, \frac{a}{x} + \frac{2}{h} \right) \right\}.$$

Since $y, z \leq h$, the product yz is $\leq h^2$, and thus

$$\frac{1}{\phi(yz)^2} \ll \frac{(\log \log h)^2}{y^2 z^2},$$

when this term appears in $S_3(g, x, a)$. In order to bound $S_3(g, x, a)$, we split the sums over y and z dyadically, defining ℓ such that $2^\ell < y \leq 2^{\ell+1}$ and $2^m < z \leq 2^{m+1}$.

Then

$$S_3(g, x, a) \ll (\log \log h)^2 \sum_{\ell, m = (\log_2(h/g))/2}^{\log_2(h/g)} \frac{C_\ell C_m}{2^{2\ell} 2^{2m}},$$

where

$$C_\ell := \# \left\{ b, y : \frac{b}{y} \in \left(\frac{ga}{x} - \frac{2g}{h}, \frac{ga}{x} + \frac{2g}{h} \right), 1 \leq b < y, y \leq 2^{\ell+1} \right\},$$

and C_m is defined identically, with m in place of ℓ . The minimum spacing of two distinct points b_1/y_1 and b_2/y_2 with denominators at most $2^{\ell+1}$ is $O(1/2^{2\ell})$, so $C_\ell \ll 2^{2\ell}(g/h) + 1$. Since $\ell \geq \frac{1}{2}(\log_2(h/g))$, $2^{2\ell}(g/h) \geq 1$, so in particular $C_\ell \ll 2^{2\ell}(g/h)$, and similarly $C_m \ll 2^{2m}(g/h)$.

Plugging this in gives

$$S_3(g, x, a) \ll (\log \log h)^2 \sum_{\ell, m = (\log_2(h/g))/2}^{\log_2(h/g)} \frac{2^{2\ell} 2^{2m} g^2}{2^{2\ell} 2^{2m} h^2} \ll \frac{g^2}{h^2} \left(\log \frac{h}{g} \right)^2 (\log \log h)^2,$$

so that

$$\begin{aligned} T_3 &\ll h(\log \log h)^2 \sum_{\substack{g,x|q \\ gx \leq h}} \frac{\mu(gx)^2 g^2}{\phi(g)^3 \phi(x)^2} \sum_{\substack{a \leq x \\ (a,x)=1}} \left(\log \frac{h}{g} \right)^2 \\ &\ll h(\log h)^2 (\log \log h)^2 \sum_{\substack{g,x|q \\ gx \leq h}} \frac{\mu(gx)^2 g^2}{\phi(g)^3 \phi(x)} \\ &\ll h(\log h)^2 (\log \log h)^2 \prod_{p \leq h^4} \left(1 + \frac{p^2}{(p-1)^3} + \frac{1}{p-1} \right), \end{aligned}$$

recalling that $q = \prod_{p \leq h^4} p$. Thus $T_3 \ll h(\log h)^4 (\log \log h)^2$. □

Putting Lemmas 2.15, 2.16, and 2.17 together completes the proof of Theorem 2.1.

3. Function field analogs: proof of Theorem 1.3

We now turn to considering analogous questions when working in $\mathbb{F}_q[t]$ rather than in \mathbb{Z} . To begin with, let us set up the situation in the function field case. Fix a finite field \mathbb{F}_q . Rather than primes in \mathbb{N} , consider monic irreducible polynomials in $\mathbb{F}_q[t]$.

The *norm* of a polynomial $F \in \mathbb{F}_q[t]$ is given by $|F| = q^{\deg F}$. We consider intervals in norm, where the interval $I(F, h)$ of degree h is defined as

$$I(F, h) := \{G \in \mathbb{F}_q[t] : |F - G| < q^h\}.$$

For a fixed monic polynomial Q , we define

$$\begin{aligned} \mathcal{C}(Q) &:= \left\{ \frac{A}{Q} \in \mathbb{F}_q[t] : |A| < |Q| \right\}, \\ \mathcal{R}(Q) &:= \left\{ \frac{A}{Q} \in \mathbb{F}_q[t] : |A| < |Q|, (A, Q) = 1 \right\}. \end{aligned}$$

For $Q = 1$, we instead for convenience define $\mathcal{C}(Q) = \{1\} = \mathcal{R}(Q)$. If $\deg Q > 0$, the set of polynomials F with $\deg F < \deg Q$ is a canonical set of representatives of $\mathbb{F}_q[t]/(Q)$; in what follows, we will identify $\{F \in \mathbb{F}_q[t] : \deg F < \deg Q\}$ with $\mathbb{F}_q[t]/(Q)$. If $Q = 1$, we will take 1 to represent the unique equivalence class of $\mathbb{F}_q[t]/(Q)$.

We consider the k -th moment of the distribution of irreducible polynomials in intervals $I(F, h)$. As in the integer case, we begin by considering the related quantity of the distribution of reduced residues modulo a square-free monic polynomial Q . That is, for Q a fixed square-free monic polynomial, we consider

$$m_k(Q; h) = \sum_{F \in \mathcal{C}(Q)} \left(\left(\sum_{\substack{G \in I(F, h) \\ (G, Q) = 1}} 1 \right) - \frac{q^h \phi(Q)}{|Q|} \right)^k. \tag{15}$$

Here we are taking the centered moment $m_k(Q; h)$ by subtracting $q^h \phi(Q)/|Q|$, which is the mean value of $\sum_{G \in I(F, h), (G, Q) = 1} 1$.

As in the integer case, we can express the moment $m_k(Q; h)$ in terms of exponential sums. For $\alpha = F/G \in \mathbb{F}_q(t)$ a rational function, let $\text{res}(\alpha)$ denote the coefficient of $1/t$ when α is written as a Laurent series with finitely many positive terms. Then define

$$e(\alpha) := e_q(\text{res}(\alpha)) = \exp(2\pi i \cdot \text{tr}(\text{res}(\alpha))/p),$$

where q is a power of the prime p and $\text{tr} : \mathbb{F}_q \rightarrow \mathbb{F}_p$ is the trace function. This exponential function, like its integer analog, satisfies the crucial property that, for a monic polynomial $F \in \mathbb{F}_q[t]$,

$$\sum_{\alpha \in \mathcal{C}(F)} e(\alpha) = \begin{cases} 1 & \text{if } F = 1, \\ 0 & \text{otherwise.} \end{cases}$$

We then have the following lemma, analogous to [Montgomery and Vaughan 1986, Lemma 2].

Lemma 3.1. *Let $Q \in \mathbb{F}_q[t]$ be square-free and let $h \in \mathbb{N}_{\geq 1}$. Define $m_k(Q; h)$ by (15). Then*

$$m_k(Q; h) = |Q| \left(\frac{\phi(Q)}{|Q|} \right)^k V_k(Q; h),$$

where

$$V_k(Q; h) := \sum_{\substack{R_1, \dots, R_k \\ |R_i| > 1 \\ R_i \text{ monic}}} \prod_{i=1}^k \frac{\mu(R_i)}{\phi(R_i)} \sum_{\substack{\rho_1, \dots, \rho_k \\ \rho_i \in \mathcal{R}(R_i) \\ \sum_i \rho_i / R_i = 0}} E\left(\frac{\rho_1}{R_1}\right) \cdots E\left(\frac{\rho_k}{R_k}\right),$$

and where, for $\alpha \in \mathbb{F}_q(t)$ a rational function,

$$E(\alpha) := \sum_{M \in I(0, h)} e(M\alpha).$$

The proof follows that of [Montgomery and Vaughan 1986, Lemma 2] very closely.

Proof. Let $\kappa(R) = 1$ when $(R, Q) = 1$, and $\kappa(R) = 0$ otherwise. Then

$$\kappa(R) = \sum_{S|(R, Q)} \mu(S) = \sum_{S|Q} \frac{\mu(S)}{|S|} \sum_{\sigma \in \mathcal{C}(S)} e(R\sigma) = \sum_{T|Q} \left(\sum_{\substack{A \in \mathcal{C}(T) \\ (A, T) = 1}} e(RA) \right) \left(\sum_{T|S|Q} \frac{\mu(S)}{|S|} \right).$$

Here the second factor is $(\phi(Q)/|Q|)(\mu(T)/|T|)$. The function $\kappa(R)$ has mean value $\phi(Q)/|Q|$, so we subtract $\phi(Q)/|Q|$ from both sides, which removes the term when $T = 1$. We then substitute $R = M + N$, and sum over M to see that

$$\sum_{\substack{|M| < q^h \\ (M+N, Q) = 1}} 1 - h \frac{\phi(Q)}{|Q|} = \frac{\phi(Q)}{|Q|} \sum_{\substack{R|Q \\ |R| > 1}} \frac{\mu(R)}{\phi(R)} \sum_{\substack{A \in \mathcal{C}(R) \\ (A, R) = 1}} E\left(\frac{A}{R}\right) e\left(\frac{NA}{R}\right).$$

The argument is completed upon raising both sides to the k -th power, summing over N , multiplying out the right-hand side, and appealing to the fact that

$$\sum_{|N| < q^d} e(N(\alpha_1 + \cdots + \alpha_k)) = \begin{cases} q^d & \text{if } \sum \alpha_i \in \mathbb{Z}, \\ 0 & \text{else.} \end{cases} \quad \square$$

One important difference between the integer setting and the function field setting is the behavior of the sums $E(\alpha)$, which are particularly well-behaved in $\mathbb{F}_q[t]$. These sums have also been studied in [Hayes 1966, Theorem 3.5].

Lemma 3.2. *Let $\alpha \in \mathbb{F}_q(t)$ be a rational function with $\deg \alpha \leq -1$. Then*

$$E(\alpha) = \begin{cases} q^h & \text{if } \deg \alpha < -h, \\ 0 & \text{if } \deg \alpha \geq -h. \end{cases}$$

Proof. Let $\mathcal{P}_h \subseteq \mathbb{F}_q[t]$ be the set of polynomials of degree less than h . Assume first that $\deg \alpha < -h$. Then, for all $M \in \mathcal{P}_h$, $\deg M\alpha = \deg M + \deg \alpha \leq h - 1 - h - 1 = -2$, so the Laurent series for $M\alpha$ has

no $1/t$ -term, and thus $\text{res}(M\alpha) = 0$. But then

$$E(\alpha) = \sum_{M \in \mathcal{P}_h} e(M\alpha) = \sum_{M \in \mathcal{P}_h} e_q(\text{res}(M\alpha)) = \sum_{M \in \mathcal{P}_h} 1 = q^h.$$

Now assume that $\deg \alpha \geq -h$. Consider the map $\text{res}_\alpha : \mathcal{P}_h \rightarrow \mathbb{F}_q$ which at a polynomial M returns the residue of $M\alpha$. This map is linear over \mathbb{F}_q , so its image is either 0 or all of \mathbb{F}_q . Let $M = t^{-\deg \alpha - 1}$. Since $-h \leq \deg \alpha \leq -1$, we have $0 \leq -\deg \alpha - 1 \leq h - 1$, so M indeed is a polynomial in \mathcal{P}_h . On the other hand, $\text{res}(M\alpha)$ is precisely the leading coefficient of α , which must be nonzero. Thus the image of res_α is nonzero, so it is all of \mathbb{F}_q . In particular, $\text{res}_\alpha(M)$ takes each value in \mathbb{F}_q equally often. Thus

$$E(\alpha) = \sum_{M \in \mathcal{P}_h} e_q(\text{res}(M\alpha))$$

is a balanced exponential sum, which has sum 0. □

This fact and other properties of the sums $E(\alpha)$ mean that the analysis in [Montgomery and Vaughan 1986] in the function field setting is more streamlined. In fact, their work automatically gives the analog of our desired bound for the third moment in the function field case.

3.1. The analog of [Montgomery and Vaughan 1986] in the function field setting. We begin with the following fundamental lemma, with an identical proof to the integer case.

Lemma 3.3 (Fundamental lemma). *Let $R_1, \dots, R_k \in \mathbb{F}_q[t]$ be square-free monic polynomials with $R = [R_1, \dots, R_k]$. Suppose, for all irreducible $P \mid R$, P divides at least two R_i 's. Let G_i be positive real-valued function defined on $\mathcal{C}(R_i)$. Then*

$$\left| \sum_{\substack{A_i \in \mathcal{C}(R_i) \\ \sum_i A_i/R_i = 0}} G_1\left(\frac{A_1}{R_1}\right) \cdots G_k\left(\frac{A_k}{R_k}\right) \right| \leq \frac{1}{|R|} \prod_{i=1}^k \left(|R_i| \sum_{A_i \in \mathcal{C}(R_i)} \left| G_i\left(\frac{A_i}{R_i}\right) \right|^2 \right)^{1/2}.$$

The proof follows [Montgomery and Vaughan 1986] very closely.

Proof. We proceed by induction on k .

Assume first that $k = 2$. Then we must have $R_1 = R_2 = R$. By Cauchy–Schwarz,

$$\left| \sum_{|A| < |R|} G_1\left(\frac{A}{R}\right) G_2\left(\frac{A}{R}\right) \right| \leq \left(\sum_{|A| < |R|} \left| G_1\left(\frac{A}{R}\right) \right|^2 \right)^{1/2} \left(\sum_{|A| < |R|} \left| G_2\left(\frac{A}{R}\right) \right|^2 \right)^{1/2},$$

which after a bit of rearranging gives the desired result.

Now assume by induction that the result holds for $j \leq k - 1$. For arbitrary k , set $D = (R_1, R_2)$, and write $D = ST$, with $S \mid R_3 \cdots R_k$ and $(T, R_3 \cdots R_k) = 1$. Furthermore, write $R_1 = DR'_1$ and $R_2 = DR'_2$. Consider any term in the sum. Since $\sum_i A_i/R_i = 0$, we have $T \mid (A_1/R_1 + A_2/R_2)$. Thus $A_1/(STR'_1) + A_2/(STR'_2)$ can be expressed as a fraction $A/(R'_1R'_2S)$.

By the Chinese remainder theorem,

$$\frac{A_1}{STR'_1} = \frac{\alpha_1}{R'_1} + \frac{\beta_1}{ST} \quad \text{and} \quad \frac{A_2}{STR'_2} = \frac{\alpha_2}{R'_2} + \frac{\beta_2}{ST},$$

where

$$\frac{\beta_2}{ST} = -\frac{\beta_1}{ST} + \frac{\gamma}{S}$$

because $T|(A_1/R_1 + A_2/R_2)$. Thus A_1/R_1 and A_2/R_2 can be written as

$$\frac{A_1}{R_1} = \frac{A'_1}{R'_1} + \frac{\delta}{D} \quad \text{and} \quad \frac{A_2}{R_2} = \frac{A'_2}{R'_2} + \frac{\sigma}{S} - \frac{\delta}{D},$$

with each rational function of degree less than 0.

Let $R^* = R'_1 R'_2 S$. For each A^* with $|A^*| < |R^*|$, A^*/R^* is uniquely of the form

$$\frac{A^*}{R^*} = \frac{A'_1}{R'_1} + \frac{A'_2}{R'_2} + \frac{\sigma}{S}.$$

Define

$$G^*\left(\frac{A^*}{R^*}\right) = \sum_{\delta \in \mathcal{C}(D)} G_1\left(\frac{A'_1}{R'_1} + \frac{\delta}{D}\right) G_2\left(\frac{A'_2}{R'_2} + \frac{\sigma}{S} - \frac{\delta}{D}\right).$$

Then the sum in question is

$$\sum_{\substack{A^* \in \mathcal{C}(R^*) \\ A_i \in \mathcal{C}(R_i) \\ A^*/R^* + \sum_{i=3}^k A_i/R_i = 0}} G^*\left(\frac{A^*}{R^*}\right) G_3\left(\frac{A_3}{R_3}\right) \cdots G_k\left(\frac{A_k}{R_k}\right).$$

Via Cauchy–Schwarz as well as the induction hypothesis, the above is

$$\leq \frac{|T|}{|R|} \left(|R^*| \sum_{A^* \in \mathcal{C}(R^*)} G^*\left(\frac{A^*}{R^*}\right)^2 \right)^{1/2} \prod_{i=3}^k \left(|R_i| \sum_{A_i \in \mathcal{C}(R_i)} G_i\left(\frac{A_i}{R_i}\right)^2 \right)^{1/2}.$$

It remains to bound the sum over G^* in terms of G_1 and G_2 . By Cauchy–Schwarz,

$$G^*\left(\frac{A^*}{R^*}\right)^2 \leq \left(\sum_{\delta \in \mathcal{C}(D)} G_1\left(\frac{A'_1}{R'_1} + \frac{\delta}{D}\right)^2 \right) \left(\sum_{\delta \in \mathcal{C}(D)} G_2\left(\frac{A'_2}{R'_2} + \frac{\sigma}{S} - \frac{\delta}{D}\right)^2 \right),$$

so summing over A^* gives

$$\sum_{A^* \in \mathcal{C}(R^*)} G^*\left(\frac{A^*}{R^*}\right)^2 \leq |S| \left(\sum_{A_1 \in \mathcal{C}(R_1)} G_1\left(\frac{A_1}{R_1}\right)^2 \right) \left(\sum_{A_2 \in \mathcal{C}(R_2)} G_2\left(\frac{A_2}{R_2}\right)^2 \right). \quad \square$$

We now present several preliminary lemmas about the sums $E(\alpha)$. The following lemma is analogous to [Montgomery and Vaughan 1986, Lemma 4].

Lemma 3.4. *For any polynomial $R \in \mathbb{F}_q[t]$,*

$$\sum_{S \in \mathcal{C}(R)} E\left(\frac{S}{R}\right)^2 = \max\{q^{2h}, |R|q^h\}.$$

Moreover, for any polynomial $R \in \mathbb{F}_q[t]$ and any rational function $\alpha \in \mathbb{F}_q(t)$,

$$\sum_{S \in \mathcal{C}(R)} E\left(\frac{S}{R} + \alpha\right)^2 \begin{cases} = \max\{q^{2h}, |R|q^h\} & \text{if } |\alpha| < q^{-h}, \\ \leq |R|q^{h-1} & \text{if } |\alpha| \geq q^{-h}. \end{cases}$$

Proof. If $\deg R \leq h$, then, for all S with $0 \neq |S| < |R|$, we have $h \geq \deg R - \deg S$, and thus $E(S/R)^2 = 0$. Meanwhile, $E(0)^2 = q^{2h}$, so in this case $\sum_{S \in \mathcal{C}(R)} E(S/R)^2 = q^{2h}$.

Now suppose $\deg R > h$. Then $E(S/R)$ is nonzero if and only if $\deg S < \deg R - h$. Thus

$$\sum_{S \in \mathcal{C}(R)} E\left(\frac{S}{R}\right)^2 = \sum_{\substack{S \in \mathcal{C}(R) \\ |S| < |R|/q^h}} E\left(\frac{S}{R}\right)^2 = \sum_{\substack{S \in \mathcal{C}(R) \\ |S| < |R|/q^h}} q^{2h} = |R|q^h,$$

which completes the first portion.

Fix a rational function α . For all S/R , $E(S/R + \alpha)$ is unchanged by replacing α with its fractional part, i.e., subtracting off the polynomial portion of α so that $|\alpha| < 1$, including the possibility that $\alpha = 0$.

If a term $E(S/R + \alpha)$ is nonzero, then $|S/R + \alpha| < q^{-h}$. We'll split into two cases, when $|\alpha| < q^{-h}$ and when $|\alpha| \geq q^{-h}$. First, if $|\alpha| < q^{-h}$, then $|S/R + \alpha| < q^{-h}$ if and only if $|S/R| < q^{-h}$. If $|R| \geq q^h$, there are $|R|/q^h$ values of S satisfying this; if not, there is one value. Thus if $|\alpha| < q^{-h}$, we have

$$\sum_{S \in \mathcal{C}(R)} E\left(\frac{S}{R} + \alpha\right)^2 E\left(\frac{S}{R} + \alpha\right) = \max(q^{2h}, |R|q^h).$$

Now assume $|\alpha| \geq q^{-h}$. If $|S/R + \alpha| < q^{-h}$, we must have $|S/R| = |\alpha| \geq q^{-h}$. Also, the first $\deg \alpha + h + 1$ terms of S/R are fixed, because they must cancel with the corresponding terms of α to yield a rational function of small enough degree. Correspondingly, the first $\deg \alpha + h + 1$ terms of S are determined. Since $|S| = |R\alpha|$, there are at most $|R\alpha| \cdot 1/(|\alpha| \cdot |q^{h+1}|) = |R|q^{-h-1}$ nonzero choices of S . Thus in this case, $\sum_{S \in \mathcal{C}(R)} E(S/R + \alpha)^2 \leq |R|q^{h-1}$. \square

The following lemma corresponds to Lemma 6 of [Montgomery and Vaughan 1986].

Lemma 3.5. *Let $R \in \mathbb{F}_q[t]$ be a polynomial, and let $\alpha, \beta \in \mathbb{F}_q(t)$ be rational functions. Then*

$$\sum_{S \in \mathcal{C}(R)} E\left(\frac{S}{R} + \alpha\right) E\left(\frac{S}{R} + \beta\right) \ll E(\alpha - \beta)q^{-h} \sum_{S \in \mathcal{C}(R)} E\left(\frac{S}{R} + \alpha\right)^2.$$

Proof. Again, we split into two cases. Assume first that $|\alpha - \beta| \geq q^{-h}$, so $E(\alpha - \beta) = 0$. Then either $|S/R + \beta| \geq q^{-h}$, or $|S/R + \alpha| \geq q^{-h}$. Thus for each S/R , either $E(S/R + \alpha) = 0$ or $E(S/R + \beta) = 0$, so the product must be 0, and thus the sum is 0.

Now assume that $|\alpha - \beta| < q^{-h}$, so $E(\alpha - \beta) = q^h$. By Lemma 3.2, if $|\alpha - \beta| < q^{-h}$, then $E(S/R + \alpha) = E(S/R + \beta)$ for all S . This gives the result. \square

We are now ready to prove the following lemma, which is analogous to [Montgomery and Vaughan 1986, Lemma 7].

Lemma 3.6. *Let $k \geq 3$, and let $R_1, \dots, R_k \in \mathbb{F}_q[t]$ be square-free polynomials with $|R_i| > 1$ for all i . Let $R = [R_1, \dots, R_k]$. Let $D = (R_1, R_2)$ and $D = ST$, with $S|R_3 \cdots R_k$ and $(T, R_3 \cdots R_k) = 1$. Write*

$R_1 = DR'_1$, $R_2 = DR'_2$, and $R^* = R'_1 R'_2 S$. Define

$$S(R_1, \dots, R_k) := \sum_{\substack{A_i \in \mathcal{R}(R_i) \\ \sum_i A_i/R_i = 0}} \prod_{i=1}^k E\left(\frac{A_i}{R_i}\right).$$

If, for some i , $|R_i| \leq q^h$, then $S(R_1, \dots, R_k) = 0$. Otherwise,

$$S(R_1, \dots, R_k) \ll |R_1 \cdots R_k| \cdot |R|^{-1} (q^h)^{k/2} (X_1 + X_2 + X_3),$$

where

$$X_1 = q^{-h/2}, \quad X_2 = \begin{cases} |D|^{-1} & \text{if } |R'_1| > q^h, \\ 0 & \text{otherwise,} \end{cases} \quad X_3 = \begin{cases} |S|^{-1/2} & \text{if } R_1 = R_2, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Assume first that, for some i , $|R_i| \leq q^h$. Then $E(A_i/R_i) = 0$ whenever $A_i \neq 0$, so in particular for all A_i with $(A_i, R_i) = 1$, so the sum is 0. Assume from now on that $|R_i| > q^h$ for all i .

We now return to the proof of the fundamental lemma. For $A^*/R^* = A'_1/R'_1 + A'_2/R'_2 + \sigma/S$, define

$$G^*\left(\frac{A^*}{R^*}\right) = \sum_{\substack{\delta \in \mathcal{C}(D) \\ (DA'_1 + \delta R'_1, R_1) = 1 \\ (DA'_2 + R'_2 T \sigma - R'_2 \delta, R_2) = 1}} E\left(\frac{A'_1}{R'_1} + \frac{\delta}{D}\right) E\left(\frac{A'_2}{R'_2} + \frac{\sigma}{S} - \frac{\delta}{D}\right).$$

For this sum to be nonempty, $(A'_1, R'_1) = (A'_2, R'_2) = 1$. Then

$$S(R_1, \dots, R_k) \leq \frac{|T|}{|R|} \left(|R^*| \sum_{A^* \in \mathcal{C}(R^*)} G^*\left(\frac{A^*}{R^*}\right)^2 \right)^{1/2} \prod_{i=3}^k \left(|R_i| \sum_{\substack{A_i \in \mathcal{R}(R_i) \\ |A_i| < |R_i|/q^h}} 1 \right)^{1/2}.$$

By Lemma 3.4, the product is $\ll |R_3 \cdots R_k| q^{hk/2-h}$. Thus it suffices to show that

$$\sum_{A^* \in \mathcal{C}(R^*)} G^*\left(\frac{A^*}{R^*}\right)^2 \ll |R_1| \cdot |R_2| \cdot |S| q^{2h} (X_1^2 + X_2^2 + X_3^2).$$

By Lemma 3.5,

$$G^*\left(\frac{A^*}{R^*}\right) \ll E\left(\frac{A^*}{R^*}\right) q^{-h} \sum_{\delta \in \mathcal{C}(D)} E\left(\frac{\delta}{D} + \frac{A'_1}{R'_1}\right),$$

so by Lemma 3.4,

$$G^*\left(\frac{A^*}{R^*}\right) \ll \begin{cases} E\left(\frac{A^*}{R^*}\right) \max\{q^h, |D|\} & \text{if } \left| \frac{A'_1}{R'_1} \right| < q^{-h}, \\ E\left(\frac{A^*}{R^*}\right) |D| q^{-1} & \text{if } \left| \frac{A'_1}{R'_1} \right| \geq q^{-h}. \end{cases}$$

Summing over A^* then gives

$$\sum_{A^* \in \mathcal{C}(R^*)} G^*\left(\frac{A^*}{R^*}\right)^2 \ll \sum_{\substack{A^* \in \mathcal{C}(R^*) \\ |A^*/R^*| < q^{-h} \\ |A'_1/R'_1| < q^{-h}}} E\left(\frac{A^*}{R^*}\right)^2 \max\{q^{2h}, |D|^2\} + \sum_{\substack{A^* \in \mathcal{C}(R^*) \\ |A^*/R^*| < q^{-h} \\ |A'_1/R'_1| \geq q^{-h}}} E\left(\frac{A^*}{R^*}\right)^2 |D|^2. \tag{16}$$

Here as in the definition of G^* , for any nonzero term we must have $(A'_1, R'_1) = (A'_2, R'_2) = 1$. In particular, $A'_1 \equiv 0 \pmod{R'_1}$ only if $R'_1 = 1$. We now split into cases based on whether or not $|R^*| > q^h$ and whether or not $|R'_1| > q^h$.

First assume that $|R^*| > q^h$ and $|R'_1| > q^h$. Then

$$\begin{aligned} \sum_{A^* \in \mathcal{C}(R^*)} G^* \left(\frac{A^*}{R^*} \right)^2 &\ll \max\{q^{2h}, |D|^2\} q^{2h} \frac{|R'_1|}{q^h} \frac{|R'_2 S|}{q^h} + |D|^2 \sum_{\substack{A^* \in \mathcal{C}(R^*) \\ |A^*/R^*| < q^{-h} \\ |A'_1/R'_1| \geq q^h}} E \left(\frac{A^*}{R^*} \right)^2 \\ &\ll \max\{q^{2h}, |D|^2\} |R^*| + |D|^2 |R^*| q^h \\ &\ll |R_1| \cdot |R_2| \cdot |S| q^{2h} (X_1^2 + X_2^2). \end{aligned}$$

Now assume that $|R^*| > q^h$ but $|R'_1| \leq q^h$. The first sum in (16) is empty unless $R'_1 = 1$ (and $A'_1 = 0$). If $R'_1 = 1$, then $R_1 = D$, so $|D| > q^h$. Equation (16) then becomes

$$\sum_{A^* \in \mathcal{C}(R^*)} G^* \left(\frac{A^*}{R^*} \right)^2 \ll q^{2h} |D|^2 + \frac{|R^*|}{q^h} q^{2h} |D|^2 = |R_1 R_2 S| q^{2h} \left(\frac{1}{|R^*|} + q^{-h} \right) \ll |R_1 R_2 S| q^{2h} (X_1^2).$$

If $R'_1 \neq 1$, then the first sum is empty, so (16) becomes

$$\sum_{A^* \in \mathcal{C}(R^*)} G^* \left(\frac{A^*}{R^*} \right)^2 \ll \frac{|R^*|}{q^h} q^{2h} |D|^2 = |R_1 R_2 S| q^{2h} (X_1^2).$$

Finally, assume that $|R^*| \leq q^h$ and thus $|R'_1| \leq q^h$. In this case the only nonzero term in (16) in either sum is when $A^* = 0$, which forces $A'_1 = A'_2 = \sigma = 0$. But then since $(A'_1, R'_1) = (A'_2, R'_2) = 1$, we also have $R'_1 = R'_2 = 1$, and thus $R_1 = R_2 = D$, which has magnitude $> q^h$. Thus

$$\sum_{A^* \in \mathcal{C}(R^*)} G^* \left(\frac{A^*}{R^*} \right)^2 \ll q^{2h} |D|^2 = |R_1 R_2 S| q^{2h} \cdot |S|^{-1} = |R_1 R_2 S| q^{2h} X_3^2. \quad \square$$

We now turn to the proof of Theorem 1.3, which corresponds to [Montgomery and Vaughan 1986, Lemma 8]. The main strategy here is a careful application of Lemma 3.6, keeping in mind that we can choose which variables play the roles of R_1 and R_2 .

Lemma 3.7. *For any fixed $k \geq 3$, for $Q \in \mathbb{F}_q[t]$ square-free, for $h \geq 1$ and $m_k(Q; h)$ defined by (15),*

$$m_k(Q; h) \ll |Q| (q^h)^{k/2} \left(\frac{\phi(Q)}{|Q|} \right)^{k/2} \left(1 + ((q^h)^{-1/2} + (q^h)^{-1/(k-2)}) \left(\frac{\phi(Q)}{|Q|} \right)^{-2k+k/2} \right).$$

Proof. We begin with the bound that

$$m_k(Q; h) \ll |Q| \left(\frac{\phi(Q)}{|Q|} \right)^k \sum_{\substack{R|Q \\ R \text{ monic}}} \sum_{\substack{R_i|Q \\ R_i \text{ monic} \\ |R_i| > 1 \\ [R_1, \dots, R_k] = R}} \frac{S(R_1, \dots, R_k)}{\phi(R_1) \cdots \phi(R_k)},$$

where

$$S(R_1, \dots, R_k) = \sum_{\substack{A_i \in \mathcal{R}(R_i) \\ \sum_i A_i/R_i = 0}} \prod_{i=1}^k E \left(\frac{A_i}{R_i} \right).$$

We apply Lemma 3.6, but while using the fact that we have flexibility in how we label R_1, \dots, R_k in our application of Lemma 3.6. For clarity, we will write \tilde{R}_1 and \tilde{R}_2 to be the R_i 's that serve as the first two in our application of Lemma 3.6. Choose \tilde{R}_1 and \tilde{R}_2 as follows.

If, for any i , $|R_i| < q^h$, then $S(R_1, \dots, R_k)$ must be 0, so assume that $|R_i| \geq q^h$ for all i . Let $R_{ij} = (R_i, R_j)$. For all i , since $R_i \mid \prod_{i \neq j} R_j$, we have $R_i \mid \prod_{i \neq j} R_{ij}$ as well. Thus for all i , there exists $j \neq i$ such that $|R_{ij}| \geq |R_i|^{1/(k-1)}$. If, for some i, j , $|R_{ij}| \geq |R_i|^{1/(k-1)}$ but $R_i \neq R_j$, then pick \tilde{R}_1 and \tilde{R}_2 to be R_i and R_j , respectively.

If no such i exists, then, for each i , there is some $j \neq i$ with $R_i = R_j$. If there exists any triple $R_i = R_j = R_l$, then pick $\tilde{R}_1 = R_i$, $\tilde{R}_2 = R_j$. If not, then the R_i 's must be equal in pairs and otherwise disjoint, and k must be even. Without loss of generality, say that $R_1 = R_2, R_3 = R_4, \dots, R_{k-1} = R_k$. Write $R = UV$, where V is the product of all primes dividing at least two R_{2i} 's, and U is the product of all primes dividing exactly one R_{2i} . Then

$$V^2 \mid \prod_{i=1}^{k/2} \left(R_{2i}, \prod_{j \neq i} R_{2j} \right),$$

so there exists some i with $\left| (R_{2i}, \prod_{j \neq i} R_{2j}) \right| \geq |V|^{4/k}$. Take \tilde{R}_1 and \tilde{R}_2 to be R_{2i} and R_{2i-1} .

Now we return to our bound on $m_k(Q; h)$. We have

$$m_k(Q; h) \ll |Q| \left(\frac{\phi(Q)}{|Q|} \right)^k (q^h)^{k/2} \sum_{\substack{R \mid Q \\ R \text{ monic}}} \frac{1}{|R|} \sum_{\substack{R_i \mid Q \\ |R_i| > 1 \\ [R_1, \dots, R_k] = R}} \frac{|R_1 \cdots R_k|}{\phi(R_1) \cdots \phi(R_k)} (X_1 + X_2 + X_3),$$

where the X_i arise by use of Lemma 3.6 as described above.

Consider the contribution from each X_i . Since $X_1 = q^{-h/2}$, the X_1 -terms contribute

$$\begin{aligned} &\ll |Q| \left(\frac{\phi(Q)}{|Q|} \right)^k (q^h)^{k/2-1/2} \sum_{\substack{R \mid Q \\ R \text{ monic}}} \frac{1}{|R|} \sum_{\substack{R_i \mid Q \\ |R_i| \geq q^h \\ [R_1, \dots, R_k] = R}} \frac{|R_1 \cdots R_k|}{\phi(R_1) \cdots \phi(R_k)} \\ &\ll |Q| \left(\frac{\phi(Q)}{|Q|} \right)^k (q^h)^{k/2-1/2} \prod_{P \mid Q} \left(1 + \frac{1}{|P|} \left(2 + \frac{1}{|P|-1} \right)^k \right) \\ &\ll |Q| (q^h)^{k/2-1/2} \left(\frac{\phi(Q)}{|Q|} \right)^{-2^k+k}. \end{aligned}$$

Now consider the X_2 -contribution. If $X_2 \neq 0$, then $|R'_1| > q^h$, and by our choice of R_1, R_2 , we have $|D| \geq |R_1|^{1/(k-1)} = |R'_1 \cdot D|^{1/(k-1)}$. But then $|D|^{-1} \leq q^{-h/(k-2)}$, so in turn $X_2 \leq q^{-h/(k-2)}$. By the same logic as for the X_1 -terms, the X_2 terms contribute $\ll |Q| (q^h)^{k/2-1/(k-2)} (\phi(Q)/|Q|)^{-2^k+k}$.

Finally, consider X_3 . If $X_3 \neq 0$, then $R_1 = R_2$. By our choice of R_1 and R_2 for the application of Lemma 3.6, in this case each R_i is equal to some R_j . If there exists some $R_i = R_1 = R_2$, with $i \geq 3$, then

$S = R_1 = R_2$, so $|S| > q^h$, and thus for these terms we get a saving of $q^{-h/2}$ and the bound for X_1 applies. If not, then k is even and the R_i 's must be equal in pairs. Let $R = UV$ as above, where U is the product of irreducibles P dividing exactly one pair of R_i 's, and V is the product of all other irreducibles P dividing R . Write $R_i = U_i V_i$, where $U_i = (R_i, U)$ and $V_i = (R_i, V)$. For fixed U, V , let $C(U, V)$ be the set of k -tuples (R_1, \dots, R_k) yielding U and V . There are at most $\tau_{k/2}(U)$ choices for U_2, U_4, \dots, U_k , where $\tau_{k/2}$ is the $k/2$ -fold divisor function. Since $V_i | V$, there are at most $\tau(V)^{k/2}$ choices for V_2, V_4, \dots, V_k . Thus $\#C(U, V) \leq \tau_{k/2}(U)d(V)^{k/2}$. In our application of Lemma 3.6 we have $|S| \geq |V|^{4/k}$, so

$$\begin{aligned} \sum_{\substack{UV|Q \\ \text{monic}}} \frac{1}{|UV|} \sum_{(R_1, \dots, R_k) \in C(U, V)} \left(\prod_{i=1}^k \frac{|R_i|}{\phi(R_i)} \right) X_3 &\ll \sum_{\substack{UV|Q \\ \text{monic}}} \frac{\tau_{k/2}(U)(|U|/\phi(U))^2 \tau(V)^{k/2} (|V|/\phi(V))^k}{|U| \cdot |V|^{1+2/k}} \\ &= \prod_{P|Q} \left(1 + \frac{k|P|}{2(|P|-1)^2} + \frac{2^{k/2}(|P|/(|P|-1))^k}{|P|^{1+2/k}} \right) \\ &\ll \left(\frac{\phi(Q)}{|Q|} \right)^{-k/2}, \end{aligned}$$

so the X_3 -terms contribute $\ll |Q|(q^h)^{k/2}(\phi(Q)/|Q|)^{k/2}$, which completes the proof. □

The final contribution of X_3 only arises when k is even, so when k is odd we have the estimate

$$m_k(Q; h) \ll |Q|((q^h)^{k/2-1/2} + (q^h)^{k/2-1/(k-2)}) \left(\frac{\phi(Q)}{|Q|} \right)^{k-2k}.$$

For $k = 3$ this implies that

$$m_3(Q; h) \ll |Q|q^h \left(\frac{\phi(Q)}{|Q|} \right)^{-5}.$$

In the case when $k = 5$, we can bound $m_5(Q; h)$ via a more involved argument.

4. The fifth moment of reduced residues in the function field setting

Our goal in this section is to prove Theorem 1.4, which is a stronger bound on $m_5(Q; h)$ when $Q = \prod_{|P| \leq q^{6h}} P$. We will also prove Corollary 1.5, bounding $R_3(q^h)$ and $R_5(q^h)$ in the ring $\mathbb{F}_q[t]$.

Lemma 3.7 already implies a bound on $m_5(Q; h)$, showing that $m_5(Q; h) \ll |Q|(q^h)^{13/6}(\phi(Q)/|Q|)^{-27}$. Our goal is a bound where the power of q^h is $2 + \varepsilon$ for all $\varepsilon > 0$; note that Conjecture 1.1 would predict a bound where the power of q^h is 2. In turn, this will allow us to prove Corollary 1.5, that $R_5(q^h) \ll q^{(2+\varepsilon)h}$.

4.1. Proof of Theorem 1.4. As in the proof of Lemma 3.7, we begin by bounding

$$m_5(Q; h) \ll |Q| \left(\frac{\phi(Q)}{|Q|} \right)^5 \sum_{\substack{R|Q \\ R \text{ monic}}} \sum_{\substack{R_i|Q \\ |R_i|>1 \\ [R_1, \dots, R_5]=R}} \frac{S(R_1, \dots, R_5)}{\phi(R_1) \cdots \phi(R_5)},$$

where

$$S(R_1, \dots, R_5) = \sum_{\substack{A_i \in \mathcal{R}(R_i) \\ \sum_i A_i/R_i=0}} \prod_{i=1}^5 E\left(\frac{A_i}{R_i}\right).$$

Our goal is to apply Lemma 3.6 to bound the size of $S(R_1, \dots, R_5)$. But, when applying this lemma, we can freely choose which of the R_i 's plays the roles of R_1 and R_2 . As in the previous section, we will denote our choice by \tilde{R}_1 and \tilde{R}_2 . If any R_i satisfies $|R_i| < q^h$, the choice is immaterial, so assume that $|R_i| \geq q^h$ for all i . If there is any triple R_i, R_j, R_ℓ with $R_i = R_j = R_\ell$, pick $\tilde{R}_1 = R_i$ and $\tilde{R}_2 = R_j$. In this case X_2 will have no contribution, and X_3 and X_1 will each be $\ll q^{-h/2}$, for a total contribution to $m_5(Q; h)$ from these terms (as in the proof of Lemma 3.7) of $\ll |Q|q^{2h}(\phi(Q)/|Q|)^{-27}$. If there is no such triple, but there exists $R_i \neq R_j$ with either $|R_i/(R_i, R_j)| < q^h$, or $|R_i/(R_i, R_j)| \geq q^h$ and $|(R_i, R_j)| \geq q^{h/2}$, then we choose $\tilde{R}_1 = R_i$ and $\tilde{R}_2 = R_j$. In this case we have $X_3 = 0$ and X_1, X_2 each contributing $\ll q^{-h/2}$, and again the total contribution to $m_5(Q; h)$ from these terms is $\ll |Q|q^{2h}(\phi(Q)/|Q|)^{-27}$. So, it remains to bound what happens in the remaining cases. We first show that in the remaining cases, up to some reordering, certain factors of R_2 and R_3 are bounded.

Lemma 4.1. *For fixed square-free $Q \in \mathbb{F}_q[t]$, let (R_1, \dots, R_5) be a tuple of divisors of Q such that*

- $|R_i| \geq q^h$ for all i ,
- no three R_i 's are equal,
- for any R_i, R_j , either $R_i = R_j$, or $|R_i/(R_i, R_j)| \geq q^h$ and $|(R_i, R_j)| < q^{h/2}$, and
- R_1, R_2 , and R_3 are all distinct.

Then

- $|R_2/(R_1, R_2)| \geq q^h$, and
- $|R_3/(R_3, R_1R_2)| \geq q^{h/2}$.

Loosely, this lemma states that in the cases that we cannot already bound by the tools of the previous section, prime factors must “spread out” among the first three R_i 's.

Remark. The bound on $|R_3/(R_3, R_1R_2)|$ above is worse than the bound on $|R_2/(R_1, R_2)|$. In order to apply Lemma 4.3 below, we will need both of them to be at least of size $q^{h/2}$, so the bound on $|R_2/(R_1, R_2)|$ is better than necessary.

However, the fact that these bounds get worse is precisely what prevents us from applying our technique to bound higher moments. If instead we applied the same argument to a 7-tuple (R_1, \dots, R_7) of divisors of Q , we would not be able to guarantee that $|R_4/(R_4, R_1R_2R_3)| \geq q^{h/2}$, even if we weaken the conditions to allow reordering. This threshold is crucial for our argument, which does not generalize to 7-tuples.

Proof. The fact that $|R_2/(R_1, R_2)| \geq q^h$, follows directly from the third assumption, since $R_1 \neq R_2$.

For the second conclusion, let $R_{123} = \gcd(R_1, R_2, R_3)$ and let $R_{13} = (R_1, R_3)/\gcd(R_1, R_2, R_3)$ and $R_{23} = (R_2, R_3)/(R_1, R_2, R_3)$, so that R_{13} is the product of all primes dividing R_1 and R_3 but not R_2 , and vice versa. Then $(R_3, R_1R_2) = R_{13}R_{23}R_{123}$. By assumption, $|(R_2, R_3)| < q^{h/2}$, so $|R_{23}R_{123}| < q^{h/2}$, and in particular $|R_{23}| < q^{h/2}$. Now assume by contradiction that $|R_3/(R_3, R_1R_2)| < q^{h/2}$. Then

$$\left| \frac{R_3}{(R_1, R_3)} \right| = \left| \frac{R_3}{R_{13}R_{123}} \right| = \left| \frac{R_3}{R_{13}R_{23}R_{123}} \right| \cdot |R_{23}| < q^{h/2} \cdot q^{h/2} = q^h,$$

which contradicts the third assumption because $R_1 \neq R_3$. □

The following auxiliary lemma provides a standard bound on τ_k , the k -fold divisor function, in the function field setting. We will also use that $\phi(F) \gg |F|/\log \log |F|$ for all $F \in \mathbb{F}_q[t]$.

Lemma 4.2. Fix $k \geq 1$. Let $M = \max_{b \geq 1} (\tau_k(t^b))^{1/b}$. Then

$$\limsup_{\deg F \rightarrow \infty} \frac{\log \tau_k(F) \log \log |F|}{\log |F|} = \log M,$$

and thus, for all $\varepsilon > 0$, we have $\tau_k(F) \ll_\varepsilon |F|^\varepsilon$.

Proof. The proof of the above lemma follows closely along the lines of [Shiu 1980]. We will show one direction of the statement, adapted to our setting; the other direction also follows very closely, so we omit it. Note first that

$$1 \leq (\tau_k(t^b))^{1/b} = \binom{b+k-1}{b}^{1/b} < \left(\frac{(b+k-1)e}{k-1} \right)^{(k-1)/b} \rightarrow 1$$

as $b \rightarrow \infty$, so M exists.

We now show that

$$\limsup_{\deg F \rightarrow \infty} \frac{\log \tau_k(F) \log \log |F|}{\log |F|} \geq \log M.$$

Fix b such that $\tau_k(t^b) = M^b$. Let

$$F = \prod_{\substack{\deg P=d \\ P \text{ irred.}}} P^b,$$

so that $\tau_k(F) = \prod_{\deg P=d} \tau_k(P^b) = (\tau_k(t^b))^{\pi(d; \mathbb{F}_q)} = M^{b\pi(d; \mathbb{F}_q)}$. We have that $\pi(d; \mathbb{F}_q) \sim q^d/d$ as $d \rightarrow \infty$, so that

$$\log |F| = bd \log q \pi(d; \mathbb{F}_q) \sim bq^d \log q,$$

and

$$\log \log |F| = d \log q + O(1).$$

Thus as $d \rightarrow \infty$,

$$\log \tau_k(F) = b\pi(d; \mathbb{F}_q) \log M \sim b \log M \cdot \frac{q^d}{d} \sim \frac{\log M \log |F|}{\log \log |F|},$$

so

$$\limsup_{\deg F \rightarrow \infty} \frac{\log \tau_k(F) \log \log |F|}{\log |F|} \geq \log M.$$

As mentioned above, the proof that

$$\limsup_{\deg F \rightarrow \infty} \frac{\log \tau_k(F) \log \log |F|}{\log |F|} \leq \log M$$

also follows the proof in [Shiu 1980] closely, so we omit it. □

The above bound implies that, for all $\varepsilon > 0$, $\tau_k(F) = |F|^{O(1/\log \log |F|)} = O_\varepsilon(|F|^\varepsilon)$.

Here we have a final preparatory lemma before the main proposition leading to the bound on $m_5(Q; h)$. In what follows, our main strategy will be carefully isolating factors of the R_i 's in order to bound the number of terms in our sum. In doing so, we will make use of the following bound.

Lemma 4.3. *Let $Q \in \mathbb{F}_q[t]$ be a square-free polynomial, and let $n \in \mathbb{N}_{\geq 2}$. Let $\mathcal{I} \subseteq \mathbb{F}_q(t)$ be an interval of size q^{-h} . That is to say, for some rational function $\alpha \in \mathbb{F}_q(t)$, let $\mathcal{I} := \{\beta \in \mathbb{F}_q(t) : |\alpha - \beta| < q^{-h}\}$. Assume in the following that $X_i, Y_i \in \mathbb{F}_q[t]$ for all i . Then, for any $\varepsilon > 0$,*

$$\sum_{\substack{Y_1, \dots, Y_n | Q \\ X_i \in \mathcal{R}(Y_i) \\ \sum_i X_i/Y_i \in \mathcal{I} \\ q^{h/2} \leq |\prod_i Y_i| \leq q^{2h}}} \frac{\mu(\prod_i Y_i)^2}{\prod_i \phi(Y_i)^2} \ll_{n, \varepsilon} q^{-h(1-\varepsilon)}.$$

Proof. For given X_1, \dots, X_n and Y_1, \dots, Y_n , let X and Y be defined so that $Y = \prod_i Y_i$ and $X/Y = \sum_i X_i/Y_i$. Then for all tuples considered in the sum, $X/Y \in \mathcal{I}$ and $q^{h/2} \leq |Y| \leq q^{2h}$. Proceed by counting the number of possibilities for X/Y satisfying this constraint, which is bounded above by the number of points in \mathcal{I} with denominator smaller than q^{2h} , and finally count the number of ways of splitting Y up into Y_1, \dots, Y_n . However, we want to also consider the weighting in the sum of $1/\phi(Y)^2$, so we start by splitting the sum up into different sizes of Y , and then applying bounds on $\phi(Y)$.

To begin with, we rewrite the sum in terms of X and Y . Note that all Y_i in our sum are relatively prime, because of the Möbius factor. Thus Y is square-free and $\phi(Y) = \prod_i \phi(Y_i)$. Moreover, a choice of X, Y , and a decomposition $Y = Y_1 \cdots Y_n$ determines X_i for each i by the Chinese remainder theorem. Our sum is thus equal to

$$\sum_{\substack{Y | Q \\ q^{h/2} \leq |Y| \leq q^{2h}}} \sum_{\substack{X \in \mathcal{R}(Y) \\ X/Y \in \mathcal{I}}} \frac{\mu(Y)^2}{\phi(Y)^2} \#\{Y_1, \dots, Y_n : Y_1 \cdots Y_n = Y\}.$$

Now split the sum up according to $|Y|$, defining $m := \deg Y$. The sum is then equal to

$$\sum_{m=h/2}^{2h} \sum_{\substack{Y | Q \\ |Y|=q^m}} \sum_{\substack{X \in \mathcal{R}(Y) \\ X/Y \in \mathcal{I}}} \frac{\mu(Y)^2}{\phi(Y)^2} \tau_n(Y) \ll_{n, \varepsilon} \sum_{m=h/2}^{2h} (q^m)^{\varepsilon/3} \sum_{\substack{Y | Q \\ |Y|=q^m}} \sum_{\substack{X \in \mathcal{R}(Y) \\ X/Y \in \mathcal{I}}} \frac{\mu(Y)^2 (\log \log |Y|)^2}{|Y|^2},$$

by Lemma 4.2 and the fact that $\phi(Y)^{-2} \ll (|Y|/\log \log |Y|)^{-2}$. We can further relax the condition that $|Y| = q^m$ to the condition that $|Y| \leq q^m$. The number of X/Y with $|Y| \leq q^m$ in the interval \mathcal{I} is $q^{2m-h} + O(1)$; since $m \geq h/2$, this is $\ll q^{2m-h}$. Thus the sum is

$$\ll_{n, \varepsilon} \sum_{m=h/2}^{2h} q^{m(\varepsilon/3)} \frac{(\log \log(q^m))^2}{q^{2m}} q^{2m-h} \ll q^{-h} \sum_{m=h/2}^{2h} q^{m(2\varepsilon/3)} \ll q^{-h(1-\varepsilon)},$$

as desired. □

We now turn to bounding the contribution to the fifth moment $m_5(Q; h)$ coming from tuples (R_1, \dots, R_5) satisfying the conclusions of Lemma 4.1.

Proposition 4.4. Fix $h \geq 1$ and let $Q \in \mathbb{F}_q[t]$ be square-free. Let \mathcal{S} be the set of tuples (R_1, \dots, R_5) such that

- $R_i \mid Q$ for all i ,
- $q^h \leq |R_i| \leq q^{2h}$ for all i ,
- $|R_2/(R_1, R_2)| \geq q^{h/2}$, and
- $|R_3/(R_3, R_1 R_2)| \geq q^{h/2}$.

Then, for all $\varepsilon > 0$,

$$\sum_{(R_1, \dots, R_5) \in \mathcal{S}} \prod_{i=1}^5 \frac{1}{\phi(R_i)} \sum_{\substack{A_i \in \mathcal{R}(R_i) \\ |A_i/R_i| < q^{-h} \\ \sum_{1 \leq i \leq 5} A_i/R_i = 0}} q^{5h} \ll q^{(2+\varepsilon)h} \frac{|Q|}{\phi(Q)}.$$

Proof. We begin by sketching an overview of the strategy. For each subset $I \subseteq [5]$, let

$$R_I = \prod_{\substack{P \mid R_i \forall i \in I \\ P \nmid R_j \forall i \notin I}} P$$

be the product of the irreducible factors dividing R_i if and only if $i \in I$. Note that these R_I 's must be pairwise relatively prime.

We start by using the constraint that $|A_1/R_1| < q^{-h}$. We will count the total number of rational functions in this interval with denominator of degree at most $2h$. For each option of A_1/R_1 , we can decompose $R_1 = \prod_{I \ni 1} R_I$, so the number of ways to decompose R_1 into these R_I -factors is $\tau_{2k-1-1}(R_1)$, which we can bound based on the degree of R_1 . We then also get $A_1/R_1 = \sum_{I \ni 1} A_I/R_I$, where the A_I 's are determined by the Chinese remainder theorem.

We will then focus on the constraint that $|A_2/R_2| < q^{-h}$. However, $(R_1, R_2) = \prod_{1,2 \in I} R_I$ has already been fixed, so the same analysis as used for R_1 applies to the remaining factors of R_2 . Crucially, $R_2/(R_1, R_2)$ remains relatively large by assumption, which will ensure that we save enough by doing this. Finally, the constraint on A_3/R_3 , using our assumption that $R_3/(R_3, R_1 R_2)$ is large enough, yields savings in the same way.

We begin by rewriting our sum in terms of the R_I . For each subset $I \subseteq [5]$, and for a fixed R_1, \dots, R_5 , we again define R_I to be the product of all primes P so that P divides R_i for each $i \in I$ and P does not divide R_j for all $j \notin I$. The R_I are a system of *relative greatest common divisors*; see [Elsholtz and Planitzer 2020] for details. For example, $R_{\{1,2\}}$ is the product of all primes dividing R_1 and R_2 , but $(R_{\{1,2\}}, R_j) = 1$ for $j = 3, 4, 5$. The polynomials R_I must satisfy the following properties, implied by the constraints on the R_i 's:

- Each R_I divides Q , and, for each $I \neq J \subseteq [5]$, $(R_I, R_J) = 1$.
- Each irreducible polynomial dividing an R_i must divide at least two of them in order for the sum over A_i to be nonempty, so $R_I = 1$ unless $|I| \geq 2$. We will always assume that $|I| \geq 2$.

- Each choice of A_i is equivalent to a choice of $A_{i,I}$ for all subsets I containing i , that is, $A_i/R_i = \sum_{I \ni i} A_{i,I}/R_I$.
- The quantity (A_i, R_i) is equal to 1 for all i if and only if $(A_{i,I}, R_I)$ is equal to 1 for all I, i .
- The constraint that, for all i , $|A_i/R_i| < q^{-h}$, implies that, for each index i ,

$$\left| \sum_{I \ni i} \frac{A_{i,I}}{R_I} \right| < q^{-h}.$$

- The constraint that $\sum_{i=1}^5 A_i/R_i = 0$ implies that, for each subset I ,

$$\sum_{i \in I} A_{i,I} = 0.$$

Finally, define ℓ_I to be the minimum element of a subset $I \subseteq [5]$. The requirement that $(R_1, \dots, R_5) \in \mathcal{S}$ implies the following:

- For all i ,

$$q^h \leq \left| \prod_{I \ni i} R_I \right| \leq q^{2h}.$$

- Since $R_2/(R_1, R_2) = \prod_{\ell_I=2} R_I$, and $R_3/(R_1, R_2, R_3) = \prod_{\ell_I=3} R_I$,

$$\left| \prod_{\ell_I=2} R_I \right| \geq q^{h/2} \quad \text{and} \quad \left| \prod_{\ell_I=3} R_I \right| \geq q^{h/2}.$$

The sum under consideration is then

$$\ll q^{5h} \sum_{\substack{R_I | Q \\ I \subseteq [5]}} \frac{\mu(\prod_I R_I)^2}{\prod_I \phi(R_I)^{|I|}} \sum_{\substack{I, i \in I \\ A_{i,I} \in \mathcal{R}(R_I)}} 1.$$

$$\begin{array}{l} q^h \leq | \prod_{I \ni i} R_I | \leq q^{2h} \\ | \prod_{\ell_I=2} R_I | \geq q^{h/2} \\ | \prod_{\ell_I=3} R_I | \geq q^{h/2} \end{array} \quad \forall i, \begin{array}{l} | \sum_{I \ni i} A_{i,I}/R_I | < q^{-h} \\ \forall I, \sum_{i \in I} A_{i,I} = 0 \end{array}$$

Note first that if m_i is the maximum element of a subset I , then $A_{m_i,I}$ is fully determined by the other $A_{i,I}$ and the fact that $\sum_{i \in I} A_{i,I} = 0$. Then for $i \in I$ with $\ell_I < i < m_I$, we will use the trivial bound on the number of options for $A_{i,I}$; namely that there are at most R_I choices for $A_{i,I}$. For the rest of this bound, we treat $A_{i,I}$ as fixed when $\ell_I < i < m_I$.

We finally consider the number of options for the remaining $A_{\ell_I,I}$, where ℓ_I is the smallest element in I , which is where the savings in the argument will come from. We will proceed by ordering the intervals I in our sum by ℓ_I ; we will first sum over options for A_I when $I = \{4, 5\}$, with $\ell_I = 4$, and then over $A_{i,I}$ for all I with $\ell_I = 3$, and so on. As we do this, we will need at each step to satisfy the constraints that, for each i ,

$$\left| \sum_{I \ni i} \frac{A_{i,I}}{R_I} \right| < q^{-h}, \tag{17}$$

where as we split up the sums over different $A_{i,I}$'s, some of the values in this sum will be fixed and others will still be free to vary in our sum. But even if some of the terms in the sum above are fixed, the remaining terms are still constrained to lie in some interval of size q^{-h} , possibly an interval centered at a nonzero rational function. In particular, the constraints in (17) are equivalent to the constraints that, for all i ,

$$\left| F_i + \sum_{\substack{I \subseteq [5] \\ \ell_I = i}} \frac{A_{i,I}}{R_I} \right| < q^{-h},$$

where F_i is a fixed rational function determined by the values of $A_{i,I}$ when $\ell_I < i < m_I$. The bounds we use are independent F_i , only requiring that the size of the interval is q^{-h} , so we can replace F_i by 0. This yields

$$\ll q^{5h} \sum_{\substack{R_I | Q \\ I \subseteq [5] \\ q^h \leq |\prod_{I \ni i} R_I| \leq q^{2h} \\ |\prod_{\ell_I=2} R_I| \geq q^{h/2} \\ |\prod_{\ell_I=3} R_I| \geq q^{h/2}}} \frac{\mu(\prod_I R_I)^2}{\prod_I \phi(R_I)^{|I|}} \prod_I \phi(R_I)^{|I|-2} \sum_{\substack{A_{\ell_I,I} \in \mathcal{R}(R_I) \\ I \subseteq [5] \\ \forall i, |\sum_{\ell_J=i} A_{i,J}/R_J| < q^{-h}}} 1. \tag{18}$$

The only terms $A_{i,I}$ that remain in (18) are of the form $A_{\ell_I,I}$, and there is only one term for each subset I , so to simplify our notation we will write $A_I := A_{\ell_I,I}$ from now on.

Consider subsets I with $\ell_I = 4$. There is only one of these, namely $\{4, 5\}$, so we rewrite the sum as

$$\ll q^{5h} \sum_{\substack{R_I | Q \\ I \subseteq [5], I \neq \{4,5\} \\ q^h \leq |\prod_{I \ni i} R_I| \leq q^{2h} \\ |\prod_{\ell_I=2} R_I| \geq q^{h/2} \\ |\prod_{\ell_I=3} R_I| \geq q^{h/2}}} \frac{\mu(\prod_I R_I)^2}{\prod_I \phi(R_I)^2} \sum_{\substack{A_I \in \mathcal{R}(R_I) \\ I \subseteq [5], I \neq \{4,5\} \\ \forall i, |\sum_{\ell_J=i} A_J/R_J| < q^{-h}}} \sum_{\substack{R_{\{4,5\}} | Q \\ A_{\{4,5\}} \in \mathcal{R}(R_{\{4,5\}})}} \frac{1}{\phi(R_{\{4,5\}})^2}.$$

In the inside sum, we have dropped the additional constraint that $A_{\{4,5\}}/R_{\{4,5\}}$ must lie in an interval of size q^{-h} , since ignoring it only increases the size of the sum. For each $R_{\{4,5\}}$, there are $\phi(R_{\{4,5\}})$ choices of $A_{\{4,5\}}$, so the inner sum becomes

$$\sum_{R_{\{4,5\}} | Q} \frac{1}{\phi(R_{\{4,5\}})} = \frac{|Q|}{\phi(Q)},$$

since Q is square-free.

Now consider subsets I with $\ell_I = 3$, i.e., $\{3, 4\}$, $\{3, 4, 5\}$, and $\{3, 5\}$. We first bookkeep by isolating these terms in the sum, yielding

$$\ll q^{5h} \frac{|Q|}{\phi(Q)} \sum_{\substack{R_I | Q \\ I \subseteq [5], \ell_I < 3 \\ q^h \leq |\prod_{\ell_I=1} R_I| \leq q^{2h} \\ q^{h/2} \leq |\prod_{\ell_I=2} R_I| \leq q^{2h}}} \frac{\mu(\prod_I R_I)^2}{\prod_I \phi(R_I)^2} \sum_{\substack{A_I \in \mathcal{R}(R_I) \\ I \subseteq [5], \ell_I < 3 \\ \forall i, |\sum_{\ell_J=i} A_J/R_J| < q^{-h}}} \sum_{\substack{\ell_I=3 \\ R_I | Q \\ A_I \in \mathcal{R}(R_I) \\ q^{h/2} \leq |\prod_{\ell_I=3} R_I| \leq q^{2h} \\ |\sum_{\ell_I=3} A_I/R_I| < q^{-h}}} \frac{\mu(\prod_{\ell_I=3} R_I)^2}{\prod_{\ell_I=3} \phi(R_I)^2}.$$

We now bound the inner sum using Lemma 4.3. The inner sum comprises three terms R_I , so apply the lemma with $n = 3$ to get that the inner sum is $\ll q^{-h(1+\varepsilon)}$.

We repeat the process, now considering subsets I with $\ell_I = 2$. Isolating these terms yields

$$\ll q^{4h+\varepsilon h} \frac{|Q|}{\phi(Q)} \sum_{\substack{R_I | Q \\ I \subseteq [5], \ell_I=1 \\ q^h \leq |\prod_{\ell_I=1} R_I| \leq q^{2h}}} \frac{\mu(\prod_{\ell_I=1} R_I)^2}{\prod_{\ell_I=1} \phi(R_I)^2} \sum_{\substack{A_I \in \mathcal{R}(R_I) \\ I \subseteq [5], \ell_I=1 \\ |\sum_{\ell_I=1} A_I/R_I| < q^{-h}}} \sum_{\substack{\ell_I=2 \\ R_I | Q \\ A_I \in \mathcal{R}(R_I) \\ q^{h/2} \leq |\prod_{\ell_I=2} R_I| \leq q^{2h} \\ |\sum_{\ell_I=2} A_I/R_I| < q^{-h}}} \frac{\mu(\prod_{\ell_I=2} R_I)^2}{\prod_{\ell_I=2} \phi(R_I)^2}.$$

Here there are seven R_I terms and seven A_I terms in the inner sum, so, again applying Lemma 4.3, the inner sum is $\ll q^{-h+\varepsilon h}$. Lastly, we address the terms with $\ell_I = 1$:

$$\ll q^{3h} q^{2\varepsilon h} \frac{|Q|}{\phi(Q)} \sum_{\substack{R_I | Q \\ I \subseteq [5], \ell_I=1 \\ q^h \leq |\prod_{\ell_I=1} R_I| \leq q^{2h}}} \frac{\mu(\prod_{\ell_I=1} R_I)^2}{\prod_{\ell_I=1} \phi(R_I)^2} \sum_{\substack{A_I \in \mathcal{R}(R_I) \\ I \subseteq [5], \ell_I=1 \\ |\sum_{\ell_I=1} A_I/R_I| < q^{-h}}} 1.$$

We apply Lemma 4.3 one final time, this time with $n = 15$, since there are 15 sets $I \subseteq [5]$ with $|I| \geq 2$ and $\ell_I = 1$. This yields

$$\ll q^{2h+3\varepsilon h} \frac{|Q|}{\phi(Q)},$$

as desired. □

We are now ready to prove a general bound on $m_5(Q; h)$.

Theorem 4.5. Fix $\varepsilon > 0$ and let $Q \in \mathbb{F}_q[t]$ be square-free. Define $m_5(Q; h)$ by (15). Then

$$m_5(Q; h) \ll |Q| q^{2h+\varepsilon} \left(\frac{|Q|}{\phi(Q)}\right)^{-4} + |Q| q^{2h} \left(\frac{|Q|}{\phi(Q)}\right)^{27}.$$

Proof. Using Lemma 3.1, we can express

$$m_5(Q; h) = |Q| \left(\frac{\phi(Q)}{|Q|}\right)^5 V_5(Q; h),$$

where

$$V_5(Q; h) = \sum_{\substack{R_1, \dots, R_5 | Q \\ |R_i| > 1 \\ R_i \text{ monic}}} \prod_{i=1}^5 \frac{\mu(R_i)}{\phi(R_i)} \sum_{\substack{A_1, \dots, A_5 \in \mathcal{R}(R_i) \\ \sum_i A_i/R_i = 0}} E\left(\frac{A_1}{R_1}\right) \cdots E\left(\frac{A_5}{R_5}\right).$$

Now apply Lemma 3.6 to bound the contribution to $V_5(Q; h)$ from many tuples R_1, \dots, R_5 . If $|R_i| < q^h$ for any i , then these terms contribute 0; assume from now on that $|R_i| \geq q^h$. If for any triple i, j, k we have $R_i = R_j = R_k$, then we apply Lemma 3.6 with $\tilde{R}_1 = R_i$ and $\tilde{R}_2 = R_j$; in this case $X_2 = 0$ and X_1 and X_3 are $O(q^{-h/2})$, so these terms contribute $O(q^{2h}(|Q|/\phi(Q))^{32})$. If there exist $R_i \neq R_j$ such that either $|R_i/(R_i, R_j)| < q^h$ or $|(R_i, R_j)| \geq q^{h/2}$, then we apply Lemma 3.6 with $\tilde{R}_1 = R_i$ and $\tilde{R}_2 = R_j$; in this case, $X_3 = 0$, and X_1 and X_2 are each $O(q^{-h/2})$, so these terms contribute $O(q^{2h}(|Q|/\phi(Q))^{32})$ as well.

Assume now that $(R_1, R_2, R_3, R_4, R_5)$ does not fall into either of the above cases. Then, for all i , $|R_i| < q^{2h}$. To see this, assume that $(R_1, R_2, R_3, R_4, R_5)$ has no i, j, k with $R_i = R_j = R_k$, and that, for all $R_i \neq R_j$, $|R_i/(R_i, R_j)| \geq q^h$ and $|(R_i, R_j)| < q^{h/2}$. Assume, relabeling if necessary, that $R_1 \geq q^{2h}$. Since $R_1 \mid \prod_{j \neq 1} (R_1, R_j)$, we must have $|(R_1, R_j)| \geq q^{h/2}$ for some $j \neq 1$. This cannot be true for some j with $R_j \neq R_1$, so we have $R_j = R_1$. At the same time, there can only be one $j \neq 1$ with $R_j = R_1$, so without loss of generality our tuple must be of the form $(R_1, R_1, R_3, R_4, R_5)$. There cannot be an additional equal pair among R_3, R_4 , and R_5 ; if there is (without loss of generality $R_3 = R_4$), then $R_5 \mid (R_1, R_5)(R_3, R_5)$, so since $|R_5| \geq q^h$ either $|(R_1, R_5)| \geq q^{h/2}$ or $|(R_3, R_5)| \geq q^{h/2}$, which along with the lack of equal triples yields a contradiction. Now consider R_3 . Note that $R_3 \mid (R_1, R_3)(R_4, R_3)(R_5, R_3)$ and $(R_3/(R_1, R_3)) \mid (R_4, R_3)(R_5, R_3)$. But by assumption, $|R_3/(R_1, R_3)| \geq q^h$ and $|(R_4, R_3)(R_5, R_3)| < (q^{h/2})^2 = q^h$, which yields a contradiction.

So, the only terms remaining are those with $|R_i| < q^{2h}$ for all i , no equal triple, and either $|(R_i, R_j)| \geq q^{h/2}$ or $|R_i/(R_i, R_j)| < q^h$ whenever $R_i \neq R_j$. By Lemma 4.1, (R_1, \dots, R_5) satisfies the constraints of Proposition 4.4. By Proposition 4.4, these terms contribute $O(q^{(2+\varepsilon)h} |Q|/\phi(Q))$ to $V_5(Q; h)$ for all $\varepsilon > 0$. Thus, for all $\varepsilon > 0$,

$$V_5(Q; h) \ll q^{(2+\varepsilon)h} \frac{|Q|}{\phi(Q)} + q^{2h} \left(\frac{|Q|}{\phi(Q)} \right)^{32},$$

so

$$m_5(Q; h) \ll |Q|q^{(2+\varepsilon)h} \left(\frac{|Q|}{\phi(Q)} \right)^{-4} + |Q|q^{2h} \left(\frac{|Q|}{\phi(Q)} \right)^{27}. \quad \square$$

As in the integer case, we particularly want to consider Q to be the product of irreducible polynomials P with $|P| \leq q^{2h}$. In this case, $|Q|/\phi(Q) \ll h$, so that we get the following corollary.

Corollary 4.6. Fix $\varepsilon > 0$ and let $Q \in \mathbb{F}_q[t]$ be given by

$$Q = \prod_{\substack{P \text{ irred.} \\ |P| \leq q^{2h}}} P.$$

Then

$$m_5(Q; h) \ll_\varepsilon |Q|q^{(2+\varepsilon)h}.$$

4.2. Proof of Corollary 1.5: bounds on $R_k(q^h)$. In this subsection, we discuss the transition from bounds on $V_k(Q; h)$, from Theorem 1.4 and Lemma 3.7, to bounds on sums of singular series in function fields, in order to prove Corollary 1.5. Much of this is similar to the integer case discussion in Section 2.

As in the integer case, for $\mathcal{D} = \{D_1, \dots, D_k\}$ a set of distinct polynomials in $\mathbb{F}_q[T]$, we define the singular series

$$\mathfrak{S}(\mathcal{D}) := \prod_{P \text{ monic, irred.}} \frac{(1 - v_P(\mathcal{D})/|P|)}{(1 - 1/|P|)^k} = \sum_{\substack{R_1, \dots, R_k \\ 1 \leq |R_i|}} \left(\prod_{i=1}^k \frac{\mu(R_i)}{\phi(R_i)} \right) \sum_{\substack{A_1, \dots, A_k \\ A_i \in \mathcal{R}(R_i) \\ \sum_i A_i/R_i = 0}} e\left(\sum_{i=1}^k \frac{A_i D_i}{R_i} \right),$$

where $\nu_P(\mathcal{D})$ is the number of equivalence classes of $\mathbb{F}_q[T]/(P)$ occupied by elements of \mathcal{D} . We also define $\mathfrak{S}_0(\mathcal{D})$, given by $\mathfrak{S}_0(\mathcal{D}) := \sum_{\mathcal{J} \subseteq \mathcal{D}} (-1)^{|\mathcal{D} \setminus \mathcal{J}|} \mathfrak{S}(\mathcal{J})$, and consider

$$R_k(q^h) := \sum_{\substack{D_1, \dots, D_k \\ D_i \text{ distinct} \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, \dots, D_k\}). \tag{19}$$

Our results on $m_k(Q; h)$ (and equivalently $V_k(Q; h)$) imply bounds on these sums of k -fold singular series, just as in the integer case in Section 2. We set Q to be the product of all monic irreducible polynomials of degree at most $2h$, so that $|Q|/\phi(Q) \ll_q h$. Just as in the integer case, we can truncate the expression for $\mathfrak{S}_0(\mathcal{D})$ to only contain terms dividing Q , with an acceptable error term. In particular, we get

$$R_k(h) = \sum_{\substack{D_1, \dots, D_k \\ D_i \text{ distinct} \\ |D_i| \leq q^h}} \sum_{\substack{R_1, \dots, R_k \\ |R_i| > 1 \\ R_i | Q}} \prod_{i=1}^k \frac{\mu(R_i)}{\phi(R_i)} \sum_{\substack{A_1, \dots, A_k \\ A_i \in \mathcal{R}(R_i) \\ \sum_i A_i/R_i = 0}} e\left(\sum_{i=1}^k \frac{D_i A_i}{R_i}\right) + O(1).$$

It will again be helpful for us to define the singular series of a k -tuple $\mathcal{D} = (D_1, \dots, D_k)$ relative to the modulus Q . Here the k -tuple can have repeated elements; since the Euler product is finite, convergence is not a concern. We define

$$\mathfrak{S}(\mathcal{D}; Q) := \prod_{\substack{P | Q \\ P \text{ monic}}} \frac{(1 - \nu_P(\mathcal{D})/|P|)}{(1 - 1/|P|)^k} = \sum_{\substack{R_1, \dots, R_k | Q \\ R_i \text{ monic}}} \left(\prod_{i=1}^k \frac{\mu(R_i)}{\phi(R_i)} \right) \sum_{\substack{A_1, \dots, A_k \\ A_i \in \mathcal{R}(R_i) \\ \sum_i A_i/R_i = 0}} e\left(\sum_{i=1}^k \frac{A_i D_i}{R_i}\right).$$

If \mathcal{D} has a repeated element, so that $\mathcal{D} = \{D, D, D_3, \dots, D_k\}$, then

$$\mathfrak{S}(\mathcal{D}; Q) = \frac{|Q|}{\phi(Q)} \mathfrak{S}(\{D, D_3, \dots, D_k\}; Q),$$

so we can remove repeated elements from \mathcal{D} at the expense of a factor of $|Q|/\phi(Q)$. We define $\mathfrak{S}_0(\mathcal{D}; Q)$ to be the alternating sum $\sum_{\mathcal{J} \subseteq \mathcal{D}} (-1)^{|\mathcal{D} \setminus \mathcal{J}|} \mathfrak{S}(\mathcal{J}; Q)$, so we have

$$R_k(q^h) = \sum_{\substack{D_1, \dots, D_k \\ D_i \text{ distinct} \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, \dots, D_k\}; Q) + O(1).$$

This is quite close to the quantity $V_k(Q; h)$, except with the added constraint that the D_i 's must be distinct. It suffices to remove this condition. To do so, we put $\delta_{ij} = 1$ if $D_i = D_j$ and 0 otherwise, so that

$$\sum_{\substack{D_1, \dots, D_k \\ D_i \text{ distinct} \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, \dots, D_k\}; Q) = \sum_{\substack{D_1, \dots, D_k \\ |D_i| \leq q^h}} \left(\prod_{1 \leq i < j \leq k} (1 - \delta_{ij}) \right) \mathfrak{S}_0(\{D_1, \dots, D_k\}; Q).$$

We can expand the product and group terms according to which D_i 's are required to be equal, noting that, for example, $\delta_{12}\delta_{23} = \delta_{13}\delta_{23}$. We can also combine terms according to symmetry; the term δ_{12} and the term δ_{34} will have identical contributions in the final sum.

Let us now proceed with analyzing $R_5(q^h)$. After some counting, we get that

$$\sum_{\substack{D_1, \dots, D_5 \\ D_i \text{ distinct} \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, \dots, D_5\}; Q) = \sum_{\substack{D_1, \dots, D_5 \\ |D_i| \leq q^h}} f((\delta_{i,j})_{i,j \in [5]}) \mathfrak{S}_0(\{D_1, \dots, D_5\}; Q),$$

where

$$f((d_{i,j})_{i,j \in [5]}) = 1 - 10\delta_{12} + 20\delta_{12}\delta_{13} + 15\delta_{12}\delta_{34} - 20\delta_{12}\delta_{13}\delta_{45} - 30\delta_{12}\delta_{13}\delta_{14} + 24\delta_{12}\delta_{13}\delta_{14}\delta_{15}.$$

We will consider the contribution from each term in f . The term 1 gives us precisely $V_5(Q; h)$, which we have already analyzed. We can then bound each of the remaining six terms by expanding \mathfrak{S}_0 into a sum of \mathfrak{S} , removing any repeated terms in the appropriate tuple, and applying Lemma 3.7 to bound $V_k(Q; h)$ for some $k < 5$. These computations are summarized in the following lemma.

Lemma 4.7. *Using the notation of this section,*

- (a) $\sum_{\substack{D_1, D_3, D_4, D_5 \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, D_1, D_3, D_4, D_5\}; Q) \ll q^{2h} (|Q|/\phi(Q))^9,$
- (b) $\sum_{\substack{D_1, D_4, D_5 \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, D_1, D_1, D_4, D_5\}; Q) \ll q^{2h} (|Q|/\phi(Q))^3 + q^h (|Q|/\phi(Q))^{10},$
- (c) $\sum_{\substack{D_1, D_3, D_5 \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, D_1, D_3, D_3, D_5\}; Q) \ll q^{2h} (|Q|/\phi(Q))^3 + q^h (|Q|/\phi(Q))^{10},$
- (d) $\sum_{\substack{D_1, D_4 \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, D_1, D_1, D_4, D_4\}; Q) \ll q^{2h} (|Q|/\phi(Q))^3 + q^h (|Q|/\phi(Q))^4,$
- (e) $\sum_{\substack{D_1, D_5 \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, D_1, D_1, D_1, D_5\}; Q) \ll q^h (|Q|/\phi(Q))^4,$
- (f) $\sum_{\substack{D_1 \\ |D_1| \leq q^h}} \mathfrak{S}_0(\{D_1, D_1, D_1, D_1, D_1\}; Q) \ll q^h (|Q|/\phi(Q))^4.$

Proof. For the sake of brevity we omit most of these computations, which are very similar, but we will show that the term corresponding to δ_{12} , in part (a), is $\ll q^{2h} (|Q|/\phi(Q))^9$.

Assume we have a tuple $\mathcal{D} = \{D_1, D_1, D_3, D_4, D_5\}$, with one repeated term. As mentioned above,

$$\mathfrak{S}(\mathcal{D}; Q) = \frac{|Q|}{\phi(Q)} \mathfrak{S}(\{D_1, D_3, D_4, D_5\}; Q).$$

Expanding \mathfrak{S}_0 and applying this relation shows that

$$\mathfrak{S}_0(\mathcal{D}; Q) = \left(\frac{|Q|}{\phi(Q)} - 2 \right) \mathfrak{S}_0(\{D_1, D_3, D_4, D_5\}; Q) + \left(\frac{|Q|}{\phi(Q)} - 1 \right) \mathfrak{S}_0(\{D_3, D_4, D_5\}; Q),$$

so in this way we can remove repeated elements from our sum. The term we want to bound is

$$\begin{aligned} \sum_{\substack{D_1, D_3, D_4, D_5 \\ |D_i| \leq q^h}} \mathfrak{S}_0(\{D_1, D_1, D_3, D_4, D_5\}; Q) \\ &= \sum_{\substack{D_1, D_3, D_4, D_5 \\ |D_i| \leq q^h}} \left(\frac{|Q|}{\phi(Q)} - 2 \right) \mathfrak{S}_0(\{D_1, D_3, D_4, D_5\}; Q) + \left(\frac{|Q|}{\phi(Q)} - 1 \right) \mathfrak{S}_0(\{D_3, D_4, D_5\}; Q) \\ &= \left(\left(\frac{|Q|}{\phi(Q)} - 2 \right) V_4(Q; h) + q^h \left(\frac{|Q|}{\phi(Q)} - 1 \right) V_3(Q; h) \right) \\ &\ll \left(\frac{|Q|}{\phi(Q)} \right)^3 q^{2h} + \left(\frac{|Q|}{\phi(Q)} \right)^9 q^{2h}, \end{aligned}$$

where in the last step the bounds follow from Lemma 3.7. □

This lemma gives the following corollary.

Corollary 4.8. *Let*

$$Q = \prod_{\substack{P \text{ irred.} \\ |P| \leq q^{6h}}} P.$$

For all $\varepsilon > 0$,

$$R_5(q^h) \ll V_5(Q; h) + q^{2h} \left(\frac{|Q|}{\phi(Q)} \right)^9 \ll q^{(2+\varepsilon)h}.$$

Performing the same analysis when $k = 3$ yields the bound:

Corollary 4.9. *Let Q be as above. Then*

$$R_3(q^h) \ll V_3(Q; h) + q^h \left(\frac{|Q|}{\phi(Q)} \right)^2 \ll q^h \left(\frac{|Q|}{\phi(Q)} \right)^8.$$

5. Numerical evidence for odd moments

Here we present several charts supporting our conjectures on the sizes of the odd moments. To begin with, we have computed $\frac{1}{6}R_3(h) = \sum_{1 \leq d_1 < d_2 < d_3 \leq h} \mathfrak{S}_0(\{d_1, d_2, d_3\})$. In Figure 1, left, $\frac{1}{6}R_3(h)$ is plotted in black. We expect $R_3(h)$, and thus also $\frac{1}{6}R_3(h)$, to be of the shape $Ah(\log h)^2$ for some constant A . We found an experimental best fit value of $A = 0.373727$, and for this A have plotted $Ah(\log h)^2$ alongside $\frac{1}{6}R_3(h)$, as a dashed gray line.

The fit of the theoretical gray dashed curve is quite close, but there are lower-order fluctuations; in Figure 1, right, we plot the difference between $\frac{1}{6}R_3(h)$ and $Ah(\log h)^2$.

Our analysis above includes relatively little discussion about the moments of the distribution of primes themselves. We have computed several third, fifth, and seventh moments of the distribution of primes, shown in Figures 2–4. Specifically, we have computed

$$\tilde{M}_k(N; N^\delta) = \frac{1}{N} \sum_{n=N}^{2N} (\psi(n + N^\delta) - \psi(n) - N^\delta)^k$$

for each of $\delta = 0.25, 0.5$ and 0.75 , and for each of $k = 3, 5, 7$. For a fixed δ and k , we plot $\tilde{M}_k(N; N^\delta)$ for values of N ranging from 1 to 10^7 , and growing exponentially.

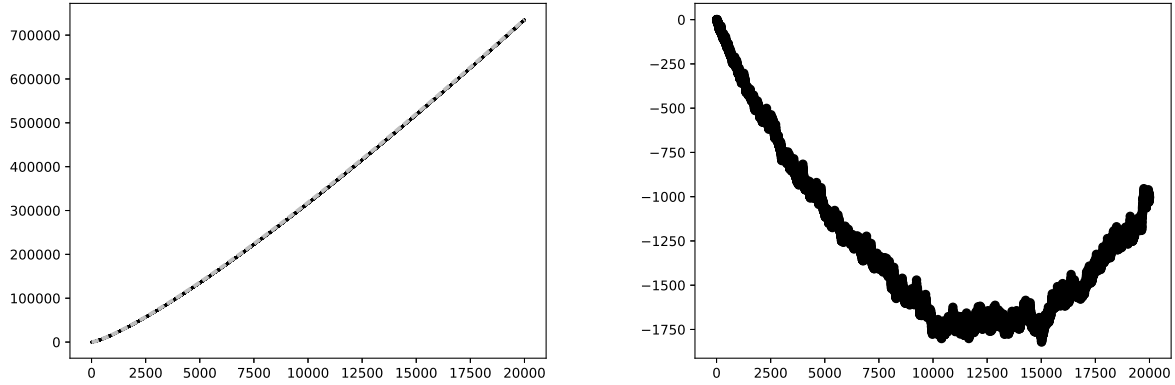


Figure 1. Left: $\frac{1}{6}R_3(h)$ for $3 \leq h \leq 20000$. Right: $\frac{1}{6}R_3(h) - Ah(\log h)^2$ for $3 \leq h \leq 20000$.

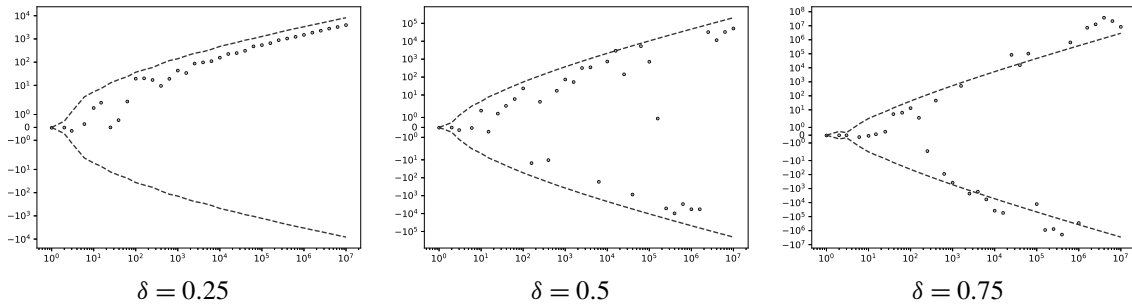


Figure 2. Plots of the third moment $M_3(N; N^\delta)$ for $N \leq 10^7$.

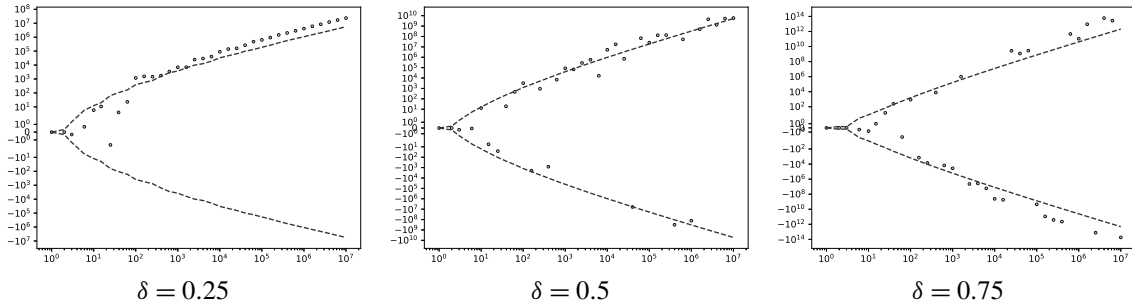


Figure 3. Plots of the fifth moment $M_5(N; N^\delta)$ for $N \leq 10^7$.

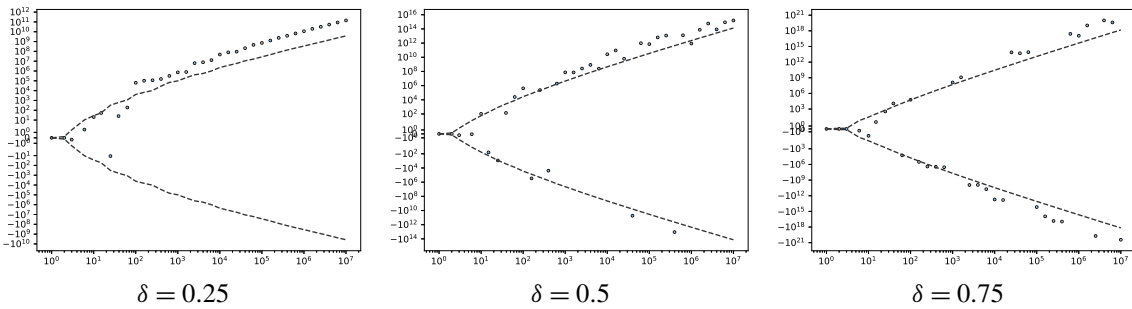


Figure 4. Plots of the seventh moment $M_7(N; N^\delta)$ for $N \leq 10^7$.

Each of the plots in Figures 2–4 is drawn with both x - and y -axes on a logarithmic scale. We expect the k -th moment to be of size approximately $O(H^{(k-1)/2}(\log(N/H))^{(k+1)/2})$, where $H = N^\delta$, so to give a sense of size, for each plot, $N^{\delta(k-1)/2}(\log N^{1-\delta})^{(k+1)/2}$ is plotted in dashed red. We have also plotted the reflection of the red dashed curve across the x -axis, since the odd moments are frequently negative.

The fit of the red line is reasonably good in all cases, but not perfect. In every case here we seem to see that the odd moments are more frequently positive than negative, but still take on negative values. For $\delta = 0.25$, the odd moments seem to be positive for sufficiently large N ; it is possible that this effect occurs for all sufficiently large N , where the threshold depends on k and δ .

6. Toy models and open problems

Throughout, we have studied the sum

$$R_k(h) = \sum_{\substack{q_1, \dots, q_k \\ 1 < q_i}} \left(\prod_{i=1}^k \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i \\ (a_i, q_i) = 1 \\ \sum a_i / q_i \in \mathbb{Z}}} \prod_{i=1}^k E\left(\frac{a_i}{q_i}\right),$$

where $E(\alpha) = \sum_{m=1}^h e(m\alpha)$. The sums $E(\alpha)$ approximately detect when $\|\alpha\| \leq 1/h$; the analogous sum in the function field case precisely detects when α has small degree. As a result, much of our understanding boils down to answering the following key question.

Question 6.1. *Let $\delta > 0$ and let $Q > 1/\delta$. What is*

$$\#\left\{q_1, \dots, q_k \in [Q, 2Q], a_i \bmod q_i : \left\| \frac{a_i}{q_i} \right\| \leq \delta, \sum_i \frac{a_i}{q_i} \in \mathbb{Z} \right\}?$$

We conjecture that the answer to this question is as follows.

Conjecture 6.2. *Let $\delta > 0$ and let $Q > 1/\delta$. Let S be the size of the set in Question 6.1. Then for any $\varepsilon > 0$,*

$$S \ll \begin{cases} Q^{k+\varepsilon} \delta^{k/2}, & k \text{ even,} \\ Q^{k+\varepsilon} \delta^{(k+1)/2}, & k \text{ odd.} \end{cases}$$

As we discussed in the Introduction, Montgomery and Vaughan [1986] considered the related problem of moments of reduced residues modulo q . Their work depends on the following answer to Question 6.1 above.

Theorem 6.3. *Let S be the size of the set in Question 6.1. Then*

$$S \ll \begin{cases} \delta^{k/2} \sum_{\substack{Q \leq r_i \leq 2Q \\ 1 \leq i \leq k/2}} \frac{r_1^2 \cdots r_{k/2}^2}{\text{lcm}(r_i)} + \delta^{k/2-1/7k} \sum_{\substack{Q \leq r_i \leq 2Q \\ 1 \leq i \leq k}} \frac{r_1 \cdots r_k}{\text{lcm}(r_i)}, & k \text{ even,} \\ \delta^{k/2-1/7k} \sum_{\substack{Q \leq r_i \leq 2Q \\ 1 \leq i \leq k}} \frac{r_1 \cdots r_k}{\text{lcm}(r_i)}, & k \text{ odd.} \end{cases}$$

The proof of the above theorem is identical to the proof in [Montgomery and Vaughan 1986]. This agrees with Conjecture 6.2 for the case when k is even, but gives a weaker bound when k is odd.

We can also consider generalizations of Question 6.1. For example, instead of specifying that $\|a_i/q_i\| \leq \delta$, we may ask that it lie in any specified interval.

Question 6.4. Let $Q > 1/\delta$ and let I_1, \dots, I_k be k intervals in $[0, 1]$ with $|I_j| \geq \delta$ for all j . What is

$$\#\left\{q_1, \dots, q_k \in [Q, 2Q], a_i \bmod q_i : \left\| \frac{a_i}{q_i} \right\| \in I_i, \sum_i \frac{a_i}{q_i} \in \mathbb{Z} \right\}?$$

Answers to these questions would give us more refined understanding of sums of singular series. The conjectures above are related to sums over $\mathfrak{S}(\{h_1, \dots, h_k\})$, where each h_i lies in the same interval $[0, h]$. We can instead ask about sums of singular series restricted to arbitrary intervals, or along arithmetic progressions. We state the following questions using smooth cutoff functions as opposed to intervals.

Question 6.5. Let Φ_1, \dots, Φ_k be smooth functions with compact support on \mathbb{R} , and let $H \in \mathbb{R}_{>0}$. What is

$$\sum_{h_1, \dots, h_k \in \mathbb{Z}} \mathfrak{S}_0(\{h_1, \dots, h_k\}) \Phi_1\left(\frac{h_1}{H}\right) \cdots \Phi_k\left(\frac{h_k}{H}\right)?$$

Question 6.6. Let Φ_1, \dots, Φ_k be smooth functions with compact support on \mathbb{R} , and let $H \in \mathbb{R}_{>0}$. For arithmetic progressions $a_1 \bmod q_1, \dots, a_k \bmod q_k$, what is

$$\sum_{\substack{h_1, \dots, h_k \in \mathbb{Z} \\ h_i \equiv a_i \pmod{q_i}}} \mathfrak{S}_0(\{h_1, \dots, h_k\}) \Phi_1\left(\frac{h_1}{H}\right) \cdots \Phi_k\left(\frac{h_k}{H}\right)?$$

Question 6.5 addresses the correlations of $\psi(x+h) - \psi(x)$ and $\psi(x+h_1+h) - \psi(x+h_1)$; in other words, the correlations of the number of primes in intervals in different places. Question 6.6 addresses the correlations of the number of primes in distinct arithmetic progressions. For both of these questions, the main term ought to come from diagonal terms where $h_1 = h_2$, for example, thus collapsing the weight function, whereas the error term ought to arise from off-diagonal contributions.

In the case when $k = 2$, Question 6.6 has been widely studied in the context of prime number races. The “Shanks–Rényi prime number race” is the following problem: Let $\pi(x; q, a)$ denote the number of primes $p \leq x$ with $p \equiv a \pmod{q}$. Then, for any n -tuple (a_1, \dots, a_n) of equivalence classes mod q that are relatively prime to q , will we have the ordering

$$\pi(x; q, a_1) > \pi(x; q, a_2) > \cdots > \pi(x; q, a_n)$$

for infinitely many integers x ? Many aspects of this question have been studied; see for example the expositions of [Granville and Martin 2006; Ford and Konyagin 2002].

Ford, Harper, and Lamzouri [2019] showed that, although any ordering appears infinitely often, for n large with respect to q , the prime number races among orderings can exhibit large biases. They rely on the fact that counts of primes in distinct progressions have negative correlations, which they arrange to produce a bias. This analysis is also connected to the work of Lemke Oliver and Soundararajan [2016],

who use averages of two-term singular series in arithmetic progressions to show bias in the distribution of consecutive primes. It is plausible that a more precise understanding of the questions above would lead to an extension of the work of Lemke Oliver and Soundararajan.

References

- [Chan 2006] T. H. Chan, “A note on primes in short intervals”, *Int. J. Number Theory* **2**:1 (2006), 105–110. MR Zbl
- [Cramér 1936] H. Cramér, “On the order of magnitude of the difference between consecutive prime numbers”, *Acta Arith.* **2**:1 (1936), 23–46. Zbl
- [Elsholtz and Planitzer 2020] C. Elsholtz and S. Planitzer, “The number of solutions of the Erdős–Straus equation and sums of k unit fractions”, *Proc. Roy. Soc. Edinburgh Sect. A* **150**:3 (2020), 1401–1427. MR Zbl
- [Ford and Konyagin 2002] K. Ford and S. Konyagin, “Chebyshev’s conjecture and the prime number race”, pp. 67–91 in *IV International Conference “Modern problems of number theory and its applications”: Current problems, Part II* (Tula, 2001), edited by V. N. Chubarikov and G. I. Arkhipov, Mosk. Gos. Univ. im. Lomonosova, Mekh.-Mat. Fak., Moscow, 2002. MR Zbl
- [Ford, Harper, and Lamzouri 2019] K. Ford, A. J. Harper, and Y. Lamzouri, “Extreme biases in prime number races with many contestants”, *Math. Ann.* **374**:1-2 (2019), 517–551. MR Zbl
- [Gallagher 1976] P. X. Gallagher, “On the distribution of primes in short intervals”, *Mathematika* **23**:1 (1976), 4–9. MR Zbl
- [Granville and Lumley 2023] A. Granville and A. Lumley, “Primes in short intervals: heuristics and calculations”, *Exp. Math.* **32**:2 (2023), 378–404. MR Zbl
- [Granville and Martin 2006] A. Granville and G. Martin, “Prime number races”, *Amer. Math. Monthly* **113**:1 (2006), 1–33. MR Zbl
- [Hayes 1966] D. R. Hayes, “The expression of a polynomial as a sum of three irreducibles”, *Acta Arith.* **11** (1966), 461–488. MR Zbl
- [Keating and Rudnick 2014] J. P. Keating and Z. Rudnick, “The variance of the number of prime polynomials in short intervals and in residue classes”, *Int. Math. Res. Not.* **2014**:1 (2014), 259–288. MR
- [Lemke Oliver and Soundararajan 2016] R. J. Lemke Oliver and K. Soundararajan, “Unexpected biases in the distribution of consecutive primes”, *Proc. Natl. Acad. Sci. USA* **113**:31 (2016), E4446–E4454. MR
- [Montgomery and Soundararajan 2002] H. L. Montgomery and K. Soundararajan, “Beyond pair correlation”, pp. 507–514 in *Paul Erdős and his mathematics, I* (Budapest, 1999), edited by G. Halász et al., Bolyai Soc. Math. Stud. **11**, János Bolyai Math. Soc., Budapest, 2002. MR Zbl
- [Montgomery and Soundararajan 2004] H. L. Montgomery and K. Soundararajan, “Primes in short intervals”, *Comm. Math. Phys.* **252**:1-3 (2004), 589–617. MR Zbl
- [Montgomery and Vaughan 1986] H. L. Montgomery and R. C. Vaughan, “On the distribution of reduced residues”, *Ann. of Math. (2)* **123**:2 (1986), 311–333. MR Zbl
- [Shiu 1980] P. Shiu, “The maximum orders of multiplicative functions”, *Quart. J. Math. Oxford Ser. (2)* **31**:122 (1980), 247–252. MR Zbl

Communicated by Andrew Granville

Received 2021-10-08 Revised 2024-05-14 Accepted 2024-06-15

vivian.kuperberg@math.ethz.ch

Departement Mathematik, ETH Zürich, Zürich, Switzerland

Efficient resolution of Thue–Mahler equations

Adela Gherga and Samir Siksek

A Thue–Mahler equation is a Diophantine equation of the form

$$F(X, Y) = a \cdot p_1^{z_1} \cdots p_v^{z_v}, \quad \gcd(X, Y) = 1$$

where F is an irreducible binary form of degree at least 3 with integer coefficients, a is a nonzero integer and p_1, \dots, p_v are rational primes. Existing algorithms for resolving such equations require computations in the field $L = \mathbb{Q}(\theta, \theta', \theta'')$, where $\theta, \theta', \theta''$ are distinct roots of $F(X, 1) = 0$. We give a new algorithm that requires computations only in $K = \mathbb{Q}(\theta)$ making it far more suited for higher degree examples. We also introduce a lattice sieving technique reminiscent of the Mordell–Weil sieve that makes it practical to tackle Thue–Mahler equations of higher degree and with larger sets of primes than was previously possible. We give several examples including one of degree 11.

Let $P(m)$ denote the largest prime divisor of an integer $m \geq 2$. As an application of our algorithm we determine all pairs (X, Y) of coprime nonnegative integers such that $P(X^4 - 2Y^4) \leq 100$, finding that there are precisely 49 such pairs.

1. Introduction

Let

$$F(X, Y) = a_0 X^d + a_1 X^{d-1} Y + \cdots + a_d Y^d \tag{1}$$

be a binary form of degree $d \geq 3$ with coefficients $a_i \in \mathbb{Z}$. Suppose F is irreducible over \mathbb{Q} . Let a be a nonzero integer and let p_1, \dots, p_v be distinct primes such that $p_i \nmid a$. The purpose of this paper is to give an efficient algorithm to solve the Thue–Mahler equation

$$F(X, Y) = a \cdot p_1^{z_1} \cdots p_v^{z_v}, \quad X, Y \in \mathbb{Z}, \gcd(X, Y) = 1, \tag{2}$$

for unknown integers X, Y , and unknown nonnegative integers z_1, \dots, z_v . The set of solutions is known to be finite by a famous result of Mahler [1933] which extends classical work of Thue [1909]. Mahler’s theorem is ineffective. The first effective bounds on the size of the solutions are due to Vinogradov and Sprindzhuk [1968] and to Coates [1970]. Vastly improved effective bounds have since been given by Bugeaud and Györy [1996a]. Evertse [1984, Corollary 2] showed that the number of solutions to (2) is at most $2 \times 7^{d^3(2v+3)}$. For $d \geq 6$, this has been improved by Bombieri [1987, Main Theorem] who showed that the number of solutions is at most $16(v+1)^2 \cdot (4d)^{26(v+1)}$.

The authors are supported by the EPSRC grant *Moduli of Elliptic curves and Classical Diophantine Problems* (EP/S031537/1). MSC2020: primary 11D59; secondary 11D61.

Keywords: Thue equation, Thue–Mahler equation, LLL, linear form in logarithms.

Besides being of independent interest, Thue–Mahler equations frequently arise in a number of contexts:

- The problem of determining all elliptic curves over \mathbb{Q} with good reduction outside a given set of primes algorithmically reduces to the problem of solving certain cubic Thue–Mahler equations (here cubic means $d = 3$). The earliest example appears to be due to Agrawal, Coates, Hunt, and van der Poorten [Agrawal et al. 1980] who used it to determine all elliptic curves over \mathbb{Q} of conductor 11. The recent paper of Bennett, Gherga and Reznitzner [Bennett et al. 2019] gives a systematic and general treatment of this approach. In fact, the link between cubic Thue–Mahler equations and elliptic curves can be used in conjunction with modularity of elliptic curves to give an algorithm for solving cubic Thue–Mahler equations as in the work of von Känel and Matschke [2023, Section 5] and of Kim [2017].
- Many Diophantine problems naturally reduce to the resolution of Thue–Mahler equations. These include the Lebesgue–Nagell equations (e.g., [Cangül et al. 2010; Soydan and Tzanakis 2016]), and Goormaghtigh’s equation (e.g., [Bennett et al. 2020]). The most striking of such applications is the reduction, due to Bennett and Dahmen [2013], of asymptotic cubic superelliptic equations to cubic Thue–Mahler equations, via the modularity of Galois representation attached to elliptic curves.

Before the current paper, the only general algorithm for solving Thue–Mahler equations was the one due to Tzanakis and de Weger [1989]. A modern implementation of this algorithm, due Hambrook [2011], has been profitably used to solve a number of low degree Thue–Mahler equations, for example in [Cangül et al. 2010; Soydan and Tzanakis 2016].

Instead of (2), we consider the equation

$$F(X, Y) = a \cdot p_1^{z_1} \cdots p_v^{z_v}, \quad X, Y \in \mathbb{Z}, \gcd(X, Y) = \gcd(a_0, Y) = 1. \quad (3)$$

Thus we have added the assumption $\gcd(a_0, Y) = 1$, where a_0 is the leading coefficient of F as in (1). This is a standard computational simplification in the subject, and is also applied by Tzanakis and de Weger. There is no loss of generality in adding this assumption in the following sense: an algorithm for solving (3) yields an algorithm for solving (2). To see this, let (X, Y) be a solution to (2) and let $b = \gcd(a_0, Y)$. Write $Y = bY'$ with $Y' \in \mathbb{Z}$. The possible values for b are the divisors of a_0 ; for each divisor b we need to solve $F(X, bY') = a \cdot p_1^{z_1} \cdots p_v^{z_v}$. Note that $F(X, bY') = bG(X, Y')$ where G has integral coefficients and leading coefficient $a'_0 = a_0/b$, which satisfies $\gcd(a'_0, Y') = 1$. The equation $bG(X, Y') = a \cdot p_1^{z_1} \cdots p_v^{z_v}$ is impossible unless $b/\gcd(a, b) = p_1^{w_1} \cdots p_v^{w_v}$ where $w_i \geq 0$, in which case

$$G(X, Y') = (a/\gcd(a, b)) \cdot p_1^{z_1 - w_1} \cdots p_v^{z_v - w_v},$$

which is now a Thue–Mahler equation of the form (3).

The approach of Tzanakis and de Weger can be summarized as follows:

- (I) First (3) is reduced to a number of ideal equations

$$(a_0X - \theta Y)\mathcal{O}_K = \mathfrak{a} \cdot \mathfrak{p}_1^{n_1} \cdots \mathfrak{p}_s^{n_s}. \quad (4)$$

Here θ is a root of the monic polynomial $a_0^{d-1}F(X/a_0, 1)$, and $K = \mathbb{Q}(\theta)$. Moreover, \mathfrak{a} is a fixed ideal of the ring of integers \mathcal{O}_K , and $\mathfrak{p}_1, \dots, \mathfrak{p}_s$ are fixed prime ideals of \mathcal{O}_K . The variables X, Y, n_1, \dots, n_s represent the unknowns.

(II) Next, each ideal equation (4) is reduced to a number of equations of the form

$$a_0X - \theta Y = \tau \cdot \delta_1^{b_1} \cdots \delta_r^{b_r} \quad (5)$$

where $\tau, \delta_1, \dots, \delta_r \in K^\times$ are fixed and X, Y, b_1, \dots, b_r are unknowns.

(III) The next step generates a very large upper bound for the exponents b_1, \dots, b_r using the theory of real, complex, and p -adic linear forms in logarithms. This bound is then considerably reduced using the LLL algorithm [Lenstra et al. 1982] applied to approximation lattices associated to these linear forms, and finally, all solutions below this reduced bound are found using the algorithm of Fincke and Pohst [1985].

To compute these approximation lattices alluded to in step (III), the algorithm of Tzanakis and de Weger relies on extensive computations in the number field $K' = \mathbb{Q}(\theta_1, \theta_2, \theta_3)$ where $\theta_1, \theta_2, \theta_3$ are distinct roots of $a_0^{d-1}F(X/a_0, 1)$, as well as p -adic completions of K' . The field K' typically has degree $d(d-1)(d-2)$, making their algorithm impractical if the degree d is large. Even if the degree d is small (say $d = 3$), we have found that the Tzanakis–de Weger algorithm runs into a combinatorial explosion of cases in step (I) if the number of primes v is large, and in step (II) if the class number h of K is large.

In this paper, we present an algorithm that builds on many of the powerful ideas in the paper of Tzanakis and de Weger but avoids computations in number fields other than the field $K = \mathbb{Q}(\theta)$ of degree d , and avoids all computations in p -adic fields or their extensions. The algorithm includes a number of refinements that circumvent the explosion of cases in steps (I) and (II). For example, to each ideal equation (4) we associate at most one equation (5); by contrast, the algorithm of Tzanakis and de Weger, typically associates h^{s-1} equations (5) to each ideal equation (4), where h is the class number of K . Moreover, inspired by the Mordell–Weil sieve (e.g., [Bruin and Stoll 2008; 2010; Bugeaud et al. 2008]), we introduce a powerful “Dirichlet sieve” that vastly improves the determination of the solutions to (5) after the LLL step, even if the remaining bound on the exponents b_i is large.

We have implemented the algorithm described in this paper in the computer algebra system Magma [Bosma et al. 1997].¹

Below we give four examples of Thue–Mahler equations solved using our implementation. Our solutions will always be subject to the assumptions

$$\gcd(X, Y) = \gcd(a_0, Y) = 1.$$

They will be given in the form $[X, Y, z_1, z_2, \dots, z_v]$. We will revisit these examples later on in the paper to illustrate the differences between our algorithm and that of Tzanakis and de Weger [1989]. We do point out that a number of recent papers also make use of our implementation, or the ideas in the present paper, to solve various Thue–Mahler equations where the degree d , or the number of primes v are large.

¹Our implementation is available from <https://github.com/adelaigherga/ThueMahler/tree/master/Code/TMSolver>.

For example in [Bennett et al. 2020; 2022; Bennett and Siksek 2023a; 2023b], our algorithm is used to solve Thue–Mahler equations of degrees 5, 20, 7 and 11 respectively.

Example 1.1. An ongoing large-scale computational project, led by Bennett, Cremona, Gherga and Sutherland, aims to provably compute all elliptic curves of conductor at most 10^6 . The method combines the approach in [Bennett et al. 2019] with our Thue–Mahler solver described in the current paper. We give an example to illustrate this application. Consider the problem of computing all elliptic curves E/\mathbb{Q} with trivial 2-torsion and conductor

$$771456 = 2^7 \cdot 3 \cdot 7^2 \cdot 41.$$

Applying Theorem 1 of [Bennett et al. 2019] results in 13 cubic Thue–Mahler equations of the form

$$a_0X^3 + a_1X^2Y + a_2XY^2 + a_3Y^3 = 3^{z_1} \cdot 7^{z_2} \cdot 41^{z_3}, \quad \gcd(X, Y) = 1,$$

whose resolution algorithmically yields the desired set of elliptic curves. The coefficients (a_0, a_1, a_2, a_3) for these 13 forms are:

$$(1, 7, 2, -2), \quad (2, 1, 0, 3), \quad (1, 4, 3, 6), \quad (3, 4, 4, 4), \quad (4, 4, 6, 3), \quad (2, 5, 0, 6), \quad (1, 7, 4, 12), \\ (3, 3, -1, 7), \quad (3, 7, 14, 14), \quad (1, 3, 17, 43), \quad (8, 12, 13, 8), \quad (4, 1, 12, -6), \quad (3, 9, 5, 19)$$

Our implementation solved all 13 of these Thue–Mahler equations and computed the corresponding elliptic curves in a total of 5.4 minutes on a single core.²

For illustration, we consider one of these 13 Thue–Mahler equations:

$$4X^3 + X^2Y + 12XY^2 - 6Y^3 = 3^{z_1} \cdot 7^{z_2} \cdot 41^{z_3}, \quad \gcd(X, Y) = 1.$$

Our implementation solved this in 41 seconds. The solutions are

$$[-3, -7, 1, 0, 1], \quad [-1, -5, 2, 2, 0], \quad [1, -1, 1, 1, 0], \\ [3, 1, 1, 2, 0], \quad [5, 11, 0, 2, 0], \quad [9, 17, 1, 2, 1], \quad [19, 23, 5, 3, 0].$$

Of the seven solutions, only $(X, Y) = (1, -1)$ gives rise to elliptic curves of conductor 771456:

$$E_1 : y^2 = x^3 - x^2 + 13655x + 2351833, \\ E_2 : y^2 = x^3 - x^2 + 3414x - 295686.$$

Example 1.2. Consider the Thue–Mahler equation

$$7X^3 + X^2Y + 29XY^2 - 25Y^3 = 2^{z_1} \cdot 3^{z_2} \cdot 17^{z_3} \cdot 37^{z_4} \cdot 53^{z_5}. \quad (6)$$

This in fact is one of the Thue–Mahler equations whose resolution, via the method of [Bennett et al. 2019], is needed to determine all elliptic curves of conductor

$$2^\alpha \cdot 3^\beta \cdot 17 \cdot 37 \cdot 53, \quad \text{where } \alpha \in \{2, 3, 4, 6, 7\} \text{ and } \beta \in \{1, 2\}.$$

²See <https://github.com/adelaigherga/ThueMahler/tree/master/GhSiData/Example1> for full computational details, as well as a list of all corresponding elliptic curves.

The class number of the cubic field associated to the cubic form in (6) is 33. As we shall see later (at the end of Section 4), our approach to dealing with the class group requires us to solve only 30 equations of the form (5), whereas in comparison the method of Tzanakis and de Weger requires the resolution of approximately 80990 equations of the form (5). For now, we merely point out that our implementation solved (6) in 2 minutes. The solutions are

$$\begin{aligned} & [19, -23, 2, 4, 0, 1, 1], \quad [13, -6, 0, 0, 1, 1, 1], \quad [-343, -463, 2, 11, 1, 0, 0], \\ & [79, -8, 0, 2, 2, 2, 0], \quad [37, -13, 2, 1, 1, 0, 2], \quad [1, 1, 2, 1, 0, 0, 0], \quad [3, 4, 0, 0, 1, 0, 0]. \end{aligned}$$

Example 1.3. Most explicit examples of the resolution of Thue–Mahler equations (3) found in the literature involve a relatively small set of primes $\{p_1, \dots, p_v\}$. The following example, aside from being an interesting Diophantine application in its own right, is intended to illustrate that our algorithm can cope with a relatively large set of primes.

For a nonzero integer m , let $P(m)$ denote the maximum prime divisor of m (where we take $P(1) = P(-1) = 0$). In this example, we solve the inequality

$$P(X^4 - 2Y^4) \leq 100, \quad \gcd(X, Y) = 1.$$

Let

$$F(X, Y) = X^4 - 2Y^4.$$

We therefore would like to solve (3) with $a = \pm 1$ and with p_1, \dots, p_v being the set of primes ≤ 100 (of which there are 25). However, it is clear that if 2 is not a fourth power modulo p , then $p \nmid (X^4 - 2Y^4)$. Thus we reduce to the much smaller set of primes $p \leq 100$ for which 2 is a fourth power. This is the set

$$\{2, 7, 23, 31, 47, 71, 73, 79, 89\}.$$

Therefore the Thue–Mahler equation we shall consider is

$$X^4 - 2Y^4 = \pm 2^{z_1} \cdot 7^{z_2} \cdot 23^{z_3} \cdot 31^{z_4} \cdot 47^{z_5} \cdot 71^{z_6} \cdot 73^{z_7} \cdot 79^{z_8} \cdot 89^{z_9}, \quad \gcd(X, Y) = 1.$$

Our implementation took roughly 3.5 days to solve this Thue–Mahler equation. There are 49 solutions (up to changing the signs of X, Y):

$$\begin{aligned} & [0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0], & [1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], \\ & [1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0], & [1, 2, 0, 0, 0, 1, 0, 0, 0, 0, 0], \\ & [1, 3, 0, 1, 1, 0, 0, 0, 0, 0, 0], & [1, 4, 0, 1, 0, 0, 0, 0, 1, 0, 0], \\ & [1, 11, 0, 1, 0, 0, 1, 0, 0, 0, 1], & [2, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0], \\ & [2, 3, 1, 0, 0, 0, 0, 0, 1, 0, 0], & [2, 27, 1, 1, 0, 2, 0, 0, 0, 1, 0], \\ & [3, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0], & [3, 2, 0, 2, 0, 0, 0, 0, 0, 0, 0], \\ & [3, 14, 0, 0, 1, 0, 1, 1, 0, 0, 0], & [4, 3, 1, 0, 0, 0, 1, 0, 0, 0, 0], \\ & [4, 5, 1, 1, 0, 0, 0, 1, 0, 0, 0], & [5, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1], \end{aligned}$$

[5, 8, 0, 1, 1, 0, 1, 0, 0, 0, 0],	[6, 5, 1, 0, 1, 0, 0, 0, 0, 0, 0],
[6, 19, 1, 0, 0, 1, 1, 0, 0, 0, 1],	[8, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1],
[10, 23, 1, 2, 0, 0, 0, 1, 0, 1, 0],	[11, 9, 0, 2, 0, 1, 0, 0, 0, 0, 0],
[11, 20, 0, 0, 0, 0, 1, 0, 1, 0, 1],	[15, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0],
[15, 13, 0, 0, 0, 0, 0, 0, 1, 0, 1],	[16, 21, 1, 0, 1, 0, 0, 0, 0, 1, 1],
[17, 5, 0, 2, 1, 0, 0, 0, 1, 0, 0],	[19, 20, 0, 4, 0, 0, 0, 0, 0, 1, 0],
[21, 11, 0, 0, 0, 1, 0, 0, 2, 0, 0],	[22, 49, 1, 0, 1, 1, 0, 0, 0, 0, 2],
[33, 13, 0, 1, 0, 0, 2, 0, 1, 0, 0],	[37, 19, 0, 0, 1, 2, 0, 0, 1, 0, 0],
[40, 13, 1, 1, 0, 1, 0, 0, 1, 1, 0],	[52, 51, 1, 2, 1, 1, 0, 0, 0, 0, 1],
[53, 44, 0, 1, 1, 1, 0, 0, 0, 1, 0],	[59, 56, 0, 0, 0, 1, 1, 1, 1, 0, 0],
[61, 48, 0, 1, 0, 0, 0, 1, 1, 0, 1],	[66, 101, 1, 0, 2, 1, 0, 0, 1, 1, 0],
[68, 43, 1, 1, 1, 2, 1, 0, 0, 0, 0],	[95, 58, 0, 1, 0, 1, 1, 0, 1, 1, 0],
[118, 101, 1, 2, 1, 0, 0, 1, 0, 0, 1],	[142, 57, 1, 1, 3, 1, 0, 0, 1, 0, 0],
[162, 137, 1, 2, 0, 0, 2, 0, 1, 0, 0],	[181, 124, 0, 0, 1, 0, 1, 0, 0, 2, 1],
[221, 295, 0, 0, 2, 0, 1, 0, 1, 1, 1],	[281, 199, 0, 1, 1, 0, 1, 1, 1, 1, 0],
[286, 283, 1, 3, 1, 0, 0, 0, 3, 0, 0],	[389, 96, 0, 4, 0, 1, 1, 0, 1, 0, 1],
[420, 437, 1, 0, 0, 1, 3, 0, 1, 0, 1].	

Example 1.4. All examples found in the literature are of Thue–Mahler equations where the form F has the property that the field $K' = \mathbb{Q}(\theta_1, \theta_2, \theta_3)$ (defined above) have small degree. As indicated above, a distinguishing feature of our algorithm is that the computations are carried out in the much smaller extension $K = \mathbb{Q}(\theta)$ (also defined above). Our last example is intended to illustrate this difference. Consider the Thue–Mahler equation,

$$5X^{11} + X^{10}Y + 4X^9Y^2 + X^8Y^3 + 6X^7Y^4 + X^6Y^5 + 6X^5Y^6 + 6X^3Y^8 + 4XY^{10} - 2Y^{11} = 2^{z_1} \cdot 3^{z_2} \cdot 5^{z_3} \cdot 7^{z_4} \cdot 11^{z_5}.$$

As usual, we let F denote the form on the left-hand side. The Galois group of F is S_{11} , therefore the field K has degree 11, whereas the field K' has degree $11 \times 10 \times 9 = 990$. Our program solved this Thue–Mahler equation in around 6.8 hours. However this time was almost entirely taken up with the computation of the class group and the units of K . Once the class group and unit computations were complete, it took only 3.6 minutes to provably determine the solutions. These are

$$[0, -1, 1, 0, 0, 0, 0], \quad [1, -1, 1, 1, 1, 0, 0], \quad [1, 1, 5, 0, 0, 0, 0], \quad [1, 2, 0, 3, 0, 1, 1].$$

1.1. Notation and organization of the paper. As before $F \in \mathbb{Z}[X, Y]$ will be a binary form of degree $d \geq 3$, irreducible in $\mathbb{Q}[X, Y]$, and with coefficients a_0, \dots, a_d as in (1). Let a be a nonzero integer and p_1, \dots, p_v

be distinct primes satisfying $p_i \nmid a$. Let

$$f(x) = a_0^{d-1} \cdot F(x/a_0, 1) = x^d + a_1x^{d-1} + a_0a_2x^{d-2} + \cdots + a_0^{d-1}a_d.$$

This is an irreducible monic polynomial with coefficients in \mathbb{Z} . Let θ be a root of f and let $K = \mathbb{Q}(\theta)$. Note that K is a number field of degree d . Write \mathcal{O}_K for the ring of integers of K . We can rewrite our Thue–Mahler equation (3) as

$$\text{Norm}(a_0X - \theta Y) = a_0^{d-1} \cdot a \cdot p_1^{z_1} \cdots p_v^{z_v}. \quad (7)$$

Note that we do not assume that $(a_0, p_i) = 1$.

The paper is organized as follows:

- (a) In Section 2 we consider the decomposition of the ideal $(a_0X - \theta Y)\mathcal{O}_K$ as a product of prime ideals. In particular, we introduce an algorithm to compare and restrict the possible valuations of all prime ideals above each of p_1, \dots, p_v .
- (b) We summarize the results of applying this algorithm in Section 3, wherein we reduce solving (7) to solving a family of ideal equations of the form (4).
- (c) In Section 4, we show that such ideal equations are either impossible due to a class group obstruction, or reduce to a single equation of the form (5). The remainder of the paper is devoted to solving equations of the form (5) where the unknowns are coprime integers X, Y and nonnegative integers b_1, \dots, b_r .
- (d) In Section 5, we recall key theorems from the theory of lower bounds for linear forms in complex and p -adic logarithms due to Matveev and to Yu.
- (e) In Section 6, with the help of these theorems, we obtain a very large upper bound on the exponents b_1, \dots, b_r in (5).
- (f) In Section 7 we show how an application of close vector algorithms allows us to obtain a substantially improved bound on the p -adic valuation of $a_0X - \theta Y$ for any prime p . This step avoids the p -adic logarithms of earlier approaches.
- (g) Section 8 uses the real and complex embeddings of K applied to (5) to obtain $d - 2$ approximate relations involving the exponents b_i . In Section 9, we set up an “approximation lattice” using these $d - 2$ approximate relations. We explain how close vector algorithms can be used to substantially reduce our bound for the exponents b_1, \dots, b_r in (5). Earlier approaches used just one of the $d - 2$ relations to construct the approximation lattice, but we explain why using just one approximate relation can fail in certain situations.
- (h) Steps (f) and (g) are applied repeatedly until no further improvements in the bounds are possible. In Section 10 we introduce an analogue of the Mordell–Weil sieve, which we call the “Dirichlet sieve” which is capable of efficiently sieving for the solutions up to the remaining bounds, thereby finally resolving the Thue–Mahler equation (3).

2. The p -part of $(a_0X - \theta Y)\mathcal{O}_K$

If \mathfrak{c} is a fractional ideal of \mathcal{O}_K , and p is a rational prime, we define the p -part of \mathfrak{c} to be the fractional ideal

$$\prod_{\mathfrak{p} \mid p} \mathfrak{p}^{\text{ord}_{\mathfrak{p}}(\mathfrak{c})}.$$

For each rational prime $p \in \{p_1, \dots, p_v\}$ of (7), we want to study the p -part of $(a_0X - \theta Y)\mathcal{O}_K$ coming from the prime ideals above p . The so-called prime ideal removal lemma in Tzanakis and de Weger compares the possible valuations of $(a_0X - \theta Y)\mathcal{O}_K$ at two prime ideals $\mathfrak{p}_1, \mathfrak{p}_2 \mid p$ to help cut down the possibilities for the p -part of $(a_0X - \theta Y)\mathcal{O}_K$. However if $\mathfrak{p}_1 \mid (a_0X - \theta Y)\mathcal{O}_K$ then this restricts the values of X and Y modulo p . Indeed, any choice of X and Y modulo p affects the valuations of $(a_0X - \theta Y)\mathcal{O}_K$ at all primes $\mathfrak{p} \mid p$. So we study all valuations at the same time, not just two of them. This enables us to give a much smaller list of possibilities for the p -part of $(a_0X - \theta Y)\mathcal{O}_K$ than in Tzanakis and de Weger, as we will see in Section 4.

Definition 2.1. Let p be a rational prime. Let L_p be a subset of the ideals \mathfrak{b} supported at the prime ideals above p . Let M_p be a subset of the set of pairs $(\mathfrak{b}, \mathfrak{p})$ where \mathfrak{b} is supported at the prime ideals above p , and $\mathfrak{p} \mid p$ is a prime ideal satisfying $e(\mathfrak{p} \mid p) = f(\mathfrak{p} \mid p) = 1$, where $e(\mathfrak{p} \mid p)$ and $f(\mathfrak{p} \mid p)$ are, respectively, the ramification index and inertial degree of \mathfrak{p} over p . We call the pair L_p, M_p *satisfactory* if for every solution (X, Y) to (3),

- (i) either the p -part of $(a_0X - \theta Y)\mathcal{O}_K$ is in L_p , or
- (ii) there is a pair $(\mathfrak{b}, \mathfrak{p}) \in M_p$ and a nonnegative integer l such that the p -part of $(a_0X - \theta Y)\mathcal{O}_K$ is equal to $\mathfrak{b}\mathfrak{p}^l$.

At this point the definition is perhaps mysterious. Lemma 2.4 and the following remark give an explanation for the definition and for the existence of finite satisfactory sets L_p, M_p . We will give an algorithm to produce (hopefully small) satisfactory sets L_p and M_p . Before that we embark on a simplification. The expression $a_0X - \theta Y$ is a linear form in two variables X, Y . It is easier to scale so that we are dealing with a linear expression in just one variable.

For a rational prime p , let

$$\mathbb{Z}_{(p)} = \{U \in \mathbb{Q} : \text{ord}_p(U) \geq 0\}.$$

Definition 2.2. Let p be a rational prime. Let $\alpha \in K$ and $\beta \in K^\times$. Let L be a subset of the ideals \mathfrak{b} supported on the prime ideals of \mathcal{O}_K above p . Let M be a subset of the set of pairs $(\mathfrak{b}, \mathfrak{p})$ where \mathfrak{b} is supported on the prime ideals above p , and where \mathfrak{p} is a prime ideal above p satisfying $e(\mathfrak{p} \mid p) = f(\mathfrak{p} \mid p) = 1$. We call L, M *adequate for* (α, β) if for every $U \in \mathbb{Z}_{(p)}$,

- (i) either the p -part of $\beta \cdot (U + \alpha)\mathcal{O}_K$ is in L , or
- (ii) there is a pair $(\mathfrak{b}, \mathfrak{p}) \in M$ and a nonnegative integer l such that the p -part of $\beta \cdot (U + \alpha)\mathcal{O}_K$ equals $\mathfrak{b}\mathfrak{p}^l$.

Lemma 2.3. Let L, M be adequate for $(-\theta/a_0, a_0)$ and let $L_p = L \cup \{1 \cdot \mathcal{O}_K\}$ and $M_p = M$. Then the pair L_p, M_p is satisfactory.

Proof. Recall that $\gcd(X, Y) = \gcd(a_0, Y) = 1$.

If $p \mid Y$ then $\text{ord}_p(Y) > 0$ for any \mathfrak{p} above p , and thus

$$\text{ord}_p(a_0X - \theta Y) = 0.$$

If $p \nmid Y$, we write

$$U = \frac{X}{Y}, \quad \alpha = \frac{-\theta}{a_0}, \quad \beta = a_0.$$

Then $U \in \mathbb{Z}_{(p)}$ and $\text{ord}_p(a_0X - \theta Y) = \text{ord}_p(\beta \cdot (U + \alpha))$ for all prime ideals \mathfrak{p} above p . Thus the p -part of $\beta \cdot (U + \alpha)$ is equal to the p -part of $\text{ord}_p(a_0X - \theta Y)$.

The lemma follows. \square

We now demystify Definitions 2.1 and 2.2.

Lemma 2.4. *Let p be a rational prime and γ a generator of K . Then there is a bound B depending only on p and γ such that the following hold:*

(a) *For any $U \in \mathbb{Z}_{(p)}$ and any pair of distinct prime ideals $\mathfrak{p}_1, \mathfrak{p}_2$ lying over p ,*

$$\text{ord}_{\mathfrak{p}_1}(U + \gamma) \leq B \quad \text{or} \quad \text{ord}_{\mathfrak{p}_2}(U + \gamma) \leq B.$$

(b) *For any $U \in \mathbb{Z}_{(p)}$ and any prime ideal \mathfrak{p} over p with $e(\mathfrak{p} \mid p) \neq 1$ or $f(\mathfrak{p} \mid p) \neq 1$,*

$$\text{ord}_{\mathfrak{p}}(U + \gamma) \leq B.$$

Proof. Let \mathfrak{p} be a prime ideal above p , and suppose that $\text{ord}_{\mathfrak{p}}(U + \gamma)$ is unbounded for $U \in \mathbb{Z}_{(p)}$. Thus there is an infinite sequence $\{U_i\} \subset \mathbb{Z}_{(p)}$ such that

$$\lim_{i \rightarrow \infty} \text{ord}_{\mathfrak{p}}(U_i + \gamma) = \infty.$$

However, $\mathbb{Z}_{(p)} \subset \mathbb{Z}_p$, where the latter is compact. Thus $\{U_i\}$ contains an infinite subsequence $\{U_{n_i}\}$ converging to, say, $U \in \mathbb{Z}_p$. Write $\phi_{\mathfrak{p}} : K \hookrightarrow \mathbb{C}_p$ for the embedding of K corresponding to the prime ideal \mathfrak{p} . It follows that $\phi_{\mathfrak{p}}(\gamma) = -U \in \mathbb{Z}_p$. Recall the assumption that $K = \mathbb{Q}(\gamma)$. Thus $K_{\mathfrak{p}}$, the topological closure of $\phi_{\mathfrak{p}}(K)$ in \mathbb{C}_p , is in fact \mathbb{Q}_p . Thus $e(\mathfrak{p} \mid p) = f(\mathfrak{p} \mid p) = 1$. This proves (b).

For (a), suppose that there is a pair of distinct primes $\mathfrak{p}_1, \mathfrak{p}_2$ above p and an infinite sequence $\{U_i\} \subset \mathbb{Z}_{(p)}$ such that

$$\lim_{i \rightarrow \infty} \text{ord}_{\mathfrak{p}_1}(U_i + \gamma) = \lim_{i \rightarrow \infty} \text{ord}_{\mathfrak{p}_2}(U_i + \gamma) = \infty. \quad (8)$$

Again, let $\{U_{n_i}\}$ be an infinite subsequence of $\{U_i\}$ converging to, say, $U \in \mathbb{Z}_p$. Then $\phi_{\mathfrak{p}_1}(\gamma) = -U = \phi_{\mathfrak{p}_2}(\gamma)$. As $K = \mathbb{Q}(\gamma)$, the embeddings $\phi_{\mathfrak{p}_1}, \phi_{\mathfrak{p}_2}$ are equal, contradicting $\mathfrak{p}_1 \neq \mathfrak{p}_2$. \square

Remark. We apply Lemma 2.4 with $\gamma = \alpha$ as in Lemma 2.3 in order to explain Definitions 2.1 and 2.2. The valuation of $\beta \cdot (U + \alpha)$ can be arbitrarily large only for those \mathfrak{p} above p that satisfy $e(\mathfrak{p} \mid p) = f(\mathfrak{p} \mid p) = 1$, and if it is sufficiently large for one such \mathfrak{p} then it is bounded for all others. Thus there must exist finite adequate sets L, M . We now turn to the task of giving an algorithm to determine such adequate sets L, M .

Lemma 2.5. *Let $\alpha \in K$ and let p be a rational prime. Let \mathfrak{p} be a prime ideal of \mathcal{O}_K above p . Suppose $U \in \mathbb{Z}_{(p)}$ and*

$$\text{ord}_{\mathfrak{p}}(U + \alpha) > \min\{0, \text{ord}_{\mathfrak{p}}(\alpha)\}. \quad (9)$$

Then the following hold:

- (i) $\text{ord}_{\mathfrak{p}}(\alpha) \geq 0$.
- (ii) *The image of α in $\mathbb{F}_{\mathfrak{p}} := \mathcal{O}_K/\mathfrak{p}$ belongs to the prime subfield \mathbb{F}_p . In particular, there is a unique $u \in \{0, \dots, p-1\}$ such that $u \equiv -\alpha \pmod{\mathfrak{p}}$.*
- (iii) *With u as in (ii), $U = pU' + u$ where $U' \in \mathbb{Z}_{(p)}$.*

Proof. Since $U \in \mathbb{Z}_{(p)}$, we have $\text{ord}_{\mathfrak{p}}(U) \geq 0$. If $\text{ord}_{\mathfrak{p}}(\alpha) < 0$, it follows that $\text{ord}_{\mathfrak{p}}(U + \alpha) = \text{ord}_{\mathfrak{p}}(\alpha)$, contradicting (9). Thus $\text{ord}_{\mathfrak{p}}(\alpha) \geq 0$, proving (i).

Write $\bar{\alpha}$ for the image of α in $\mathbb{F}_{\mathfrak{p}}$, and suppose this does not belong to the prime subfield \mathbb{F}_p . In particular $\text{ord}_{\mathfrak{p}}(\alpha) = 0$. However, the image \bar{U} of U in $\mathbb{F}_{\mathfrak{p}}$ does belong to \mathbb{F}_p . Thus $U \not\equiv -\alpha \pmod{\mathfrak{p}}$, or equivalently $\text{ord}_{\mathfrak{p}}(U + \alpha) = 0$, contradicting (9). We deduce that $\bar{\alpha} \in \mathbb{F}_p$, and thus (ii) holds.

Now, let u be as in (ii). By (9), we have $\text{ord}_{\mathfrak{p}}(U + \alpha) > 0$, and thus $\bar{U} = -\bar{\alpha} = \bar{u}$. But $\bar{U}, \bar{u} \in \mathbb{F}_p$. Therefore, $\text{ord}_{\mathfrak{p}}(U - u) > 0$, and so $U = pU' + u$ for some $U' \in \mathbb{Z}_{(p)}$. \square

Algorithm 2.6. *Given p a rational prime, $\alpha \in K$ satisfying $K = \mathbb{Q}(\alpha)$, and $\beta \in K^\times$, to compute L, M adequate for (α, β) :*

Step (a) *Let*

$$\mathcal{B} = \{\mathfrak{p} \mid p : \text{ord}_{\mathfrak{p}}(\alpha) \geq 0 \text{ and the image of } \alpha \text{ in } \mathbb{F}_{\mathfrak{p}} \text{ belongs to } \mathbb{F}_p\},$$

and

$$\mathfrak{b} = \prod_{\mathfrak{p} \mid p} \mathfrak{p}^{\text{ord}_{\mathfrak{p}}(\beta) + \min\{0, \text{ord}_{\mathfrak{p}}(\alpha)\}}.$$

Step (b) *If $\mathcal{B} = \emptyset$ then return $L = \{\mathfrak{b}\}$, $M = \emptyset$ and terminate the algorithm.*

Step (c) *If \mathcal{B} consists of a single prime ideal \mathfrak{p}' satisfying $e(\mathfrak{p}' \mid p) = f(\mathfrak{p}' \mid p) = 1$ then return $L = \emptyset$, $M = \{(\mathfrak{b}, \mathfrak{p}')\}$ and terminate the algorithm.*

Step (d) *Let*

$$\mathcal{U} = \{0 \leq u \leq p-1 : \text{there is some } \mathfrak{p} \in \mathcal{B} \text{ such that } \alpha \equiv -u \pmod{\mathfrak{p}}\}.$$

Loop through the elements $u \in \mathcal{U}$. For each u , use Algorithm 2.6 to compute adequate L_u, M_u for the pair $((u + \alpha)/p, p\beta)$.

Step (d1) *If $\mathcal{U} = \{0, 1, 2, \dots, p-1\}$ then return*

$$L = \bigcup_{u \in \mathcal{U}} L_u, \quad M = \bigcup_{u \in \mathcal{U}} M_u, \quad (10)$$

and terminate the algorithm.

Step (d2) *Else, return*

$$L = \{\mathfrak{b}\} \cup \bigcup_{u \in \mathcal{U}} L_u, \quad M = \bigcup_{u \in \mathcal{U}} M_u, \quad (11)$$

and terminate the algorithm.

Remarks. • Algorithm 2.6 is recursive. If the hypotheses of (b) and (c) fail then the algorithm replaces the linear form $\beta \cdot (U + \alpha)$ with a number of linear forms

$$p\beta \cdot (U' + (u + \alpha)/p) = \beta \cdot (pU' + u + \alpha).$$

In essence we are replacing $\mathbb{Z}_{(p)}$ with a number of the cosets of $p\mathbb{Z}_{(p)}$. The algorithm is then applied to each of these linear forms individually.

- Note that the number of prime ideals \mathfrak{p} above p is bounded by the degree $[K : \mathbb{Q}]$. In particular, $\#\mathcal{U} \leq [K : \mathbb{Q}]$. Therefore the number of branches at each iteration of the algorithm is bounded independently of p .

Proposition 2.7. *Suppose $\beta \in K^\times$, $\alpha \in K$ and moreover, $K = \mathbb{Q}(\alpha)$. Then Algorithm 2.6 terminates in finite time and produces adequate L, M for (α, β) .*

Proof. Let \mathcal{B} and \mathfrak{b} be as in Step (a). Observe that, for any \mathfrak{p} above p ,

$$\text{ord}_{\mathfrak{p}}(\beta \cdot (U + \alpha)) \geq \text{ord}_{\mathfrak{p}}(\beta) + \min\{0, \text{ord}_{\mathfrak{p}}(\alpha)\} = \text{ord}_{\mathfrak{p}}(\mathfrak{b}).$$

It follows that \mathfrak{b} divides the p -part of $\beta \cdot (U + \alpha)$. Lemma 2.5 tells us that

$$\text{ord}_{\mathfrak{p}}(\beta \cdot (U + \alpha)) = \text{ord}_{\mathfrak{p}}(\beta) + \min\{0, \text{ord}_{\mathfrak{p}}(\alpha)\} = \text{ord}_{\mathfrak{p}}(\mathfrak{b}) \quad (12)$$

for all prime ideals \mathfrak{p} lying above p , except possibly for $\mathfrak{p} \in \mathcal{B}$. If $\mathcal{B} = \emptyset$ (i.e., the hypothesis of (b) is satisfied), then the p -part of $\beta \cdot (U + \alpha)$ is \mathfrak{b} , and hence the pair $L = \{\mathfrak{b}\}$, $M = \emptyset$ is adequate for (α, β) . If $\mathcal{B} = \{\mathfrak{p}'\}$ where $e(\mathfrak{p}' | p) = f(\mathfrak{p}' | p) = 1$ (i.e., the hypothesis of step (c) is satisfied) then the p -part of $\beta \cdot (U + \alpha)$ has the form $\mathfrak{b} \cdot \mathfrak{p}'^l$ for some $l \geq 0$. Hence $L = \emptyset$, $M = \{(\mathfrak{b}, \mathfrak{p}')\}$ are adequate for (α, β) .

Suppose the hypotheses of steps (b) and (c) fail. Let \mathcal{U} be as in step (d). If $U \equiv u \pmod{p}$ for some $u \in \mathcal{U}$, then $U = pU' + u$ for some $U' \in \mathbb{Z}_{(p)}$. Thus $\beta \cdot (U + \alpha) = p\beta \cdot (U' + (u + \alpha)/p)$, and so naturally the p -parts of $\beta \cdot (U + \alpha)$ and $p\beta \cdot (U' + (u + \alpha)/p)$ agree. In (d1), \mathcal{U} represents all of the congruence classes modulo p , and this justifies (10). In (d2), \mathcal{U} represents some of the congruence classes. If $u \notin \mathcal{U}$, then by Lemma 2.5, the equality (12) holds for all prime ideals \mathfrak{p} above p , and hence \mathfrak{b} is the p -part of $\beta \cdot (U + \alpha)$. This justifies (11).

Next we show that the algorithm terminates in finitely many steps. Suppose otherwise. Then there will be an infinite sequence of $u_i \in \{0, 1, \dots, p-1\}$ and pairs (α_i, β_i) with

$$\alpha_0 = \alpha, \quad \beta_0 = \beta, \quad \alpha_{i+1} = \frac{u_i + \alpha_i}{p}, \quad \beta_{i+1} = p\beta_i.$$

Let us denote by \mathcal{B}_i the set \mathcal{B} for the pair (α_i, β_i) . It is easy to see from the definition that $\mathcal{B}_{i+1} \subseteq \mathcal{B}_i$. Suppose \mathfrak{p} belongs to infinitely many of the \mathcal{B}_i . Then, infinitely often, $\text{ord}_{\mathfrak{p}}(\alpha_i) \geq 0$. However,

$$\alpha = \alpha_0 = -u_0 - u_1p - u_2p^2 - \dots - u_{i-1}p^{i-1} + p^i\alpha_i.$$

Let $\mu = -u_0 - u_1p - \dots \in \mathbb{Z}_p$. Let $\phi_{\mathfrak{p}} : K \hookrightarrow \mathbb{C}_p$ be the embedding corresponding to \mathfrak{p} . Then $\phi_{\mathfrak{p}}(\alpha) = \mu$. Since $K = \mathbb{Q}(\alpha)$, the embedding $\phi_{\mathfrak{p}}$ is determined by the image of α . Since $\phi_{\mathfrak{p}} \neq \phi_{\mathfrak{p}'}$ whenever $\mathfrak{p} \neq \mathfrak{p}'$, we see that \mathcal{B}_i consists of at most one prime for i sufficiently large. For such a prime, we must have $\phi_{\mathfrak{p}}(K) = \mathbb{Q}_p$ so that $e(\mathfrak{p} | p) = f(\mathfrak{p} | p) = 1$. Thus for sufficiently large i the algorithm must terminate at Step (b) or Step (c), giving a contradiction. \square

Following Lemma 2.3, we thus let $L_p = L \cup \{1 \cdot \mathcal{O}_K\}$ and $M_p = M$, where we compute L and M using Algorithm 2.6 with $\alpha = -\theta/a_0$ and $\beta = a_0$.

Refinements. Let L_p, M_p be a satisfactory pair (for example, produced by Algorithm 2.6). We explain here some obvious refinements that will reduce or simplify these sets, whilst maintaining the satisfactory property:

- If some pair $(\mathfrak{b}, \mathfrak{p})$ is in M_p then we may replace this with the pair $(\mathfrak{b}', \mathfrak{p})$ where

$$\mathfrak{b}' = \frac{\mathfrak{b}}{\mathfrak{p}^{\text{ord}_{\mathfrak{p}}(\mathfrak{b})}}.$$

- If some \mathfrak{b} is contained in L_p , and some $(\mathfrak{b}', \mathfrak{p})$ is contained in M_p with $\mathfrak{b}' | \mathfrak{b}$ and $\mathfrak{b}/\mathfrak{b}' = \mathfrak{p}^w$ for some $w \geq 0$, then we may delete \mathfrak{b} from L_p .
- Suppose $p | a_0$. Observe that if $\mathfrak{b} \in L_p$ is the p -part of $(a_0X - \theta Y)\mathcal{O}_K$ then $\text{ord}_p(\text{Norm}(a_0X - \theta Y)) = \text{ord}_p(\text{Norm}(\mathfrak{b}))$. However, it is clear that $\text{ord}_p(\text{Norm}(a_0X - \theta Y)) \geq (d - 1) \text{ord}_p(a_0)$. Thus, we may delete \mathfrak{b} from L_p if $\text{ord}_p(\text{Norm}(\mathfrak{b})) < (d - 1) \text{ord}_p(a_0)$.

3. An equation in ideals

Let (X, Y) be a solution of (3). For every $p \in P := \{p_1, \dots, p_v\}$ we let L_p, M_p be a corresponding satisfactory pair. From Definition 2.1, we see that there is some partition $P = P_1 \cup P_2$ such that for every $p \in P_1$, the p -part of $(a_0X - \theta Y)\mathcal{O}_K$ equals $\mathfrak{b}p^l$ for some $(\mathfrak{b}, \mathfrak{p}) \in M_p$ and $l \geq 0$, and for every $p \in P_2$, it equals some $\mathfrak{b} \in L_p$. Let $P = P_1 \cup P_2$ be a partition of P and write

$$P_1 = \{q_1, \dots, q_s\}, \quad P_2 = \{q_{s+1}, \dots, q_v\}.$$

Let \mathcal{Z}_{P_1, P_2} be the set of all pairs (α, S) such that there are $(\mathfrak{b}_i, \mathfrak{p}_i) \in M_{q_i}$ for $1 \leq i \leq s$, and $\mathfrak{b}_j \in L_{q_j}$ for $s + 1 \leq j \leq v$ satisfying

$$\alpha = \mathfrak{b}_0 \cdot \mathfrak{b}_1 \mathfrak{b}_2 \cdots \mathfrak{b}_s \cdot \mathfrak{b}_{s+1} \cdots \mathfrak{b}_v, \quad S = \{\mathfrak{p}_1, \dots, \mathfrak{p}_s\},$$

where \mathfrak{b}_0 denotes an ideal of \mathcal{O}_K of norm

$$R = \left| \frac{a \cdot a_0^{d-1}}{\gcd(\text{Norm}(\mathfrak{b}_1 \cdots \mathfrak{b}_v), a \cdot a_0^{d-1})} \right|.$$

Let

$$\mathcal{Z} := \bigcup_{P_1 \subseteq P} \mathcal{Z}_{P_1, P-P_1}.$$

Proposition 3.1. *Let (X, Y) be a solution to (3). Then there is some $(\mathfrak{a}, S) \in \mathcal{Z}$ such that*

$$(a_0 X - \theta Y) \mathcal{O}_K = \mathfrak{a} \cdot \mathfrak{p}_1^{n_1} \cdots \mathfrak{p}_s^{n_s}, \quad (13)$$

where $S = \{\mathfrak{p}_1, \dots, \mathfrak{p}_s\}$, and n_1, \dots, n_s are nonnegative integers. Moreover, the set S has the following properties:

- (i) $e(\mathfrak{p}_i | q_i) = f(\mathfrak{p}_i | q_i) = 1$ for $1 \leq i \leq s$.
- (ii) Let $1 \leq i \leq s$. Let p be the unique rational prime below \mathfrak{p}_i . Then $\mathfrak{p}_j \nmid p$ for all $1 \leq j \leq s$ with $j \neq i$.

Proof. The claims follow from the definitions of \mathcal{Z} and M_p . □

To solve (3), we will solve (13) for every possible choice of $(\mathfrak{a}, S) \in \mathcal{Z}$.

Remark. Observe that for any $(\mathfrak{a}, S) \in \mathcal{Z}$, there may be several possibilities for \mathfrak{b}_0 . Let \mathcal{R} denote the set of all ideals \mathfrak{b}_0 having norm R for some $(\mathfrak{a}, S) \in \mathcal{Z}$. Here, we provide a simple refinement to cut down the number of ideals in \mathcal{R} . In particular, we apply Algorithm 2.6 and Lemma 2.3 to each rational prime p dividing R , generating the corresponding sets M_p and L_p . For each $\mathfrak{b}_0 \in \mathcal{R}$, if the p -part of \mathfrak{b}_0 cannot be made up of any of the elements of M_p or L_p , we may remove \mathfrak{b}_0 from \mathcal{R} . Moreover, if this process yields $\mathcal{R} = \emptyset$, we may remove (\mathfrak{a}, S) from \mathcal{Z} .

4. Making the ideals principal

From now on we fix $(\mathfrak{a}, S) \in \mathcal{Z}$ and we focus on a solution of (13). The method of Tzanakis and de Weger [1989] reduces (13) to at most $(m/2) \cdot h^s$ S -unit equations, where m is the number of roots of unity and h is the class number of K . Our method, explained below, gives at most only $m/2$ S -unit equations.

Given an ideal \mathfrak{b} of \mathcal{O}_K , we denote its class in the class group $\text{Cl}(K)$ by $[\mathfrak{b}]$.

Lemma 4.1. *Let*

$$\phi : \mathbb{Z}^s \rightarrow \text{Cl}(K), \quad (m_1, \dots, m_s) \mapsto [\mathfrak{p}_1]^{m_1} \cdots [\mathfrak{p}_s]^{m_s} :$$

- (a) *If $[\mathfrak{a}]^{-1}$ is not in the image of ϕ then (13) has no solutions.*
- (b) *Suppose $[\mathfrak{a}]^{-1} = \phi(\mathbf{r})$, where $\mathbf{r} = (r_1, \dots, r_s)$. Let ζ be a generator of the roots of unity in K , and suppose it has order m . Let $\delta_1, \dots, \delta_r$ be a basis for the group of S -units \mathcal{O}_S^\times modulo the torsion*

subgroup $\langle \zeta \rangle$. Let α be a generator of the principal ideal $\mathfrak{a} \cdot \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_s^{r_s}$. Let (X, Y) satisfy (13). Then, after possibly replacing (X, Y) by $(-X, -Y)$, we have

$$a_0X - \theta Y = \tau \cdot \delta_1^{b_1} \cdots \delta_r^{b_r}, \tag{14}$$

where $\tau = \zeta^a \cdot \alpha$ with $0 \leq a \leq \frac{m}{2} - 1$, and $b_1, \dots, b_r \in \mathbb{Z}$.

Proof. Note that if (13) has a solution $\mathbf{n} = (n_1, \dots, n_s)$ then

$$\phi(\mathbf{n}) = [\mathfrak{a}]^{-1}.$$

This proves (a). For (b), suppose $[\mathfrak{a}]^{-1} = \phi(r_1, \dots, r_s)$. Thus $\mathfrak{a} \cdot \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_s^{r_s}$ is principal and we let α be a generator. Then the fractional ideal $((a_0X - \theta Y)/\alpha)\mathcal{O}_K$ is supported on $S = \{\mathfrak{p}_1, \dots, \mathfrak{p}_s\}$. Hence $(a_0X - \theta Y)/\alpha \in \mathcal{O}_S^\times$. Now $\zeta, \delta_1, \dots, \delta_r$ is a set of generators for the S -unit group, where $\delta_1, \dots, \delta_r$ is in fact a basis for $\mathcal{O}_S^\times/\langle \zeta \rangle$. Thus (14) holds for some $0 \leq a \leq m - 1$, and $b_1, \dots, b_r \in \mathbb{Z}$. However $\zeta^{m/2} = -1$. Thus we can suppose $0 \leq a \leq m/2 - 1$ by replacing (X, Y) by $(-X, -Y)$ if necessary. \square

Lemma 4.2. *The denominator of the fractional ideal $\tau\mathcal{O}_K$ is supported on the set of prime ideals $S = \{\mathfrak{p}_1, \dots, \mathfrak{p}_s\}$.*

Proof. This follows immediately from (14) since $\delta_1, \dots, \delta_r$ are S -units. \square

We have reduced the task of solving our original Thue–Mahler equation (3) to solving equations of the form

$$a_0X - \theta Y = \tau \cdot \delta_1^{b_1} \cdots \delta_r^{b_r}, \tag{15}$$

subject to the conditions

$$X, Y \in \mathbb{Z}, \quad \gcd(X, Y) = 1, \quad \gcd(a_0, Y) = 1, \quad b_i \in \mathbb{Z}. \tag{16}$$

For technical reasons we would like to exclude the case $b_1 = b_2 = \dots = b_r = 0$; of course we can trivially test if this case leads to a solution. Hence we shall henceforth suppose in addition to (16) that

$$\max\{|b_1|, \dots, |b_r|\} \geq 1. \tag{17}$$

We will tackle each of these equations (15) separately. In [Tzanakis and de Weger 1989], the authors work in the field generated by three conjugates of θ and its completions. This is fine theoretically but difficult computationally. We will work with that extension theoretically simply to obtain a bound for

$$B := \max\{|b_1|, \dots, |b_r|\}. \tag{18}$$

(The reason for restriction (17) is that in Section 6 we work with $\log B$). To reduce the bound, we will need to carry out certain computations; these will take place only in $K, \mathbb{R}, \mathbb{C}$, but not in extensions of K , and certainly not in extensions of \mathbb{Q}_p .

To obtain our initial bound for B we shall mostly follow ideas found in [Bugeaud and Győry 1996b; Bugeaud et al. 2008; Gallegos-Ruiz 2011]. However, we have a key advantage that will allow us to obtain

	# \mathcal{Z}	number of $(\tau, \delta_1, \dots, \delta_r)$	rank frequencies
Example 1.1	16	16	(1, 2), (2, 6), (3, 6), (4, 2)
Example 1.2	32	30	(2, 8), (3, 12), (4, 8), (5, 2)
Example 1.3	4096	4096	(10, 4096)
Example 1.4	2	2	(9, 1), (10, 1)

Table 1. This table gives the sizes of the set \mathcal{Z} and the number of resulting $(\tau, \delta_1, \dots, \delta_r)$ for Examples 1.1–1.4. The last column is a list of pairs (r, t) meaning there are t cases where the S -unit rank is r .

sharper bounds: namely we assume knowledge of the S -unit basis $\delta_1, \dots, \delta_r$ rather than working with estimates for the size of a basis.

4.1. Convention on the choice of S -unit basis. As before let ζ be a generator for the roots of unity in \mathcal{O}_K^\times . We note that the unit group \mathcal{O}_K^\times is a subgroup of the S -unit group \mathcal{O}_S^\times . Moreover, it is saturated in the sense that the quotient $\mathcal{O}_S^\times/\mathcal{O}_K^\times$ is torsion-free. Hence there is a basis $\delta_1, \dots, \delta_r$ for $\mathcal{O}_S^\times/\langle\zeta\rangle$ such that $\delta_1, \dots, \delta_{u+v-1}$ is a basis for $\mathcal{O}_K^\times/\langle\zeta\rangle$, where (u, v) is the signature of K . We shall in fact work with such a basis.

4.2. Examples continued. Table 1 gives some data for Examples 1.1–1.4. At this stage of the algorithm, we would like to stress the main difference between our approach and that of Tzanakis and de Weger [1989]. When dealing with

$$(a_0X - \theta Y)\mathcal{O}_K = \alpha \cdot \mathfrak{p}_1^{n_1} \cdots \mathfrak{p}_s^{n_s},$$

they write each exponent n_i as $n_i = k_i h_i + m_i$, where h_i is the order of \mathfrak{p}_i in the class group of \mathcal{O}_K and $0 \leq m_i \leq h_i - 1$. Now $\mathfrak{p}_i^{h_i}$ is principal and we may write it as $(\beta_i)\mathcal{O}_K$. It follows from (13) that $\alpha \cdot \mathfrak{p}_1^{m_1} \cdots \mathfrak{p}_s^{m_s}$ is principal. Clearly $h_i \mid h$ and there are at most h^s possibilities for this ideal. In the worst case, we expect around h^{s-1} ideals to be principal, and so of the form $(\tau)\mathcal{O}_K$. This results in a huge explosion of cases when h is nontrivial, as it is often the case that $h_i = h$. For instance, in our Example 1.2, the class number is 33. There are 32 possibilities for (α, S) in \mathcal{Z} . The 32 possible values for $s = \#S$ are

0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4,

where all possible ideals in S have order 33 in \mathcal{O}_K . Following the method of Tzanakis and de Weger, we would need to check 2672672 ideals if they are principal, and this approach leads to approximately 80990 equations of the form (15). With our approach, as we need only to deal with 32 ideal equations, we expect to deal with at most 32 equations of the form (15). Indeed, in doing so, we obtain merely 30 equations of the form (15).

5. Lower bounds for linear forms in logarithms

In this section, we state theorems of Matveev [2000] and of Yu [2007] for lower bounds for linear forms in complex and p -adic logarithms. In the next section, we will use these results to obtain bounds for

b_1, \dots, b_r . We begin by establishing some notation, as well as some key results which we will need for the lower bound:

L a number field.

D the degree $[L : \mathbb{Q}]$.

M_L the set of all places of L .

M_L^∞ the subset of infinite places.

M_L^0 the subset of finite places.

v a place of L .

D_v the local degree $[L_v : \mathbb{Q}_v]$.

$|\cdot|_v$ the usual normalized absolute value associated to v :

- If v is infinite and associated to a real or complex embedding σ of L , then $|\alpha|_v = |\sigma(\alpha)|$.
- If v is finite and above the rational prime p , then $|p|_v = p^{-1}$.

$\|\cdot\|_v = |\cdot|_v^{D_v}$.

$h(\cdot)$ the absolute logarithmic height, defined in (22).

In the above notation, the product formula may be stated as

$$\prod_{v \in M_L} \|\alpha\|_v = 1 \quad (19)$$

for all $\alpha \in L^\times$. In particular, if v is infinite, corresponding to a real or complex embedding σ of L , then

$$\|\alpha\|_v = \begin{cases} |\sigma(\alpha)| & \text{if } \sigma \text{ is real,} \\ |\sigma(\alpha)|^2 & \text{if } \sigma \text{ is complex.} \end{cases} \quad (20)$$

If v is finite, and \mathfrak{P} is the prime ideal corresponding to v , then for $\alpha \in L^\times$ we have

$$\|\alpha\|_v = \text{Norm}(\mathfrak{P})^{-\text{ord}_{\mathfrak{P}}(\alpha)}; \quad (21)$$

this easily follows from $D_v = e(\mathfrak{P} | p)f(\mathfrak{P} | p)$, where $e(\mathfrak{P} | p)$ and $f(\mathfrak{P} | p)$ are respectively the ramification index and the inertial degree of \mathfrak{P} .

For $\alpha \in L$, we define the absolute logarithmic height $h(\alpha)$ by

$$h(\alpha) = \frac{1}{[L : \mathbb{Q}]} \sum_{v \in M_L} D_v \log \max\{1, |\alpha|_v\} = \frac{1}{[L : \mathbb{Q}]} \sum_{v \in M_L} \log \max\{1, \|\alpha\|_v\}. \quad (22)$$

Lemma 5.1. *The absolute logarithmic height of an algebraic number α is independent of the number field L containing α . Moreover, if α and β are Galois conjugates, then $h(\alpha) = h(\beta)$.*

For proofs of the following two lemmata, see [Bugeaud et al. 2008, Lemma 4.1] and [Gallegos-Ruiz 2011, Lemma 3.2].

Lemma 5.2. For $\alpha_1, \dots, \alpha_n \in L$ we have

$$h(\alpha_1 \cdots \alpha_n) \leq h(\alpha_1) + \cdots + h(\alpha_n), \quad h(\alpha_1 + \cdots + \alpha_n) \leq \log n + h(\alpha_1) + \cdots + h(\alpha_n).$$

For any $\alpha \in L^\times$, we have $h(\alpha) = h(\alpha^{-1})$. Moreover, for any place $v \in M_L$,

$$\log \|\alpha\|_v \leq [L : \mathbb{Q}] \cdot h(\alpha). \tag{23}$$

Lemma 5.3. Let L be a number field of degree D . Let \mathfrak{S} be a finite set of finite places of L . Let $\varepsilon \in \mathcal{O}_{\mathfrak{S}}^\times$. Let $\eta \in M_L$ be a place of L chosen so that $\|\varepsilon\|_\eta$ is minimal. Then $\|\varepsilon\|_\eta \leq 1$ and

$$h(\varepsilon) \leq \frac{(\#M_L^\infty + \#\mathfrak{S})}{D} \cdot \log(\|\varepsilon^{-1}\|_\eta).$$

5.1. Lower bounds for linear forms in \mathfrak{P} -adic logarithms. Let L be a number field of degree D . Let \mathfrak{P} be a prime ideal of \mathcal{O}_L and let p be the rational prime below \mathfrak{P} . Let $v \in M_L^0$ correspond to \mathfrak{P} . Let $\alpha_1, \dots, \alpha_n \in L^\times$. Write $e = \exp(1)$.

Let

$$\begin{aligned} h_j &:= \max \left\{ h(\alpha_j), \frac{1}{16e^2 D^2} \right\}, \quad j = 1, \dots, n; \\ c_1(n, D) &:= (16eD)^{2n+2} \cdot n^{5/2} \cdot \log(2nD) \cdot \log(2D), \\ c_2(n, \mathfrak{P}) &:= e(\mathfrak{P} | p)^n \cdot \frac{p^{f(\mathfrak{P} | p)}}{f(\mathfrak{P} | p) \cdot \log p}, \\ c_3(n, D, \mathfrak{P}, \alpha_1, \dots, \alpha_n) &:= c_1(n, D) \cdot c_2(n, \mathfrak{P}) \cdot h_1 \cdots h_n. \end{aligned} \tag{24}$$

We shall make use of the following theorem of Yu [2007].

Theorem 5.4 (K. Yu). Let b_1, \dots, b_n be rational integers and let

$$B = \max\{|b_1|, \dots, |b_n|\},$$

and suppose $B \geq 3$. Let

$$\Lambda = \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1,$$

and suppose $\Lambda \neq 0$. Then

$$\log \|\Lambda^{-1}\|_v < c_3(n, D, \mathfrak{P}, \alpha_1, \dots, \alpha_n) \cdot \log B.$$

Proof. Let

$$c_4(n, D, \mathfrak{P}) := \frac{c_1(n, D) \cdot c_2(n, \mathfrak{P})}{n \cdot f(\mathfrak{P} | p) \cdot \log p}.$$

As stated in [Yu 2007, page 190], a consequence of Yu’s Main Theorem is

$$\text{ord}_{\mathfrak{P}}(\Lambda) < n \cdot c_4(n, D, \mathfrak{P}) \cdot h_1 \cdots h_n \cdot \log B.$$

By (21) we have

$$\log \|\Lambda^{-1}\|_v = \log(\text{Norm}(\mathfrak{P})) \cdot \text{ord}_{\mathfrak{P}}(\Lambda) = f(\mathfrak{P} | p) \cdot \log p \cdot \text{ord}_{\mathfrak{P}}(\Lambda).$$

The theorem follows. □

5.2. Lower bounds for linear forms in real or complex logarithms. We continue with the above notation.

Let

$$h'_j = \sqrt{h(\alpha_j)^2 + \frac{\pi^2}{D^2}}, \quad j = 1, \dots, n. \quad (25)$$

The following theorem is a version of Matveev's bound [2000] for linear forms in logarithms.

Theorem 5.5 (Matveev). *Let v be an infinite place of L . Suppose $\Lambda \neq 0$. Let*

$$c_5(n, D, \alpha_1, \dots, \alpha_n) = 6 \cdot 30^{n+4} \cdot (n+1)^{5.5} \cdot D^{n+2} \cdot \log(eD) \cdot h'_1 \cdots h'_n.$$

Then

$$\log \|\Lambda^{-1}\|_v \leq c_5(n, D, \alpha_1, \dots, \alpha_n) \cdot (\log(en) + \log B).$$

Proof. This in fact follows from a version of Matveev's theorem derived in [Bugeaud et al. 2006]. Let σ be a real or complex embedding of L corresponding to v . Let

$$h''_j = \max\{Dh(\alpha_j), |\log(\sigma(\alpha_j))|, 0.16\},$$

where $\log(\sigma(\alpha_j))$ is the principal determination of the logarithm (i.e., the imaginary part of \log lies in $(-\pi, \pi]$). Let

$$c_6(n, D) = 3 \cdot 30^{n+4} \cdot (n+1)^{5.5} \cdot D^2 \cdot \log(eD).$$

Then Theorem 9.4 of [Bugeaud et al. 2006] asserts that

$$\log |\Lambda|_v \geq -c_6(n, D) \cdot h''_1 \cdots h''_n \cdot (\log(en) + \log B).$$

Since $\|\Lambda\| = |\Lambda|^{D_v}$, where D_v is either 1 or 2, we have

$$\log \|\Lambda^{-1}\| \leq 2 \cdot c_6(n, D) \cdot h''_1 \cdots h''_n \cdot (\log(en) + \log B).$$

Thus it is sufficient to show that $h''_j \leq Dh'_j$. However,

$$\log(\sigma(\alpha_j)) = \log|\sigma(\alpha_j)| + i\theta$$

where $-\pi < \theta \leq \pi$. But by (23) and $\|\cdot\|_v = |\cdot|_v^{D_v}$, we have

$$\log|\sigma(\alpha_j)| = \frac{1}{D_v} \log\|\alpha_j\|_v \leq \log\|\alpha_j\|_v \leq D \cdot h(\alpha_j).$$

Thus

$$|\log(\sigma(\alpha_j))| \leq \sqrt{D^2 \cdot h(\alpha_j)^2 + \pi^2} = D \cdot h'_j.$$

It is now clear that $h''_j \leq D \cdot h'_j$. □

6. The S-unit equation

We now return to the task of studying the solutions of (15) satisfying (16), (17). Here, we use the theorems of Matveev and Yu (recalled in the previous section) to establish bounds for b_1, \dots, b_r , following the ideas of [Bugeaud and Györy 1996b; Bugeaud et al. 2008; Gallegos-Ruiz 2011], and taking care to keep our constants completely explicit and as small as possible.

We begin by establishing the following notation:

θ, K as defined in Section 1.

d the degree $[K : \mathbb{Q}] \geq 3$.

S a set $\{\mathfrak{p}_1, \dots, \mathfrak{p}_s\}$ of prime ideals of K satisfying conditions (a), (b) of Proposition 3.1.

s the cardinality $\#S$ of the set S .

$\delta_1, \dots, \delta_r$ a basis for the S -unit group \mathcal{O}_S^\times modulo torsion, also appearing in (15).

τ a nonzero element of K , appearing in (15).

X, Y, b_1, \dots, b_r a solution to (15) satisfying (16), (17).

$\varepsilon = \delta_1^{b_1} \cdots \delta_r^{b_r}$. Thus (15) can be rewritten as $a_0X - \theta Y = \tau \cdot \varepsilon$.

$\mu = \tau \cdot \varepsilon = a_0X - \theta Y$.

$B = \max\{|b_1|, \dots, |b_r|\}$.

$\theta_1, \theta_2, \theta_3$ three conjugates of θ chosen below, with $\theta = \theta_1$.

L the extension $\mathbb{Q}(\theta_1, \theta_2, \theta_3) \supseteq K$.

D the degree $[L : \mathbb{Q}]$.

σ_i the embedding $K \hookrightarrow L, \theta \mapsto \theta_i$.

$\mu_i, \varepsilon_i, \tau_i, \delta_{j,i}$ the images of $\mu, \varepsilon, \tau, \delta_j$ under the embedding σ_i .

ξ_1, ξ_2, ξ_3 defined in (26).

We want to write down an S -unit equation starting with (15). For this we will need to work with three conjugates of θ . Let $d = [K : \mathbb{Q}]$, and let $\theta_1, \dots, \theta_d$ be the conjugates of θ in some splitting field $M \supseteq K$. We shall not need M explicitly, but we assume that we are able to compute the Galois group G of M/\mathbb{Q} as a transitive permutation group on the conjugates θ_i . From this we are able to list all subgroups (up to conjugacy), and for each subgroup determine if it fixes at least three conjugates of θ . Let H be a subgroup of G fixing at least three conjugates of θ with index $[G : H]$ as small as possible. Let $L = M^H$ be the fixed field of H . Then $L = \mathbb{Q}(\theta_1, \theta_2, \theta_3)$ for some three conjugates of θ (after a possible reordering of conjugates) and it has the property that its degree is minimal amongst all extensions generated by three conjugates. Write

$$D := [G : H] = [L : \mathbb{Q}].$$

Again we shall not need the field L explicitly, but only its degree D , which we can deduce from the Galois group. We identify $\theta = \theta_1$, and so can think of $K \subseteq L$.

Write $\mu = a_0X - \theta Y$. Let $\varepsilon = \delta_1^{b_1} \cdots \delta_r^{b_r}$. Then $\mu = \tau \cdot \varepsilon$. Let $\mu_i, \varepsilon_i, \tau_i, \delta_{j,i}$ be the images of $\mu, \varepsilon, \tau, \delta_j$ under the embeddings $\sigma_i : K \hookrightarrow L, \theta \mapsto \theta_i$. We observe the following Siegel identity:

$$(\theta_3 - \theta_2)\mu_1 + (\theta_1 - \theta_3)\mu_2 + (\theta_2 - \theta_1)\mu_3 = 0.$$

Let

$$\xi_1 = (\theta_3 - \theta_2) \cdot \tau_1, \quad \xi_2 = (\theta_1 - \theta_3) \cdot \tau_2, \quad \xi_3 = (\theta_2 - \theta_1) \cdot \tau_3. \tag{26}$$

Then

$$\xi_1 \varepsilon_1 + \xi_2 \varepsilon_2 + \xi_3 \varepsilon_3 = 0. \tag{27}$$

Note that ε_1 is an S -unit in K and $\varepsilon_2, \varepsilon_3$ are Galois conjugates of ε . This equation will serve as our S -unit equation. We would like to rewrite (27) in a manner that makes it convenient to apply Theorems 5.4 and 5.5. Observe that (27) can be rewritten as

$$\frac{\xi_3 \varepsilon_3}{\xi_1 \varepsilon_1} = \left(\frac{-\xi_2}{\xi_1} \right) \left(\frac{\varepsilon_2}{\varepsilon_1} \right) - 1. \tag{28}$$

Let

$$\alpha_j := \frac{\delta_{j,2}}{\delta_{j,1}} \quad (j = 1, \dots, r), \quad \alpha_{r+1} := \frac{-\xi_2}{\xi_1}, \quad b_{r+1} = 1.$$

Then

$$\frac{\xi_3 \varepsilon_3}{\xi_1 \varepsilon_1} = \Lambda \tag{29}$$

where Λ is the “linear form”

$$\Lambda := \alpha_1^{b_1} \dots \alpha_{r+1}^{b_{r+1}} - 1.$$

We assume that we know θ, τ and $\delta_1, \dots, \delta_r$ explicitly and can therefore compute their absolute logarithmic heights. We will use this to estimate the heights of other algebraic numbers, such as ξ_i, α_j , without computing their minimal polynomials. By Lemmas 5.1 and 5.2,

$$h(\xi_i) \leq c_7, \quad c_7 := \log 2 + 2h(\theta) + h(\tau).$$

Lemma 6.1. *Let*

$$c_8 := 2Dc_7.$$

Let v be any place of L . Then

$$\log \|(\varepsilon_3/\varepsilon_1)^{-1}\|_v \leq \log \|\Lambda^{-1}\|_v + c_8.$$

Proof. Note that

$$\log \|(\varepsilon_3/\varepsilon_1)^{-1}\|_v = \log \|\Lambda^{-1}\|_v + \log \|\xi_3/\xi_1\|_v.$$

By Lemma 5.2

$$\log \|\xi_3/\xi_1\|_v \leq D \cdot h(\xi_3/\xi_1) \leq D \cdot (h(\xi_3) + h(\xi_1)). \quad \square$$

By definition $B = \max\{|b_1|, \dots, |b_r|\}$. However, by (17), and since $b_{r+1} = 1$, we have

$$B = \max\{|b_1|, \dots, |b_r|, |b_{r+1}|\}.$$

We now apply Matveev’s theorem in order to obtain a bound for B .

Lemma 6.2. *Let*

$$h_j^* := \sqrt{4h(\delta_j)^2 + \frac{\pi^2}{D^2}} \quad \text{for } j = 1, \dots, r \text{ and } h_{r+1}^* := \sqrt{4c_7^2 + \frac{\pi^2}{D^2}}.$$

Let

$$c_9 = 6 \cdot 30^{r+5} \cdot (r+2)^{5.5} \cdot D^{r+3} \cdot \log(eD) \cdot h_1^* \cdots h_{r+1}^* \quad \text{and} \quad c_{10} = c_8 + c_9 \cdot \log(e(r+1)).$$

Let v be an infinite place of L . Then

$$\log \|(\varepsilon_3/\varepsilon_1)^{-1}\|_v \leq c_{10} + c_9 \cdot \log B.$$

Proof. We will apply Theorem 5.5 with $n = r + 1$. Observe that

$$h(\alpha_j) \leq 2h(\delta_j) \quad \text{for } j = 1, \dots, r \text{ and } h(\alpha_{r+1}) \leq 2c_7.$$

Thus $h'_j \leq h_j^*$, where h'_j is defined in (25). By Theorem 5.5,

$$\log \|\Lambda^{-1}\|_v \leq c_9 \cdot (\log(e(r+1)) + \log B).$$

Lemma 6.1 completes the proof. □

We also apply Yu's theorem.

Lemma 6.3. *Let*

$$h_j^\dagger := \max \left\{ 2h(\delta_j), \frac{1}{16e^2 D^2} \right\} \quad \text{for } j = 1, \dots, r,$$

and

$$h_{r+1}^\dagger := \max \left\{ 2c_7, \frac{1}{16e^2 D^2} \right\}.$$

Let T be the set of rational primes p below the primes $\mathfrak{p} \in S$. Let

$$c_{11} := \max_{p \in T} \max \left\{ \frac{u^{r+1} \cdot p^v}{v \cdot \log p} : u, v \text{ are positive integers and } uv \leq D/d \right\}$$

and

$$c_{12} := c_1(r+1, D) \cdot c_{11} \cdot h_1^\dagger \cdots h_{r+1}^\dagger.$$

Let v be a finite place of L . Then

$$\log \|(\varepsilon_3/\varepsilon_1)^{-1}\|_v \leq c_8 + c_{12} \log B.$$

Proof. Of course we may suppose that $\|(\varepsilon_3/\varepsilon_1)^{-1}\|_v \neq 1$. Let \mathfrak{P} be the prime ideal of \mathcal{O}_L corresponding to v . We will deduce the lemma from Theorem 5.4 combined with Lemma 6.1. For this, it suffices to show that $c_3(r+1, D, \mathfrak{P}, \alpha_1, \dots, \alpha_{r+1}) \leq c_{12}$. Observe that $h_j \leq h_j^\dagger$ for $j = 1, \dots, r+1$. Thus it is enough to show that $c_2(r+1, \mathfrak{P}) \leq c_{11}$.

Let $K_i = \mathbb{Q}(\theta_i) \subseteq L$. Recall that ε is an S -unit and that ε_i is the image of ε under the map $\sigma_i : K \rightarrow K_i$, $\theta \mapsto \theta_i$. As $\|(\varepsilon_3/\varepsilon_1)^{-1}\|_v \neq 1$, we see that $\text{ord}_{\mathfrak{P}}(\varepsilon_i) \neq 0$ for $i = 1$ or 3 . Thus \mathfrak{P} must be a prime above $\mathfrak{p}_i := \sigma_i(\mathfrak{p})$ of \mathcal{O}_{K_i} for some $\mathfrak{p} \in S$, where $i = 1$ or 3 . In particular \mathfrak{P} is above some rational prime $p \in T$. However, $e(\mathfrak{p} | p) = f(\mathfrak{p} | p) = 1$ for all $\mathfrak{p} \in S$. Thus $e(\mathfrak{P} | p) = e(\mathfrak{P} | \mathfrak{p}_i)$ and $f(\mathfrak{P} | p) = f(\mathfrak{P} | \mathfrak{p}_i)$ for $i = 1$ or 3 . Let $u = e(\mathfrak{P} | p)$ and $v = f(\mathfrak{P} | p)$. We see that $uv = e(\mathfrak{P} | \mathfrak{p}_i) f(\mathfrak{P} | \mathfrak{p}_i) \leq [L : K_i] = D/d$. Now $c_2(r+1, \mathfrak{P}) \leq c_{11}$ follows from the definitions of c_2 and c_{11} . □

Lemma 6.4. *Let*

$$c_{13} := \frac{\#M_K^\infty + 2 \cdot \#S}{d},$$

and

$$c_{14} := 2h(\tau) + c_{13} \cdot \max(c_8, c_{10}), \quad c_{15} := c_{13} \cdot \max(c_9, c_{12}).$$

Then

$$h(\mu_3/\mu_1) \leq c_{14} + c_{15} \cdot \log B.$$

Proof. Let \mathfrak{S} be the prime ideals appearing in the support of $\varepsilon_3/\varepsilon_1$. We will show that $\#\mathfrak{S} \leq (2D/d) \cdot \#S$. Indeed, ε_i belongs to $K_i = \mathbb{Q}(\theta_i)$ and its support in K_i is contained in $\sigma_i(S)$. Now a prime belonging to $\sigma_i(S)$ has at most $[L : K_i] = D/d$ primes above it in L . Thus

$$\#\mathfrak{S} \leq (D/d) \cdot \#\sigma_1(S) + (D/d) \cdot \#\sigma_3(S) \leq (2D/d) \cdot \#S$$

as required. Moreover, since $[L : K] = D/d$, we have $\#M_L^\infty \leq (D/d) \cdot \#M_K^\infty$.

Let $\eta \in M_L$ be the place of L such that $\|\varepsilon_3/\varepsilon_1\|_\eta$ is minimal. By Lemma 5.3

$$h(\varepsilon_3/\varepsilon_1) \leq \frac{\#M_L^\infty + \#\mathfrak{S}}{D} \cdot \|(\varepsilon_3/\varepsilon_1)^{-1}\|_\eta.$$

From the above inequalities for $\#M_L^\infty$ and $\#\mathfrak{S}$, we deduce that

$$h(\varepsilon_3/\varepsilon_1) \leq c_{13} \cdot \|(\varepsilon_3/\varepsilon_1)^{-1}\|_\eta.$$

We now apply Lemmas 6.2 and 6.3 to obtain

$$h(\varepsilon_3/\varepsilon_1) \leq c_{13} \cdot (\max(c_8, c_{10}) + \max(c_9, c_{12}) \cdot \log B).$$

Finally, observe that $\mu_3/\mu_1 = (\tau_3/\tau_1) \cdot (\varepsilon_3/\varepsilon_1)$ and thus

$$h(\mu_3/\mu_1) \leq 2h(\tau) + h(\varepsilon_3/\varepsilon_1). \quad \square$$

We shall henceforth suppose $(X, Y) \neq (\pm 1, 0)$. As $\gcd(X, Y) = 1$, this is equivalent to $Y \neq 0$.

Lemma 6.5. *Let v be a place of L . Let*

$$\kappa_v = \begin{cases} 1 & \text{if } v \in M_L^0, \\ \left(\frac{1}{2}\right)^{D_v} & \text{if } v \in M_L^\infty. \end{cases}$$

Then

$$\max\{\|\mu_1\|_v, \|\mu_3\|_v\}^2 \geq \kappa_v^2 \cdot \min\{1, \|\theta_1 - \theta_3\|_v\}^2 \cdot \max\{1, \|\mu_1\|_v\} \cdot \max\{1, \|\mu_3\|_v\}.$$

Proof. There is nothing to prove unless, $\|\mu_i\|_v \leq 1$ for both $i = 1, 3$. In this case, it is enough to show that

$$\max\{\|\mu_1\|_v, \|\mu_3\|_v\} \geq \kappa_v \cdot \|\theta_1 - \theta_3\|_v. \tag{30}$$

Suppose first that v is finite and let \mathfrak{P} be the prime ideal of \mathcal{O}_L corresponding to v . Then (30) is equivalent to

$$\min\{\text{ord}_{\mathfrak{P}}(\mu_1), \text{ord}_{\mathfrak{P}}(\mu_3)\} \leq \text{ord}_{\mathfrak{P}}(\theta_1 - \theta_3).$$

Let $k = \min\{\text{ord}_{\mathfrak{P}}(\mu_1), \text{ord}_{\mathfrak{P}}(\mu_3)\}$. Then $\mathfrak{P}^k \mid \mu_i$ for $i = 1, 3$. Recall that $\mu_i = a_0X - \theta_iY$. Thus \mathfrak{P}^k divides both

$$(\theta_1 - \theta_3)a_0X = \theta_1\mu_3 - \theta_3\mu_1 \quad \text{and} \quad (\theta_1 - \theta_3)Y = \mu_3 - \mu_1.$$

However $\gcd(X, Y) = \gcd(a_0, Y) = 1$, thus $\mathfrak{P}^k \mid (\theta_1 - \theta_3)$ as desired.

Next suppose v is infinite. As $(\theta_1 - \theta_3)Y = \mu_3 - \mu_1$, and $Y \neq 0$ is a rational integer, we have

$$\|\theta_1 - \theta_3\|_v \leq \|\mu_1 - \mu_3\|_v = |\mu_1 - \mu_3|_v^{D_v} \leq 2^{D_v} \cdot \max\{|\mu_1|_v, |\mu_3|_v\}^{D_v} \leq 2^{D_v} \cdot \max\{\|\mu_1\|_v, \|\mu_3\|_v\}.$$

This completes the proof. \square

Lemma 6.6. *Let*

$$c_{16} := c_{14} + 2 \log 2 + 2h(\theta) + h(\tau).$$

Then

$$h(\varepsilon) \leq c_{16} + c_{15} \cdot \log B. \quad (31)$$

Proof. First note that

$$\begin{aligned} h(\mu_3/\mu_1) &= \frac{1}{D} \sum_{v \in M_L} \log \max\{1, \|\mu_3/\mu_1\|_v\} \\ &= \frac{1}{D} \sum_{v \in M_L} \log \max\{1, \|\mu_3/\mu_1\|_v\} + \frac{1}{D} \sum_{v \in M_L} \log \|\mu_1\|_v \quad (\text{from (19)}) \\ &= \frac{1}{D} \sum_{v \in M_L} \log \max\{\|\mu_1\|, \|\mu_3\|_v\} \\ &\geq \frac{1}{2}(h(\mu_1) + h(\mu_3)) + \frac{1}{D} \sum_{v \in M_L} (\log \kappa_v + \log \min\{1, \|(\theta_1 - \theta_3)\|_v\}) \end{aligned}$$

by Lemma 6.5. However μ_1, μ_3 are conjugates of μ , thus $h(\mu_1) = h(\mu_3) = h(\mu)$. Moreover,

$$\frac{1}{D} \sum_{v \in M_L} \log \kappa_v = -\frac{\log 2}{D} \sum_{v \in M_L^\infty} D_v = -\log 2.$$

Thus

$$\begin{aligned} h(\mu) &\leq h(\mu_3/\mu_1) + \log 2 - \frac{1}{D} \sum_{v \in M_L} \log \min\{1, \|(\theta_1 - \theta_3)\|_v\} \\ &= h(\mu_3/\mu_1) + \log 2 + \frac{1}{D} \sum_{v \in M_L} \log \max\{1, \|(\theta_1 - \theta_3)^{-1}\|_v\} \\ &= h(\mu_3/\mu_1) + \log 2 + h((\theta_1 - \theta_3)^{-1}) \\ &\leq h(\mu_3/\mu_1) + 2 \log 2 + 2h(\theta), \end{aligned}$$

by Lemmas 5.1 and 5.2. But $\varepsilon = \tau^{-1}\mu$, thus

$$h(\varepsilon) \leq h(\mu_3/\mu_1) + 2 \log 2 + 2h(\theta) + h(\tau).$$

Applying Lemma 6.4 completes the proof. \square

It is worthwhile to take stock for a moment. The inequality (31) relates the height of $\varepsilon = \delta_1^{b_1} \cdots \delta_r^{b_r}$ to $B = \max\{|b_1|, \dots, |b_r|\}$. The constants c_7, \dots, c_{16} are given explicitly in terms of $\theta, \tau, \delta_1, \dots, \delta_r$ (all belonging to K), the prime ideals of K belonging to S , the signature of K , and the degree D , which can be deduced from the Galois group of the minimal polynomial of θ . We do not need the field L explicitly.

Lemma 6.7. *Let U be any subset of $S \cup M_K^\infty$ of size r . Let \mathcal{M} be the $r \times r$ -matrix*

$$\mathcal{M} = (\log \|\delta_j\|_v)_{v \in U, 1 \leq j \leq r}.$$

The matrix \mathcal{M} is invertible. Let c_{17} be the largest of the absolute values of the entries of \mathcal{M}^{-1} . Then

$$B \leq 2d \cdot c_{17} \cdot h(\varepsilon).$$

Proof. The determinant of \mathcal{M} is in fact

$$\pm \left(\prod_{v \in U} D_v \right) \cdot R(\delta_1, \dots, \delta_r)$$

where $R(\delta_1, \dots, \delta_r)$ is the regulator of system of S -units $\delta_1, \dots, \delta_r$, and therefore does not vanish; see [Bugeaud and Györy 1996b, Section 3]. Consider the vectors $\mathbf{b} := [b_j]_{j=1, \dots, r}$ and $\mathbf{u} := [\log \|\varepsilon\|_v]_{v \in U}$ in \mathbb{R}^r . As $\varepsilon = \delta_1^{b_1} \cdots \delta_r^{b_r}$ we see that $\mathbf{u} = \mathcal{M}\mathbf{b}$ and so $\mathbf{b} = \mathcal{M}^{-1}\mathbf{u}$. It follows, for $j = 1, \dots, r$ that

$$|b_j| \leq c_{17} \cdot \sum_{v \in U} |\log \|\varepsilon\|_v| \leq c_{17} \cdot \sum_{v \in M_K} \log \max\{1, \|\varepsilon\|_v\} + \log \max\{1, \|\varepsilon^{-1}\|_v\} = 2d \cdot c_{17} \cdot h(\varepsilon)$$

as required. □

Remark. Observe that there are $r + 1$ possibilities for the set U . To compute c_{17} in practice, we iterate across all such sets and select c_{17} as the smallest possible value across each of the associated $r + 1$ matrices \mathcal{M} .

Proposition 6.8. *Let*

$$c_{18} := 2d \cdot c_{17} \cdot c_{16}, \quad c_{19} := 2d \cdot c_{17} \cdot c_{15}, \quad c_{20} := 2c_{18} + \max\{2c_{19} \log c_{19}, 4e^2\}.$$

Then

$$B \leq c_{20}. \tag{32}$$

Proof. Combining Lemma 6.7 with (31) we have

$$B \leq 2d \cdot c_{17} \cdot (c_{16} + c_{15} \cdot \log B) \leq c_{18} + c_{19} \log B.$$

The proposition follows from a result of Pethő and de Weger [1998, Lemma B.1 in Appendix B]. □

6.1. Example 1.4 continued. We give some further details for Example 1.4. Here $a_0 = 5$ and the minimal polynomial for θ is

$$x^{11} + x^{10} + 20x^9 + 25x^8 + 750x^7 + 625x^6 + 18750x^5 + 468750x^3 + 7812500x - 19531250.$$

The field $K = \mathbb{Q}(\theta)$ has degree 11, and signature $(1, 5)$. In this case we have two possibilities for $(\tau, \delta_1, \dots, \delta_r)$; one has S -unit rank $r = 9$ and the other has S -unit rank $r = 10$. We take a closer look

at one of these two possibilities, with $r = 10$. The set S is composed of the following five primes of ramification degree 1 and inertial degree 1:

$$p_1 = \langle 11, 3 + \theta \rangle, \quad p_2 = \langle 7, 1 + \theta \rangle, \quad p_3 = \langle 5, \phi \rangle, \quad p_4 = \langle 3, 5 + \theta \rangle, \quad p_5 = \langle 2, 1 + \theta \rangle, \quad (33)$$

where

$$\phi = \frac{1}{5^9} (4\theta^{10} + 9\theta^9 + 185\theta^8 + 425\theta^7 + 4625\theta^6 + 13750\theta^5 + 131250\theta^4 + 750000\theta^3 + 3203125\theta^2 + 26953125\theta + 5859375).$$

The bound for B given by Proposition 6.8 is

$$B \leq 1.57 \times 10^{222}.$$

7. Controlling the valuations of $a_0X - \theta Y$

In this section and the next, we will suppose that we have a bound

$$B \leq \mathcal{B}_\infty \quad (34)$$

and we will explain a method for replacing this bound with what is hopefully a smaller bound. Our subsequent constants will depend on \mathcal{B}_∞ . Initially we may take $\mathcal{B}_\infty = c_{20}$ by Proposition 6.8. However, if we succeed in obtaining a smaller bound for B , we may replace \mathcal{B}_∞ by that bound and repeat the process.

We shall replace the reduction step using linear forms in p -adic logarithms as in the paper of Tzanakis and de Weger [1989]. In particular we will eliminate all computations with completions of extensions of K , as these are extremely tedious and error-prone.

7.1. The bounds \mathcal{B}_∞ , \mathcal{B}_1 and \mathcal{B}_2 . Henceforth \mathcal{B}_∞ , \mathcal{B}_1 and \mathcal{B}_2 will denote the known bounds for the ∞ -norm, 1-norm and 2-norm of our exponent vector $\mathbf{b} = [b_j]_{j=1,\dots,r}$:

$$\mathcal{B} := \|\mathbf{b}\|_\infty \leq \mathcal{B}_\infty, \quad \|\mathbf{b}\|_1 \leq \mathcal{B}_1, \quad \|\mathbf{b}\|_2 \leq \mathcal{B}_2. \quad (35)$$

Initially, thanks to Proposition 6.8, we can make the assignments

$$\mathcal{B}_\infty = c_{20}, \quad \mathcal{B}_2 = \sqrt{r} \cdot \mathcal{B}_\infty, \quad \mathcal{B}_1 = r \cdot \mathcal{B}_\infty \quad (\text{initial values for } \mathcal{B}_\infty, \mathcal{B}_1, \mathcal{B}_2). \quad (36)$$

However, as we progress in our algorithm, we will update the values of \mathcal{B}_∞ , \mathcal{B}_1 , \mathcal{B}_2 so that (35) is still satisfied.

Given a lattice $L \subseteq \mathbb{Z}^r$ and a vector $\mathbf{w} \in \mathbb{Z}^r$, we denote by $D(L, \mathbf{w})$ the shortest length of a vector belonging to the coset $\mathbf{w} + L$. This value can be computed using a closest vector algorithm. Indeed, for $\mathbf{v} \in \mathbb{Z}^r$, write $\mathbf{c}(L, \mathbf{v})$ for the closest vector in L to \mathbf{v} (if there is more than one at the closest distance, choose any of them).

Lemma 7.1. $D(L, \mathbf{w}) = \|\mathbf{w} + \mathbf{c}(L, -\mathbf{w})\|_2$.

Proof. Let $\mathbf{l} \in L$ and suppose

$$\|\mathbf{w} + \mathbf{l}\|_2 < \|\mathbf{w} + \mathbf{c}(L, -\mathbf{w})\|_2.$$

Then

$$\|l - (-w)\|_2 < \|c(L, -w) - (-w)\|_2.$$

Thus l is a vector belonging to L that is strictly closer to $-w$ than $c(L, -w)$ giving a contradiction. \square

Our first goal is to use the bounds (35) to deduce bounds on the valuations $\text{ord}_{\mathfrak{p}}(a_0X - \theta Y)$ for $\mathfrak{p} \in S$.

Proposition 7.2. *Let $\mathfrak{p} \in S$ and let p be the rational prime below \mathfrak{p} . Let $k \geq 1$. Then there is some $\theta_0 \in \mathbb{Z}$ such that*

$$\theta \equiv \theta_0 \pmod{\mathfrak{p}^k}.$$

Write

$$\mathfrak{a} := (p\mathcal{O}_K)/\mathfrak{p}, \quad \tau\mathcal{O}_K = \mathcal{T}_1/\mathcal{T}_2, \tag{37}$$

where \mathcal{T}_1 and \mathcal{T}_2 are coprime ideals. The following hold:

(i) If $\text{gcd}(\mathfrak{a}^k, \theta - \theta_0) \neq \text{gcd}(\mathfrak{a}^k, \mathcal{T}_1)$ then

$$\text{ord}_{\mathfrak{p}}(a_0X - \theta Y) \leq k - 1. \tag{38}$$

Suppose $\text{gcd}(\mathfrak{a}^k, \theta - \theta_0) = \text{gcd}(\mathfrak{a}^k, \mathcal{T}_1)$. Let

$$\mathfrak{b} := \mathfrak{a}^k / \text{gcd}(\mathfrak{a}^k, \mathcal{T}_1).$$

Let

$$k' := \max_{\mathfrak{q} | \mathfrak{b}} \left\lceil \frac{\text{ord}_{\mathfrak{q}}(\mathfrak{b})}{e(\mathfrak{q} | p)} \right\rceil;$$

this satisfies $\mathfrak{b} \cap \mathbb{Z} = p^{k'}\mathbb{Z}$ and therefore $(\mathbb{Z}/p^{k'}\mathbb{Z})^\times$ naturally injects into $(\mathcal{O}_K/\mathfrak{b})^\times$. Given $u \in K^\times$ whose support is coprime with \mathfrak{b} , denote its image in $(\mathcal{O}_K/\mathfrak{b})^\times / (\mathbb{Z}/p^{k'}\mathbb{Z})^\times$ by \bar{u} . Let

$$\phi : \mathbb{Z}^r \rightarrow (\mathcal{O}_K/\mathfrak{b})^\times / (\mathbb{Z}/p^{k'}\mathbb{Z})^\times, \quad (n_1, \dots, n_r) \mapsto \bar{\delta}_1^{n_1} \dots \bar{\delta}_r^{n_r}.$$

Write

$$\tau_0 := \frac{(\theta_0 - \theta)}{\tau}.$$

Then the support of τ_0 is coprime with \mathfrak{b} :

(ii) If $\bar{\tau}_0$ does not belong to $\text{Image}(\phi)$ then (38) holds.

(iii) Suppose $\bar{\tau}_0 = \phi(w)$ for some $w \in \mathbb{Z}^r$. Let $L = \text{Ker}(\phi)$ and suppose $D(L, w) > \mathcal{B}_2$. Then (38) holds.

Proof. Let \mathfrak{p} , p , and k be as in the statement of the proposition. We suppose that

$$\text{ord}_{\mathfrak{p}}(a_0X - \theta Y) \geq k \tag{39}$$

and show that this leads to a contradiction under the hypotheses of any of (i), (ii), (iii). Recall $k \geq 1$.

From the proof of Lemma 2.3, we know $p \nmid Y$.

Since $e(\mathfrak{p} | p) = f(\mathfrak{p} | p) = 1$, we have $\mathcal{O}_K/\mathfrak{p}^k \cong \mathbb{Z}/p^k$. Thus there is some $\theta_0 \in \mathbb{Z}$ such that $\theta - \theta_0 \equiv 0 \pmod{\mathfrak{p}^k}$. However, $a_0X - \theta Y \equiv 0 \pmod{\mathfrak{p}^k}$ and so therefore $a_0X - \theta_0Y \equiv 0 \pmod{\mathfrak{p}^k}$. However

$a_0X - \theta_0Y \in \mathbb{Z}$. Thus, recalling that $e(\mathfrak{p} | p) = 1$, we have $a_0X - \theta_0Y \equiv 0 \pmod{p^k}$. From (15)

$$Y(\theta_0 - \theta) \equiv \tau \cdot \delta_1^{b_1} \cdots \delta_r^{b_r} \pmod{(p\mathcal{O}_K)^k}.$$

Note that the prime \mathfrak{p} belongs to the support of the δ_i . However the other primes $\mathfrak{p}' | p$, $\mathfrak{p}' \neq \mathfrak{p}$ do not belong to the support of the δ_i . We now eliminate \mathfrak{p} ; as in the statement of the proposition we take $\mathfrak{a} := (p\mathcal{O}_K)/\mathfrak{p}$. Then

$$Y(\theta_0 - \theta) \equiv \tau \cdot \delta_1^{b_1} \cdots \delta_r^{b_r} \pmod{\mathfrak{a}^k}.$$

Observe that \mathfrak{a} is coprime to the support of the δ_i and Y . Recall $\tau\mathcal{O}_K = \mathcal{T}_1/\mathcal{T}_2$, where $\mathcal{T}_1, \mathcal{T}_2$ are coprime integral ideals. By Lemma 4.2, the ideal \mathcal{T}_2 is supported on S and therefore coprime to \mathfrak{a} . We therefore have a contradiction if $\gcd(\theta_0 - \theta, \mathfrak{a}^k) \neq \gcd(\mathcal{T}_1, \mathfrak{a}^k)$. This proves (i). Suppose $\gcd(\theta_0 - \theta, \mathfrak{a}^k) = \gcd(\mathcal{T}_1, \mathfrak{a}^k)$. Then

$$Y \cdot \tau_0 \equiv \delta_1^{b_1} \cdots \delta_r^{b_r} \pmod{\mathfrak{b}},$$

where Y, τ_0 , and δ_i all have support disjoint from \mathfrak{b} . As in the proposition, k' is the smallest positive integer such that $\mathfrak{b} | p^{k'}$, and thus $(\mathbb{Z}/p^{k'}\mathbb{Z})^\times$ is a subgroup of $(\mathcal{O}_K/\mathfrak{b})^\times$ containing the image of Y . Therefore $\bar{Y} = \bar{1}$ and $\phi(\mathfrak{b}) = \bar{\tau}_0$. If $\bar{\tau}_0 \notin \text{Image}(\phi)$, then we have a contradiction, and so our original assumption (39) is false. This proves (ii).

Suppose now that $\bar{\tau}_0 \in \text{Image}(\phi)$ and write $\bar{\tau}_0 = \phi(\mathfrak{w})$ with $\mathfrak{w} \in \mathbb{Z}^r$. Then $\mathfrak{b} \in \mathfrak{w} + L$. Thus $\|\mathfrak{b}\|_2 \geq D(L, \mathfrak{w})$. If $D(L, \mathfrak{w}) > \mathcal{B}_2$ then $\|\mathfrak{b}\|_2 > \mathcal{B}_2$ and we contradict (35). This proves (iii). \square

Remarks. • θ_0 can be easily computed using Hensel’s lemma.

• To apply the proposition in practice, it is necessary to compute the abelian group structure of $(\mathcal{O}_K/\mathfrak{b})^\times$ for ideals \mathfrak{b} of very large norm (but supported on the primes above p). For this we may apply the algorithms in [Cohen 2000, Section 4.2].

• $\mathfrak{c}(-\mathfrak{w}, L)$ (and therefore $D(\mathfrak{w}, L)$) can be computed using a closest vector algorithm such as Fincke and Pohst [1985].

• To effectively apply Proposition 7.2 in practice, we need to guess a value of k such that $D(L, \mathfrak{w}) > \mathcal{B}_2$. We expect $D(L, \mathfrak{w})$ to be around $I^{1/r}$, where I is the index $[\mathbb{Z}^r : L]$. Let us make two simplifying assumptions: the first is that ϕ is surjective, and the second is that $\gcd(\mathfrak{a}^k, \mathcal{T}_1) = 1$ so that $\mathfrak{b} = \mathfrak{a}^k$ and $k' = k$. Then

$$I = \frac{\#(\mathcal{O}_K/\mathfrak{a}^k)^\times}{\#(\mathbb{Z}/p^k\mathbb{Z})^\times} \approx \frac{\text{Norm}(\mathfrak{a})^k}{p^k} = p^{(d-2)k},$$

where d is the degree of K . Thus we should expect a contradiction if $p^{(d-2)k/r}$ is much bigger than \mathcal{B}_2 , or equivalently

$$k \gg \frac{r \log \mathcal{B}_2}{(d-2) \log p}.$$

This heuristic gives a good guide for which values of k to try.

7.2. Example 1.4 continued. We continue giving details for Example 1.4, and in particular for the tuple $(\tau, \delta_1, \dots, \delta_{10})$ alluded to on page 690. In Section 6 we noted that $B \leq 1.57 \times 10^{222}$. Thus we take

$$B_\infty = 1.57 \times 10^{222}, \quad B_2 = \sqrt{10} \cdot B_\infty \approx 4.96 \times 10^{222}. \tag{40}$$

We let $\mathfrak{p} = \mathfrak{p}_1 = \langle 11, 3 + \theta \rangle$ which is a prime above 11. The above heuristic suggests that we choose k to be larger than

$$\frac{10 \log B_2}{(11 - 2) \cdot \log 11} \approx 237.60.$$

Our program tries $k = 238$. It turns out (in the notation of Proposition 7.2) that $\gcd(\mathfrak{a}^k, \theta - \theta_0) = \gcd(\mathfrak{a}^k, \mathcal{T}_1) = 1$, thus $\mathfrak{b} = \mathfrak{a}^k$, and moreover $k' = k = 238$. The map ϕ is surjective, and thus L does indeed have index

$$I = \frac{\#(\mathcal{O}_K/\mathfrak{a}^k)^\times}{\#(\mathbb{Z}/\mathfrak{p}^k\mathbb{Z})^\times} = 2^7 \times 3^2 \times 5 \times 11^{2133} \times 61 \times 7321 \approx 5.02 \times 10^{2230}.$$

We do not give L as its basis vectors are naturally huge. However, we find that

$$D(L, \mathbf{w}) \approx 1.14 \times 10^{223}.$$

This is much larger than B_2 and we therefore know from Proposition 7.2 that $\text{ord}_{\mathfrak{p}}(a_0X - \theta Y) \leq k - 1 = 237$.

It is interesting to note that $I^{1/10} \approx 1.18 \times 10^{223}$ which is rather close to $D(L, \mathbf{w})$. If instead we take $k = 237$, we find that $I^{1/10} \approx 1.36 \times 10^{222}$ and $D(L, \mathbf{w}) \approx 9.55 \times 10^{221}$ which is somewhat less than B_2 . We have generally found the above heuristic to be remarkably accurate in predicting a good choice for k .

Now let $\mathfrak{p}_1, \dots, \mathfrak{p}_5$ be the primes of S as in (33), where $\mathfrak{p}_1 = \mathfrak{p}$ as above. Proposition 7.2 gives upper bounds 237, 292, 354, 518, 821 for $\text{ord}_{\mathfrak{p}_j}(a_0X - \theta Y)$ with $j = 1, \dots, 5$ respectively.

8. Linear forms in real logarithms

In this section, we determine bounds on linear forms in logarithms which we will subsequently use in Section 9 to successively reduce the large upper bound B_∞ established in Section 6.

We let $s := \#S$ and write

$$S = \{\mathfrak{p}_1, \dots, \mathfrak{p}_s\}.$$

Using Proposition 7.2, we suppose that we have obtained, for $1 \leq j \leq s$, integers k_j such that

$$\text{ord}_{\mathfrak{p}_j}(a_0X - \theta Y) \leq k_j - 1. \tag{41}$$

Recall that

$$\delta_1^{b_1} \dots \delta_r^{b_r} = \varepsilon = (a_0X - \theta Y)/\tau. \tag{42}$$

We write $k'_j := \text{ord}_{\mathfrak{p}_j}(\tau)$, and $k''_j := k_j - 1 - k'_j$. We obtain

$$-k'_j \leq \text{ord}_{\mathfrak{p}_j}(\varepsilon) \leq k''_j. \tag{43}$$

8.1. Updating B_1 and B_2 . Recall that B_∞, B_1, B_2 are respectively the known bounds for $\|\mathbf{b}\|_\infty, \|\mathbf{b}\|_1, \|\mathbf{b}\|_2$ as in (35). Initially we take these as in (36). In practice, we are often able to update B_1 and B_2 with

a smaller bound after each iteration of Proposition 7.2. Let (u, v) be the signature of K . Since r is the rank of the S -unit group \mathcal{O}_S^\times , we have

$$r = u + v - 1 + s.$$

Recall our convention (page 681) on the choice of S -unit basis $\delta_1, \dots, \delta_r$: we suppose that the basis is chosen so that $\delta_1, \dots, \delta_{u+v-1}$ is in fact a basis for the unit group modulo torsion. Thus $\log \|\delta_i\|_v = 0$ for all $v \in M_K^0$ and $1 \leq i \leq u + v - 1$. Let \mathcal{M}_0 denote the $s \times s$ matrix

$$\mathcal{M}_0 = (\log \|\delta_j\|_v)_{v \in S, u+v \leq j \leq r}.$$

In Lemma 6.7 let $U = \{v_1, \dots, v_r\}$ where v_1, \dots, v_{u+v-1} are any $u + v - 1$ elements of M_K^∞ and the remainder are the elements of S . Then, in the notation of Lemma 6.7,

$$\mathcal{M} = \left(\begin{array}{c|c} * & * \\ \hline 0 & \mathcal{M}_0 \end{array} \right).$$

Since \mathcal{M} is invertible by Lemma 6.7, it follows that \mathcal{M}_0 is invertible. We partition our exponent vector \mathbf{b} as

$$\mathbf{b} = [\mathbf{b}' \mid \mathbf{b}''], \quad \mathbf{b}' = [b_i]_{i=1, \dots, u+v-1}, \quad \mathbf{b}'' = [b_i]_{i=u+v, \dots, r}.$$

Write $\mathbf{u}'' := [\log \|\varepsilon\|_v]_{v \in S}$ in \mathbb{R}^s . By the above, we have $\mathbf{u}'' = \mathcal{M}_0 \mathbf{b}''$ and thus $\mathbf{b}'' = \mathcal{M}_0^{-1} \mathbf{u}''$. That is, for $1 \leq i \leq s$,

$$b_{u+v-1+i} = m_{i1} \log \|\varepsilon\|_{p_1} + \dots + m_{is} \log \|\varepsilon\|_{p_s},$$

where $\mathcal{M}_0^{-1} = [m_{ij}]$. It follows that

$$|b_{u+v-1+i}| \leq |m_{i1}| \cdot |\log \|\varepsilon\|_{p_1}| + \dots + |m_{is}| \cdot |\log \|\varepsilon\|_{p_s}|. \tag{44}$$

Applying Proposition 7.2 to any \mathfrak{p}_j for $1 \leq j \leq s$ and using (43) and (21), we obtain

$$|\log \|\varepsilon\|_{\mathfrak{p}_j}| \leq \log(\text{Norm}(\mathfrak{p}_j)) \cdot \max\{|k'_j|, |k''_j|\}.$$

Write

$$\rho'_i := \sum_{j=1}^s |m_{i,j}| \cdot \log(\text{Norm}(\mathfrak{p}_j)) \cdot \max\{|k'_j|, |k''_j|\}, \quad \rho_i = \min\{\mathcal{B}_\infty, \rho'_i\}. \tag{45}$$

From equation (44) it follows that $|b_{u+v-1+i}| \leq \rho'_i$ for $1 \leq i \leq s$. However, $\max\{|b_i|\}_{i=1}^r = \|\mathbf{b}\|_\infty \leq \mathcal{B}_\infty$, so we know that $|b_{u+v-1+i}| \leq \mathcal{B}_\infty$. We deduce that

$$|b_{u+v-1+i}| \leq \rho_i, \quad 1 \leq i \leq s. \tag{46}$$

Hence

$$\|\mathbf{b}\|_1 = \|\mathbf{b}'\|_1 + \|\mathbf{b}''\|_1 \leq (u + v - 1)\mathcal{B}_\infty + \rho_1 + \dots + \rho_s$$

and

$$\|\mathbf{b}\|_2 = \sqrt{\|\mathbf{b}'\|_2^2 + \|\mathbf{b}''\|_2^2} \leq \sqrt{(u + v - 1)\mathcal{B}_\infty^2 + \rho_1^2 + \dots + \rho_s^2}.$$

We now update our values for \mathcal{B}_1 and \mathcal{B}_2 :

$$\mathcal{B}_1 = (u + v - 1)\mathcal{B}_\infty + \rho_1 + \dots + \rho_s, \tag{47}$$

$$\mathcal{B}_2 = \sqrt{(u + v - 1)\mathcal{B}_\infty^2 + \rho_1^2 + \dots + \rho_s^2}. \tag{48}$$

Note, since by (45) we have $\rho_i \leq \mathcal{B}_\infty$, these new values for \mathcal{B}_1 and \mathcal{B}_2 are bounded above by the old values given in (36). In practice we usually find that these give significantly better bounds for $\|\mathbf{b}\|_1, \|\mathbf{b}\|_2$.

8.2. Embeddings and improving the initial bound (34). To improve our initial bound (34), we rely on the inequality $B \leq 2c_{17} \cdot d \cdot h(\varepsilon)$ furnished by Lemma 6.7. However $h(\varepsilon) = h(\varepsilon^{-1})$ and so

$$B \leq 2c_{17} \sum_{v \in M_K} \log \max\{1, \|\varepsilon^{-1}\|_v\}.$$

Since ε is an S -unit, for $v \notin M_K^\infty \cup S$, we have $\|\varepsilon\|_v = 1$. Thus

$$B \leq 2c_{17} \sum_{v \in M_K^\infty \cup S} \log \max\{1, \|\varepsilon^{-1}\|_v\}. \tag{49}$$

Therefore, to obtain a better bound for B , it is enough to gain good control on the contributions to the sum on the right-hand side of (49).

Lemma 8.1. *Let*

$$c_{21} = \sum_{i=1}^s \max\{0, k_i''\} \cdot \log(\text{Norm}(\mathfrak{p}_i)).$$

Then

$$B \leq 2c_{17} \left(c_{21} + \sum_{v \in M_K^\infty} \log \max\{1, \|\varepsilon^{-1}\|_v\} \right). \tag{50}$$

Proof. From (43) and (21) we have

$$\sum_{v \in S} \log \max\{1, \|\varepsilon^{-1}\|_v\} \leq c_{21}.$$

The lemma now follows from (49). □

We shall write

$$M_K^\infty = M_K^\mathbb{R} \cup M_K^\mathbb{C},$$

where $M_K^\mathbb{R}$ and $M_K^\mathbb{C}$ are respectively the sets of real and complex places. Recall that (u, v) denotes the signature of K . Thus we have embeddings

$$\sigma_1, \dots, \sigma_u, \quad \sigma_{u+1}, \dots, \sigma_{u+v}, \overline{\sigma_{u+1}}, \dots, \overline{\sigma_{u+v}}$$

of K , where σ_i are real embeddings for $1 \leq i \leq u$, and $\sigma_{u+i}, \overline{\sigma_{u+i}}$ are pairs of complex conjugate embeddings. Let

$$\mathcal{E}_K^\mathbb{R} := \{\sigma_1, \dots, \sigma_u\}, \quad \mathcal{E}_K^\mathbb{C} := \{\sigma_{u+1}, \dots, \sigma_{u+v}\}. \tag{51}$$

For the membership of $\mathcal{E}_K^{\mathbb{C}}$, we are making an arbitrary choice of a member from each pair of conjugate complex embeddings, but that is unimportant. Note that $M_K^{\mathbb{R}}$ is in one-to-one correspondence with $\mathcal{E}_K^{\mathbb{R}}$ and $M_K^{\mathbb{C}}$ is in one-to-one correspondence with $\mathcal{E}_K^{\mathbb{C}}$. We consider the contribution to the sum (50) coming from $\nu \in M_K^{\mathbb{C}}$, or equivalently from $\sigma \in \mathcal{E}_K^{\mathbb{C}}$.

Let $\Im(z)$ denote the imaginary part of a complex number z .

Lemma 8.2. *Let*

$$c_{22} = 2 \sum_{\sigma \in \mathcal{E}_K^{\mathbb{C}}} \log \max \left\{ 1, \frac{|\sigma(\tau)|}{|\Im(\sigma(\theta))|} \right\}.$$

Then

$$B \leq 2c_{17} \left(c_{21} + c_{22} + \sum_{\sigma \in \mathcal{E}_K^{\mathbb{R}}} \log \max \{ 1, |\sigma(\varepsilon)|^{-1} \} \right). \quad (52)$$

Proof. Note that (50) can be rewritten as

$$B \leq 2c_{17} \left(c_{21} + 2 \sum_{\sigma \in \mathcal{E}_K^{\mathbb{C}}} \log \max \{ 1, |\sigma(\varepsilon)|^{-1} \} + \sum_{\sigma \in \mathcal{E}_K^{\mathbb{R}}} \log \max \{ 1, |\sigma(\varepsilon)|^{-1} \} \right).$$

Let $\sigma \in \mathcal{E}_K^{\mathbb{C}}$. Then as $a_0X - \theta Y = \tau \cdot \varepsilon$, we have

$$|\sigma(\varepsilon)| = \frac{1}{|\sigma(\tau)|} \cdot |\sigma(a_0X - \theta Y)| \geq \frac{1}{|\sigma(\tau)|} \cdot |Y| \cdot |\Im(\sigma(\theta))| \geq \frac{|\Im(\sigma(\theta))|}{|\sigma(\tau)|},$$

because of our assumption $|Y| \neq 0$. The lemma follows. \square

The following is immediate.

Proposition 8.3. *If K is totally imaginary then*

$$B \leq 2c_{17}(c_{21} + c_{22}).$$

8.3. The nontotally complex case. Suppose now that K has one or more real embeddings. Recall that the signature of K is (u, v) . Thus $u \geq 1$.

Lemma 8.4. *If $u = 1$ we let $c_{23} := 1$. If $u \geq 2$ we let*

$$c_{23} := \min \left\{ \frac{|\sigma(\theta) - \sigma'(\theta)|}{|\sigma(\tau)| + |\sigma'(\tau)|} : \sigma, \sigma' \in \mathcal{E}_K^{\mathbb{R}}, \sigma \neq \sigma' \right\}.$$

Then there is at most one $\sigma \in \mathcal{E}_K^{\mathbb{R}}$ such that $|\sigma(\varepsilon)| < c_{23}$.

Proof. Suppose otherwise. Then there are $\sigma, \sigma' \in \mathcal{E}_K^{\mathbb{R}}$ with $\sigma \neq \sigma'$ such that $|\sigma(\varepsilon)| < c_{23}$ and $|\sigma'(\varepsilon)| < c_{23}$. As $a_0X - \theta Y = \tau \cdot \varepsilon$ we find that

$$|a_0X - \sigma(\theta)Y| < c_{23} \cdot |\sigma(\tau)|, \quad |a_0X - \sigma'(\theta)Y| < c_{23} \cdot |\sigma'(\tau)|.$$

Thus

$$|\sigma(\theta) - \sigma'(\theta)| \cdot |Y| < c_{23} \cdot (|\sigma(\tau)| + |\sigma'(\tau)|).$$

Recall our assumption that $Y \neq 0$. This inequality now contradicts our definition of c_{23} . \square

Lemma 8.5. *Let*

$$c_{24} := c_{21} + c_{22} + (u - 1) \log \max\{1, c_{23}^{-1}\}, \quad c_{25} := \exp(c_{24}) \quad \text{and} \quad c_{26} := \frac{1}{2c_{17}}.$$

Suppose $B > 2c_{17} \cdot c_{24}$. *Let* $\sigma \in \mathcal{E}_K^{\mathbb{R}}$ *be chosen so that* $|\sigma(\varepsilon)|$ *is minimal. Then*

$$|\sigma(\varepsilon)| \leq c_{25} \cdot \exp(-c_{26} \cdot B). \tag{53}$$

Proof. From Lemma 8.4, we have

$$|\sigma'(\varepsilon)| \geq c_{23}$$

for all $\sigma' \in \mathcal{E}_K^{\mathbb{R}}$ with $\sigma' \neq \sigma$. From (52) we deduce that

$$B \leq 2c_{17}(c_{21} + c_{22} + (u - 1) \log \max\{1, c_{23}^{-1}\} + \log \max\{1, |\sigma(\varepsilon)|^{-1}\}) = 2c_{17}(c_{24} + \log \max\{1, |\sigma(\varepsilon)|^{-1}\}).$$

It follows that

$$\log \max\{1, |\sigma(\varepsilon)|^{-1}\} \geq \frac{1}{2c_{17}}B - c_{24} = c_{26} \cdot B - c_{24}.$$

The hypothesis $B > 2c_{17} \cdot c_{24}$ forces the right-hand side to be positive, and so the left-hand side must simply be $\log |\sigma(\varepsilon)|^{-1}$. After exponentiating and rearranging, we obtain (53). □

8.4. Approximate relations. As in Lemma 8.5 we shall let $\sigma \in \mathcal{E}_K^{\mathbb{R}}$ be the real embedding that makes $|\sigma(\varepsilon)|$ minimal. Recall that the signature of K is (u, v) ; we keep the assumption that $u \geq 1$. Let

$$n := r + v. \tag{54}$$

In this section we introduce additional unknown integers b_{r+1}, \dots, b_n , closely related to the exponents b_1, \dots, b_r found in (15). We shall use Lemma 8.5 to write down $d - 2$ linear forms in b_1, \dots, b_n with real coefficients, whose values are very small. We shall later give a method, based on standard ideas originally due to de Weger, that uses these “approximate relations” to reduce our bound for $B = \max(|b_1|, \dots, |b_r|, 1)$.

We label the elements of $\mathcal{E}_K^{\mathbb{R}}$ and $\mathcal{E}_K^{\mathbb{C}}$ as in (51), where $\sigma_1 = \sigma$. Write

$$\theta_j = \sigma_j(\theta), \quad \tau_j = \sigma_j(\tau), \quad \varepsilon_j = \sigma_j(\varepsilon), \quad \delta_{i,j} = \sigma_j(\delta_i), \quad 1 \leq j \leq u + v, 1 \leq i \leq r.$$

Let $2 \leq j \leq u + v$ and write

$$z_j := \frac{a_0X - \theta_1Y}{a_0X - \theta_jY}.$$

Observe that

$$\begin{aligned} Y(\theta_1 - \theta_j) &= (a_0X - \theta_jY) - (a_0X - \theta_1Y) \\ &= (a_0X - \theta_jY) \cdot (1 - z_j) \\ &= \tau_j \cdot \delta_{1,j}^{b_1} \cdots \delta_{r,j}^{b_r} \cdot (1 - z_j). \end{aligned} \tag{55}$$

In the following lemma, as always, \log denotes the principal determination of the logarithm (i.e., the imaginary part of \log lies in $(-\pi, \pi]$).

Lemma 8.6. *Let*

$$c_{27} := \frac{|\tau_1| \cdot c_{25}}{\min\{|\tau_i| : \sigma_i \in \mathcal{E}_K^{\mathbb{R}}, \sigma_i \neq \sigma\} \cdot c_{23}}, \quad c_{28} := \frac{|\tau_1| \cdot c_{25}}{\min\{|\mathfrak{S}(\theta_i)| : \sigma_i \in \mathcal{E}_K^{\mathbb{C}}\}},$$

and

$$c_{29}(j) := \begin{cases} |\tau_1| \cdot c_{25}/(|\tau_j| \cdot c_{23}) & \text{for } 2 \leq j \leq u, \\ |\tau_1| \cdot c_{25}/|\mathfrak{S}(\theta_j)| & \text{for } u + 1 \leq j \leq u + v. \end{cases}$$

Define

$$c_{30} := \max\{2c_{17} \cdot c_{24}, \log(2c_{27})/c_{26}, \log(2c_{28})/c_{26}\}$$

and suppose $B > c_{30}$. Then

$$|\log(1 - z_j)| \leq 2c_{29}(j) \cdot \exp(-c_{26} \cdot B) \quad \text{for } 2 \leq j \leq u + v.$$

Proof. Let $2 \leq j \leq u + v$. If $\sigma_j \in \mathcal{E}_K^{\mathbb{R}}$, Lemma 8.4 yields

$$|a_0X - \theta_jY| = |\tau_j| \cdot |\varepsilon_j| \geq |\tau_j| \cdot c_{23}.$$

Conversely, if $\sigma_j \in \mathcal{E}_K^{\mathbb{C}}$, following the proof of Lemma 8.2, we have

$$|a_0X - \theta_jY| = |\tau_j| \cdot |\varepsilon_j| \geq |\mathfrak{S}(\theta_j)|.$$

Now, by Lemma 8.5 we have

$$|a_0X - \theta_1Y| = |\tau_1| \cdot |\varepsilon_1| \leq |\tau_1| \cdot c_{25} \cdot \exp(-c_{26} \cdot B);$$

it is in invoking this lemma that we have made use of the assumption $B > 2c_{17} \cdot c_{24}$. Thus

$$|z_j| \leq c_{29}(j) \cdot \exp(-c_{26} \cdot B).$$

Our assumption $B > c_{30} \geq \log(2c_{29}(j))/c_{26}$ gives $|z_j| < \frac{1}{2}$. From the standard Maclaurin expansion for $\log(1 - x)$ we conclude that $|\log(1 - z_j)| \leq 2 \cdot |z_j|$, completing the proof. \square

To ease notation, let

$$w := u + v - 2. \tag{56}$$

We now give our first set of w approximate relations. These only involve our original unknown exponents b_1, \dots, b_r found in (15).

Lemma 8.7. *Suppose $B > c_{30}$ holds. Let $1 \leq j \leq w$. Let*

$$\beta_j := \log \left| \frac{(\theta_1 - \theta_2) \cdot \tau_{j+2}}{(\theta_1 - \theta_{j+2}) \cdot \tau_2} \right|, \quad \alpha_{1,j} := \log \left| \frac{\delta_{1,j+2}}{\delta_{1,2}} \right|, \dots, \alpha_{r,j} := \log \left| \frac{\delta_{r,j+2}}{\delta_{r,2}} \right|.$$

Then

$$|\beta_j + b_1\alpha_{1,j} + \dots + b_r\alpha_{r,j}| \leq 2(c_{29}(2) + c_{29}(j + 2)) \cdot \exp(-c_{26} \cdot B). \tag{57}$$

Proof. From (55),

$$\frac{(\theta_1 - \theta_2) \cdot \tau_{j+2}}{(\theta_1 - \theta_{j+2}) \cdot \tau_2} \cdot \left(\frac{\delta_{1,j+2}}{\delta_{1,2}}\right)^{b_1} \cdots \left(\frac{\delta_{r,j+2}}{\delta_{r,2}}\right)^{b_r} = \frac{1 - z_2}{1 - z_{j+2}}.$$

Taking absolute values and then logs gives

$$|\beta_j + b_1\alpha_{1,j} + \cdots + b_r\alpha_{r,j}| \leq |\log|1 - z_2|| + |\log|1 - z_{j+2}||.$$

For a complex number z , we have

$$|\log|z|| \leq |\log z|$$

since $\log|z|$ is the real part of $\log z$. The lemma now follows from Lemma 8.6. □

In essence, in the above lemma, we have made use of the fact that

$$K^\times \rightarrow \mathbb{R}, \quad \phi \mapsto \log|\sigma(\phi)| \tag{58}$$

is a homomorphism for each embedding σ of K , and applied this to the approximate multiplicative relation (55) to obtain an approximate (additive) relation (57). If σ is complex, then σ and its conjugate $\bar{\sigma}$ induce the same homomorphism (58), and thus we need only consider the embeddings $\sigma_1, \dots, \sigma_{u+v}$. Note that although these are $u + v$ embeddings, we have obtained only $u + v - 2$ approximate relations so far. That is, we have had to sacrifice embeddings because we wanted to eliminate the two unknowns, X and Y . For σ real, $\log|\sigma(\phi)|$ determines $\sigma(\phi)$ up to signs. However, if σ is complex, then (58) loses the argument of $\sigma(\phi)$. Thus we should consider another homomorphism

$$K^\times \rightarrow \mathbb{R}/\mathbb{Z}\pi, \quad \phi \mapsto \Im(\log \sigma(\phi)) \tag{59}$$

where $\Im(z)$ denotes the imaginary part of a complex number z . Observe that $\Im(\log \sigma(\phi))$ denotes the argument of $\sigma(\phi)$ which naturally lives in $\mathbb{R}/\mathbb{Z}2\pi$, whilst here we use $\mathbb{R}/\mathbb{Z}\pi$ as the codomain. In practice, we have found that using $\mathbb{R}/\mathbb{Z}2\pi$ introduces extra factors but only results in negligible improvements to the bounds. Applying these homomorphisms to (55) allows us to obtain additional approximate relations. Since there are v complex embeddings, we obtain an additional v approximate relations. However since these homomorphisms are into $\mathbb{R}/\mathbb{Z}\pi$, the approximate relations are only valid after shifting by an appropriate multiple of π ; thus for each complex embedding σ_{u+j} , we will need an additional parameter b_{r+j} .

Recall that $w = u + v - 2$.

Lemma 8.8. *Let $1 \leq j \leq v$. Let*

$$\begin{aligned} \beta_{w+j} &:= \Im\left(\log\left(\frac{\theta_1 - \theta_{u+j}}{\tau_{u+j}}\right)\right), \\ \alpha_{1,w+j} &:= -\Im(\log \delta_{1,u+j}), \dots, \alpha_{r,w+j} := -\Im(\log \delta_{r,u+j}), \quad \alpha_{r+j,w+j} := \pi. \end{aligned}$$

Suppose $B > c_{30}$ holds. Then there is some $b_{r+j} \in \mathbb{Z}$ such that

$$|\beta_{w+j} + b_1\alpha_{1,w+j} + \cdots + b_r\alpha_{r,w+j} + b_{r+j}\alpha_{r+j,w+j}| \leq 2c_{29}(u + j) \cdot \exp(-c_{26} \cdot B). \tag{60}$$

Moreover,

$$|b_{r+j}| \leq |b_1| + \cdots + |b_r| + \frac{\pi + 1}{\pi}. \tag{61}$$

Proof. From (55), and Lemma 8.6,

$$|\log(Y) + \log((\theta_1 - \theta_{u+j})/\tau_{u+j}) - b_1 \log \delta_{1,u+j} - \cdots - b_r \log \delta_{r,u+j} + b' \cdot \pi i| \leq 2c_{29}(u+j) \cdot \exp(-c_{26} \cdot B),$$

for some $b' \in \mathbb{Z}$. Thus

$$|\Im(\log(Y)) + \beta_{w+j} + b_1 \alpha_{1,w+j} + \cdots + b_r \alpha_{r,w+j} + b' \pi| \leq 2c_{29}(u+j) \cdot \exp(-c_{26} \cdot B).$$

Recall that $Y \in \mathbb{Z} \setminus \{0\}$, so $\Im(\log(Y))$ is either 0 or π depending on whether Y is positive or negative. We take $b_{r+j} = b'$ in the former case and $b_{r+j} = b' + 1$ in the latter case. This gives (60).

It remains to prove (61). Our assumption $B > c_{30}$ gives

$$2c_{29}(u+j) \cdot \exp(-c_{26} \cdot B) < 1.$$

Moreover, $|\beta_{w+j}| \leq \pi$ and $|\alpha_{i,w+j}| \leq \pi$ for $0 \leq i \leq r$. From (60),

$$\begin{aligned} \pi \cdot |b_{r+j}| &= |\alpha_{r+j,w+j}| \cdot |b_{r+j}| \\ &\leq |\beta_{w+j}| + |\alpha_{1,w+j}| \cdot |b_1| + \cdots + |\alpha_{r,w+j}| \cdot |b_r| + 1 \\ &\leq \pi(1 + |b_1| + \cdots + |b_r|) + 1. \end{aligned}$$

The lemma follows. □

Summing up, Lemmas 8.7 and 8.8 give us $(u + v - 2) + v = d - 2$ approximate relations (57), (60) in integer unknowns b_1, \dots, b_{r+v} .

9. Reduction of bounds

We do not know which real embedding $\sigma \in \mathcal{E}_K^{\mathbb{R}}$ makes $|\sigma(\varepsilon)|$ minimal. So the procedure described below for reducing the bound (34) needs to be repeated for each possible choice of embedding σ in $\mathcal{E}_K^{\mathbb{R}}$. Thus, for every possible choice of $\sigma \in \mathcal{E}_K^{\mathbb{R}}$, we let $\sigma_1 = \sigma$ and we choose an ordering of the other embeddings as in (51). Given a real number γ , we denote by $[\gamma]$ the nearest integer to γ , with the convention that $[k + 1/2] = k + 1$ for $k \in \mathbb{Z}$. Let n be as in (54) and observe that

$$n = (s + 1) + d - 2,$$

where we recall that $s = \#S$. Let C be a positive integer to be chosen later. Let \mathbf{I}_m and $\mathbf{0}_{i,j}$ be the $m \times m$ identity matrix and $i \times j$ zero matrix, respectively. Let M be the $n \times n$ matrix

$$M := \begin{bmatrix} \mathbf{0}_{w,s+1} & [C\alpha_{1,1}] \cdots [C\alpha_{1,w}] & [C\alpha_{1,w+1}] \cdots [C\alpha_{1,d-2}] \\ \vdots & \vdots & \vdots \\ [C\alpha_{w,1}] \cdots [C\alpha_{w,w}] & [C\alpha_{w,w+1}] \cdots [C\alpha_{w,d-2}] \\ \cdots & \cdots & \cdots \\ [C\alpha_{w+1,1}] \cdots [C\alpha_{w+1,w}] & [C\alpha_{w+1,w+1}] \cdots [C\alpha_{w+1,d-2}] \\ \vdots & \vdots & \vdots \\ \mathbf{I}_{s+1} & \vdots & \vdots \\ \cdots & \cdots & \cdots \\ [C\alpha_{r,1}] \cdots [C\alpha_{r,w}] & [C\alpha_{r,w+1}] \cdots [C\alpha_{r,d-2}] \\ \cdots & \cdots & \cdots \\ \mathbf{0}_{v,s+1} & \mathbf{0}_{v,w} & [C\pi] \cdot \mathbf{I}_v \end{bmatrix}$$

and let L be the sublattice of \mathbb{Z}^n spanned by the rows of M . Recall that $\mathcal{B}_\infty, \mathcal{B}_1, \mathcal{B}_2$ are respectively the known bounds for $\|\mathbf{b}\|_\infty, \|\mathbf{b}\|_1, \|\mathbf{b}\|_2$. Let

$$\begin{aligned} \mathbf{w} &:= (\underbrace{0, 0, \dots, 0}_{s+1}, [C\beta_1], \dots, [C\beta_{d-2}]) \in \mathbb{Z}^n, \\ \mathcal{A}_1 &:= \frac{1 + \mathcal{B}_1}{2}, \quad \mathcal{A}_2 := \frac{2\pi(1 + \mathcal{B}_1) + 1}{2\pi}, \\ \mathcal{B}_3 &:= \sum_{j=1}^w (c_{29}(2) + c_{29}(j + 2))^2 + \sum_{j=1}^v c_{29}(u + j)^2, \\ \mathcal{B}_4 &:= \mathcal{A}_1 \sum_{j=1}^w (c_{29}(2) + c_{29}(j + 2)) + \mathcal{A}_2 \sum_{j=1}^v c_{29}(u + j), \text{ and} \\ \mathcal{B}_5 &:= \sqrt{\mathcal{B}_2^2 - w\mathcal{B}_\infty^2 + w\mathcal{A}_1^2 + v\mathcal{A}_2^2}. \end{aligned}$$

By (48), we observe that

$$\mathcal{B}_2^2 = (w + 1)\mathcal{B}_\infty^2 + \rho_1^2 + \dots + \rho_s^2$$

so that $\mathcal{B}_2^2 - w\mathcal{B}_\infty^2 \geq 0$ and thus the argument of the square root in the above definition of \mathcal{B}_5 is positive.

Write

$$\mathbf{b}_e := (b_1, b_2, \dots, b_r, b_{r+1}, \dots, b_{r+v}),$$

where b_{r+1}, \dots, b_{r+v} are as in Lemma 8.8. We think of this as the ‘‘extended exponent vector’’. Note that the number of entries in \mathbf{b}_e is

$$r + v = u + v + s - 1 + v = d + s - 1 = n. \tag{62}$$

If \mathbf{b}_e is known, then the solution is known.

Lemma 9.1.
$$\|\mathbf{b}_e\|_2 \leq \sqrt{\mathcal{B}_2^2 + v \left(\mathcal{B}_1 + \frac{\pi + 1}{\pi} \right)^2}.$$

Proof. This follows immediately from (61) and the definitions of $\mathcal{B}_1, \mathcal{B}_2$. □

Proposition 9.2. *Suppose $\mathbf{b}_e \cdot M \neq -\mathbf{w}$. Let*

$$\mathcal{D} := \begin{cases} D(L, \mathbf{w}) & \text{if } \mathbf{w} \notin L, \\ \min_{\substack{\mathbf{x} \in L \\ \mathbf{x} \neq \mathbf{0}}} \|\mathbf{x}\|_2 & \text{if } \mathbf{w} \in L. \end{cases} \tag{63}$$

Suppose $\mathcal{D} > \mathcal{B}_5$. Then

$$B \leq \max \left(c_{30}, \frac{1}{c_{26}} \cdot \log \left(\frac{2C \cdot \mathcal{B}_3}{\sqrt{\mathcal{B}_3(\mathcal{D}^2 - \mathcal{B}_5^2) + \mathcal{B}_4^2 - \mathcal{B}_4}} \right) \right). \tag{64}$$

Proof. If $B \leq c_{30}$ then (64) holds. We will therefore suppose that $B > c_{30}$. Thus, inequalities (57) and (60) hold. Write

$$\mathbf{w} + \mathbf{b}_e \cdot M = (b_{u+v-1}, b_{u+v}, \dots, b_r, \Theta_1, \Theta_2, \dots, \Theta_{d-2}),$$

where we take this equality as the definition of $\Theta_1, \dots, \Theta_{d-2}$. That is, for $1 \leq j \leq w$,

$$\Theta_j = [C\beta_j] + b_1[C\alpha_{1,j}] + \dots + b_r[C\alpha_{r,j}].$$

Hence, again for $1 \leq j \leq w$,

$$\begin{aligned} |\Theta_j| &\leq \frac{1}{2} + \frac{1}{2}|b_1| + \dots + \frac{1}{2}|b_r| + C \cdot |\beta_j + b_1\alpha_{1,j} + \dots + b_r\alpha_{r,j}| \\ &\leq \frac{1}{2}(1 + \mathcal{B}_1) + 2C \cdot (c_{29}(2) + c_{29}(j+2)) \cdot \exp(-c_{26} \cdot B) \\ &\leq \mathcal{A}_1 + (c_{29}(2) + c_{29}(j+2)) \cdot \eta, \end{aligned}$$

where the second inequality follows by (34) and (57), and

$$\eta := 2C \cdot \exp(-c_{26} \cdot B).$$

Recall that $w = u + v - 2$. For $1 \leq j \leq v$,

$$\Theta_{w+j} = [C\beta_{w+j}] + b_1[C\alpha_{1,w+j}] + \dots + b_r[C\alpha_{r,w+j}] + b_{r+j}[C\pi].$$

Thus, for $1 \leq j \leq v$,

$$\begin{aligned} |\Theta_{w+j}| &\leq \frac{1}{2} + \frac{1}{2}|b_1| + \dots + \frac{1}{2}|b_r| + \frac{1}{2}|b_{r+j}| + C \cdot |\beta_{w+j} + b_1\alpha_{1,w+j} + \dots + b_r\alpha_{r,w+j} + b_{r+j}\pi| \\ &\leq \frac{2\pi + 1}{2\pi} + \mathcal{B}_1 + 2C \cdot c_{29}(u+j) \cdot \exp(-c_{26} \cdot B) \\ &\leq \mathcal{A}_2 + c_{29}(u+j) \cdot \eta, \end{aligned}$$

where the second inequality follows from (34), (60), and (61).

By assumption, $\mathbf{w} + \mathbf{b}_e \cdot M \neq \mathbf{0}$; hence

$$\begin{aligned} \mathcal{D}^2 &\leq \|\mathbf{w} + \mathbf{b}_e \cdot M\|_2^2 \\ &= b_{u+v-1}^2 + \dots + b_r^2 + \Theta_1^2 + \dots + \Theta_{d-2}^2 \\ &\leq b_{u+v-1}^2 + \dots + b_r^2 + w\mathcal{A}_1^2 + v\mathcal{A}_2^2 + 2\mathcal{B}_4\eta + \mathcal{B}_3\eta^2. \end{aligned}$$

However, $|b_{u+v-1}| \leq \|\mathbf{b}\|_\infty \leq \mathcal{B}_\infty$. Moreover, by (46) we have $|b_{u+v-1+i}| \leq \rho_i$ for $i = 1, \dots, s$, where ρ_i is given in (45). It follows that

$$\begin{aligned} \mathcal{D}^2 &\leq \mathcal{B}_\infty^2 + \rho_1^2 + \dots + \rho_s^2 + w\mathcal{A}_1^2 + v\mathcal{A}_2^2 + 2\mathcal{B}_4\eta + \mathcal{B}_3\eta^2 \\ &= \mathcal{B}_2^2 - w\mathcal{B}_\infty^2 + w\mathcal{A}_1^2 + v\mathcal{A}_2^2 + 2\mathcal{B}_4\eta + \mathcal{B}_3\eta^2 \quad (\text{from (48)}) \\ &= \mathcal{B}_5^2 + 2\mathcal{B}_4\eta + \mathcal{B}_3\eta^2 \\ &= \mathcal{B}_5^2 + \mathcal{B}_3 \left(\eta + \frac{\mathcal{B}_4}{\mathcal{B}_3} \right)^2 - \frac{\mathcal{B}_4^2}{\mathcal{B}_3}. \end{aligned}$$

Recall our assumption that $\mathcal{B}_5 < \mathcal{D}$. Thus

$$\frac{\mathcal{B}_4^2}{\mathcal{B}_3} < \mathcal{D}^2 - \mathcal{B}_5^2 + \frac{\mathcal{B}_4^2}{\mathcal{B}_3} \leq \mathcal{B}_3 \left(\eta + \frac{\mathcal{B}_4}{\mathcal{B}_3} \right)^2,$$

and so

$$0 < \frac{\sqrt{\mathcal{B}_3(D^2 - \mathcal{B}_5^2) + \mathcal{B}_4^2} - \mathcal{B}_4}{\mathcal{B}_3} \leq \eta.$$

However $\eta = 2C \cdot \exp(-c_{26} \cdot B)$. This yields the bound

$$B \leq \frac{1}{c_{26}} \cdot \log \left(\frac{2C \cdot \mathcal{B}_3}{\sqrt{\mathcal{B}_3(D^2 - \mathcal{B}_5^2) + \mathcal{B}_4^2} - \mathcal{B}_4} \right),$$

which gives (64). □

Heuristic. It remains to decide on a reasonable choice for C . We expect that the determinant of the matrix M is approximately C^{d-2} . Thus the distance between adjacent vectors in L is expected to be in the region of $C^{(d-2)/n}$, and so we anticipate (very roughly) that $\mathcal{D} \sim C^{(d-2)/n}$. We would like $\mathcal{D} > \mathcal{B}_5$. Therefore it is reasonable to choose $C \gg \mathcal{B}_5^{n/(d-2)}$. If, for a particular choice of C , the condition $\mathcal{D} > \mathcal{B}_5$ fails, then we simply try again with a larger choice of C .

Remarks. Our approach is somewhat unusual in that it uses all $d - 2$ available approximate relations to reduce the initial bound. In contrast, it is much more common to use one relation (e.g., [Tzanakis and de Weger 1989, Section 16]) to reduce the bound. In most examples, we have found that both approaches give similar reductions in the size of the bound and that there is no advantage in using one over the other. However in some examples the approach of using only one relation fails spectacularly. Here are two such scenarios:

(i) Suppose δ_1 (say) belongs to a proper subfield K' of K . Now let σ_2, σ_3 be distinct embeddings of K that agree on K' . Then, in the notation of Lemma 8.7 we find $\alpha_{1,3} = \log|\sigma_3(\delta_1)/\sigma_2(\delta_1)| = 0$, and so the coefficient of the unknown b_1 is zero in the approximate relation (57). Therefore the one relation (57) on its own fails to provide any information on the size of b_1 . In practice, the lattice constructed in [Tzanakis and de Weger 1989, Section 16] from this one relation will contain the tiny vector $(1, 0, \dots, 0)$, and this will result in the computational failure of the closest vector algorithm.

(ii) We continue to suppose that δ_1 belongs to the proper subfield K' as above. Let σ_{u+1} be a complex embedding of K that extends a real embedding of K' , and suppose for simplicity that $\sigma_{u+1}(\delta_1)$ is positive. Then again, $\alpha_{1,w+1} = 0$ in the approximate relation (60), and so that relation on its own fails to control b_1 .

In the above two examples, we chose to illustrate the possible failure of the approach of using one relation by imposing $\delta_1 \in K'$ where K' is a subfield of K . However, a similar failure occurs (and is more difficult to find) if the S -unit basis $\delta_1, \dots, \delta_r$ is multiplicatively dependent over K' , meaning that there is some nontrivial $(c_1, \dots, c_r) \in \mathbb{Z}^r$ such that $\delta_1^{c_1} \cdots \delta_r^{c_r} \in K'$.

In Proposition 9.2, we require $\mathbf{b}_e \cdot M \neq -\mathbf{w}$. Of course, if M is nonsingular, we can simply check whether $\mathbf{b}_e = -\mathbf{w} \cdot M^{-1}$ yields a solution, and therefore there is no harm in making this assumption. In all our examples, M has been nonsingular, and we expect that by choosing C large enough we can ensure that this happens. However, if M is singular then the equation $\mathbf{b}_e \cdot M = -\mathbf{w}$ either has no solutions or the solutions belong to the translate of a sublattice of \mathbb{Z}^n whose rank is the corank of the matrix M . A glance

at the matrix M reveals that this corank is at most $w = u + v - 2$. If this case was to ever arise in practice, we would need to enumerate all vectors \mathbf{b}_e satisfying $\mathbf{b}_e \cdot M = -\mathbf{w}$ and the bound in Lemma 9.1 and test if they lead to solutions.

9.1. Example 1.4 continued. We continue giving details for the tuple $(\tau, \delta_1, \dots, \delta_{10})$ alluded to on page 690. The values of \mathcal{B}_∞ and \mathcal{B}_2 are given in (40). The set S consists of five primes $\mathfrak{p}_1, \dots, \mathfrak{p}_5$ given by (33). By applying Proposition 7.2 we had (page 694) obtained bounds 237, 292, 354, 518, 821 for $\text{ord}_{\mathfrak{p}_j}(a_0X - \theta Y)$ with $i = 1, \dots, 5$ respectively; these are the values denoted $k_j - 1$ in (41). Now $\text{ord}_{\mathfrak{p}_j}(\tau) = 0, 0, 1, 0, 0$ respectively for $j = 1, \dots, 5$. Letting ε be as in (42), we may take $(-k'_j, k''_j)$ in (43) to be $(0, 237), (0, 292), (-1, 353), (0, 518), (0, 821)$ respectively. This allows us to compute the constant c_{21} defined in Lemma 8.1. We find that $c_{21} \approx 2842.79$. The field K has signature $(u, v) = (1, 5)$ and thus there is only one possibility for $\sigma \in \mathcal{E}_K^{\mathbb{R}}$. We therefore take $\sigma_1 = \sigma$ to be the unique real embedding. For illustration, we give the values of constants appearing in Section 8:

$$\begin{aligned} c_{22} &\approx 30.31, & c_{23} &= 1, & c_{24} &\approx 2873.10, & c_{25} &\approx 5.91 \times 10^{1247}, & c_{26} &\approx 0.35, \\ c_{27} &\approx 5.91 \times 10^{1247}, & c_{28} &\approx 1.40 \times 10^{1246}, & c_{29}(2) &\approx 1.25 \times 10^{1246}, \\ c_{29}(3) &\approx 8.30 \times 10^{1245}, & c_{29}(4) &\approx 7.83 \times 10^{1245}, & c_{29}(5) &\approx 9.21 \times 10^{1245}, \\ c_{29}(6) &\approx 1.40 \times 10^{1246}, & c_{30} &\approx 8290.02. \end{aligned}$$

Lemma 8.7 gives $w = u + v - 2 = 4$ approximate relations and Lemma 8.8 gives another $v = 5$ relations. Therefore we have $d - 2 = 9$ relations altogether, and $n = \#S + d - 2 = 15$. Therefore the matrix M is 15×15 and the lattice L belongs to \mathbb{Z}^{15} . We find that

$$\begin{aligned} \mathcal{B}_1 &\approx 7.85 \times 10^{222}, & \mathcal{A}_1 &\approx 3.92 \times 10^{222}, & \mathcal{A}_2 &\approx 7.85 \times 10^{222}, & \mathcal{B}_3 &\approx 2.59 \times 10^{2493}, \\ \mathcal{B}_4 &\approx 7.57 \times 10^{1469}, & \mathcal{B}_5 &\approx 1.93 \times 10^{223}. \end{aligned}$$

In accordance with the above heuristic, our program chooses

$$C = [\mathcal{B}_5^{n/(d-2)}] \approx 1.39 \times 10^{372}.$$

The matrix M and the lattice L are too huge to reproduce here, but we point out that

$$[\mathbb{Z}^{15} : L] \approx 2.66 \times 10^{3357}; \quad \mathcal{D} \approx 7.23 \times 10^{223}.$$

In this case, $\mathbf{w} \notin L$ so that \mathcal{D} is computed using $D(L, \mathbf{w})$. Hence the hypothesis $\mathcal{D} > \mathcal{B}_5$ of Proposition 9.2 is satisfied. We may therefore apply Proposition 9.2 to obtain a new bound for B given by (64):

$$B \leq 9270.82.$$

We now start again with $\mathcal{B}_\infty = 9270$ and repeat the previous steps, first for obtaining bounds for $\text{ord}_{\mathfrak{p}_j}(a_0X - \theta Y)$ and then for writing down the lattice L and applying Proposition 9.2. Table 2 illustrates the results.

Iteration	B_∞	bounds for $\text{ord}_{p_j}(a_0X - \theta Y)$ with $1 \leq j \leq 5$				
0	1.57×10^{222}	237	292	354	518	821
1	9270	4	5	8	10	15
2	251	3	3	5	6	10
3	190	2	3	5	6	9
4	180	2	3	5	6	9
5	180	2	3	5	6	9

Table 2. We successively reduce the bounds for B and for $\text{ord}_{p_j}(a_0X - \theta Y)$, where $j = 1, \dots, 5$.

Note that at the fifth iteration we fail to obtain any improvement on the bounds, and so we stop there. Recall that $r = 10$ and that $B = \max(|b_1|, \dots, |b_{10}|)$, where b_1, \dots, b_{10} are the exponents in (15). Our final bound is $B \leq 180$. The set of possible integer tuples (b_1, \dots, b_{10}) satisfying this bound has size

$$(2 \times 180 + 1)^{10} = 362^{10} \approx 3.86 \times 10^{25}.$$

The huge size of this region does not allow brute force enumeration of the solutions. Instead, one can reduce the number of tuples to consider by using the bounds we have obtained for $\text{ord}_{p_j}(a_0X - \theta Y)$. We let $\kappa_j = 2, 3, 5, 6, 9$ for $j = 1, \dots, 5$, respectively. We know that $0 \leq \text{ord}_{p_j}(a_0X - \theta Y) \leq \kappa_j$, and so there are $\kappa_j + 1$ possibilities for the $\text{ord}_{p_j}(a_0X - \theta Y)$. Let (k_1, \dots, k_5) be some tuple of integers satisfying $0 \leq k_j \leq \kappa_j$. The condition $\text{ord}_{p_j}(a_0X - \theta Y) = k_j$ simply defines a hyperplane of codimension 1 in the space of possible (b_1, \dots, b_{10}) . Imposing all five conditions $\text{ord}_{p_j}(a_0X - \theta Y) = k_j$ with $j = 1, \dots, 5$ cuts down the dimension from 10 to 5. Thus we expect that the search region should (very roughly) have size

$$(\kappa_1 + 1) \cdots (\kappa_5 + 1) \cdot 362^5 \approx 3.13 \times 10^{16}.$$

This is still way beyond brute force enumeration and motivates our next section.

10. Sieving

In order to resolve the Thue–Mahler equation

$$F(X, Y) = a \cdot p_1^{z_1} \cdots p_v^{z_v}, \quad X, Y \in \mathbb{Z}, \text{gcd}(X, Y) = \text{gcd}(a_0, Y) = 1,$$

we have first reduced the problem to that of resolving a number of equations of the form (15), subject to the restrictions (16), (17). Recall that $B = \max\{|b_1|, \dots, |b_r|\}$, where $\mathbf{b} = (b_1, \dots, b_r) \in \mathbb{Z}^r$ denotes the vector of unknown exponents to solve for. For each such equation (15), we have used the theory of linear forms in logarithms to obtain a bound for B , and moreover, we have explained how to repeatedly reduce this bound. During each of these iterations, we have simultaneously reduced the bounds on the ∞ -norm, the 1-norm, and the 2-norm of \mathbf{b} . Let us denote the final bound for the ∞ -norm of \mathbf{b} by B'_f and write B_f for the final bound on the 2-norm of \mathbf{b} . Thus

$$\|\mathbf{b}\|_2 \leq B_f, \quad \|\mathbf{b}\|_\infty \leq B'_f. \tag{65}$$

We have also explained how to obtain and reduce the bounds on $\text{ord}_{\mathfrak{p}}(a_0X - \theta Y)$ for $\mathfrak{p} \in S$. Suppose that at the end of this process, our bounds are

$$0 \leq \text{ord}_{\mathfrak{p}}(a_0X - \theta Y) \leq \kappa_{\mathfrak{p}} \quad \text{for } \mathfrak{p} \in S. \tag{66}$$

Unfortunately, in our high-rank examples (i.e., when the S -unit rank r is large) the final bound \mathcal{B}'_f is often too large to allow for brute force enumeration of solutions. Instead, we shall sieve for solutions using both the primes \mathfrak{p} in S , and also rational primes p whose support in \mathcal{O}_K is disjoint from S . The objective of the sieve is to show that the solutions \mathbf{b} belong to a union of a certain (hopefully small) number of cosets $\mathbf{w} + L$, where the L are sublattices of \mathbb{Z}^r . As the sieve progresses, the determinants of the lattices L will grow. The larger the determinant, the fewer vectors we expect belonging to $\mathbf{w} + L$ and satisfying $\|\mathbf{b}\|_2 \leq \mathcal{B}_f$, and the easier it should be to find these vectors using the algorithm of Fincke and Pohst [1985]. The following lemma is a helpful guide to when Fincke and Pohst should be applied.

Lemma 10.1. *Let L be a sublattice of \mathbb{Z}^r . Suppose $\lambda(L) > 2\mathcal{B}_f$, where $\lambda(L)$ denotes the length of the shortest nonzero vector in L . Then there is at most one vector \mathbf{b} in the coset $\mathbf{w} + L$ satisfying $\|\mathbf{b}\|_2 \leq \mathcal{B}_f$. Moreover, any such \mathbf{b} is equal to $\mathbf{w} + \mathbf{c}(L, -\mathbf{w})$.*

Proof. Suppose there are vectors $\mathbf{b}_1, \mathbf{b}_2 \in \mathbf{w} + L$ both satisfying $\|\mathbf{b}_i\|_2 \leq \mathcal{B}_f$. Then $\mathbf{b}_1 - \mathbf{b}_2 \in L$ and $\|\mathbf{b}_1 - \mathbf{b}_2\|_2 \leq 2\mathcal{B}_f$. As $\lambda(L) > 2\mathcal{B}_f$ we see that $\mathbf{b}_1 = \mathbf{b}_2$. The second part follows from Lemma 7.1. \square

We continue sieving until the lattices L satisfy $\lambda(L) > 2\mathcal{B}_f$. We then apply the Fincke–Pohst algorithm to determine $\mathbf{c}(L, -\mathbf{w})$ and check whether the vector $\mathbf{b} = \mathbf{w} + \mathbf{c}(L, -\mathbf{w})$ leads to a solution.

10.1. Sieving using the primes of S . To recap, we seek solutions (X, Y, \mathbf{b}) to

$$a_0X - \theta Y = \tau \cdot \delta_1^{b_1} \cdots \delta_r^{b_r},$$

subject to the conditions

$$X, Y \in \mathbb{Z}, \quad \gcd(X, Y) = 1, \quad \gcd(a_0, Y) = 1, \quad b_i \in \mathbb{Z},$$

and such that

$$\|\mathbf{b}\|_2 \leq \mathcal{B}_f, \quad \text{and} \quad 0 \leq \text{ord}_{\mathfrak{p}}(a_0X - \theta Y) \leq \kappa_{\mathfrak{p}} \quad \text{for every } \mathfrak{p} \in S.$$

In particular, this last inequality (66) asserts that $\text{ord}_{\mathfrak{p}}(a_0X - \theta Y)$ belongs to a certain set of values $0, 1, \dots, \kappa_{\mathfrak{p}}$. The following proposition reduces this list somewhat, and for any k in this reduced list, yields a vector \mathbf{w}_k and a sublattice L_k of \mathbb{Z}^r such that $\mathbf{b} \in \mathbf{w}_k + L_k$ whenever $\text{ord}_{\mathfrak{p}}(a_0X - \theta Y) = k$.

Proposition 10.2. *Let $\mathfrak{p} \in S$. Let $\theta_0 \in \mathbb{Z}$ satisfy $\theta_0 \equiv \theta \pmod{\mathfrak{p}^{\kappa_{\mathfrak{p}}}}$. Let \mathfrak{a} and \mathcal{T}_1 be as in (37). Define*

$$\eta : \mathbb{Z}^r \rightarrow \mathbb{Z}, \quad \eta(n_1, \dots, n_r) = n_1 \text{ord}_{\mathfrak{p}}(\delta_1) + \cdots + n_r \text{ord}_{\mathfrak{p}}(\delta_r), \quad L'' = \text{Ker}(\eta).$$

Let

$$\mathcal{K}'' := (\text{ord}_{\mathfrak{p}}(\tau) + \text{Image}(\eta)) \cap \{0 \leq k \leq \kappa_{\mathfrak{p}} : \gcd(\mathfrak{a}^k, \theta - \theta_0) = \gcd(\mathfrak{a}^k, \mathcal{T}_1)\}.$$

For $k \in \mathcal{K}''$, let \mathbf{w}_k'' be any vector in \mathbb{Z}^r satisfying $\eta(\mathbf{w}_k'') = k - \text{ord}_p(\tau)$. If $k \in \mathcal{K}''$ satisfies $k \geq 1$, we will let ϕ and τ_0 be as in Proposition 7.2 (these depend on k). Let

$$\mathcal{K}' := \begin{cases} \{k \in \mathcal{K}'' : \bar{\tau}_0 \in \text{Image}(\phi)\} & \text{if } 0 \notin \mathcal{K}'', \\ \{0\} \cup \{k \in \mathcal{K}'' : \bar{\tau}_0 \in \text{Image}(\phi)\} & \text{if } 0 \in \mathcal{K}''. \end{cases}$$

For $k \in \mathcal{K}'$ with $k \geq 1$, we let $L_k' = \text{Ker}(\phi)$ and \mathbf{w}_k' be any vector in \mathbb{Z}^r satisfying $\phi(\mathbf{w}_k') = \bar{\tau}_0$. Let $\mathbf{w}'_0 = \mathbf{0}$, $L'_0 = \mathbb{Z}^r$, and

$$\mathcal{K} := \{k \in \mathcal{K}' : (\mathbf{w}_k'' + L'') \cap (\mathbf{w}_k' + L_k') \neq \emptyset\}.$$

For $k \in \mathcal{K}$, write

$$L_k := L'' \cap L_k'$$

and choose any $\mathbf{w}_k \in \mathbb{Z}^r$ such that

$$\mathbf{w}_k + L_k := (\mathbf{w}_k'' + L'') \cap (\mathbf{w}_k' + L_k').$$

Let $k = \text{ord}_p(a_0X - \theta Y)$. Then $k \in \mathcal{K}$ and $\mathbf{b} \in \mathbf{w}_k + L_k$.

Proof. By (66), the valuation $k := \text{ord}_p(a_0X - \theta Y)$ satisfies $0 \leq k \leq \kappa_p$. Moreover, by Proposition 7.2, part (i), we have $\gcd(\mathfrak{a}^k, \theta - \theta_0) = \gcd(\mathfrak{a}^k, \mathcal{T}_1)$. By (15), we know $k \in \text{ord}_p(\tau) + \eta(\mathbf{b})$ and thus $k \in \mathcal{K}''$ and $\mathbf{b} \in \mathbf{w}_k'' + L''$. In particular, the proposition follows in the case $k = 0$. We therefore suppose $k \geq 1$. By Proposition 7.2, part (ii) and its proof, it follows that $k \in \mathcal{K}'$ and $\mathbf{b} \in \mathbf{w}_k' + L_k'$, completing the proof. \square

Remark. For each prime $p \in S$, Proposition 10.2 yields a number of cosets $\mathbf{w}_k + L_k$ and tells us that \mathbf{b} belongs to one of them. Note that L'' is a subgroup of \mathbb{Z}^r of rank $r - 1$. Moreover, the subgroup L_k' has finite index in \mathbb{Z}^r . Therefore $L_k := L_k' \cap L''$ has rank $r - 1$. From the remarks following Proposition 7.2 (where L_k' is called L) we expect L_k' to have index $p^{(d-2)k}$ in \mathbb{Z}^r . In particular, the larger the value of k , the larger the index of L_k' . Of course the number of cosets is bounded above by $\kappa_p + 1$.

10.2. Sieving with other primes. Given a prime ideal \mathfrak{q} of \mathcal{O}_K , write $\mathcal{O}_{\mathfrak{q}}$ for the localization of \mathcal{O}_K at \mathfrak{q} ,

$$\mathcal{O}_{\mathfrak{q}} = \{\alpha \in K : \text{ord}_{\mathfrak{q}}(\alpha) \geq 0\}.$$

Now let q be a rational prime. Define

$$\mathcal{O}_q = \bigcap_{\mathfrak{q} | q} \mathcal{O}_{\mathfrak{q}} = \{\alpha \in K : \text{ord}_{\mathfrak{q}}(\alpha) \geq 0 \text{ for all } \mathfrak{q} | q\}.$$

The group of invertible elements \mathcal{O}_q^\times consists of all $\alpha \in K$ such that $\text{ord}_{\mathfrak{q}}(\alpha) = 0$ for all prime ideals $\mathfrak{q} | q$.

Let $\tau, \delta_1, \dots, \delta_r$ be as in (15). Let q be a rational prime coprime to the supports of $\tau, \delta_1, \dots, \delta_r$. Thus $\tau, \delta_1, \dots, \delta_r$ all belong to \mathcal{O}_q^\times . Let

$$\mathfrak{A}_q := (\mathcal{O}_q / q\mathcal{O}_q)^\times.$$

This is canonically isomorphic to $(\mathcal{O}_K / q\mathcal{O}_K)^\times$. Let

$$\mu : \mathbb{F}_q^\times \hookrightarrow \mathfrak{A}_q, \quad \alpha + q\mathbb{Z} \mapsto \alpha + q\mathcal{O}_q$$

be the natural map, and let

$$\mathfrak{B}_q := \mathfrak{A}_q / \mu(\mathbb{F}_q^\times)$$

be the cokernel of μ . We denote the induced homomorphism $\mathcal{O}_q^\times \rightarrow \mathfrak{B}_q$ by

$$\pi_q : \mathcal{O}_q^\times \rightarrow \mathfrak{B}_q, \quad \beta \mapsto (\beta + q\mathcal{O}_q) \cdot \mu(\mathbb{F}_q^\times).$$

Define

$$\phi_q : \mathbb{Z}^r \rightarrow \mathfrak{B}_q, \quad (m_1, \dots, m_r) \mapsto \pi_q(\delta_1)^{m_1} \cdots \pi_q(\delta_r)^{m_r}.$$

Proposition 10.3. *Let*

$$R_q = \{a_0u - \theta : u \in \{0, 1, \dots, q - 1\}\} \cup \{a_0\} \quad \text{and} \quad S_q = \{\pi_q(r)/\pi_q(\tau) : r \in R_q \cap \mathcal{O}_q^\times\} \subseteq \mathfrak{B}_q.$$

Let

$$T_q = S_q \cap \phi_q(\mathbb{Z}^r) \quad \text{and} \quad L_q = \text{Ker}(\phi_q).$$

Finally, let $W_q \subset \mathbb{Z}^r$ be a set of size $\#T_q$ such that for every $t \in T_q$ there is some $\mathbf{w} \in W_q$ with $\phi_q(\mathbf{w}) = t$. Then $\mathbf{b} \in W_q + L_q$.

Proof. Since $\tau, \delta_1, \dots, \delta_r \in \mathcal{O}_q^\times$, we have $a_0X - \theta Y \in \mathcal{O}_q^\times$. We want to determine the possibilities for the image of the algebraic integer $a_0X - \theta Y$ in \mathfrak{B}_q . Since X and Y are coprime, q divides at most one of X, Y . If $q \nmid Y$ then

$$a_0X - \theta Y \equiv v \cdot (a_0u - \theta) \pmod{q\mathcal{O}_q}$$

for some $u \in \{0, 1, \dots, q - 1\}$ and some $v \in \mathbb{F}_q^\times$. If $q \mid Y$ then $q \nmid X$ and

$$a_0X - \theta Y \equiv a_0v \pmod{q\mathcal{O}_q}$$

for some $v \in \mathbb{F}_q^\times$. We conclude that $a_0X - \theta Y \equiv v \cdot r \pmod{q\mathcal{O}_q}$ where $v \in \mathbb{F}_q^\times$ and $r \in R_q$. Moreover, since $a_0X - \theta Y \in \mathcal{O}_q^\times$ we see that $r \in R_q \cap \mathcal{O}_q^\times$. Now

$$\pi_q(a_0X - \theta Y) = \pi_q(r)\pi_q(v) = \pi_q(r).$$

It follows that $\pi_q(a_0X - \theta Y)/\pi_q(\tau) \in S_q$. However

$$\phi_q(\mathbf{b}) = \pi_q(\delta_1)^{b_1} \cdots \pi_q(\delta_r)^{b_r} = \pi_q(a_0X - \theta Y)/\pi_q(\tau),$$

where the first equality follows from the definition of ϕ_q and the second from (15). Thus $\phi_q(\mathbf{b}) = t$ for some $t \in T_q$. By definition of W_q , there is some $\mathbf{w} \in W_q$ with $\phi_q(\mathbf{w}) = t = \phi_q(\mathbf{b})$, thus $\mathbf{b} - \mathbf{w} \in L_q$. \square

Heuristic. It is appropriate that we heuristically “measure” the quality of information that Proposition 10.3 gives us about the solutions. A priori, $\phi_q(\mathbf{b})$ could be any element in $\phi_q(\mathbb{Z}^r) \subseteq \mathfrak{B}_q$. However, the lemma tells us that $\phi_q(\mathbf{b})$ belongs to T_q . We want to estimate the ratio $\#T_q/\#\phi_q(\mathbb{Z}^r)$; the smaller this ratio is, the better the information is. It is convenient to suppose that q is unramified in \mathcal{O}_K . Thus

$$\mathcal{O}_q/q\mathcal{O}_q \cong \mathcal{O}_K/q\mathcal{O}_K \cong \bigoplus_{\mathfrak{q} \mid q} \mathcal{O}_K/\mathfrak{q}.$$

Each summand $\mathcal{O}_K/\mathfrak{q}$ is a finite field of cardinality $\text{Norm}(\mathfrak{q})$. By definition

$$\mathfrak{A}_q := (\mathcal{O}_q/q\mathcal{O}_q)^\times \cong \prod_{\mathfrak{q} \mid q} (\mathcal{O}_K/\mathfrak{q})^\times.$$

Thus

$$\#\mathfrak{A}_q = \prod_{\mathfrak{q}|q} (\text{Norm}(\mathfrak{q}) - 1), \quad \#\mathfrak{B}_q = \frac{1}{q-1} \cdot \prod_{\mathfrak{q}|q} (\text{Norm}(\mathfrak{q}) - 1).$$

Moreover $\prod_{\mathfrak{q}|q} \text{Norm}(\mathfrak{q}) = q^d$ where $d = [K : \mathbb{Q}]$ is the degree of the original Thue–Mahler equation. Thus $\#\mathfrak{B}_q \approx q^{d-1}$. However $S_q \subseteq \phi(\mathbb{Z}^r) \subseteq \mathfrak{B}_q$ has at most $q + 1$ elements, and so $\#S_q/\#\mathfrak{B}_q \lesssim 1/q^{d-2}$. Now

$$\frac{\#T_q}{\#\phi_q(\mathbb{Z}^r)} = \frac{\#S_q \cap \phi_q(\mathbb{Z}^r)}{\#\phi_q(\mathbb{Z}^r)}.$$

It is reasonable to expect that the elements of S_q are uniformly distributed among the elements of \mathfrak{B}_q and so we expect $\#T_q/\#\phi_q(\mathbb{Z}^r) \lesssim 1/q^{d-2}$.

10.3. The sieve. We will sieve with the primes $\mathfrak{p} \in S$ as in Proposition 10.2 and also with additional rational primes q as in Proposition 10.3. We would like to choose a suitable set \mathfrak{S} of such primes q . The most expensive computation we will need to do for $q \in \mathfrak{S}$ is to compute, for each $t \in T_q$, some $\mathbf{w} \in \mathbb{Z}^r$ such that $\phi_q(\mathbf{w}) = t$. This involves a discrete logarithm computation in the group \mathfrak{B}_q , and to do this quickly we need \mathfrak{B}_q to be a product of cyclic factors that have relatively small order. We therefore like to avoid those q where there are $\mathfrak{q} | q$ that have large norm. In all our examples we found it enough to take \mathfrak{S} to be the set of primes $q \leq 500$, where each $\mathfrak{q} | p$ satisfies $\text{Norm}(\mathfrak{q}) \leq 10^{10}$ and where the support of q is disjoint from the supports of $\tau, \delta_1, \dots, \delta_r$.

Procedure 10.4. $\text{Solutions}(L_c, \mathbf{w}_c, S_c, \mathfrak{S}_c)$.

Input: L_c sublattice of \mathbb{Z}^r , $\mathbf{w}_c \in \mathbb{Z}^r$, $S_c \subseteq S$, $\mathfrak{S}_c \subseteq \mathfrak{S}$.

Output: Set of solutions (X, Y, \mathbf{b}) to (15), (16) satisfying $\mathbf{b} \in \mathbf{w}_c + L_c$ and $\|\mathbf{b}\|_2 \leq \mathcal{B}_f$.

1. **IF** $\lambda(L_c) > 2\mathcal{B}_f$ or $(S_c = \emptyset$ and $\mathfrak{S}_c = \emptyset)$ **THEN**
2. Apply Fincke–Pohst to find all vectors in $\mathbf{b} \in \mathbf{w}_c + L$ satisfying $\|\mathbf{b}\|_2 \leq \mathcal{B}_f$.
3. Keep only those \mathbf{b} that lead to solutions (X, Y) on (15), (16).
4. **RETURN:** Set of (X, Y, \mathbf{b}) .
5. **END.**
6. **ELSE**
7. **IF** $S_c \neq \emptyset$ **THEN**
8. Choose $\mathfrak{p} \in S_c$. Let $S'_c = S_c \setminus \{\mathfrak{p}\}$.
9. Compute \mathcal{K} as in Proposition 10.2.
10. For each $k \in \mathcal{K}$ compute \mathbf{w}_k, L_k as in Proposition 10.2.
11. Let \mathcal{K}^* be the subset of $k \in \mathcal{K}$ such that $(\mathbf{w}_k + L_k) \cap (\mathbf{w}_c + L_c) \neq \emptyset$.
12. For each $k \in \mathcal{K}^*$ let $L_{c,k} = L_c \cap L_k$.
13. For each $k \in \mathcal{K}^*$ choose $\mathbf{w}_{c,k} \in \mathbb{Z}^r$ so that $\mathbf{w}_{c,k} + L_{c,k} = (\mathbf{w}_k + L_k) \cap (\mathbf{w}_c + L_c)$.
14. **RETURN:** $\bigcup_{k \in \mathcal{K}^*} \text{Solutions}(L_{c,k}, \mathbf{w}_{c,k}, S'_c, \mathfrak{S}_c)$.
15. **END.**
16. **ELSE**

17. Choose $q \in \mathfrak{S}_c$. Let $\mathfrak{S}'_c = \mathfrak{S}_c \setminus \{q\}$.
18. Compute W_q, L_q as in Proposition 10.3.
19. Let $L'_c = L_c \cap L_q$.
20. Let $W_q^* = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ be the subset of $\mathbf{w} \in W_q$ such that $(\mathbf{w} + L_q) \cap (\mathbf{w}_c + L_c) \neq \emptyset$.
21. For $i = 1, \dots, m$ choose $\mathbf{w}_{c,i}$ such that $\mathbf{w}_{c,i} + L'_c = (\mathbf{w} + L_q) \cap (\mathbf{w}_c + L_c)$.
22. **RETURN:** $\bigcup_{i=1}^m \text{Solutions}(L'_c, \mathbf{w}_{c,i}, \emptyset, \mathfrak{S}'_c)$.
23. **END.**
24. **ENDIF**
25. **ENDIF**

Let us explain how Procedure 10.4 works. The procedure starts with a coset $\mathbf{w}_c + L_c$ and sets $S_c \subseteq S$ and $\mathfrak{S}_c \subseteq \mathfrak{S}$ (the subscript c stands for “cumulative”). The objective is to return all solutions (X, Y, \mathbf{b}) to (15), (16) with $\mathbf{b} \in \mathbf{w}_c + L_c$ and satisfying $\|\mathbf{b}\|_2 \leq \mathcal{B}_f$. The primes in S_c and \mathfrak{S}_c are used, via Propositions 10.2 and 10.3, to replace $\mathbf{w}_c + L_c$ by a union of cosets of sublattices of L_c .

We now explain lines 1–5 of the procedure. If $\lambda(L) > 2\mathcal{B}_f$, then by Lemma 10.1, the coset $\mathbf{w}_c + L_c$ has at most one vector \mathbf{b} that satisfies $\|\mathbf{b}\|_2 \leq \mathcal{B}_f$, and this maybe found by the algorithm of Fincke and Pohst. If $S_c = \emptyset$ and $\mathfrak{S}_c = \emptyset$, then we have run out of sieving primes and we simply apply the Fincke–Pohst algorithm to determine all $\mathbf{b} \in \mathbf{w}_c + L_c$ such that $\|\mathbf{b}\|_2 \leq \mathcal{B}_f$. We test all resulting \mathbf{b} to see if they lead to solutions (X, Y, \mathbf{b}) and return the set of solutions. We end here. In both these cases, no further branching of the procedure occurs.

If we have reached line 6, then either S_c is nonempty or \mathfrak{S}_c is nonempty. We first treat the case where S_c is nonempty (lines 8–14). We choose $\mathfrak{p} \in S_c \subseteq S$ to sieve with and let $S'_c = S_c \setminus \{\mathfrak{p}\}$. Here we apply Proposition 10.2. This gives a finite set \mathcal{K} of values k and lattice cosets $\mathbf{w}_k + L_k$ such that $\mathbf{b} \in \mathbf{w}_k + L_k$ for some $k \in \mathcal{K}$. However, the \mathbf{b} we are interested in belong to $\mathbf{w}_c + L_c$. We let \mathcal{K}^* be those values $k \in \mathcal{K}$ such that $(\mathbf{w}_c + L_c) \cap (\mathbf{w}_k + L_k) \neq \emptyset$. It is now clear that every \mathbf{b} we seek belongs to $(\mathbf{w}_c + L_c) \cap (\mathbf{w}_k + L_k)$ for some $k \in \mathcal{K}^*$. However $(\mathbf{w}_c + L_c) \cap (\mathbf{w}_k + L_k) = \mathbf{w}_{c,k} + L_{c,k}$ where $L_{c,k} = L_c \cap L_k$, for a suitable coset representative $\mathbf{w}_{c,k}$. We apply the procedure to the set $(L_{c,k}, \mathbf{w}_{c,k}, S'_c, \mathfrak{S}_c)$ for each $k \in \mathcal{K}^*$ to compute those \mathbf{b} belonging to $\mathbf{w}_{c,k} + L_{c,k}$ and return the union.

If however $S_c = \emptyset$, then (lines 17–22) we choose a prime $q \in \mathfrak{S}_c \subseteq \mathfrak{S}$ to sieve with, and we let $\mathfrak{S}'_c = \mathfrak{S}_c \setminus \{q\}$. Now we apply Proposition 10.3. This gives a lattice L_q and a finite set W_q such that $\mathbf{b} \in \mathbf{w}_q + L_q$. Therefore there is some $\mathbf{w} \in W_q$ such that $\mathbf{b} \in (\mathbf{w} + L_q) \cap (\mathbf{w}_c + L_c)$. We let W_q^* be the subset of those $\mathbf{w} \in W_q$ such that $(\mathbf{w} + L_q) \cap (\mathbf{w}_c + L_c) \neq \emptyset$, and write $W_q^* = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$. Now $\mathbf{b} \in (\mathbf{w}_i + L_q) \cap (\mathbf{w}_c + L_c)$ for some $i = 1, \dots, m$. Write $L'_c = L_c \cap L_q$. Then $(\mathbf{w}_i + L_q) \cap (\mathbf{w}_c + L_c)$ is a coset of L'_c for $i = 1, \dots, m$, and we choose $\mathbf{w}_{c,i}$ so that $\mathbf{w}_{c,i} + L'_c = (\mathbf{w}_i + L_q) \cap (\mathbf{w}_c + L_c)$. It is therefore enough to find the \mathbf{b} belonging to each one of these $\mathbf{w}_{c,i} + L'_c$. Thus we apply the procedure to $(L'_c, \mathbf{w}_{c,i}, \emptyset, \mathfrak{S}'_c)$ for $i = 1, \dots, m$, collect the solutions and return their union (line 22).

Remarks. • To compute the solutions to (15) satisfying (16), it is clearly enough to apply the above procedure to $(\mathbb{Z}^t, \mathbf{0}, S, \mathfrak{S})$.

\mathfrak{p}	\mathcal{K}	$\det(L_k)$ with $k \in \mathcal{K}$
\mathfrak{p}_1	$\{0, 1\}$	1, 6616761038619033600
\mathfrak{p}_2	$\{0, 1, 2, 3\}$	1, 2114272224838656, 3442909640611645594437761516544, 5606480875148980721912830543593855583743968256
\mathfrak{p}_3	$\{0, 1, 2, 3, 4, 5\}$	1, 1, 3800066789376, 14496104390625000000000000, 55298249781131744384765625000000000000, 210946082233931520022451877593994140625000000000000
\mathfrak{p}_4	$\{0, 1, 2, 3, 6\}$	1, 504631296, 21722722606780416, 935091979414469275815936, 54375352676603537816702220559499682956095667933184
\mathfrak{p}_5	$\{0, 1, 5\}$	1, 57600, 1062532458645670173081600

Table 3. This table gives the sets \mathcal{K} and the determinants of the sublattices $L_k \subset \mathbb{Z}^{10}$ with $k \in \mathcal{K}$ as in Proposition 10.2. Observe that the sublattices L_k all have rank $r - 1 = 9$.

- Recall that $\delta_1, \dots, \delta_r$ is a basis for the S -units (modulo torsion); in particular this allows us to identify the S -units (modulo torsion) with \mathbb{Z}^r . Let $\mathfrak{p} \in S$ and η be as in Proposition 10.2. Note that L_k is a subgroup of finite index in $\text{Ker}(\eta)$. Now $\text{Ker}(\eta)$ itself corresponds to the $(S \setminus \{\mathfrak{p}\})$ -units, and therefore has rank $r - 1$. Therefore L_k has rank $r - 1$. That is, if we apply the procedure to $(\mathbb{Z}^r, \mathbf{0}, S, \mathfrak{S})$, then at depth $\#S + 1$ (when the set S has been entirely depleted), the lattice L_c will have rank $r - \#S$ which is the unit rank. Beyond this depth, the rank remains constant but the determinant of the lattice grows.
- The reader will note that we have not specified how to choose the next prime $\mathfrak{p} \in S$ or $q \in \mathfrak{S}$. In our implementation we order the primes in $\mathfrak{p} \in S$ by the size of their norms; from largest to smallest. The reason is that the primes $\mathfrak{p} \in S$ of large norm lead to lattices of large determinants and we therefore expect few short vectors. Once S is exhausted, the choices we make for the next $q \in \mathfrak{S}$ actually depend on the cumulative lattice L_c . We choose the prime $q \in \mathfrak{S}_c$ that minimizes $\#W_q/[L_c : L_c \cap L_q]$. Our justification for this is that we are replacing one coset of L_c with a union of cosets of $L_c \cap L_q$. The number of such cosets is bounded by $\#W_q$. The function $q \mapsto \#W_q/[L_c : L_c \cap L_q]$ estimates the “relative change in density” between the old lattice and the new union for that particular choice of q .

10.4. Example 1.4 continued. Recall that $B'_f = 180$. Following the remark in Section 8, we find that $B_f \approx 402.67$. Consider the information given by Proposition 10.2. Recall there are five possibilities for $\mathfrak{p} \in S$, ordered as $\mathfrak{p}_1, \dots, \mathfrak{p}_5$, in order of decreasing norm. Table 2 yields 2, 3, 5, 6, 9 for $\kappa_{\mathfrak{p}_j}$ with $j = 1, \dots, 5$, respectively.

We take \mathfrak{S} to be the set of rational primes $q < 200$ coprime to the prime ideals in S and such that every prime ideal factor of $q\mathcal{O}_K$ has norm $\leq 10^{10}$. This is done in order to keep our computations fast, as previously explained. However, of this set, our program only needs to use the primes 23 and 71, selected in that order using the heuristic detailed in the above remarks. For $q = 23$ and $q = 71$, Proposition 10.3

gives a lattice $L_q \subset \mathbb{Z}^{10}$ (now of rank 10) and a set W_q such that $\mathbf{b} \in W_q + L_q$. We find that

$$[\mathbb{Z}^{10} : L_{71}] = 3253933989048960000 \approx 3.26 \times 10^{18}, \quad \text{with } \#W_{71} = 71,$$

and

$$[\mathbb{Z}^{10} : L_{23}] = 41191874887680 \approx 4.12 \times 10^{13}, \quad \text{with } \#W_{23} = 23.$$

Observe that in Procedure 10.4 branching occurs at lines 14, 20. Thus we obtain “paths” through the algorithm depending on the choice of $k \in \mathcal{K}^*$ (line 14) or the choice of $\mathbf{w}_i \in W_q^*$ (line 20). A path “dies” if the criterion of line 1 is satisfied, or if \mathcal{K}^* (defined in line 11) is empty, or if W_q^* (defined in line 20) is empty. Our program needs to check a total of 98 paths. Five of these terminate at line 1 with the condition $\lambda(L_c) > 2\mathcal{B}_f$ being satisfied, and the remaining 93 paths terminate at line 11 with $\mathcal{K}^* = \emptyset$. Of the 5 paths that terminate at line 1, three of these yield a vector $\mathbf{b} \in W_c + L_c$ satisfying $\|\mathbf{b}\|_2 \leq \mathcal{B}_f$. These three vectors are

$$(-1, -1, -2, 0, 2, 0, 3, 0, 1, 1), \quad (0, 0, -1, 1, 1, 1, 1, 1, 0, 0), \quad \text{and} \quad (-1, -1, -2, 3, 2, 5, 0, 0, 0, 0).$$

These vectors respectively lead to the solutions

$$F(1, 2) = 3^3 \cdot 7 \cdot 11, \quad F(1, -1) = 2 \cdot 3 \cdot 5, \quad F(1, 1) = 2^5.$$

Acknowledgements

The authors are grateful to Mike Bennett, Rafael von Känel and Benjamin Matschke for stimulating discussions. The authors are indebted to the referee for many pertinent corrections and improvements.

References

- [Agrawal et al. 1980] M. K. Agrawal, J. H. Coates, D. C. Hunt, and A. J. van der Poorten, “Elliptic curves of conductor 11”, *Math. Comp.* **35**:151 (1980), 991–1002. MR Zbl
- [Bennett and Dahmen 2013] M. A. Bennett and S. R. Dahmen, “Klein forms and the generalized superelliptic equation”, *Ann. of Math. (2)* **177**:1 (2013), 171–239. MR Zbl
- [Bennett and Siksek 2023a] M. A. Bennett and S. Siksek, “Differences between perfect powers: prime power gaps”, *Algebra Number Theory* **17**:10 (2023), 1789–1846. MR Zbl
- [Bennett and Siksek 2023b] M. A. Bennett and S. Siksek, “Differences between perfect powers: the Lebesgue–Nagell equation”, *Trans. Amer. Math. Soc.* **376**:1 (2023), 335–370. MR Zbl
- [Bennett et al. 2019] M. A. Bennett, A. Gherga, and A. Rechnitzer, “Computing elliptic curves over \mathbb{Q} ”, *Math. Comp.* **88**:317 (2019), 1341–1390. MR Zbl
- [Bennett et al. 2020] M. A. Bennett, A. Gherga, and D. Kreso, “An old and new approach to Goormaghtigh’s equation”, *Trans. Amer. Math. Soc.* **373**:8 (2020), 5707–5745. MR
- [Bennett et al. 2022] M. A. Bennett, A. Gherga, V. Patel, and S. Siksek, “Odd values of the Ramanujan tau function”, *Math. Ann.* **382**:1-2 (2022), 203–238. MR
- [Bombieri 1987] E. Bombieri, “On the Thue–Mahler equation”, pp. 213–243 in *Diophantine approximation and transcendence theory* (Bonn, Germany, 1985), edited by G. Wüstholz, Lecture Notes in Math. **1290**, Springer, 1987. MR Zbl
- [Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, “The Magma algebra system, I: The user language”, *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl
- [Bruin and Stoll 2008] N. Bruin and M. Stoll, “Deciding existence of rational points on curves: an experiment”, *Exp. Math.* **17**:2 (2008), 181–189. MR Zbl

- [Bruin and Stoll 2010] N. Bruin and M. Stoll, “The Mordell–Weil sieve: proving non-existence of rational points on curves”, *LMS J. Comput. Math.* **13** (2010), 272–306. MR Zbl
- [Bugeaud and Györy 1996a] Y. Bugeaud and K. Györy, “Bounds for the solutions of Thue–Mahler equations and norm form equations”, *Acta Arith.* **74**:3 (1996), 273–292. MR Zbl
- [Bugeaud and Györy 1996b] Y. Bugeaud and K. Györy, “Bounds for the solutions of unit equations”, *Acta Arith.* **74**:1 (1996), 67–80. MR
- [Bugeaud et al. 2006] Y. Bugeaud, M. Mignotte, and S. Siksek, “Classical and modular approaches to exponential Diophantine equations, I: Fibonacci and Lucas perfect powers”, *Ann. of Math. (2)* **163**:3 (2006), 969–1018. MR Zbl
- [Bugeaud et al. 2008] Y. Bugeaud, M. Mignotte, S. Siksek, M. Stoll, and S. Tengely, “Integral points on hyperelliptic curves”, *Algebra Number Theory* **2**:8 (2008), 859–885. MR Zbl
- [Cangül et al. 2010] İ. N. Cangül, M. Demirci, G. Soydan, and N. Tzanakis, “On the Diophantine equation $x^2 + 5^a \cdot 11^b = y^n$ ”, *Funct. Approx. Comment. Math.* **43**:2 (2010), 209–225. MR Zbl
- [Coates 1970] J. Coates, “An effective p -adic analogue of a theorem of Thue, II: The greatest prime factor of a binary form”, *Acta Arith.* **16** (1970), 399–412. MR Zbl
- [Cohen 2000] H. Cohen, *Advanced topics in computational number theory*, Grad. Texts in Math. **193**, Springer, 2000. MR Zbl
- [Evertse 1984] J.-H. Evertse, “On equations in S -units and the Thue–Mahler equation”, *Invent. Math.* **75**:3 (1984), 561–584. MR Zbl
- [Fincke and Pohst 1985] U. Fincke and M. Pohst, “Improved methods for calculating vectors of short length in a lattice, including a complexity analysis”, *Math. Comp.* **44**:170 (1985), 463–471. MR Zbl
- [Gallegos-Ruiz 2011] H. R. Gallegos-Ruiz, “ S -integral points on hyperelliptic curves”, *Int. J. Number Theory* **7**:3 (2011), 803–824. MR Zbl
- [Hambrook 2011] K. D. Hambrook, *Implementation of a Thue–Mahler equation solver*, Ph.D. thesis, University of British Columbia, 2011, available at <http://hdl.handle.net/2429/38244>.
- [von Känel and Matschke 2023] R. von Känel and B. Matschke, *Solving S -unit, Mordell, Thue, Thue–Mahler and generalized Ramanujan–Nagell equations via the Shimura–Taniyama conjecture*, Mem. Amer. Math. Soc. **1419**, Amer. Math. Soc., Providence, RI, 2023. MR Zbl
- [Kim 2017] D. Kim, “A modular approach to cubic Thue–Mahler equations”, *Math. Comp.* **86**:305 (2017), 1435–1471. MR Zbl
- [Lenstra et al. 1982] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, “Factoring polynomials with rational coefficients”, *Math. Ann.* **261**:4 (1982), 515–534. MR Zbl
- [Mahler 1933] K. Mahler, “Zur Approximation algebraischer Zahlen, I”, *Math. Ann.* **107**:1 (1933), 691–730. MR Zbl
- [Matveev 2000] E. M. Matveev, “An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers, II”, *Izv. Ross. Akad. Nauk Ser. Mat.* **64**:6 (2000), 125–180. In Russian; translated in *Izv. Math.* **64**:6 (2000), 1217–1269. MR Zbl
- [Pethő and de Weger 1998] A. Pethő and B. M. M. de Weger, “Appendix B: Two useful lemmata”, pp. 229–230 in *The algorithmic resolution of Diophantine equations*, Lond. Math. Soc. Stud. Texts **41**, Cambridge Univ. Press, 1998. MR Zbl
- [Soydan and Tzanakis 2016] G. Soydan and N. Tzanakis, “Complete solution of the Diophantine equation $x^2 + 5^a \cdot 11^b = y^n$ ”, *Bull. Hellenic Math. Soc.* **60** (2016), 125–151. MR Zbl
- [Thue 1909] A. Thue, “Über Annäherungswerte algebraischer Zahlen”, *J. Reine Angew. Math.* **135** (1909), 284–305. MR Zbl
- [Tzanakis and de Weger 1989] N. Tzanakis and B. M. M. de Weger, “On the practical solution of the Thue equation”, *J. Number Theory* **31**:2 (1989), 99–132. MR Zbl
- [Vinogradov and Sprindzhuk 1968] A. I. Vinogradov and V. G. Sprindzhuk, “The representation of numbers by binary forms”, *Mat. Zametki* **3** (1968), 369–376. In Russian; translated in *Math. Notes* **3**:4 (1968), 235–239. MR Zbl
- [Yu 2007] K. Yu, “ p -adic logarithmic forms and group varieties, III”, *Forum Math.* **19**:2 (2007), 187–280. MR Zbl

Communicated by Antoine Chambert-Loir

Received 2022-07-28

Revised 2024-04-09

Accepted 2024-06-15

adelagherga@gmail.com

Tutte Institute for Mathematics and Computing, Ottawa, Ontario, Canada

s.siksek@warwick.ac.uk

Mathematics Institute, University of Warwick, Coventry, United Kingdom

Automorphisms of del Pezzo surfaces in characteristic 2

Igor Dolgachev and Gebhard Martin

We classify the automorphism groups of del Pezzo surfaces of degrees 1 and 2 over an algebraically closed field of characteristic 2. This finishes the classification of automorphism groups of del Pezzo surfaces in all characteristics.

Introduction	715
1. Notation	716
2. Del Pezzo surfaces of degree ≥ 3	717
3. Del Pezzo surfaces of degree 2	718
4. Automorphism groups of del Pezzo surfaces of degree 2	727
5. Del Pezzo surfaces of degree 1	736
6. Automorphism groups of del Pezzo surfaces of degree 1	745
References	761

Introduction

This is a continuation of our paper [Dolgachev and Martin 2024], where we finished the classification of the automorphism groups of del Pezzo surfaces over an algebraically closed field of positive characteristic $p \neq 2$. In this paper, we treat the remaining case when the characteristic equals 2.

As we explained in the Introduction to [loc. cit.], the remaining part of the classification concerns del Pezzo surfaces of degrees 1 and 2. The cases of odd and even positive characteristic are drastically different since, in the latter case, the anticanonical map (resp. the antibicanonical map) is a separable Artin–Schreier cover of degree 2 but not a Kummer cover as in the cases of odd characteristic. So, no plane quartic curves (and no canonical genus-4 curves with vanishing theta characteristic) appear as branch curves.

Instead, in characteristic 2, the branch curve B of the anticanonical (resp. antibicanonical) map is not necessarily smooth plane conic (resp. a cubic in \mathbb{P}^3). The ramification curve R is a purely inseparable cover of B . Theorems 3.4 and 5.6 give normal forms for del Pezzo surfaces of degree 2 and 1 depending on the singularities of R and B .

Although plane quartics and canonical curves of genus 4 disappear in characteristic 2, their familiar attributes, like 28 bitangent lines or 120 tritangent planes, persist. We call them *fake bitangents* and *fake tritangent planes*. They are defined to be lines in the plane (resp. planes in the 3-dimensional space) that split under the anticanonical (resp. antibicanonical) map.

MSC2020: 14E07, 14G17, 14J26, 14J50.

Keywords: del Pezzo surfaces, automorphisms, characteristic 2, Cremona group.

It is well known that the blow-up of the anticanonical base point on a del Pezzo surface of degree 1 yields a rational elliptic surface with only irreducible fibers and, conversely, the contraction of a section of a rational elliptic surface with only irreducible fibers yields a del Pezzo surface of degree 1. Thus, the normal forms of Theorem 5.6 also give normal forms for all rational elliptic surfaces with only irreducible fibers.

Quite surprisingly, in characteristic 2, also every del Pezzo surface of degree 2 has a canonically associated rational elliptic surface. This surface is obtained by blowing up the base points of the preimage of the pencil of lines through the *strange point* of the branch locus B . We study the properties of this *strange fibration* in Section 3.4.

Using these geometric observations, we classify the automorphism groups of all del Pezzo surfaces of degree 2 and 1 in characteristic 2. The following result is proved in Theorems 4.3 and 6.8.

Theorem. *A finite group G is realized as the automorphism group $\text{Aut}(X)$ of a del Pezzo surface X of degree 1 or 2 over an algebraically closed field k of characteristic $\text{char}(k) = 2$ if and only if G is listed in, respectively, Table 9 (page 760) or Table 4 (page 736).*

Table 4 (resp. Table 9) also gives the conjugacy classes in $W(E_7)$ (resp. $W(E_8)$) of all elements of $\text{Aut}(X)$ for all del Pezzo surfaces X of degree 2 (resp. degree 1). We refer to [Dolgachev and Martin 2024] for a general discussion of the history of the problem and its relationship to the classification of conjugacy classes of finite subgroups of the planar Cremona group. Also, the reader finds there some general facts about del Pezzo surfaces, e.g., the relationship with the Weyl groups of roots systems and some classification results from group theory.

1. Notation

We recall the notation for some finite groups we will encounter in this article. Throughout, p is a prime number, and q is a power of p . Unless stated otherwise, \mathbb{k} denotes an algebraically closed field of characteristic 2.

- C_n is the cyclic group of order n .
- \mathfrak{S}_n and \mathfrak{A}_n are the symmetric and alternating groups on n letters.
- Q_8 is the quaternion group of order 8.
- D_{2n} is the dihedral group of order $2n$.
- $n^k = (\mathbb{Z}/n\mathbb{Z})^k$. In particular, $n = n^1 = \mathbb{Z}/n\mathbb{Z}$.
- p_{\pm}^{1+2n} is the extra special group. For odd p the sign $+$ ($-$) defines a group of exponent p (p^2). For $p = 2$, the sign distinguishes the type of the quadratic forms on $2^{2n} = \mathbb{F}_2^{2n}$ defined by the extension.
- $\text{GL}_n(q) = \text{GL}(n, \mathbb{F}_q)$.
- $\text{PGL}_n(q) = \text{GL}_n(q)/\mathbb{F}_q^*$. Its order is $N = q^{1/2n(n-1)}(q^n - 1) \cdots (q^2 - 1)$.
- $\text{SL}_n(q) = \{g \in \text{GL}_n(q) : \det(g) = 1\}$. This is a subgroup of $\text{GL}_n(q)$ of index $(q - 1)$.
- $\text{L}_n(q) = \text{PSL}_n(q)$ is the image of $\text{SL}_n(q)$ in $\text{PGL}_n(q)$. Its order is $N/(q - 1, n)$.

- For odd n , $O_n(q)$ is the subgroup of $GL_n(q)$ that preserves a nondegenerate quadratic form F .
- For even n , $O_n^+(q)$ (resp. $O_n^-(q)$) is the subgroup of $GL_n(q)$ that preserves a nondegenerate quadratic form F of Witt defect 0 (resp. 1).
- $SO_n^\pm(q)$ is the subgroup of $O_n^\pm(q)$ of elements with determinant 1.
- $PSO_n^\pm(q)$ is the quotient of $SO_n^\pm(q)$ by its center.
- $Sp_{2n}(q)$ is the subgroup of $SL_q(2n)$ preserving the standard symplectic form on \mathbb{F}_q^{2n} . Its order is $q^{n^2}(q^{2n-1} - 1) \cdots (q^2 - 1)$.
- $\mathbb{S}p_{2n}(q) = Sp_{2n}(q)/(\pm 1)$.
- $SU_n(q^2)$ is the subgroup of $SL_n(q^2)$ of matrices preserving the hermitian form $\sum_{i=1}^n x_i^{q+1}$. Its order is $q^{(1/2)n(n-1)}(q^n - (-1)^n)(q^{n-1} - (-1)^{n-1}) \cdots (q^3 + 1)(q^2 - 1)$. We have $SU_2(q^2) = SL_2(q)$.
- $PSU_n(q^2) = SU_n(q^2)/C$, where C is a cyclic group of order $(q+1, n)$ of diagonal Hermitian matrices. The simple group $PSU_n(q^2)$ is denoted by $U_n(q)$ in [Conway et al. 1985].
- $\mathcal{H}_3(3)$ is the Heisenberg group of 3×3 upper triangular matrices with entries in \mathbb{F}_3 .
- $A.B$ is a group that contains a normal subgroup A with quotient group B .
- $A : B$ is the semidirect product $A \rtimes B$.

2. Del Pezzo surfaces of degree ≥ 3

For the convenience of the reader, we first recall the classification of automorphism groups of del Pezzo surfaces of degree at least 3.

2.1. Degree ≥ 5 . For del Pezzo surfaces of degree at least 5, the description of $\text{Aut}(X)$ is characteristic-free. We refer the reader to [Dolgachev and Martin 2024, Section 3; Dolgachev 2012] for details.

2.2. Quartic del Pezzo surfaces. Starting from degree 4, the classification of automorphism groups depends on the characteristic. As in the other characteristics, a quartic del Pezzo surface X is a blow-up of five points in \mathbb{P}^2 no three of which are colinear. Moreover, the anticanonical linear system $|-K_X| = |\mathcal{O}_{\mathbb{P}^2}(3) - p_1 - p_2 - p_3 - p_4 - p_5|$ embeds X into \mathbb{P}^4 as a complete intersection of two quadrics.

Since $p = 2$, these quadrics cannot be diagonalized. Instead, as shown in [Dolgachev and Duncan 2019], one can choose the normal forms

$$(ab + b + 1)t_2^2 + at_3^2 + t_2t_3 + t_3t_4 = bt_1^2 + (ab + a + 1)t_2^2 + t_1t_3 + t_2t_4 = 0, \quad (1)$$

where a, b are parameters such that the binary form $\Delta = uv(u+v)(u+av)(bu+v)$ has five distinct roots.

As in the case $p \neq 2$, the automorphism group $\text{Aut}(X)$ contains a normal subgroup H isomorphic to 2^4 , and the quotient $G = \text{Aut}(X)/H$ is isomorphic to a subgroup of \mathfrak{S}_5 . The classification is summarized in Table 1 on the next page. The first column refers to the values of the parameters a and b in (1) above. The conjugacy classes of elements of $\text{Aut}(X)$ can be obtained by combining [Dolgachev and Duncan 2019, Table 2] and [Carter 1972, Table 5].

name	Aut(X)	order	id	2A ₁	4A ₁	A ₂	A ₂ +2A ₁	A ₃	A ₃ +A ₁	A ₄	D ₄	D ₄ (a ₁)	D ₅
(ϕ, ϕ)	2 ⁴ : \mathfrak{A}_5	960	1	70	5	80	80		120	384	160	60	
(ζ_3, ζ_3)	(same as (ϕ, ϕ))												
(i, i)	(does not exist)												
(a, a)	2 ⁴ :2 ²	64	1	22	5				24			12	
general	2 ⁴	16	1	10	5								

Table 1. Automorphism groups of quartic del Pezzo surfaces; see Section 2.2.

2.3. Cubic surfaces. The classification of automorphism groups of cubic surfaces in characteristic 2 was achieved in [Dolgachev and Duncan 2019, Table 7]. For the convenience of the reader, we recall it here:

name	Aut(X)	order	id	2A ₁	4A ₁	A ₂	A ₂ +2A ₁	2A ₂	3A ₂	A ₃ +A ₁	A ₄	A ₅ +A ₁	D ₄	D ₄ (a ₁)	D ₅	E ₆	E ₆ (a ₁)	E ₆ (a ₂)
I/3C	PSU ₄ (2)	25920	1	270	45	240	2160	480	80	3240	5184	1440	1440	540		4320	5760	720
II/5A	(same as V)																	
III/12A	(same as I)																	
IV/3A	$\mathcal{H}_3(3):2$	54	1		9			24	2									18
V/4B	2 ³ : \mathfrak{S}_4	192	1	30	13			32		72		32		12				
VI/6E	(same as V)																	
VII/8A	(does not exist)																	
VIII/3D	\mathfrak{S}_3	6	1		3			2										
IX/4A	(same as V)																	
X/2B	2 ⁴	16	1	10	5													
XI/2A	2	2	1		1													
XII/1A	1	1	1															

3. Del Pezzo surfaces of degree 2

3.1. The anticanonical map. We start by describing the geometry of del Pezzo surfaces of degree $d = 2$ over an algebraically closed field \mathbb{k} of characteristic $p = 2$. We refer to [Demazure 1980] for the basic facts from the theory of del Pezzo surfaces over fields of any characteristic. It is known that the anticanonical linear system $| -K_X |$ has no base points and defines a finite morphism $f : X \rightarrow \mathbb{P}^2$ of degree 2.

If $p \neq 2$, the map f is automatically separable and its branch curve is a smooth plane quartic. So any automorphism of X induces an automorphism of the quartic, and, conversely, any automorphism of the quartic can be lifted to two automorphisms of X that differ by the deck transformation, classically called the Geiser involution.

If $p = 2$, the structure of f , being a morphism of degree 2, is more complicated. Nevertheless, as a first step, we observe that f is still always separable.

Proposition 3.1. *The anticanonical linear system $| -K_X |$ defines a finite separable morphism $f : X \rightarrow \mathbb{P}^2$ of degree 2.*

Proof. Assume that f is not separable. Then, since $\deg(f) = 2$, f is purely inseparable. Hence, f is a homeomorphism in the étale topology, which is absurd since $H_{\text{ét}}^2(X, \mathbb{Z}_\ell)$ has rank 8 (because X is the blow-up of seven points in the plane), while $H_{\text{ét}}^2(\mathbb{P}^2, \mathbb{Z}_\ell)$ has rank 1. □

Let

$$R(X, -K_X) = \bigoplus_{n=0}^{\infty} H^0(X, \mathcal{O}_X(-nK_X))$$

be the graded anticanonical ring of X . By the Riemann–Roch theorem, $\dim_{\mathbb{k}} R(X, -K_X)_1 = 3$ and $\dim_{\mathbb{k}} R(X, -K_X)_2 = 7$. One can show that $R(X, -K_X)$ is generated by $R(X, -K_X)_1$ and one element from $R(X, -K_X)_2$ that does not belong to the symmetric square of $R(X, -K_X)_1$. Let x, y, z be elements of $R(X, -K_X)_1$ and $w \in R(X, -K_X)_2$, which together generate $R(X, -K_X)$. Then, the relation between the generators is of the form

$$w^2 + A(x, y, z)w + B(x, y, z) = 0, \tag{2}$$

where A and B are homogeneous forms of degree 2 and 4, respectively. In particular, via (2), we can view X as a surface of degree 4 in the weighted projective space $\mathbb{P}(1, 1, 1, 2)$, and the anticanonical map is the projection of this surface onto the x, y, z -coordinates.

If $p \neq 2$, we can complete the square, get rid of A , and obtain the standard equation of a del Pezzo surface of degree 2. The curve $V(B(x, y, z))$ is the smooth plane quartic we mentioned in the Introduction. The Geiser involution just negates w .

In our case, when $p = 2$, we cannot get rid of A , for otherwise the map would become inseparable. Also, the coefficient B is not uniquely determined, since replacing w with $w + Q$ for any quadratic form Q changes B to $B + AQ + Q^2$, without changing the isomorphism class of the surface. Taking $Q = A$, we obtain the analog of the Geiser involution, so we keep the name for this involution.

The nonuniqueness of B becomes more natural if we take the following different point of view: By [Ekedahl 1988, Proposition 1.11], the double cover f is a torsor under a group scheme $\alpha_{\mathcal{L},s}$ of order 2 over \mathbb{P}^2 , defined by the exact sequence of fppf-sheaves

$$0 \rightarrow \alpha_{\mathcal{L},s} \rightarrow \mathcal{L} \xrightarrow{\phi} \mathcal{L}^{\otimes 2} \rightarrow 0$$

for some line bundle \mathcal{L} and a global section s . The homomorphism of sheaves ϕ is locally given by $a \mapsto a_U^2 + a_U s_U$, so s cuts out the branch locus of f . By [loc. cit., Proposition 1.7], we have $\omega_X \cong f^*(\mathcal{O}_{\mathbb{P}^2}(-3) \otimes \mathcal{L}^{-1})$; hence $\mathcal{L} \cong \mathcal{O}_{\mathbb{P}^2}(2)$ and $s = A$. The $\alpha_{\mathcal{L},s}$ -torsor corresponding to f is defined by a cohomology class in $H_{\text{fppf}}^1(\mathbb{P}^2, \alpha_{\mathcal{L},s})$. Since $H_{\text{fppf}}^1(\mathbb{P}^2, \mathcal{L}) = H^1(\mathbb{P}^2, \mathcal{L}) = 0$, we have

$$H_{\text{fppf}}^1(\mathbb{P}^2, \alpha_{\mathcal{L},s}) \cong H^0(\mathbb{P}^2, \mathcal{L}^{\otimes 2}) / \wp(H^0(\mathbb{P}^2, \mathcal{L})),$$

where $\wp = H^0(\phi)$. The ternary form B is a representative of this space, and hence it is defined only up to a transformation of the form $B \mapsto B + Q^2 + AQ$, where Q is a quadratic form in x, y, z .

By writing the equation of X locally as $w_U^2 + a_U w_U + b_U$, and taking partial derivatives, we see that the differentials $w_U da_U + db_U$ restricted to $V(A)$ glue together to define a global section α of $\Omega_{\mathbb{P}^2}^1 \otimes \mathcal{L}^{\otimes 2} \otimes \mathcal{O}_{V(A)}$. This section vanishes if and only if X is singular. So, in our case, when X is assumed to be smooth, we obtain the following.

Proposition 3.2. *In (2), the equations $A = 0$, $wA_x + B_x = 0$, $wA_y + B_y = 0$, $wA_z + B_z = 0$ have no common solutions.*

3.2. Normal forms. Recall that X is given by an equation of the form $w^2 + Aw + B = 0$, where A is a quadratic ternary form and B is quartic ternary form. We say that $V(A)$ is the branch curve of the cover, and its preimage $R = f^{-1}(V(A))$ under the anticanonical map $f : X \rightarrow \mathbb{P}^2$ will be called the *ramification curve*.

Remark 3.3. We use the notation A_{2n} for singularities of curves whose formal completion is isomorphic to the unibranch singularity $y^2 + x^{2n+1} = 0$. If $n = 1$, this is an ordinary cusp singularity. These are exactly the curve singularities that can occur on reduced purely inseparable double covers of smooth curves in characteristic 2. Indeed, after passing to formal completions, such a double cover is given by an equation of the form $y^2 + ux^m$, where $u \in k[[x]]$ is a unit. Now, we can apply a substitution of the form $y \mapsto y + f$ for a suitable power series f to assume that m is odd and then replace x by λx , where λ is an m -th root of u^{-1} , which exists by Hensel's lemma. In other words, the singularity defined by $y^2 + ux^m$ is of type A_{2n} , where $2n + 1$ is the smallest odd power of x that occurs in ux^m .

The following theorem gives normal forms for the cover $f : X \rightarrow \mathbb{P}^2$. In total, we obtain six normal forms, corresponding to the six possible combinations of singularities of $V(A)$ and R .

Theorem 3.4. *Every del Pezzo surface of degree 2 in characteristic 2 is isomorphic to a quartic surface in $\mathbb{P}(1, 1, 1, 2)$ given by an equation of the form*

$$w^2 + A(x, y, z)w + B(x, y, z),$$

where (A, B) is one of the forms shown in Table 2. The parameters satisfy the following conditions:

(1a) $\lambda \neq 0$, $\lambda^2 + a + b + c + d + e \neq 0$, $b^2 + a \neq 0$, $d^2 + e \neq 0$.

(1b) $b^2 + a \neq 0$, $d^2 + e \neq 0$.

(1c) $d^2 + e \neq 0$.

(2a) $a \neq 0$, $b \neq 0$.

(2b) $a \neq 0$.

(3) None.

In terms of these normal forms, the singularities of the irreducible components of R_{red} are as follows:

(1a) Three A_2 -singularities, over $[0 : 1 : 0]$, $[0 : 0 : 1]$ and $[1 : 1 : 1]$.

(1b) An A_4 -singularity over $[0 : 0 : 1]$ and an A_2 -singularity over $[0 : 1 : 0]$.

(1c) An A_6 -singularity over $[0 : 0 : 1]$.

(2a) Two A_2 -singularities, over $[1 : 0 : 0]$ and $[0 : 1 : 0]$.

(2b) Two A_2 -singularities, over $[1 : 0 : 0]$ and $[0 : 0 : 1]$.

(3) An A_2 -singularity over $[0 : 0 : 1]$.

name	A	B	B_1	B_0	# of parameters
(1a)	x^2+yz	$x B_1+B_0$	$\lambda yz(y+z)$	$ay^4+by^3z+cy^2z^2+dyz^3+ez^4$	6
(1b)	x^2+yz	$x B_1+B_0$	y^2z	$ay^4+by^3z+cy^2z^2+dyz^3+ez^4$	5
(1c)	x^2+yz	$x B_1+B_0$	y^3	$by^3z+cy^2z^2+dyz^3+ez^4$	4
(2a)	xy	$B_1+B_0^2$	xz^3+yz^3	$ax^2+by^2+cz^2+dxz+eyz$	5
(2b)	xy	$B_1+B_0^2$	xz^3+y^3z	$ax^2+cz^2+dxz+eyz$	4
(3)	x^2	$x B_1+B_0$	z^3+ayz^2	$y^3z+by^2z^2+cz^4$	3

Table 2. Forms of (A, B) in Theorem 3.4.

Proof. Since $f : X \rightarrow \mathbb{P}^2$ is separable, A is nonzero. Hence, up to projective equivalence, there are three possibilities for A , corresponding to (1), (2), and (3). Now, we study those cases separately. The conditions on the parameters will follow from Proposition 3.2 by computing partial derivatives, a task which we will leave to the reader.

(1) $A = x^2 + yz$. Applying a substitution of the form $w \mapsto w + Q$ for a suitable quadratic form Q allows us to assume that $B = x B_1 + B_0$ for homogeneous forms B_0 and B_1 in y and z .

Let $x = uv$, $y = u^2$, $z = v^2$ define the Veronese isomorphism between $V(A)$ and \mathbb{P}^1 . Substituting in B , we get that R is isomorphic to the double cover of \mathbb{P}^1 given by the equation

$$w^2 + uvB_1(u^2, v^2) + B_0(u^2, v^2) = 0.$$

By taking the partials, we find that R is singular exactly over the roots of B_1 .

After applying a suitable substitution that preserves A , we can move these roots to special positions. Note that the substitution $w \mapsto w + Q$ of the first paragraph does not change the position of these singularities, so we can still assume that $B = x B_1 + B_0$.

If the roots are distinct, we get case (a), if there are two distinct roots, we get case (b), and if there is only a single root, we get case (c). Note that in cases (b) and (c), the substitution $y \mapsto \lambda y$, $z \mapsto \lambda^{-1}z$ preserves the location of the roots and scales B_1 , which is why we can assume that $x B_1$ occurs with coefficient 1. Finally, in case (c), we can apply a substitution of the form $z \mapsto z + \lambda^2 y$, $x \mapsto x + \lambda y$ for a suitable λ to assume that $B_0(1, 0) = 0$.

(2) $A = xy$. After applying a substitution of the form $w \mapsto w + Q$ for a suitable quadratic form Q , we may assume that B does not contain monomials divisible by xy . This allows us to write

$$B = (a_1x^3 + a_2y^3)z + (a_3x + a_4y)z^3 + B_0(x, y, z)^2.$$

Note that the preimages R_1 and R_2 of $V(x)$ and $V(y)$ on X are members of $|-K_X|$; hence they must be reduced.

Restricted to $V(x)$, the equation becomes $w^2 + a_2y^3z + a_4yz^3$, so R_1 is singular over $[0 : \sqrt{a_4} : \sqrt{a_2}]$. Similarly, R_2 is singular over $[\sqrt{a_3} : 0 : \sqrt{a_1}]$. Note that these points must be distinct, for otherwise X is singular over $[0 : 0 : 1]$ by Proposition 3.2.

If these two points are distinct and different from $[0 : 0 : 1]$, we can apply a suitable substitution that preserves A to move them to $[0 : 1 : 0]$ and $[1 : 0 : 0]$. Then, we can repeat the substitution of the first paragraph and, after rescaling, arrive at case (a).

If the two points are distinct and one of them is $[0 : 0 : 1]$, we can assume without loss of generality that the other one lies on $V(y)$ and move it to $[1 : 0 : 0]$. After repeating the substitution of the first paragraph and rescaling, we may assume that B_1 is as in case (b). Finally, after applying a substitution of the form $z \mapsto z + \lambda y$, $w \mapsto w + \lambda z^2 + \lambda^2 yz + \lambda^3 y^2$ for a suitable λ , we may assume that $B_0(0, 1, 0) = 0$.

(3) $A = x^2$. Applying a substitution of the form $w \mapsto w + Q$ for a suitable quadratic form Q allows us to assume that $B = xB_1 + B_0$ for homogeneous forms B_0 and B_1 in y and z .

Let R' be the preimage of $V(x)$. As in case (2), since $R' \in |-K_X|$, R' must be reduced. Restricted to $V(x)$, the double cover becomes

$$w^2 + B_0(y, z) = 0;$$

hence R' is singular over the common zero of $B_{0,y}$ and $B_{0,z}$. We can assume that this zero lies at $[0 : 0 : 1]$, that is, that yz^3 does not occur in B_0 and y^3z occurs with nonzero coefficient. After rescaling, we may assume that y^3z occurs with coefficient 1.

Applying a substitution of the form $z \mapsto z + \lambda_1 x + \lambda_2 y$, $y \mapsto y + \lambda_3 x$ for suitable λ_i and repeating the substitution of the first paragraph, we can eliminate the monomials y^3 and y^2z in B_1 and the monomial y^4 in B_0 . Computing partials, we see that X is singular if and only if $B_1(0, 1) = 0$. Hence, after rescaling, we may assume that B is as claimed. \square

3.3. Fake bitangents and odd theta characteristics. It is known that a del Pezzo surface X of degree 2 contains 56 (-1) -curves (see [Dolgachev 2012, Section 8.7], where the proof is characteristic-free). They come in pairs $E_i + E'_i \in |-K_X|$, with $E_i \cdot E'_i = 2$. The Geiser involution γ switches the two curves in a pair. The image of each pair under any birational morphism $\pi : X \rightarrow \mathbb{P}^2$ given by the blow-up of seven points $p_1, \dots, p_7 \in \mathbb{P}^2$ is either the union of a line through two points p_i, p_j and the conic through the remaining five points, or a cubic passing through p_1, \dots, p_7 with a double point at some p_i (and one curve of the pair is contracted by π). The image of $E_i + E'_i$ under the anticanonical map f is a line ℓ .

If $p \neq 2$, each of the resulting 28 lines is a bitangent line to the branch quartic curve and, conversely, every bitangent to the branch quartic gives rise to a pair of (-1) -curves. A bitangent line intersects the branch curve at two points, not necessarily distinct, whose sum is an odd theta characteristic of the curve. It is known that the number of odd theta characteristics on a smooth curve of genus 3 is equal to 28.

For arbitrary p , we still have the following.

Lemma 3.5. *The preimage $f^{-1}(\ell)$ of a line ℓ is a sum of two (-1) -curves if and only if $f^{-1}(\ell)$ is reducible.*

Proof. Since f has degree 2 and ℓ is integral, the curve $f^{-1}(\ell)$ is reducible if and only if it has two irreducible components L_1 and L_2 . These components satisfy $L_1 + L_2 \in |-K_X|$. Since L_i maps birationally to ℓ , we have $p_a(L_i) = 0$; hence $L_1 \cdot L_2 = 2$ by adjunction. We have $L_1^2 = L_2^2$, because the

two curves are interchanged by the covering involution, so the equality

$$2 = K_X^2 = (L_1 + L_2)^2 = L_1^2 + L_2^2 + 2L_1 \cdot L_2$$

implies that L_i is a (-1) -curve. The converse is clear. □

So, even if $p = 2$, we have 28 splitting lines, which we call *fake bitangent lines* in analogy with the situation in the other characteristics. For the rest of this section, we assume $p = 2$. Since the anticanonical map is étale outside the branch curve $V(A)$, the intersection $E_i \cap E'_i$ lies on the ramification curve R . Let $\mathcal{L} = \mathcal{O}_R(E_i) \cong \mathcal{O}_R(E'_i)$. It is an invertible sheaf on R of degree 2. We have

$$\mathcal{L}^{\otimes 2} \cong \mathcal{O}_R(E_i + E'_i) \cong \mathcal{O}_R(-K_X).$$

Since $B \in |\mathcal{O}_{\mathbb{P}^2}(2)|$, we have $R \in |-2K_X|$. By the adjunction formula

$$\omega_R \cong \mathcal{O}_R(-2K_X + K_X) \cong \mathcal{L}^{\otimes 2}.$$

Invertible sheaves \mathcal{L} on R that satisfy this property are called *invertible theta characteristics*. They are called *even*, *odd*, or *vanishing* according to whether their space of global sections is even-dimensional, odd-dimensional, or at least 2-dimensional, respectively. We note that, on singular curves, there can be theta characteristics which are not invertible; see [Barth 1977; Beauville 1977]. In the following, we will only discuss invertible theta characteristics, so we drop the “invertible” from the notation.

Let $\Theta(R)$ be the set of isomorphism classes of theta characteristics on R and let $J(R)$ be the identity component of the Picard scheme of R , also called the generalized Jacobian of R .

Lemma 3.6. *The generalized Jacobian $J(R)$ of R is isomorphic to \mathbb{G}_a^3 .*

Proof. Since R is of arithmetic genus 3, $J(R)$ is a commutative group scheme of dimension 3. As R_{red} has only unbranched singularities, [Bosch et al. 1990, Propositions 5 and 9] shows that $J(R)$ is unipotent. Finally, we have a factorization of the absolute Frobenius $F : R \rightarrow V(A) \rightarrow R$. Note that $J(V(A))$ is trivial, even if $V(A)$ is nonreduced, since $H^1(V(A), \mathcal{O}_{V(A)}) = 0$. Since F^* is multiplication by p on $J(R)$, we obtain that $J(R)$ is p -torsion, and hence isomorphic to \mathbb{G}_a^3 . □

In particular, $J(R)(k)$ is an infinite 2-torsion group and it acts on $\Theta(R)$ via tensor products. It is easy to check that $\Theta(R)$ is a torsor under $J(R)(k)$ via this action. This already shows that the problem of finding (fake) bitangents using theta characteristics on R in characteristic 2 is much more subtle than it is in the other characteristics. Let us give an example that further illustrates this point.

Example 3.7. Assume that $V(A)$ is a smooth conic.

Consider $\pi : R \rightarrow V(A) \xrightarrow{\sim} \mathbb{P}^1$. We have $\pi^*\mathcal{O}_{\mathbb{P}^1}(2) = (f|_R)^*\mathcal{O}_{V(A)}(1) = (\omega_X)|_R$, so $\mathcal{L} := \pi^*\mathcal{O}_{\mathbb{P}^1}(1)$ is a theta characteristic on R . Moreover, we have $h^0(R, \pi^*\mathcal{O}_{\mathbb{P}^1}(1)) = 2$, so \mathcal{L} is a vanishing theta characteristic. In fact, this is the unique vanishing theta characteristic on R : Indeed, let \mathcal{L}' be another vanishing theta characteristic. Then, the Riemann–Roch formula yields

$$h^0(R, \mathcal{L} \otimes \mathcal{L}') - h^0(R, \omega_R \otimes \mathcal{L}^{-1} \otimes \mathcal{L}'^{-1}) = 2.$$

Since $h^0(R, \mathcal{L}) \geq 2$ and $h^0(R, \mathcal{L}') \geq 2$, we have $h^0(R, \mathcal{L} \otimes \mathcal{L}') \geq 3$, so $h^0(R, \omega_R \otimes \mathcal{L}^{-1} \otimes \mathcal{L}'^{-1}) \neq 0$. Since R is integral and $\omega_R \otimes \mathcal{L}^{-1} \otimes \mathcal{L}'^{-1}$ has degree 0, this implies that $\mathcal{L} \cong \mathcal{L}'$.

Next, let ℓ be any line in \mathbb{P}^2 such that $f^{-1}(\ell)$ meets R in two distinct smooth points. Then, $f^{-1}(\ell \cap V(A))_{\text{red}}$ defines an effective theta characteristic \mathcal{L} on R . By the previous paragraph, we have $h^0(R, \mathcal{L}) = 1$; hence all the infinitely many theta characteristics arising in this way are odd. It would be interesting to find an abstract characterization of the fake bitangent lines among the odd theta characteristics of R .

Nevertheless, we can find explicit equations of fake bitangent lines using the following result.

Lemma 3.8. *Let $C \rightarrow \mathbb{P}^1$ be an Artin–Schreier double cover given by an equation of the form*

$$w^2 + f(u, v)w + g(u, v) = 0,$$

where f and g are homogeneous polynomials of degree n and $2n$, respectively, and $f \neq 0$. Then, C is reducible if and only if there exists a homogeneous polynomial h of degree n with $g(u, v) = f(u, v)h(u, v) + h(u, v)^2$.

Proof. If there exists an h as in the assertion, then $w^2 + fw + g = (w + f + h)(w + h)$, so C is obviously reducible.

Conversely, assume that C is reducible. Then, C has exactly two irreducible components and these components are interchanged by the substitution $w \mapsto w + f$. In other words, we can write $w^2 + fw + g = h'(h' + f)$, where h' is a weighted homogeneous polynomial of degree n . This is only possible if h' is of the form $h' = w + h$ for some h homogeneous of degree n in the variables u and v . Then, $w^2 + fw + g = (w + h)(w + h + f) = w^2 + fw + h^2 + fh$; hence $g = fh + h^2$, as claimed. \square

Finally, for later use, we record some simple restrictions on the possible positions of fake bitangent lines with respect to the singularities of R .

Proposition 3.9. *Let ℓ be a fake bitangent line that passes through the image P of a singular point of an irreducible component of R_{red} . Then, $V(A)$ is smooth and ℓ is tangent to $V(A)$ at P .*

Proof. Write $f^{-1}(\ell) = L_1 + L_2$. Since R_{red} is singular at $f^{-1}(P)$, L_i and R have intersection multiplicity at least 2 in $f^{-1}(P)$. Since $R \in |-2K_X|$ and $L_1 + L_2 \in |-K_X|$, we have $(L_1 + L_2) \cdot R = 2K_X^2 = 4$; hence $L_1 + L_2$ and R meet only in $f^{-1}(P)$. Therefore, their images in \mathbb{P}^2 meet only in P . If $V(A)$ is smooth, this implies that ℓ is tangent to $V(A)$ in P . If $V(A)$ is the union of two lines, this implies that ℓ passes through their intersection. However, in this case, L_i and R have intersection multiplicity at least 3 in $f^{-1}(P)$, which is absurd. Finally, if $V(A)$ is a double line, then $R_{\text{red}} \in |-K_X|$ and $2 = K_X^2 = (L_1 + L_2) \cdot R_{\text{red}} \geq 4$, a contradiction. \square

Remark 3.10. We note there are del Pezzo surfaces for which fake bitangents satisfying the properties of Proposition 3.9 exist. See Proposition 5.10 for a classification in terms of the normal forms of Theorem 3.4.

3.4. Strange elliptic fibrations. To each del Pezzo surface X of degree 2 in characteristic 2 with branch locus $V(A)$ of the anticanonical map $f : X \rightarrow \mathbb{P}^2$, there is a naturally associated point P_X in \mathbb{P}^2 : if $V(A)$ is smooth, we let P_X be the strange point of $V(A)$, if $V(A)$ is the union of two lines, we let P_X be their intersection, and if $V(A)$ is a double line, we let P_X be the image of the singular point of $f^{-1}(V(A))_{\text{red}}$. We call P_X the *strange point* of X and note that the action of $\text{Aut}(X)$ fixes P_X .

The pencil \mathcal{P} of lines through P_X is $\text{Aut}(X)$ -invariant as well. Its preimage \mathcal{C} in X is an $\text{Aut}(X)$ -invariant pencil of curves of arithmetic genus 1 with two base points if $V(A)$ is smooth and with one base point of multiplicity 2 if $V(A)$ is singular. We let $\pi : Y \rightarrow X$ be the blow-up of the two (possibly infinitely near) base points of \mathcal{C} . Then, \mathcal{C} defines a relatively minimal genus-1 fibration $\phi : Y \rightarrow \mathbb{P}^1$. Since the map $X \rightarrow \mathbb{P}^2$ is separable and a general line in the pencil is not contained in $V(A)$, its preimage on Y is a smooth elliptic curve. Thus, the genus-1 fibration is an elliptic fibration. We call it the *strange elliptic fibration* associated to X .

By construction, the group $\text{Aut}(X)$ lifts to a subgroup of $\text{Aut}(Y)$ and we will use this in Proposition 4.2 to find restrictions on the possible structure of $\text{Aut}(X)$. To make the most of this connection, we will now describe the singular fibers of the elliptic fibration $\phi : Y \rightarrow \mathbb{P}^1$. We employ Kodaira’s notation: we say that a fiber isomorphic to an irreducible cuspidal cubic curve is of type II, a fiber that consists of two smooth rational curves intersecting nontransversally at one point is of type III, and a fiber that consists of three smooth rational curves intersecting at one point is of type IV.

We use the normal forms of Theorem 3.4, so that $A = x^2 + yz$, xy , or x^2 and $P_X = [1 : 0 : 0]$ in the first case and $P_X = [0 : 0 : 1]$ in the other two cases. In the first case, we let $\ell_{[t_0:t_1]}$ be the line $V(t_0y + t_1z)$ and in the other two cases, we let $\ell_{[t_0:t_1]}$ be the line $V(t_0x + t_1y)$. The fiber of ϕ corresponding to $\ell_{[t_0:t_1]}$ is denoted by $F_{[t_0:t_1]}$.

Proposition 3.11. *The generic fiber of the strange elliptic fibration associated to X is a supersingular elliptic curve. Its singular fibers are of type II, III, or IV and its Mordell–Weil group is torsion-free. Namely:*

(1) *If $A = x^2 + yz$, then the following hold:*

- *The fiber $F_{[t_0:t_1]}$ is smooth if and only if $t_0y + t_1z \nmid B_1$.*
- *The fiber $F_{[t_0:t_1]}$ is of type III if $\ell_{[t_0:t_1]}$ is a fake bitangent and of type II otherwise.*
- *The line $\ell_{[1:0]}$ is a fake bitangent if and only if $e = 0$.*
- *The line $\ell_{[0:1]}$ is a fake bitangent if and only if $a = 0$.*
- *The line $\ell_{[1:1]}$ is a fake bitangent if and only if $a + b + c + d + e = 0$.*

(2a) *If $A = xy$ and $B_1 = xz^3 + yz^3$, then the following hold:*

- *The fiber $F_{[t_0:t_1]}$ is smooth if and only if $[t_0 : t_1] \neq [1 : 0], [0 : 1], [1 : 1]$.*
- *$F_{[1:0]}$ and $F_{[0:1]}$ are of type II.*
- *$F_{[1:1]}$ is of type IV if $\ell_{[1:1]}$ is a fake bitangent and of type III otherwise.*
- *The curve $\ell_{[1:1]}$ is a fake bitangent if and only if $c = d^2 + e^2$.*

(2b) If $A = xy$ and $B_1 = xz^3 + y^3z$, then the following hold:

- The fiber $F_{[t_0:t_1]}$ is smooth if and only if $[t_0 : t_1] \neq [1 : 0], [0 : 1]$.
- The curve $F_{[0:1]}$ is of type II.
- The curve $F_{[1:0]}$ is of type III.

(3) If $A = x^2$, then the following hold:

- The fiber $F_{[t_0:t_1]}$ is smooth if and only if $[t_0 : t_1] \neq [1 : 0]$.
- The curve $F_{[1:0]}$ is of type III.

Proof. We study each case separately.

(1) In this case, $A = x^2 + yz$.

First, consider $\ell_{[1:t]} = V(y + tz)$. Plugging $y = tz$ into the equation of X , we obtain

$$w^2 + (x^2 + tz^2)w + xB_1(tz, z) + (at^4 + bt^3 + ct^2 + dt + e)z^4,$$

with $B_1(tz, z) \in \{\lambda t(t+1)z^3, t^2z^3, t^3z^3\}$. If $y + tz \nmid B_1$, then $B_1(tz, z) \neq 0$, so taking partials with respect to x and w shows that a singular point must satisfy $x = z = 0$, which is absurd. If $y + tz \mid B_1$, then $B_1(tz, z) = 0$ and $F_{[1:t]}$ is singular over $[t : t : 1]$. Similarly, one checks that $F_{[0:1]}$ is singular.

The equation

$$w^2 + (x^2 + tz^2)w + xB_1(tz, z) + (at^4 + bt^3 + ct^2 + d + e)z^4$$

shows that $F_{[1:t]}$ is a double cover of \mathbb{P}^1 branched over a single point. Hence, if $F_{[1:t]}$ is smooth, then it is supersingular, and if it is singular and irreducible, it is a cuspidal cubic.

Finally, consider the curve $F_{[1:0]}$ given by

$$w^2 + x^2w + ez^4.$$

By Lemma 3.8, it is clear that $F_{[1:0]}$ is reducible if and only if $e = 0$. The calculation for $F_{[1:1]}$ and $F_{[0:1]}$ is similar.

(2a) In this case, $A = xy$ and $B_1 = xz^3 + yz^3$.

First, consider $\ell_{[1:t]} = V(x + ty)$ with $t \neq 0, 1$. Plugging $x = ty$ into the equation of X , we obtain

$$w^2 + ty^2w + (t+1)yz^3 + B_0(ty, y, z)^2.$$

Then, taking partials shows that $F_{[1:t]}$ is smooth. Since it is a double cover of \mathbb{P}^1 branched over a single point, it is supersingular.

Next, consider $F_{[1:1]}$, whose image in X is given by

$$w^2 + y^2w + ((a+b)y^2 + (d+e)yz + cz^2)^2.$$

This curve is singular over $[0 : 0 : 1]$, so $F_{[1:1]}$ has one irreducible component contracted by $Y \rightarrow X$. By Lemma 3.8, the image of $F_{[1:1]}$ in X is reducible if and only if $c = d^2 + e^2$.

Finally, the curves $F_{[1:0]}$ and $F_{[0:1]}$ are isomorphic to their images in X and these images are irreducible and cuspidal by Theorem 3.4.

(2b) In this case, $A = xy$ and $B_1 = xz^3 + y^3z$.

First, consider $\ell_{[1:t]} = V(x + ty)$ with $t \neq 0$. Plugging $x = ty$ into the equation of X , we obtain

$$w^2 + ty^2w + tyz^3 + y^3z + B_0(ty, y, z)^2.$$

As in the previous cases, taking partials shows that $F_{[1:t]}$ is smooth and supersingular.

The curve $F_{[0:1]}$ is isomorphic to its image in X , since $f^{-1}(\ell_{[0:1]})$ is smooth over the point $[0 : 0 : 1]$. Hence, $F_{[0:1]}$ is of type II. On the other hand, the curve $F_{[1:0]}$ is of type III, since its image in X has multiplicity 2 over $[0 : 0 : 1]$.

(3) In this case $A = x^2$.

First, consider $\ell_{[t:1]} = V(tx + y)$. Plugging $y = tx$ into the equation of X , we obtain

$$w^2 + x^2w + xz^3 + (bt^2 + at)x^2z^2 + cz^4.$$

Then, taking partials shows that $F_{[t:1]}$ is smooth. Since it is a double cover of \mathbb{P}^1 branched over a single point, it is supersingular.

The curve $F_{[1:0]}$ is of type III, by the same argument as in the previous case.

The Mordell–Weil group of the fibration is torsion-free by [Oguiso and Shioda 1991, Main Theorem], since the lattice spanned by fiber components is of rank at most 4 in each case. □

Remark 3.12. The classification of singular fibers of rational elliptic surfaces with a section in characteristic 2 can be found in [Lang 2000]. Lang shows that in the cases where the general fiber is a supersingular elliptic curve, the number of singular fibers is at most 3, which agrees with what we observed in the case of strange elliptic fibrations. Proposition 3.11 shows that the singular fibers that occur on strange genus-1 fibrations are of type $9A, 9B, 10A, 10B, 10C$ or 11 in Lang’s terminology.

4. Automorphism groups of del Pezzo surfaces of degree 2

4.1. Preliminaries. Recall once more from Section 3.1 that a del Pezzo surface X of degree 2 is a surface of degree 4 in $\mathbb{P}(1, 1, 1, 2)$ given by an equation of the form

$$w^2 + A(x, y, z)w + B(x, y, z) = 0.$$

Since this is the anticanonical model of X and ω_X^{-n} admits a natural $\text{Aut}(X)$ -linearization for all n , we obtain that $\text{Aut}(X)$ is isomorphic to the subgroup of $\text{Aut}(\mathbb{P}(1, 1, 1, 2))$ of automorphisms that preserve X .

The structure of the group $\text{Aut}(\mathbb{P}(1, 1, 1, 2))$ is well known. The vector space $\mathbb{k}[x, y, z]_2$ of quadratic forms is a normal subgroup of $\text{Aut}(\mathbb{P}(1, 1, 1, 2))$ that acts via $(x, y, z, w) \mapsto (x, y, z, w + Q)$. The quotient by this subgroup is the group of transformations that change (x, y, z) linearly and multiply w by a scalar. Since the transformation $(x, y, z, w) \mapsto (\lambda x, \lambda y, \lambda z, \lambda^2 w)$ is the identity, this quotient is

isomorphic to $\mathrm{GL}_3(\mathbb{k})/\mu_2(\mathbb{k})$. Since we are in characteristic 2, the subgroup $\mu_2(\mathbb{k})$ is trivial. This gives an isomorphism

$$\mathrm{Aut}(\mathbb{P}(1, 1, 1, 2)) \cong \mathbb{k}[x, y, z]_2 : \mathrm{GL}_3(\mathbb{k}).$$

We denote elements of this group by $(Q, g) \in \mathbb{k}[x, y, z]_2 \times \mathrm{GL}_3(\mathbb{k})$ where the semidirect product structure is

$$(Q, g) \circ (Q', g') = (g^*(Q') + Q, gg').$$

Using this description of $\mathrm{Aut}(\mathbb{P}(1, 1, 1, 2))$, it is straightforward to calculate the subgroup of automorphisms preserving X . We obtain

$$\mathrm{Aut}(X) \cong \{(Q, g) : g^*(A) = A, g^*(B) = B + AQ + Q^2\}.$$

The kernel of the homomorphism

$$\mathrm{Aut}(X) \rightarrow \mathrm{GL}_3(\mathbb{k}), \quad (Q, g) \mapsto g$$

is generated by the Geiser involution γ . We let $G(X)$ be the image of $\mathrm{Aut}(X)$ in $\mathrm{GL}_3(\mathbb{k})$.

Lemma 4.1. *The homomorphism $G(X) \rightarrow \mathrm{GL}_3(\mathbb{k}) \rightarrow \mathrm{PGL}_3(\mathbb{k})$ is injective.*

Proof. Let $g \in G(X)$ be in the kernel of this homomorphism. Then, $g = \lambda I_3$ for some $\lambda \in \mathbb{k}^\times$. On the other hand, by definition of $G(X)$, we have $g^*(A) = A$. Since A has degree 2, this implies $\lambda^2 = 1$. Hence, $\lambda = 1$. \square

We recall from [Dolgachev and Martin 2024, Section 1] that a choice of a blow-up $X \rightarrow \mathbb{P}^2$ of seven points defines an injective homomorphism

$$\rho : \mathrm{Aut}(X) \rightarrow W(E_7). \quad (3)$$

The image of the Geiser involution is equal to $-\mathrm{id}_{E_7}$. It is known that $W(E_7) = \langle -\mathrm{id}_{E_7} \rangle \times W(E_7)^+$, where $W(E_7)^+ \subseteq W(E_7)$ is the kernel of the determinant map.

In particular, to determine $\mathrm{Aut}(X)$, it suffices to determine $G(X)$ and both groups are isomorphic to subgroups of $W(E_7)$ via ρ . This puts severe restrictions on the possible structure of $G(X)$. Finally, we can use the strange genus-1 fibrations of the previous section to get information on $G(X)$.

Proposition 4.2. *Let $\phi : Y \rightarrow \mathbb{P}^1$ be the strange elliptic fibration associated to X . Choose an exceptional curve E of $Y \rightarrow X$ as the zero section of ϕ and let C be the second exceptional curve. Then, there is a homomorphism $\varphi : \mathrm{Aut}(X) \rightarrow \mathrm{Aut}(Y)$ that satisfies the following properties:*

- (1) φ is injective.
- (2) $\varphi(\gamma)$ preserves every fiber of ϕ .
- (3) If $V(A)$ is smooth, then C is a section of ϕ . We have $\varphi(\gamma) = t_C \circ \iota$, where ι is the negation automorphism and t_C is translation by C .
- (4) If $V(A)$ is singular, then C is a component of a reducible fiber of ϕ . We have $\varphi(\gamma) = \iota$ and φ factors through the stabilizer of the pair (E, C) .

Proof. The surface Y is obtained by blowing up X at two points that are uniquely determined by $V(A)$, and hence is stable under the action of $\text{Aut}(X)$. This shows existence and injectivity of the homomorphism φ . The fibration ϕ is induced by the pencil of lines in \mathbb{P}^2 through the strange point of X . Since γ preserves these lines, it preserves the fibers of ϕ .

If $V(A)$ is smooth, then E and C are interchanged by $\varphi(\gamma)$. The automorphism $t_{-C} \circ \varphi(\gamma) \circ \iota$ maps E to E and $-C$ to $-C$. It is well known that every fixed point of a nontrivial automorphism of an elliptic curve is a torsion point. On the other hand, by Proposition 3.11, ϕ has no torsion sections, so $t_{-C} \circ \varphi(\gamma) \circ \iota = \text{id}$, which yields claim (3).

If $V(A)$ is singular, then C is a (-2) -curve which meets E ; hence it is the identity component of a reducible fiber of ϕ . Since $\varphi(\gamma)$ is an involution that preserves E , we have $\varphi(\gamma) = \iota$ and we obtain claim (4). □

4.2. Classification.

Theorem 4.3. *Every del Pezzo surface of degree 2 in characteristic 2 such that, in the decomposition $\text{Aut}(X) \cong 2 \times G(X)$, the group $G(X)$ is nontrivial is isomorphic to a surface of degree 4 in $\mathbb{P}(1, 1, 1, 2)$ given by an equation of the form*

$$w^2 + Aw + B,$$

where $(A, B, G(X))$ is one of the forms shown in Table 3. The parameters satisfy the following conditions:

- (1ai) $\lambda \neq 0, \lambda^2 + c \neq 0, b^2 + a \neq 0, (b, c) \neq (\lambda, a)$.
- (1a ii) $\lambda \neq 0, \lambda^2 + a \neq 0$.
- (1ci) $e \neq 0$.
- (2ai) $a \neq 0, (c, d) \neq (0, 0)$.
- (2a ii) $a \neq 0, b \neq 0, a \neq b$.
- (2a iii) $a \neq 0$.
- (3i) $c \neq 0$.
- (3ii) None.

Proof. We use the normal forms of Theorem 3.4 and the description of $\text{Aut}(X)$ and $G(X)$ given in the beginning of the current section. We go through the cases of Theorem 4.3.

(1a) Here, X is given by an equation of the form

$$w^2 + (x^2 + yz)w + \lambda xyz(y + z) + B_0,$$

with

$$B_0 = ay^4 + by^3z + cy^2z^2 + dyz^3 + ez^4$$

and the cusps lie over $[0 : 1 : 0]$, $[0 : 0 : 1]$, and $[1 : 1 : 1]$. Let $(Q, g) \in \text{Aut}(X)$ be an automorphism of X . Then, g preserves the three points lying under the cusps. Moreover, if g fixes the three cusps, then it fixes

name	A	B	B_1	B_0	$G(X)$	# of parameters
(1ai)	x^2+yz	$x B_1+B_0$	$\lambda yz(y+z)$	$ay^4+by^3z+cy^2z^2+byz^3+az^4$	2	4
(1aaii)	x^2+yz	$x B_1+B_0$	$\lambda yz(y+z)$	$ay^4+\lambda y^3z+ay^2z^2+\lambda yz^3+az^4$	S_3	2
(1ci)	x^2+yz	$x B_1+B_0$	y^3	$by^3z+cy^2z^2+ez^4$	2^3	3
(2ai)	xy	$B_1+B_0^2$	xz^3+yz^3	$ax^2+ay^2+cz^2+dxz+dyz$	2	3
(2aaii)	xy	$B_1+B_0^2$	xz^3+yz^3	ax^2+by^2	3	2
(2aiiii)	xy	$B_1+B_0^2$	xz^3+yz^3	ax^2+ay^2	6	1
(3i)	x^2	$x B_1+B_0$	z^3	y^3z+cz^4	3	1
(3ii)	x^2	$x B_1+B_0$	z^3	y^3z	9	0

Table 3. Forms of $(A, B, G(X))$ in Theorem 4.3.

$V(A)$ pointwise; hence g is trivial in $\text{PGL}_3(\mathbb{k})$, so, by Lemma 4.1, g is the identity and (Q, g) coincides with the Geiser involution. Hence, $G(X)$ acts faithfully on $\{[0 : 1 : 0], [0 : 0 : 1], [1 : 1 : 1]\}$.

Note that $G(X)$ contains an involution if and only if X admits an equation where this involution is given by $y \leftrightarrow z$. This involution is in $G(X)$ if and only if there exists a quadratic form Q such that

$$Q^2 + (x^2 + yz)Q = g^* B_0 + B_0. \tag{4}$$

Since $Q^2 + (x^2 + yz)Q$ contains a nonzero monomial divisible by x^2 as soon as it is nonzero and $g^* B_0 + B_0$ does not contain such a monomial, we must have $Q \in \{0, x^2 + yz\}$ and (4) holds if and only if $a = e$ and $b = d$, as claimed.

Next, note that $G(X)$ contains an automorphism g of order 3 if and only if g is given by $x \mapsto x + z, y \mapsto z, z \mapsto y + z$ and there exists a quadratic form Q such that

$$Q^2 + (x^2 + yz)Q = \lambda yz^2(y + z) + g^* B_0 + B_0. \tag{5}$$

By the same argument as in the previous paragraph, we have $Q \in \{0, x^2 + yz\}$ and (5) holds if and only if $a = c = e$ and $b = d = \lambda$. In particular, note that these conditions imply the conditions of the previous paragraph in this case; hence $G(X) = S_3$.

(1b) In this case, $V(A)$ is smooth and R has two nonisomorphic singularities. Then, $g \in G(X)$ must fix the images of them on $V(A)$. Since an automorphism of order 2 of \mathbb{P}^1 has only one fixed point, we may assume that the order of g is odd. By Proposition 3.9, the line ℓ through the images of the singularities is not a fake bitangent. Its preimage E in X is an integral curve of arithmetic genus 1 and the Geiser involution has two fixed points on E . Hence, either E is smooth and ordinary, or nodal. In both cases, there is no nontrivial automorphism of odd order that commutes with the involution; hence g fixes ℓ pointwise. Since g also fixes the strange point P on $V(A)$ and the projection from P is inseparable, g fixes $V(A)$ pointwise; hence g is the identity. We conclude that $G(X) = \{1\}$.

(1c) Here, X is given by an equation of the form

$$w^2 + (x^2 + yz)w + xy^3 + B_0,$$

with

$$B_0 = by^3z + cy^2z^2 + dyz^3 + ez^4.$$

The singularity of R lies over $[0 : 0 : 1]$. An element $g \in G(X)$ of odd order has at least two fixed points on $V(A)$ and then the same argument as in the previous case shows that g is the identity. Therefore, $G(X)$ is a 2-group that acts on $V(A) \cong \mathbb{P}^1$ with a fixed point. In particular, $G(X)$ is isomorphic to a subgroup of $\mathbb{G}_a(\mathbb{k})$, and hence is isomorphic to 2^n for some $n \geq 0$.

We may assume that g acts as $x \mapsto x + \alpha y$, $y \mapsto y$, $z \mapsto z + \alpha^2 y$. Then g lifts to $\text{Aut}(X)$ if and only if there exists a quadratic form Q such that

$$(x^2 + yz)Q + Q^2 = \alpha y^4 + g^*(B_0) + B_0.$$

Since the right-hand side does not contain a monomial divisible by x^2 , we get, as in the previous cases, $Q = x^2 + yz$ or $Q = 0$. Comparing coefficients yields the system of equations

$$\begin{aligned} d\alpha^2 &= 0, \\ d\alpha^4 &= 0, \\ e\alpha^8 + d\alpha^6 + c\alpha^4 + b\alpha^2 + \alpha &= 0. \end{aligned}$$

So, if $d \neq 0$, then $\alpha = 0$ and $G(X)$ is trivial. If $d = 0$, there are eight possibilities for α , one for each root of $e\alpha^8 + c\alpha^4 + b\alpha^2 + \alpha$. All the roots are distinct since the derivative of this polynomial is 1. Here, we also use that $e \neq 0$ by Theorem 3.4. Thus $G(X) \cong 2^3$.

(2a) Here, X is given by an equation of the form

$$w^2 + xyw + xz^3 + yz^3 + B_0^2,$$

with

$$B_0 = ax^2 + by^2 + cz^2 + dxz + eyz.$$

The singularities of the irreducible components of R lie over $[1 : 0 : 0]$ and $[0 : 1 : 0]$. Let $(Q, g) \in \text{Aut}(X)$. Then, g preserves these two points and the intersection of $V(x)$ and $V(y)$. Moreover, by Propositions 4.2 and 3.11, g preserves the line $V(x + y)$.

Assume that g has odd order. Then, g preserves the three lines $V(x)$, $V(y)$, and $V(x + y)$; hence it is of the form $(x, y, z) \mapsto (x, y, \alpha z)$. The quadratic form Q satisfies

$$Q^2 + xyQ = g^*(B_1 + B_0^2) + B_1 + B_0^2.$$

The right-hand side does not contain monomials divisible by xy ; hence $Q \in \{0, xy\}$. Now, $g^*B_1 + B_1 = 0$ implies that $\alpha^3 = 1$, and if $\alpha \neq 1$, then $g^*B_0^2 + B_0^2 = 0$ holds if and only if $c = d = e = 0$.

Assume that g has order a power of 2. If g does not swap the points $[1 : 0 : 0]$ and $[0 : 1 : 0]$, then it acts diagonally; hence it is the identity. Therefore, we may assume that g swaps these two points and $g^2 = \text{id}$. Hence, g acts as $x \leftrightarrow y$. The quadratic form Q satisfies

$$Q^2 + xyQ = g^*(B_1 + B_0^2) + B_1 + B_0^2,$$

and hence $a = b$ and $d = e$.

(2b) Here, X is given by an equation of the form

$$w^2 + xyw + xz^3 + y^3z + B_0^2,$$

with

$$B_0 = ax^2 + cz^2 + dxz + eyz$$

and the singularities of the irreducible components of R map to $[1 : 0 : 0]$ and $[0 : 0 : 1]$. Let $(Q, g) \in \text{Aut}(X)$.

If g has odd order, then there is a g -invariant line ℓ through $[1 : 0 : 0]$ and we may assume that $\ell \not\subseteq V(A)$. By the same argument as in case (1b), ℓ is fixed pointwise. Then, every line through $[0 : 0 : 1]$ is g -invariant. Since $g^*A = A$, this means that g acts as $(x, y, z) \mapsto (x, y, \alpha z)$. An automorphism of this form satisfies $g^*B_1 = B_1$ if and only if $\alpha = 1$, so g is trivial.

If g has order a power of 2, then by Proposition 3.11, $\varphi((Q, g)) \in \text{Aut}(Y)$ preserves the two singular fibers of $\phi : Y \rightarrow \mathbb{P}^1$; hence $\varphi((Q, g))$ acts trivially on the base of ϕ . The 2-Sylow subgroup of automorphisms of the geometric generic fiber of ϕ is the quaternion group Q_8 and $\varphi((Q, g))$ commutes with the unique involution $\varphi(\gamma)$ in Q_8 . This implies that $(Q, g) \in \langle \gamma \rangle$, so g is trivial.

(3) If $V(A)$ is a double line, then X is given by an equation of the form

$$w^2 + x^2w + xB_1 + B_0,$$

with $B_1 = z^3 + ayz^2$ and $B_0 = y^3z + by^2z^2 + cz^4$. The singularity of R_{red} lies over $[0 : 0 : 1]$. Let $(Q, g) \in \text{Aut}(X)$. Then, g is of the form

$$(x, y, z) \mapsto (x, \alpha x + \beta y, \gamma x + \delta y + \epsilon z),$$

with $\beta, \epsilon \neq 0$ and Q satisfies the equation

$$Q^2 + x^2Q = x(g^*B_1 + B_1) + g^*B_0 + B_0. \quad (6)$$

The monomials y^3z, xz^3, xyz^2, xy^2z and xy^3 do not appear on the left-hand side; hence their coefficients on the right-hand side must be zero. This yields the conditions

$$\begin{aligned} \epsilon &= \beta^{-3}, & \delta &= a(\beta + \beta^6), \\ \beta^9 &= 1, & \alpha &= a^2(1 + \beta), \\ & & \gamma &= a^3(1 + \beta^2). \end{aligned}$$

So, the order of g is equal to the order of β in k^\times ; hence it is equal to 1, 3 or 9. Now, we calculate that if $\beta^3 = 1$, then g^3 acts as

$$(x, y, z) \mapsto (x, y, a^3(\beta + \beta^2)x + z).$$

Hence, if $a \neq 0$, then g is the identity.

So, assume that $a = 0$, so that, in particular, $\alpha = \delta = \gamma = 0$. Equation (6) becomes

$$Q^2 + x^2Q = (\epsilon^3 + 1)xz^3 + (\beta^3\epsilon + 1)y^3z + b(\beta^2\epsilon^2 + 1)y^2z^2 + c(\epsilon^4 + 1)z^4.$$

On the left-hand side, the coefficients of z^4 and y^2z^2 are the squares of the coefficients of x^2z^2 and x^2yz , respectively. Since the latter monomials do not appear on the right-hand side, the coefficients of the former monomials must vanish. Therefore, we get the four conditions

$$\begin{aligned} \epsilon^3 + 1 = 0, & \quad b(\beta^2\epsilon^2 + 1) = 0, \\ \beta^3\epsilon + 1 = 0, & \quad c(\epsilon^4 + 1) = 0, \end{aligned}$$

Hence, if $b \neq 0$, then $\beta = \epsilon = 1$, so $G(X)$ is trivial. If $b = 0$ and $c \neq 0$, then $\epsilon = 1$ and $\beta^3 = 1$, and so $G(X) \cong C_3$. If $b = c = 0$, then $\epsilon = \beta^{-3}$ and $\beta^9 = 1$; hence $G(X) \cong C_9$. \square

Remark 4.4. With our choice of normal form in Theorem 4.3, the map $g \mapsto (0, g)$ defines an explicit section of the surjection $\text{Aut}(X) \rightarrow G(X)$ in every case.

Remark 4.5. The group 2^4 that appears in Theorem 4.3 occurs as a group of automorphisms of a del Pezzo surface of degree 4 in all characteristics [Dolgachev and Duncan 2019]. In characteristic 0, there is a unique conjugacy class of subgroups isomorphic to 2^4 in the Cremona group. One can prove, using the theory of birational links, that in characteristic 2, the two subgroups of $\text{Cr}_{\mathbb{k}}(2)$ are not conjugate.

Remark 4.6. The fact that 2 and 3 are the only primes that divide the order of $\text{Aut}(X)$ can be proven without the classification. It is known that 2, 3, 5, and 7 are the only primes that divide the order of $W(E_7)$. To exclude the primes 5 and 7, one can use the Lefschetz fixed-point formula and the known traces of elements of $W(E_7)$ acting on the root lattice of type E_7 to get a contradiction with the possible structure of the set of fixed points of an element of the group $G(X)$.

4.3. Conjugacy classes and comparison with the classification in characteristic 0. In this section, we determine the conjugacy classes in $W(E_7)$ of the elements of the groups that occur in Theorem 4.3 and, whenever possible, compare the surfaces in Theorem 4.3 with their counterparts in characteristic 0 (see [Dolgachev 2012, Table 8.9]). To do this, we use the following result.

Lemma 4.7. *Let X be a del Pezzo surface of degree 2 in characteristic 2. Let X' be a geometric generic fiber of a lift of X to characteristic 0 and let $\text{sp} : \text{Aut}(X') \rightarrow \text{Aut}(X)$ be the specialization map. Then, sp is injective and preserves conjugacy classes.*

Proof. Let $\mathcal{X} \rightarrow S$ be a lift of X with geometric generic fiber X' . The map sp sends an automorphism $g \in \text{Aut}(X')$ to the special fiber of the closure of g considered as a point of the relative automorphism scheme $\text{Aut}_{\mathcal{X}/S}$. To see that this is well-defined, we have to explain why $\text{Aut}_{\mathcal{X}/S}$ is proper over S .

By passing to the anticanonical model of \mathcal{X} in $\mathbb{P}_S(1, 1, 1, 2)$, the scheme $\text{Aut}_{\mathcal{X}/S}$ is identified with the stabilizer of \mathcal{X} under the action of $\text{Aut}_{\mathbb{P}_S(1,1,1,2)}$ on the space $\mathcal{H}_{dP2,S}$ of smooth quartic hypersurfaces in $\mathbb{P}_S(1, 1, 1, 2)$. To check that this stabilizer is proper, it suffices to show that the shear map

$$\begin{aligned} (\mathbb{G}_a^3 : \text{GL}_3) \times \mathcal{H}_{dP2} &\rightarrow \mathcal{H}_{dP2} \times \mathcal{H}_{dP2}, \\ (Q, g, w^2 + Aw + B) &\mapsto (w^2 + Aw + B, w^2 + (g^*A + 2Q)w + g^*B + g^*AQ + Q^2), \end{aligned}$$

is a proper morphism of schemes over $\text{Spec } \mathbb{Z}$. For this, it suffices to check that the individual shear maps for the \mathbb{G}_a^3 -action and the GL_3 -action are proper. We check this using the valuative criterion. So, let R be an arbitrary discrete valuation ring with field of fractions K and let $F = w^2 + Aw + B$ and $F' = w^2 + A'w + B'$ be equations of smooth quartics in $\mathbb{P}_R(1, 1, 1, 2)$ with $A, A' \in R[x, y, z]_2$ and $B, B' \in R[x, y, z]_4$.

Given $Q \in K[x, y, z]_2$ sending F to F' , we have the two conditions

$$\begin{aligned} A + 2Q - A' &= 0, \\ B + AQ + Q^2 - B' &= 0. \end{aligned}$$

Comparing the valuations of the coefficients of Q^2 and $AQ + B - B'$ in the second equation shows that $Q \in R[x, y, z]_2$. This proves the valuative criterion of properness for the shear map of the \mathbb{G}_a^3 -action on \mathcal{H}_{dP2} .

Given $g \in \text{GL}_3(K)$ sending F to F' , we have the two conditions

$$g^*A = A', \quad g^*B = B'.$$

Replacing g by its Smith normal form, we may assume that g acts as

$$x \mapsto \pi^{e_x}x, \quad y \mapsto \pi^{e_y}y, \quad z \mapsto \pi^{e_z}z,$$

where π is a uniformizer of R . Thus, if a monomial $x^i y^j z^k$ appears in A or B with unit coefficient, then $ie_x + je_y + ke_z = 0$. We leave it to the reader to check that, because of the smoothness of F and F' modulo π , there are enough such monomials to check that $e_x = e_y = e_z$, that is, that $g \in \text{GL}_3(R)$. This is the valuative criterion of properness for the shear map of the GL_3 -action on \mathcal{H}_{dP2} .

Since $H^0(X, T_X) = 0$, the scheme $\text{Aut}_{X/S}$ is discrete; hence the specialization map is injective. As $H^1(X, \mathcal{O}_X) = H^2(X, \mathcal{O}_X)$, the relative Picard scheme $\text{Pic}_{X/S}$ is constant over S . Now, specialization of line bundles is sp-equivariant and compatible with the intersection pairing; hence sp preserves conjugacy classes. \square

Remark 4.8. It would be interesting to determine all integers a_1, \dots, a_n and d such that the action of $\text{Aut}_{\mathbb{P}(a_1, \dots, a_n)}$ on the space \mathcal{H}_d of smooth hypersurfaces of degree d in $\mathbb{P}(a_1, \dots, a_n)$ is proper. This would be a generalization of [Katz and Sarnak 1999, Proposition 11.8.2] to the weighted case.

By Theorem 4.3, we have $|\text{Aut}(X)| \leq 18$, so types I, \dots , V and of [Dolgachev 2012, Table 8.9] do not have a reduction modulo 2 which is a del Pezzo surface. Similarly, type VII of that table has no analog in characteristic 2.

The surface of type (3ii) of Theorem 4.3 is a reduction modulo 2 of the surface of type VI in [Dolgachev 2012, Table 8.9]. Hence, we call this surface type VI. Since the conjugacy classes of elements of $\text{Aut}(X)$ are the same as the ones of the lift, the entry for type VI in Table 4 is the same as the one in [Dolgachev and Martin 2024, Table 7].

The equations of the surfaces of type (2aiii) of Theorem 4.3 define smooth surfaces in characteristic 0 and the automorphisms $x \leftrightarrow y$ and $z \mapsto \zeta_3 z$ make sense in characteristic 0. Hence, these surfaces lift to characteristic 0 as del Pezzo surfaces with an action of 2×6 . As explained above, del Pezzo surfaces of degree 2 with an automorphism group of order bigger than 18 do not have a smooth reduction modulo 2; hence these lifts are of type VIII [Dolgachev 2012, Table 8.9], so we also call the surfaces of type (2aiii) type VIII. As in the previous case, the conjugacy classes are the same as in [Dolgachev and Martin 2024, Table 7].

As for the surfaces of type (1aii), we rewrite their equations using the substitution $x \mapsto x + y + z$ as

$$w^2 + (x^2 + y^2 + z^2 - yz)w + \lambda xyz(z - y) + a(y^2 + z^2 - yz)^2.$$

This equation defines a lift of X to characteristic 0 and the $\text{Aut}(X)$ -action lifts as well, since it is generated by the Geiser involution $\gamma : w \mapsto -w$, the involution $y \leftrightarrow z$ and the automorphism g of order 3 given by

$$x \mapsto -x, \quad y \mapsto z, \quad z \mapsto z - y.$$

Hence, all surfaces of type (1aii) are reductions modulo 2 of surfaces of type IX in [Dolgachev 2012, Table 8.9]. In particular, we can read off the conjugacy classes of elements of $\text{Aut}(X)$ from [Dolgachev and Martin 2024, Table 7].

The surfaces of type (1ci) are the characteristic-2 analogs of type X from [Dolgachev 2012, Table 8.9]. We claim that every involution on a surface X of type (1ci) which is different from the Geiser involution is of conjugacy class $3A_1/4A_1$. It suffices to check this for the surface given by

$$w^2 + (x^2 + yz)w + xy^3 + z^4,$$

where $G(X)$ acts as $g_\alpha : x \mapsto x + \alpha y, z \mapsto \alpha^2 x + z$, with $\alpha^8 = \alpha$. After using the substitution $z \mapsto \alpha x + y + z, y \mapsto \alpha^6 x + \alpha^6 y$, the equation of X becomes

$$w^2 + (x^2 + xy + y^2 + \alpha^6(y + x)z + \alpha^4(x^2 + y^2))w + \alpha^4(x^3 y + x^2 y^2 + xy^3) + \alpha^3(x^4 + y^4 + z^4)$$

and the involution g_α acts as $x \leftrightarrow y$. Then, the above equation makes sense in characteristic 0 and defines a lift of X together with the involution g_α . In particular, by [Dolgachev and Martin 2024, Table 7], the conjugacy class of g_α is $3A_1$ or $4A_1$.

The equations of types (2aii) and (3i) make sense in characteristic 0, where they define a lift of the surface together with the C_3 -action. These lifts must be of type XI from [Dolgachev 2012, Table 8.9].

Similarly, the equations of types (1ai) and (2ai) define lifts to characteristic 0 together with the C_2 -action. Hence, these lifts are of type XII from [Dolgachev 2012, Table 8.9].

We summarize the classification of automorphism groups of del Pezzo surfaces of degree 2 in Table 4. In the first column, we give the name of the corresponding family, both in the notation of Theorem 4.3 and in the notation of [Dolgachev 2012, Table 8.9]. The second and third columns give the group $\text{Aut}(X)$ and its size. In the remaining columns, we list the number of elements of a given Carter conjugacy class in $\text{Aut}(X)$.

name	Aut(X)	order	id	3A ₁	4A ₁	7A ₁	2A ₂	3A ₂	2A ₃	2A ₃ +A ₁	A ₅ +A ₂	A ₆	D ₄ (a ₁)	D ₄ (a ₁)+A ₁
I–V	(do not exist)													
VI/(3ii)	18	18	1			1		2						
VII	(does not exist)													
VIII/(2aiii)	2×6	12	1	1	1	1		2					2	
IX/(1aii)	2×S ₃	12	1	3	3	1	2							
X/(1ci)	2 ⁴	16	1	7	7	1								
XI/(2aii), (3i)	6	6	1			1		2						
XII/(1ai), (2ai)	2 ²	4	1	1	1	1								
XIII	2	2	1			1								

name	Aut(X)	order	D ₅	D ₅ +A ₁	D ₆ (a ₂)+A ₁	E ₆	E ₆ (a ₁)	E ₆ (a ₂)	E ₇	E ₇ (a ₁)	E ₇ (a ₂)	E ₇ (a ₄)
I–V	(do not exist)											
VI/(3ii)	18	18						6		6		2
VII	(does not exist)											
VIII/(2aiii)	2×6	12						2				2
IX/(1aii)	2×S ₃	12			2							
X/(1ci)	2 ⁴	16										
XI/(2aii), (3i)	6	6										2
XII/(1ai), (2ai)	2 ²	4										
XIII	2	2										

Table 4. Automorphism groups of del Pezzo surfaces of degree 2.

5. Del Pezzo surfaces of degree 1

5.1. The antibicanonical map. As in the case of degree 2, we start by describing the geometry of del Pezzo surfaces of degree $d = 1$ and we refer to [Demazure 1980] for characteristic-free facts on del Pezzo surfaces. Recall that the antibicanonical system $|-2K_X|$ defines a finite morphism $f : X \rightarrow Q$ onto a quadratic cone $Q \subseteq \mathbb{P}^3$. As in degree 2, it turns out that this map is always separable, even in characteristic 2.

Proposition 5.1. *The antibicanonical linear system $|-2K_X|$ defines a finite separable morphism $f : X \rightarrow Q$ of degree 2.*

Proof. If f is not separable, then $p = 2$ and f is purely inseparable. But then f is a homeomorphism in the étale topology. This is impossible, since $H_{\text{ét}}^2(X, \mathbb{Z}_\ell)$ has rank 9 (because X is the blow-up of eight points in the plane), while $H_{\text{ét}}^2(Q, \mathbb{Z}_\ell)$ has rank 1. □

Let

$$R(X, -K_X) = \bigoplus_{n=0}^{\infty} H^0(X, \mathcal{O}_X(-nK_X))$$

be the graded anticanonical ring of X . By the Riemann–Roch theorem, we have

$$\dim_{\mathbb{k}} R(X, -K_X)_1 = 2,$$

$$\dim_{\mathbb{k}} R(X, -K_X)_2 = 4,$$

$$\dim_{\mathbb{k}} R(X, -K_X)_3 = 7.$$

Thus, we can choose u, v from $R(X, -K_X)_1$, $x \in R(X, -K_X)_2 \setminus S^2(R(X, -K_X)_1)$, and $y \in R(X, -K_X)_3 \setminus S^3(R(X, -K_X)_1) + R(X, -K_X)_1 \otimes R(X, -K_X)_2$ and obtain the following relation between the generators:

$$y^2 + y(a_1x + a_3) + x^3 + a_2x^2 + a_4x + a_6 = 0, \tag{7}$$

where a_k denotes a binary form of degree k in u and v . In particular, via (7), we can view X as a surface of degree 6 in the weighted projective space $\mathbb{P}(1, 1, 2, 3)$, the anticanonical map is the projection of this surface onto the u -, v -coordinates, and the antibicanonical map is the projection onto the u^2 -, uv -, v^2 -, x -coordinates.

If $p \neq 2$, we can replace y with $y + \frac{1}{2}(a_1x + a_3)$ to assume that $a_1 = a_3 = 0$. The surface X is a double cover of a quadratic cone $Q \cong \mathbb{P}(1, 1, 2)$. The branch curve $B = V(x^3 + a_2x^2 + a_4x + a_6)$ is a curve of degree 6 not passing through the vertex of Q . It is a smooth curve of genus 4 with a vanishing theta characteristic g_3^1 defined by the ruling of Q . If we blow up the vertex of Q , we obtain a surface isomorphic to the rational minimal ruled surface F_2 . The preimage of the curve B is a curve in the linear system $|6\mathfrak{f} + 3\mathfrak{e}|$, where \mathfrak{f} and \mathfrak{e} are the standard generators of $\text{Pic}(F_2)$, with $\mathfrak{f}^2 = 0$ and $\mathfrak{e}^2 = -2$. The curve B is its canonical model in \mathbb{P}^3 .

In our case, when the characteristic $p = 2$, the analog of B is the curve $V(a_1x + a_3)$ in Q . In particular, Proposition 5.1 tells us that $a_1x + a_3 \neq 0$ and there is no way of removing these terms. Moreover, the curve B always passes through the vertex of Q and its strict transform on F_2 is in $|3\mathfrak{f}|$ if $a_1 = 0$ and in $|3\mathfrak{f} + \mathfrak{e}|$ if $a_1 \neq 0$. The analog of the involution $y \mapsto -y$, classically called the Bertini involution, is the involution β defined by replacing y with $y + a_1x + a_3$. As in the classical case, we call this β *Bertini involution*.

By calculating the partial derivatives in (7), the smoothness of X yields the following restrictions on the a_i :

Proposition 5.2. *In (7), the smoothness of X is equivalent to the condition that the equations*

$$\begin{aligned} a_1x + a_3 &= 0, \\ x^2 + a_1y + a_4 &= 0, \\ a_{1,u}xy + a_{3,u}y + a_{2,u}x^2 + a_{4,u}x + a_{6,u} &= 0, \\ a_{1,v}xy + a_{3,v}y + a_{2,v}x^2 + a_{4,v}x + a_{6,v} &= 0, \end{aligned}$$

with $a_{i,u} := \partial a_i / \partial u$ and $a_{i,v} := \partial a_i / \partial v$ have no common solutions on X .

5.2. Normal forms. In this section, we find normal forms for del Pezzo surfaces of degree 1 in characteristic 2. In total, we will have 14 different normal forms, corresponding to the 14 possible combinations of singularities of the ramification curve R and the branch curve B . First, we simplify the equations of the branch curve.

Lemma 5.3. *Let X be a del Pezzo surface of degree 1 given by (7). Then, after a suitable change of coordinates, we may assume that the equation $a_1x + a_3$ of B is one of the following:*

- (1) $ux + v^3$; (2) ux ; (3) $uv(u + v)$; (4) u^2v ; (5) u^3 .

Proof. If $a_1 \neq 0$, we may assume that $a_1 = u$ after applying a linear substitution in u and v . Then, a substitution of the form $x \mapsto x + b_2$ for a suitable binary form b_2 of degree 2 in u and v allows us to set $a_3 = \lambda v^3$. Then, rescaling v , we can assume $\lambda \in \{0, 1\}$.

If $a_1 = 0$, we get three cases according to the number of distinct roots of a_3 . The equation can be normalized by applying a linear substitution in u and v to get cases (3), (4), and (5). \square

If we consider $\mathbb{P}(1, 1, 2)$ as a quadratic cone Q in \mathbb{P}^3 , these five normal forms for $a_1x + a_3$ correspond to the cases where B is a twisted cubic, a union of a line and a conic, a union of three lines, a union of a double line and a simple line, or a triple line, respectively. Later, we will use automorphisms of $\mathbb{P}(1, 1, 2)$ that preserve the equation of B and the form of (7) in order to move the images of the singular points of R to special positions. In the following lemma, we describe this group of automorphisms.

Lemma 5.4. *Let $H \subseteq \text{Aut}(\mathbb{k}[u, v, x]) \subseteq \text{Aut}(\mathbb{k}[u, v, x, y])$ be the subgroup of automorphisms that preserve $a_1x + a_3$, act on x as $x \mapsto x + b_2$ for some binary quadratic form b_2 in u and v , and that map (7) to one of the same form, with possibly different a_2, a_4 , and a_6 . Then, H consists of substitutions of the form*

$$u \mapsto \alpha u + \beta v, \quad v \mapsto \gamma u + \delta v, \quad x \mapsto x + b_2,$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{k}$ such that $\alpha\delta + \beta\gamma \neq 0$, and:

- (1) If $a_1x + a_3 = ux + v^3$, then $\alpha = 1, \beta = 0, \delta^3 = 1, b_2 = \gamma^3u^2 + \gamma^2\delta uv + \gamma\delta^2v^2$. In particular, $H \cong \mathbb{k}^+ : 3$.
- (2) If $a_1x + a_3 = ux$, then $\alpha = 1, \beta = b_2 = 0$. In particular, $H \cong \mathbb{k}^+ : \mathbb{k}^\times$.
- (3) If $a_1x + a_3 = uv(u + v)$, then $\alpha\gamma(\alpha + \gamma) = \beta\delta(\beta + \delta) = 0, \alpha^2\delta + \beta\gamma^2 = \alpha\delta^2 + \beta^2\gamma = 1$. In particular, $H \cong \mathbb{k}[u, v]_2 : (3 \times \mathfrak{S}_3)$.
- (4) If $a_1x + a_3 = u^2v$, then $\beta = \gamma = 0, \delta = \alpha^{-2}$. In particular, $H \cong \mathbb{k}[u, v]_2 : \mathbb{k}^\times$.
- (5) If $a_1x + a_3 = u^3$, then $\beta = 0, \alpha^3 = 1$. In particular, $H \cong \mathbb{k}[u, v]_2 : (\mathbb{k} : \mathbb{k}^\times \times 3)$.

For the convenience of the reader, we record the effect of a general substitution on the remaining a_i in (7). The proof is a straightforward calculation.

Lemma 5.5. *A substitution of the form*

$$\begin{aligned} u &\mapsto \alpha u + \beta v, & x &\mapsto x + b_2, \\ v &\mapsto \gamma u + \delta v, & y &\mapsto y + b_1x + b_3, \end{aligned}$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{k}$ and $b_i \in \mathbb{k}[u, v]_i$ such that $\alpha\delta + \beta\gamma \neq 0$, changes the coefficients (a_2, a_4, a_6) in (7) as follows:

$$\begin{aligned} a_2 &\mapsto \sigma^* a_2 + \sigma^* a_1 b_1 + b_1^2 + b_2, \\ a_4 &\mapsto \sigma^* a_4 + \sigma^* a_3 b_1 + \sigma^* a_1 b_1 b_2 + \sigma^* a_1 b_3 + b_2^2, \\ a_6 &\mapsto \sigma^* a_6 + \sigma^* a_4 b_2 + \sigma^* a_3 b_3 + \sigma^* a_2 b_2^2 + \sigma^* a_1 b_2 b_3 + b_3^2 + b_2^3, \end{aligned}$$

where $\sigma^* a_i := a_i(\alpha u + \beta v, \gamma u + \delta v)$.

Now, we are ready to describe the normal forms for del Pezzo surfaces of degree 1.

name	a_1x+a_3	a_2	a_4	a_6	# of parameters
(1a)	$ux+v^3$	av^2	$bu^4+cu^2v^2+dv^4$	$eu^6+fu^4v^2+gu^2v^4+hv^6$	8
(1b)	$ux+v^3$	av^2	$cu^2v^2+dv^4$	$eu^6+fu^4v^2+gu^2v^4+hv^6$	7
(1c)	$ux+v^3$	av^2	dv^4	$eu^6+fu^4v^2+gu^2v^4+hv^6$	6
(1d)	$ux+v^3$	av^2	cu^2v^2	$eu^6+fu^4v^2+gu^2v^4+hv^6$	6
(1e)	$ux+v^3$	av^2	0	$eu^6+fu^4v^2+gu^2v^4+hv^6$	5
(2a)	ux	av^2	v^4	$bu^6+du^4v^2+eu^3v^3+fu^2v^4+guv^5+hv^6$	7
(2b)	ux	av^2	v^4	$bu^6+du^4v^2+fu^2v^4+guv^5+hv^6$	6
(2c)	ux	av^2	v^4	$bu^6+du^4v^2+eu^3v^3+fu^2v^4+hv^6$	6
(2d)	ux	av^2	v^4	$cu^5v+du^4v^2+fu^2v^4+hv^6$	5
(2e)	ux	av^2	0	$bu^6+du^4v^2+eu^3v^3+fu^2v^4+euv^5+hv^6$	6
(2f)	ux	av^2	0	$bu^6+du^4v^2+fu^2v^4+uv^5+hv^6$	5
(3)	$uv(u+v)$	auv	$bu^3v+(b+c)u^2v^2+cu^2v^3$	$du^5v+eu^3v^3+fu^2v^5$	6
(4)	u^2v	0	$au^3v+bu^2v^2+cu^2v^3$	$du^5v+eu^3v^3+uv^5$	5
(5)	u^3	0	$au^3v+bu^2v^2+cu^2v^3$	uv^5+dv^6	4

Table 5. Forms of $(a_1, a_2, a_3, a_4, a_6)$ in Theorem 5.6.

Theorem 5.6. Every del Pezzo surface of degree 1 in characteristic 2 is isomorphic to a surface of degree 6 in $\mathbb{P}(1, 1, 2, 3)$ given by an equation of the form

$$y^2 + y(a_1(u, v)x + a_3(u, v)) + x^3 + a_2(u, v)x^2 + a_4(u, v)x + a_6(u, v) = 0, \tag{8}$$

where $(a_1, a_2, a_3, a_4, a_6)$ is one of the forms shown in Table 5. Moreover, the parameters satisfy the conditions summarized in Table 6, where

$$\Delta := a_3^4 + a_1^3 a_3^3 + a_1^4 (a_4^2 + a_1 a_3 a_4 + a_2 a_3^2 + a_1^2 a_6).$$

In Table 6, we also describe the singularities of the irreducible components of the reduction R_{red} of the ramification curve R .

Remark 5.7. The conditions on the parameters that guarantee the smoothness of X are equivalent to the conditions that (8) is the Weierstrass equation of an elliptic fibration with only irreducible fibers. We will study this fibration later in Section 5.4. The homogeneous polynomial Δ appearing in Theorem 5.6 is the discriminant of this fibration.

Proof of Theorem 5.6. By Lemma 5.3, there are, up to choice of coordinates, five possible equations for B . We will now give normal forms in each case.

(1) $a_1x + a_3 = ux + v^3$. Here, the ramification curve R is given by the two equations

$$\begin{aligned} ux + v^3 &= 0, \\ y^2 + x^3 + a_2x^2 + a_4x + a_6 &= 0. \end{aligned}$$

name	conditions on the parameters	singularities of the irreducible components of R_{red}
(1a)	Δ has only simple roots $v^8 + dv^6 + cv^4 + bv^2$ has four distinct roots	A_2 over $[1 : v : v^3]$ with $v^8 + dv^6 + cv^4 + bv^2 = 0$
(1b)	Δ has only simple roots, $c, d \neq 0$	A_4 over $[1 : 0 : 0]$ $2A_2$ over $[1 : v : v^3]$ with $v^4 + dv^2 + c = 0$
(1c)	Δ has only simple roots, $d \neq 0$	A_6 over $[1 : 0 : 0]$ A_2 over $[1 : d^{1/2} : d^{3/2}]$
(1d)	Δ has only simple roots, $c \neq 0$	$2A_4$ over $[1 : 0 : 0]$ and $[1 : c^{1/4} : c^{3/4}]$
(1e)	$e \neq 0$	A_8 over $[1 : 0 : 0]$
(2a)	$u^{-4}\Delta$ has only simple roots, $e, g, (g^2 + a + h) \neq 0$	$3A_2$ over $[0 : 1 : 1]$, $[1 : 0 : 0]$ and $[g^{1/2} : e^{1/2} : 0]$
(2b)	$b, g, (g^2 + a + h) \neq 0$	A_4 over $[1 : 0 : 0]$ A_2 over $[0 : 1 : 1]$
(2c)	$b, e, (a + h) \neq 0$	$3A_2$ over $[0 : 1 : 1]$, $[1 : 0 : 0]$ and $[0 : 1 : 0]$
(2d)	$c, (a + h) \neq 0$	A_4 over $[0 : 1 : 0]$ A_2 over $[0 : 1 : 1]$
(2e)	$u^{-6}\Delta$ has only simple roots, $e \neq 0$	$3A_2$ over $[0 : 1 : 0]$, $[1 : 0 : 0]$ and $[1 : 1 : 0]$
(2f)	$u^{-6}\Delta$ has only simple roots	A_4 over $[1 : 0 : 0]$ A_2 over $[0 : 1 : 0]$
(3)	$d, f \neq 0, (d + e + f) \notin \{0, 1\}$	$3A_2$ over $[1 : 0 : 0]$, $[0 : 1 : 0]$ and $[1 : 1 : 0]$
(4)	$d \neq 0$	$2A_2$ over $[1 : 0 : 0]$ and $[0 : 1 : 0]$
(5)	—	A_2 over $[0 : 1 : 0]$

Table 6. Conditions for the parameters in Theorem 5.6.

One checks that the curve R is smooth at the points with $u = 0$. On the affine chart $u = 1$, it is given in \mathbb{A}^2 by the single equation

$$y^2 + v^9 + a_2(1, v)v^6 + a_4(1, v)v^3 + a_6(1, v),$$

so it has singularities over the roots of the derivative F' of $F := v^9 + a_2(1, v)v^6 + a_4(1, v)v^3 + a_6(1, v)$. After applying an element of H in Lemma 5.4, we may assume that 0 is the root of highest multiplicity of F' .

Now, substitutions as in Lemma 5.5 that fix u, v , and x do not change the location of the points that lie under singularities of R and thus, by Lemma 5.5, we can assume that $a_2 = av^2$, $a_4 = bu^4 + cu^2v^2 + dv^4$, $a_6 = eu^6 + fu^4v^2 + gu^2v^4 + hv^6$. With this notation, the polynomial F' becomes $v^8 + dv^6 + cv^4 + bv^2$ and the conditions of Proposition 5.2 boil down to $v^8 + dv^6 + cv^4 + bv^2$ and

$$\Delta(1, v) = v^{12} + v^9 + (d^2 + a)v^8 + dv^7 + hv^6 + cv^5 + (c^2 + g)v^4 + bv^3 + fv^2 + b^2 + e$$

not having a common solution. The former is the derivative of the latter; hence we want that the latter has only simple zeroes.

Now, if F' has four distinct roots, we are in case (a). If F' has less than four distinct roots, we may assume $b = 0$. If F' has exactly three roots, then we are in case (b). If $b = 0$, the polynomial F' has exactly two roots if and only if either $c = 0$ and $d \neq 0$, which is case (c), or $d = 0$ and $c \neq 0$, which is case (d). Finally, F' has a single root if and only if $b = c = d = 0$, which is case (e).

(2) $a_1x + a_3 = ux$. Here, the ramification curve has two components R_1 and R_2 . The curve R_1 is given by

$$\begin{aligned} u &= 0, \\ y^2 + x^3 + a_2x^2 + a_4x + a_6 &= 0. \end{aligned}$$

This curve has a unique singularity, which is of type A_2 and located over $[0 : 1 : a_4(0, 1)^{1/2}]$. Rescaling v , we may assume that $a_4(0, 1) \in \{0, 1\}$.

The curve R_2 is given by

$$\begin{aligned} x &= 0, \\ y^2 + a_6 &= 0. \end{aligned}$$

This curve has singularities over the points $[u : v : 0]$, where the derivatives of a_6 by u and v both vanish.

First, assume that $a_4(0, 1) = 1$ and one of the singularities of R_2 does not lie over $[0 : 1 : 0]$. Then, using a substitution in v as in Lemma 5.4, we can assume that one of them lies over $[1 : 0 : 0]$. Substitutions as in Lemma 5.5 which fix u , v , and x do not change the location of these points and, after applying one of them, we may assume that $a_2 = av^2$, $a_4 = v^4$, and $a_6 = bu^6 + du^4v^2 + eu^3v^3 + fu^2v^4 + guv^5 + hv^6$. If $e, g \neq 0$, this is case (a), if $e = 0$ and $g \neq 0$, this is case (b), and if $e \neq 0$ and $g = 0$, this is case (c). The conditions of Proposition 5.2 boil down to $\Delta(1, v) = v^8 + hv^6 + gv^5 + fv^4 + ev^3 + dv^2 + b$ having only simple roots and $g^2 \neq a + h$. In particular, $(e, g) \neq (0, 0)$.

If $a_4(0, 1) = 1$, R_2 has a unique singularity, and this singularity lies over $[0 : 1 : 0]$, then the only odd monomial in a_6 is u^5v . A substitution of the form $v \mapsto v + \mu u$ and substitutions as in the previous paragraph allow us to assume that $a_2 = av^2$, $a_4 = v^4$, and $a_6 = cu^5v + du^4v^2 + fu^2v^4 + hv^6$. The conditions of Proposition 5.2 become $a + h \neq 0$ and $c \neq 0$. This is case (d).

If $a_4(0, 1) = 0$, then Proposition 5.2 implies that R_2 is smooth over $[0 : 1 : 0]$. Hence, we can assume that one of the singularities of R_2 lies over $[1 : 0 : 0]$. Using a substitution as in Lemma 5.5 which fixes u , v and x , we may assume that $a_2 = av^2$, $a_4 = 0$, and $a_6 = bu^6 + du^4v^2 + eu^3v^3 + fu^2v^4 + guv^5 + hv^6$. Since R_2 is smooth over $[0 : 1 : 0]$, we have $g \neq 0$. If $e \neq 0$, we can scale v so that $g = e$. This is case (e). If $e = 0$, we scale v so that $g = 1$. This is case (f).

(3) $a_1x + a_3 = uv(u + v)$. The curve B has the three irreducible components B_1 , B_2 , and B_3 , given by $V(u)$, $V(v)$, and $V(u + v)$, respectively. The corresponding components R_1 , R_2 , and R_3 of R are given

by

$$\begin{aligned} & y^2 + x^3 + a_2(0, v)x^2 + a_4(0, v)x + a_6(0, v), \\ & y^2 + x^3 + a_2(u, 0)x^2 + a_4(u, 0)x + a_6(u, 0), \\ & y^2 + x^3 + a_2(u, u)x^2 + a_4(u, u)x + a_6(u, u), \end{aligned}$$

respectively. The singular points of R_1 , R_2 , and R_3 lie over $[0 : 1 : a_4(0, 1)^{1/2}]$, $[1 : 0 : a_4(1, 0)^{1/2}]$, and $[1 : 1 : a_4(1, 1)^{1/2}]$, respectively.

A substitution as in Lemma 5.4 which fixes u and v allows us to set $a_4(0, 1) = a_4(1, 0) = a_4(1, 1) = 0$, that is, that $a_4 = bu^3v + (b+c)u^2v^2 + cuv^3$ for some $b, c \in \mathbb{k}$. Then, a substitution as in Lemma 5.5 which fixes u , v , and x allows us to set $a_2 = auv$ and $a_6 = du^5v + eu^3v^3 + fuv^5$. The conditions of Proposition 5.2 become $d \neq 0$, $f \neq 0$ and $d + e + f \notin \{0, 1\}$.

(4) $a_1x + a_3 = u^2v$. The curve B has two irreducible components B_1 and B_2 , given by $V(u)$ and $V(v)$, respectively. The corresponding components R_1 and R_2 of R are given by

$$\begin{aligned} & y^2 + x^3 + a_2(0, v)x^2 + a_4(0, v)x + a_6(0, v), \\ & y^2 + x^3 + a_2(u, 0)x^2 + a_4(u, 0)x + a_6(u, 0), \end{aligned}$$

respectively. The singular points of R_1 and R_2 lie over $[0 : 1 : a_4(0, 1)^{1/2}]$ and $[1 : 0 : a_4(1, 0)^{1/2}]$, respectively.

A substitution as in Lemma 5.5, which fixes u and v , allows us to set $a_4(0, 1) = a_4(1, 0)$ and gives that a_2 is a square. Then, a substitution with $b_2 = b_3 = 0$ allows us to eliminate a_2 . Finally, a substitution with $b_1 = b_2 = 0$ allows us to assume that a_6 contains no squares. If we write $a_6 = du^5v + eu^3v^3 + fuv^5$, then the conditions of Proposition 5.2 becomes $d \neq 0$ and $f \neq 0$, and we can rescale f to 1.

(5) $a_1x + a_3 = u^3$. The curve R is given by

$$y^2 + x^3 + a_2(0, v)x^2 + a_4(0, v)x + a_6(0, v)$$

and it is singular over $[0 : 1 : a_4(0, 1)^{1/2}]$.

We apply the same substitutions as in the previous case to remove a_2 . Then, we apply a substitution as in Lemma 5.5 with $b_2 = b_1^2$ to remove the v^4 -term in a_4 . Next, using a substitution that fixes u , v , and x with $b_1 = 0$, we eliminate the squares in a_6 , write $a_6 = du^5v + eu^3v^3 + fuv^5$, and rescale f to 1. After that, a substitution of the form $v \mapsto v + \lambda u$, and eliminating the square again, allows us to set $d = 0$. Next, a substitution as in Lemma 5.5 which fixes u and v , with $b_1 = \lambda u$, $b_2 = \lambda^2 u^2$, and $b_3 = \mu u^3$ for suitable λ and μ allows us to eliminate the u^4 -term in a_4 without changing a_6 . Finally, we apply a substitution with $b_3 = ev^3$ and rename the parameters to assume that $a_6 = uv^5 + dv^6$. The conditions of Proposition 5.2 are fulfilled for every choice of parameters. \square

5.3. Fake tritangent planes and odd theta characteristics. It is known that a del Pezzo surface X of degree 1 contains 240 (-1) -curves (see [Dolgachev 2012, Section 8.7], where the proof is characteristic-free). They come in pairs $E_i + E'_i \in |-2K_X|$ with $E_i \cdot E'_i = 3$. The Bertini involution β swaps the two

curves in a pair. The image of $E_i + E'_i$ under the antibicanonical map f is a plane section of Q not passing through the vertex.

If $p \neq 2$, each of the resulting 120 planes is a tritangent plane to the branch sextic curve and, conversely, every tritangent plane to the branch sextic gives rise to a pair of (-1) -curves $E_i + E'_i$ with $E_i + E'_i \in |-2K_X|$. A tritangent plane intersects the branch curve in twice a positive divisor of degree 3. This divisor is an odd theta characteristic of the curve. It is known that the number of odd theta characteristics on a smooth curve of genus 4 is equal to 120.

For arbitrary p , we still have the following.

Lemma 5.8. *The preimage $f^{-1}(C)$ of an integral conic $C = V(x + b_2)$ is a sum of two (-1) -curves if and only if it is reducible.*

Proof. Since f has degree 2 and C is integral, the curve $f^{-1}(C)$ is reducible if and only if it has two irreducible components L_1 and L_2 . These components satisfy $L_1 + L_2 \in |-2K_X|$, $L_1 \cdot L_2 = 3$, and $L_1^2 = L_2^2$. Via adjunction, this easily implies that L_1 and L_2 are (-1) -curves. The converse is clear. \square

So, even if $p = 2$, we have 120 splitting conics and we call the corresponding planes in \mathbb{P}^3 *fake tritangent planes* in analogy with the situation in the other characteristics. For the rest of this section, we assume $p = 2$.

Since the antibicanonical map is étale outside the branch curve $V(A)$, the intersection $E_i \cap E'_i$ lies on the ramification curve R . Let $\mathcal{L} = \mathcal{O}_R(E_i) \cong \mathcal{O}_R(E'_i)$. It is an invertible sheaf on C of degree 2. We have

$$\mathcal{L}^{\otimes 2} \cong \mathcal{O}_R(E_i + E'_i) \cong \mathcal{O}_R(-2K_X).$$

Since $B \in |\mathcal{O}_{\mathbb{P}(1,1,2)}(3)|$, we have $R \in |-3K_X|$. By the adjunction formula, we have

$$\omega_R \cong \mathcal{O}_R(-3K_X + K_X) \cong \mathcal{L}^{\otimes 2}.$$

As in the case of degree 2, invertible sheaves on R that satisfy this property are called invertible theta characteristics. Let $\Theta(R)$ be the set of isomorphism classes of such invertible theta characteristics on R and let $J(R)$ be the generalized Jacobian of R . As in Lemma 3.6, one can prove that $J(R)$ is a product of additive groups.

Lemma 5.9. *The generalized Jacobian $J(R)$ of R is isomorphic to \mathbb{G}_a^4 .*

Thus, as in degree 2, finding fake tritangent planes using theta characteristics on R is subtle in characteristic 2. We refer to Example 3.7 for an example in degree 2 that further illustrates this point and leave it to the reader to find a similar example in degree 1.

5.4. Rational elliptic surfaces. Equation (7) can also serve as the Weierstrass equation of the rational surface with a genus-1 fibration $\phi : Y \rightarrow \mathbb{P}^1$ obtained by blowing up the base point p_0 of $|-K_X|$. Since X is a del Pezzo surface, all members of $|-K_X|$ are irreducible; hence so are all fibers of ϕ . The discriminant of ϕ is

$$\Delta = a_3^4 + a_1^3 a_3^3 + a_1^4 (a_4^2 + a_1 a_3 a_4 + a_2 a_3^2 + a_1^2 a_6).$$

The singular fibers of ϕ lie over the zeroes of Δ . Moreover, the Bertini involution, which is given by $\beta : y \mapsto y + (a_1x + a_3)$, induces the inversion on the group structure of each fiber. In particular, for $[u_0 : v_0] \in \mathbb{P}^1$, if $a_1(u_0, v_0)x + a_3(u_0, v_0) = 0$, the corresponding fiber F of ϕ is cuspidal, if $a_1(u_0, v_0) = 0$ and $a_3(u_0, v_0) \neq 0$, then F is smooth and supersingular, and in the other cases, F is either nodal, or smooth and ordinary, according to whether $\Delta(u_0, v_0)$ is zero or not. Applying these observations to the normal forms of Theorem 5.6, we obtain the following information on ϕ .

Proposition 5.10. *Let X be a del Pezzo surface of degree 1 given by one of the normal forms in Theorem 5.6. Then, the associated genus-1 fibration ϕ is elliptic and all its fibers are irreducible. The discriminant Δ and the singular fibers of ϕ are given in Table 7.*

Remark 5.11. As in Remark 3.12, we point out the connection to Lang’s classification of singular fibers on rational elliptic surfaces: our normal forms for del Pezzo surfaces of degree 1 yield normal forms for all rational elliptic surfaces with a section in characteristic 2 whose fibers are all irreducible.

name	Δ	nodal fibers over the	cuspidal fibers over
(1a)	$v^{12} + u^3v^9 + (d^2 + a)u^4v^8 + du^5v^7 + hu^6v^6 + cu^7v^5 + (c^2 + g)u^8v^4 + bu^9v^3 + fu^{10}v^2 + (b^2 + e)u^{12}$	12 roots of Δ	–
(1b)	$v^{12} + u^3v^9 + (d^2 + a)u^4v^8 + du^5v^7 + hu^6v^6 + cu^7v^5 + (c^2 + g)u^8v^4 + fu^{10}v^2 + eu^{12}$	12 roots of Δ	–
(1c)	$v^{12} + u^3v^9 + (d^2 + a)u^4v^8 + du^5v^7 + hu^6v^6 + gu^8v^4 + fu^{10}v^2 + eu^{12}$	12 roots of Δ	–
(1d)	$v^{12} + u^3v^9 + au^4v^8 + hu^6v^6 + cu^7v^5 + (c^2 + g)u^8v^4 + fu^{10}v^2 + eu^{12}$	12 roots of Δ	–
(1e)	$v^{12} + u^3v^9 + au^4v^8 + hu^6v^6 + gu^8v^4 + fu^{10}v^2 + eu^{12}$	12 roots of Δ	–
(2a)	$u^4(v^8 + u^2(bu^6 + du^4v^2 + eu^3v^3 + fu^2v^4 + gsv^5 + hv^6))$	8 roots of $u^{-4}\Delta$	[0 : 1]
(2b)	$u^4(v^8 + u^2(bu^6 + du^4v^2 + fu^2v^4 + gsv^5 + hv^6))$	8 roots of $u^{-4}\Delta$	[0 : 1]
(2c)	$u^4(v^8 + u^2(bu^6 + du^4v^2 + eu^3v^3 + fu^2v^4 + hv^6))$	8 roots of $u^{-4}\Delta$	[0 : 1]
(2d)	$u^4(v^8 + u^2(cu^5v + du^4v^2 + fu^2v^4 + hv^6))$	8 roots of $u^{-4}\Delta$	[0 : 1]
(2e)	$u^6(bu^6 + du^4v^2 + eu^3v^3 + fu^2v^4 + euv^5 + hv^6)$	If $h \neq 0$: 6 roots of $u^{-6}\Delta$ if $h = 0$: 5 roots of $u^{-7}\Delta$	[0 : 1]
(2f)	$u^6(bu^6 + du^4v^2 + fu^2v^4 + uv^5 + hv^6)$	if $h \neq 0$: 6 roots of $u^{-6}\Delta$ if $h = 0$: 5 roots of $u^{-7}\Delta$	[0 : 1]
(3)	$u^4v^4(u + v)^4$	–	[1 : 0], [0 : 1], [1 : 1]
(4)	u^8v^4	–	[1 : 0], [0 : 1]
(5)	u^{12}	–	[0 : 1]

Table 7. The discriminant Δ and the singular fibers of ϕ for Proposition 5.10.

6. Automorphism groups of del Pezzo surfaces of degree 1

This section consists of three parts. In the first part, we collect various restrictions on the group $G(X) = \text{Aut}(X)/\langle\beta\rangle$ arising from the geometry of X . In the second part, we give an explicit description of $\text{Aut}(X)$ in terms of (7) and use it to classify all surfaces where $G(X)$ is nontrivial and to determine the group $\text{Aut}(X)$ in every case. In the third part, we compare our classification with the classification in characteristic 0 from [Dolgachev 2012, Table 8.14] and use this to determine the conjugacy classes of all elements in $\text{Aut}(X)$ (see Table 9 on page 760). Throughout, we assume $p = 2$.

6.1. Restrictions on $G(X)$. Since the elliptic fibration $\phi : Y \rightarrow \mathbb{P}^1$ associated to X is obtained by blowing up the base point of $|-K_X|$, we can identify $\text{Aut}(X)$ with the subgroup of $\text{Aut}(Y)$ preserving a chosen section. Let $r : \text{Aut}(X) \rightarrow \text{Aut}(\mathbb{P}^1)$ be the natural homomorphism defined by the action of $\text{Aut}(X)$ on the coordinates $[u : v]$ of the base of ϕ . Since ϕ is the unique relatively minimal smooth proper model of its generic fiber F_η , the kernel $K = \text{Ker}(r)$ is isomorphic to the group of automorphisms of the elliptic curve F_η . In particular, K contains the Bertini involution β and it can contain more automorphisms only if the j -invariant of F_η is equal to $0 = 1728$, in which case K is a subgroup of $Q_8 : 3 \cong \text{SL}_2(\mathbb{F}_3)$.

Let P be the image of r . Evidently, P is a finite subgroup of $\text{Aut}(\mathbb{P}^1)$ that leaves invariant the set S_1 of points $p = [u_i : v_i]$ corresponding to the singular fibers. It also leaves invariant the set S_2 of the projections of singular points of the irreducible components of the ramification curve R .

The following proposition shows what kind of groups can be expected to occur for P . We use the known classification of finite subgroups of $\text{Aut}(\mathbb{P}^1) \cong \text{PGL}_2(\mathbb{k}) \cong \text{SL}_2(\mathbb{k})$ [Dolgachev and Martin 2024, Theorem 2.5].

Proposition 6.1. *The group P is isomorphic to $G_{\xi,A}$ or D_{2n} .*

Proof. Since $\text{SL}_2(2) \cong \mathfrak{S}_3 \cong D_6$, it suffices to show that $\text{SL}_2(\mathbb{F}_q) \not\subseteq P$ for $q = 2^m$ and $m \geq 2$. Since the set S_2 has cardinality at most 4 and P preserves S_2 , every homogeneous polynomial F with simple roots along S_2 is P -semi-invariant of degree at most 4. On the other hand, by [Neusel and Smith 2002, Theorem 6.1.8], the ring $\mathbb{k}[u, v]^{\text{SL}_2(\mathbb{F}_q)}$ is generated over \mathbb{F}_q by the Dickson polynomials L and $d_{2,1}$ of degrees $q + 1$ and $q^2 - q$, respectively. If $\text{SL}_2(q) \subseteq P$, then F is also a semi-invariant polynomial for $\text{SL}_2(q)$ and if $q \neq 2$, then $\text{SL}_2(\mathbb{F}_q)$ is simple, so $F \in \mathbb{k}[u, v]^{\text{SL}_2(q)} = \mathbb{k}[L, d_{2,1}]$. Hence, $q = 2$, as claimed. □

We recall from [Dolgachev and Martin 2024, Section 1.3] that the image of the Bertini involution β under the injective homomorphism $\rho : \text{Aut}(X) \rightarrow W(E_8)$ is equal to $-\text{id}_{E_8}$. However, in contrast to the situation in degree 2, the extension $W(E_8) \rightarrow W(E_8)/(-\text{id}_{E_8}) \cong O_8^+(2)$ does not split. The semidirect product $W(E_8) = 2 \cdot \text{GO}_8^+(2)$ corresponds to a nontrivial homomorphism $O_8^+(2) \rightarrow C_2$, whose kernel is a simple group $O_8(2)$, where we use the ATLAS notation.

Therefore, in order to determine $\text{Aut}(X)$, it is not enough to determine the image $G(X)$ of the homomorphism $\text{Aut}(X) \rightarrow \text{Aut}(X)/\langle\beta\rangle$, and thus the calculation of $\text{Aut}(X)$ is more complicated than in the case of del Pezzo surfaces of degree 2.

Let us summarize the restrictions on $\text{Aut}(X)$ and $G(X)$ that we have collected by now.

Theorem 6.2. *Let X be a del Pezzo surface of degree 1 in characteristic 2. Let $G(X)$ be the image of the homomorphism $\text{Aut}(X) \rightarrow \text{Aut}(\mathbb{P}(1, 1, 2))$, let K be the kernel of the homomorphism $r : \text{Aut}(X) \rightarrow \text{Aut}(\mathbb{P}^1)$, let P be the image of r , and let $\phi : Y \rightarrow \mathbb{P}^1$ be the elliptic fibration associated to X . Then, the following hold:*

- (i) $\text{Aut}(X)$ is a central extension of $G(X)$ by $\langle \beta \rangle \cong C_2$.
- (ii) $\text{Aut}(X)$ is an extension of P by K .
- (iii) $\text{Aut}(X)$ is a subgroup of $W(E_8)$.
- (iv) $G(X)$ is a subgroup of $O_8^+(2)$.
- (v) K is the automorphism group of the generic fiber of ϕ .
- (vi) P is isomorphic to $G_{\xi, A}$ or D_{2n} .
- (vii) P preserves the set S_1 of points lying under singular fibers of ϕ . Moreover, it preserves the decomposition of S_1 into subsets corresponding to isomorphic fibers.
- (viii) P preserves the set S_2 of points lying under the singularities of R . Moreover, it preserves the decomposition of S_2 into subsets of isomorphic singularities.
- (ix) The j -function of ϕ is P -invariant.

This yields the following preliminary restrictions on $\text{Aut}(X)$ and $G(X)$.

Corollary 6.3. *Let X be a del Pezzo surface of degree 1 in characteristic 2 given by one of the normal forms in Theorem 5.6.*

- (i) In case (1), $G(X)$ is a subgroup of A_4 .
- (ii) In cases (2a), (2b), (2c), and (2d), $G(X)$ is a subgroup of 2^3 .
- (iii) In cases (2e) and (2f), $G(X)$ is a subgroup of C_5 or C_2 .
- (iv) In case (3), K is a subgroup of $\text{SL}_2(3)$ and P is a subgroup of \mathfrak{S}_3 .
- (v) In case (4), K is a subgroup of $\text{SL}_2(3)$ and P is cyclic of order 1, 3, 5, 7, 9, or 15.
- (vi) In case (5) K is a subgroup of $\text{SL}_2(3)$ and $P \cong G_{\xi, A}$, where ξ is a primitive n -th root of unity with $n \in \{1, 3, 5, 7, 9, 15\}$.

Proof. In case (1), the generic fiber of ϕ is ordinary; hence $K = \langle \beta \rangle$ and $G(X) \cong P$. The fibration ϕ has 12 nodal fibers; hence the j -function has 12 poles, so $|P| \mid 12$. Since P is isomorphic to $G_{\xi, A}$ or D_{2n} with n odd, this implies that P is isomorphic to a subgroup of A_4 .

In cases (2a)–(2f), we also have $K = \langle \beta \rangle$ and $G(X) \cong P$. In cases (2a), (2b), (2c), and (2d), the fibration ϕ has eight nodal fibers; hence $|P| \mid 8$. This implies that P is elementary abelian of order 1, 2, 4 or 8. In cases (2e) and (2f), the fibration ϕ has five or six nodal fibers. If it has five nodal fibers, then $|P| \mid 5$; hence P is a subgroup of C_5 . If it has six nodal fibers, then P is either a subgroup of C_2 or

isomorphic to the dihedral group D_6 . In the latter case, P acts without fixed point on \mathbb{P}^1 , which is impossible, since ϕ admits a unique cuspidal fiber.

In case (3), we have $K \subseteq \mathrm{SL}_2(3)$, since the generic fiber of ϕ is supersingular. Since ϕ has three singular fibers, P is isomorphic to a subgroup of \mathfrak{S}_3 .

In case (4), we also have $K \subseteq \mathrm{SL}_2(3)$. Since one of the components of R is reduced and the other is not, P acts trivially on S_2 , hence with two fixed points on \mathbb{P}^1 . So, P is cyclic of odd order. Moreover, P is a subgroup of $\mathrm{O}_8^+(2)$. In particular, P admits a faithful representation of dimension at most 8. Hence, if we denote Euler's totient function by φ , then $\varphi(|P|) \leq 8$. Thus, P is of order 1, 3, 5, 7, 9 or 15.

In case (5), we have $K \subseteq \mathrm{SL}_2(3)$ and the action of P on \mathbb{P}^1 fixes the point lying under the unique singular fiber of ϕ ; hence $P \cong G_{\xi,A}$. The order of ξ can be bounded by the same argument as in the previous paragraph. □

In particular, we get upper bounds on the size of $\mathrm{Aut}(X)$ in every case. Further information on the 2-groups that can occur in case (5) can be obtained using the following remark.

Remark 6.4. Since the maximal powers of 2 that divide $|W(E_8)|$ and $|W(D_8)|$ are both 2^{14} , and since $W(D_8)$ is a subgroup of $W(E_8)$, the 2-Sylow subgroups P in $W(E_8)$ are isomorphic to the 2-Sylow subgroups in $W(D_8) = 2^7 : \mathfrak{S}_8$. Hence, P is isomorphic to $2^7 : (\mathfrak{S}_8)_2$, where 2^7 acts on \mathbb{Z}^8 by an even number of sign changes and $(\mathfrak{S}_8)_2$ is a 2-Sylow subgroup of \mathfrak{S}_8 acting as permutations on \mathbb{Z}^8 . The group $(\mathfrak{S}_8)_2$ is isomorphic to the symmetry group of a binary tree of depth 3, considered as a subgroup of \mathfrak{S}_8 via the permutation it induces on the leaves of the tree. An equivalent description is as the wreath product $D_8 \wr C_2$, where $D_8 \times D_8$ is a subgroup of $\mathfrak{S}_4 \times \mathfrak{S}_4 \subset \mathfrak{S}_8$. The Bertini involution β corresponds to the element $(-1, \mathrm{id})$ that changes all signs. The 2-groups that can occur in Corollary 6.3 are isomorphic to subgroups of P .

In the following example, we apply this remark to give an explicit description of the group 2_+^{1+6} , which will occur in our classification.

Example 6.5. With notation as in the previous remark, let $G \subseteq P$ be a subgroup containing β such that $G/\langle\beta\rangle$ is an elementary abelian 2-group and such that $\beta \in Q_8 \subseteq G$. Then, each element of G is of the form (σ, τ) , where $\mathrm{ord}(\tau) \leq 2$ and either τ preserves the set of coordinates whose sign is changed by σ and then (σ, τ) has order 1 or 2, or τ swaps this set with the set of coordinates whose sign is not changed and then (σ, τ) has order 4. In particular, in the latter case, τ has cycle type $(2, 2, 2, 2)$. Since $Q_8 \subseteq G$, the image of $G \rightarrow (\mathfrak{S}_8)_2$ contains a subgroup H of order 4 generated by involutions of cycle type $(2, 2, 2, 2)$. The centralizer C of H is of order 8 and its nontrivial elements are involutions of cycle type $(2, 2, 2, 2)$. The kernel of $G \rightarrow (\mathfrak{S}_8)_2$ consists of sign changes σ that are compatible with all $\tau \in H$ in the sense that $(\sigma, \tau)^2 \in \langle\beta\rangle$. One checks that the group N of all such compatible sign changes has order 16 and that all elements of N are also compatible with C . Then, G is a subgroup of the resulting extension M of C by N .

We have $M/\langle\beta\rangle = 2^6$. This is a quadratic space over \mathbb{F}_2 with the quadratic form $q : M/\langle\beta\rangle \rightarrow \langle\beta\rangle$ defined as $q(x) = \tilde{x}^2$, where \tilde{x} is a lift of x to M . The subspace $N/\langle\beta\rangle$ is totally isotropic of dimension 3

and the description of M in the previous paragraph shows that q is nondegenerate. Hence, by [Aschbacher 2000, (23.10)], M is isomorphic to the extra-special 2-group of 2_+^{1+6} .

6.2. Classification. Recall that X is a hypersurface of degree 6 in $\mathbb{P}(1, 1, 2, 3)$ given by (7). An automorphism of $\mathbb{P}(1, 1, 2, 3)$ is induced by a substitution of the form

$$\begin{aligned} u &\mapsto \alpha u + \beta v, & x &\mapsto \epsilon x + b_2, \\ v &\mapsto \gamma u + \delta v & y &\mapsto \zeta y + b_1 x + b_3, \end{aligned}$$

where $\alpha, \beta, \gamma, \delta, \epsilon, \zeta \in \mathbb{k}$, $b_i \in \mathbb{k}[u, v]_i$, and $\alpha\delta + \beta\gamma, \epsilon, \zeta \neq 0$. The substitutions that induce the identity on $\mathbb{P}(1, 1, 2, 3)$ are the ones with $\beta, \gamma, b_1, b_2, b_3 = 0$ and $\gamma = \alpha, \epsilon = \alpha^2, \zeta = \alpha^3$.

Since X is anticanonically embedded into $\mathbb{P}(1, 1, 2, 3)$, all automorphisms of X are induced by the substitutions as above that map (7) to a multiple of itself. Clearly, we can represent every such automorphism by a substitution with $\zeta = 1$. Then, the substitution does not change the coefficient of y^2 in (7); hence $\epsilon^3 = 1$. Therefore, we may assume $\epsilon = 1$ as well. In particular, using Lemma 5.5, we obtain the following description of $\text{Aut}(X)$, where we write σ for the substitution

$$\begin{aligned} u &\mapsto \alpha u + \beta v, \\ v &\mapsto \gamma u + \delta v, \end{aligned}$$

and $\sigma^* a_i := a_i(\alpha u + \beta v, \gamma u + \delta v)$.

Lemma 6.6. *Let X be a del Pezzo surface of degree 1 given by (7). Then, $\text{Aut}(X)$ can be identified with the group of 4-tuples (b_1, b_2, b_3, σ) , where $b_i \in \mathbb{k}[u, v]_i$ and $\sigma \in \text{GL}_2(\mathbb{k})$ such that*

$$\begin{aligned} \sigma^* a_1 + a_1 &= 0, \\ \sigma^* a_2 + a_2 &= a_1 b_1 + b_1^2 + b_2, \\ \sigma^* a_3 + a_3 &= a_1 b_2, \\ \sigma^* a_4 + a_4 &= a_3 b_1 + a_1 b_3 + b_2^2, \\ \sigma^* a_6 + a_6 &= a_4 b_2 + a_3(b_3 + b_1 b_2) + a_2 b_2^2 + a_1(b_2 b_3 + b_1 b_2^2) + b_3^2 + b_2^3 + b_1^2 b_2^2 \end{aligned}$$

and where the composition is given by

$$(b_1, b_2, b_3, \sigma) \circ (b'_1, b'_2, b'_3, \sigma') = (\sigma'^* b_1 + b'_1, \sigma'^* b_2 + b'_2, \sigma'^* b_3 + b'_3 + \sigma'^* b_1 b'_2, \sigma \circ \sigma')$$

In particular, there is a homomorphism $\text{Aut}(X) \rightarrow H \subseteq \text{Aut}(\mathbb{P}(1, 1, 2))$, where H is the group from Lemma 5.4.

Lemma 6.7. *The kernel of the homomorphism $\text{Aut}(X) \rightarrow H$ is generated by the Bertini involution.*

Proof. Let (b_1, b_2, b_3, σ) be in the kernel. Then, $\sigma = \text{id}$ and $b_2 = 0$. The conditions $\sigma^* a_2 = a_2 + a_1 b_1 + b_1^2$, $\sigma^* a_4 = a_4 + a_3 b_1 + a_1 b_3$, and $\sigma^* a_6 = a_6 + a_3 b_3 + b_3^2$ show that $(b_1, b_3) \in \{(0, 0), (a_1, a_3)\}$, so we recover our explicit description of the Bertini involution. □

Now, we use the normal forms of Theorem 5.6 to classify all del Pezzo surfaces X of degree 1 with nontrivial $G(X)$.

name	a_1x+a_3	a_2	a_4	a_6	$G(X)$	$\text{Aut}(X)$	# of parameters
(1ai)	$ux+v^3$	av^2	$bu^4+(b+1)u^2v^2$	$eu^6+fu^4v^2+(a+b+b^2+f)u^2v^4+bv^6$	2	4	4
(1aii)	$ux+v^3$	0	bu^4	eu^6+hv^6	3	6	3
(1aiii)	$ux+v^3$	av^2	u^4	$eu^6+au^4v^2+v^6$	2^2	Q_8	2
(1aiv)	$ux+v^3$	0	u^4	eu^6+v^6	A_4	$SL_2(3)$	1
(1di)	$ux+v^3$	av^2	u^2v^2	$eu^6+fu^4v^2+(a+f)u^2v^4$	2	4	3
(1ei)	$ux+v^3$	0	0	eu^6+hv^6	3	6	2
(2ai)	ux	av^2	v^4	$bu^6+(efg^{-1}+e^{3/2}g^{-1/2}+e^3g^{-3})u^4v^2+eu^3v^3+fu^2v^4+guv^5+e^{-1/2}g^{3/2}v^6$	2	2^2	5
(2di)	ux	av^2	v^4	$cu^5v+du^4v^2+fu^2v^4$	2^3	2^4	4
(2ei)	ux	av^2	0	$bu^6+(e+f)u^4v^2+eu^3v^3+fu^2v^4+euv^5+ev^6$	2	2^2	4
(2fi)	ux	0	0	bu^6+uv^5	5	10	1
(3i)	$uv(u+v)$	auv	$bu^3v+bu^2v^3$	$du^5v+eu^3v^3+duv^5$	2	2^2	4
(3ii)	$uv(u+v)$	auv	$a^{1/2}u^3v+a^{1/2}uv^3$	$(e+e^{1/2})u^5v+eu^3v^3+(e+e^{1/2})uv^5$	\mathfrak{S}_3	$2 \times \mathfrak{S}_3$	2
(3iii)	$uv(u+v)$	0	0	$du^5v+eu^3v^3+duv^5$	6	2×6	2
(3iv)	$uv(u+v)$	0	$bu^3v+\xi_3bu^2v^2+\xi_3^2bu^2v^3$	$(e+e^{1/2})u^5v+eu^3v^3+(e+e^{1/2})uv^5$	3	6	2
(3v)	$uv(u+v)$	0	0	$(e+e^{1/2})u^5v+eu^3v^3+(e+e^{1/2})uv^5$	$3 \times \mathfrak{S}_3$	$6 \times \mathfrak{S}_3$	1
(4i)	u^2v	0	0	$du^5v+eu^3v^3+uv^5$	3	6	2
(5i)	u^3	0	$au^3v+bu^2v^2$	uv^5+dv^6	2^6	2_+^{1+6}	3
(5ii)	u^3	0	0	uv^5+dv^6	$2^6:3$	$2_+^{1+6}:3$	1
(5iii)	u^3	0	0	uv^5	$2^6:15$	$2_+^{1+6}:15$	0

Table 8. Forms of $(a_1, a_2, a_3, a_4, a_6, G(X), \text{Aut}(X))$ in Theorem 6.8.

Theorem 6.8. Every del Pezzo surface of degree 1 in characteristic 2 such that $G(X)$ is nontrivial is isomorphic to a surface of degree 6 in $\mathbb{P}(1, 1, 2, 3)$ given by an equation of the form

$$y^2 + (a_1x + a_3)y + x^3 + a_2x^2 + a_4x + a_6,$$

where $(a_1, a_2, a_3, a_4, a_6, G(X), \text{Aut}(X))$ is one of the forms in Table 8.

Here, $\mathfrak{S}_3, D_8, Q_8,$ and 2_+^{1+6} , denote the symmetric group on three letters, the dihedral group of order 8, the quaternion group, and the even extra-special group of order 128, respectively. In each case, the parameters have to satisfy the conditions of Theorem 5.6 and the obvious genericity conditions that keep them from specializing to other subcases.

Proof. We use the normal forms of Theorem 5.6 and let H be the group of Lemma 5.4. By Lemma 6.6, we have $G(X) \subseteq H$. We apply Lemma 6.6 to calculate $\text{Aut}(X)$.

(1a) Let $(b_2, \sigma) \in H$. If $(b_2, \sigma) \in G(X)$, then σ permutes the roots of the polynomial $F' := v^8 + dv^6 + cv^4 + bv^2$, since these are determined by the singularities of R . We have

$$\sigma^*F' = \delta^2v^8 + dv^6 + \delta(\gamma^2d + c)v^4 + \delta^2(\gamma^4d + b)v^2 + \gamma^8 + \gamma^6d + \gamma^4c + \gamma^2b.$$

If $d \neq 0$, this is a multiple of F' if and only if $\delta = 1$ and $\gamma = 0$; hence σ is the identity and $G(X)$ is trivial. If $d = 0$, it is a multiple of F' if and only if

$$\gamma^8 + \gamma^4c + \gamma^2b = 0 \tag{9}$$

and $\delta = 1$ or $c = 0$.

So, assume first that $c \neq 0$ and $\delta = 1$. If $(b_2, \sigma) \in G(X)$, then there exist polynomials b_1 and b_3 such that $\sigma^*a_2 = a_2 + a_1b_1 + b_1^2 + b_2$ and $\sigma^*a_4 = a_4 + a_3b_1 + a_1b_3 + b_2^2$. In our case, this means

$$\begin{aligned} 0 &= \gamma^2 au^2 + ub_1 + b_1^2 + \gamma^3 u^2 + \gamma^2 uv + \gamma v^2, \\ 0 &= \gamma^2 cu^4 + v^3 b_1 + ub_3 + \gamma^6 u^4 + \gamma^4 u^2 v^2 + \gamma^2 v^4; \end{aligned}$$

hence $b_1 = \lambda u + \gamma^2 t$ with $\lambda^2 + \lambda = \gamma^2 a + \gamma^3$ and $\gamma^4 = \gamma$, and $b_3 = (\gamma^2 c + \gamma^3)u^3 + \gamma uv^2 + \lambda v^3$. If $\gamma \neq 0$, then $\gamma^4 = \gamma$ implies $\gamma^3 = 1$. Modifying the equation of X by an element of H , we may assume that $\gamma = 1$. Plugging this into (9), we obtain $c = b + 1$. Hence, $b_1 = \lambda u + v$ with $\lambda^2 + \lambda = a$ and $b_3 = bu^3 + uv^2 + v^3$. Plugging this into the equation for σ^*a_6 and comparing coefficients in Lemma 6.6, we obtain the conditions $h = b$ and $g = a + b + b^2 + f$. Since γ is uniquely determined by (9), we have $G(X) \cong C_2$. The square of any lift of a nontrivial element of $G(X)$ to $\text{Aut}(X)$ is the Bertini involution; hence $\text{Aut}(X) \cong C_4$.

Next, assume that $c = 0$. If $(b_2, \sigma) \in G(X)$, then there exist polynomials b_1 and b_3 such that $\sigma^*a_2 = a_2 + a_1b_1 + b_1^2 + b_2$ and $\sigma^*a_4 = a_4 + a_3b_1 + a_1b_3 + b_2^2$. In our case, this means

$$\begin{aligned} 0 &= \gamma^2 au^2 + (1 + \delta^2)av^2 + ub_1 + b_1^2 + \gamma^3 u^2 + \gamma^2 \delta uv + \gamma \delta^2 v^2, \\ 0 &= v^3 b_1 + ub_3 + \gamma^6 u^4 + \gamma^4 \delta^2 u^2 v^2 + \gamma^2 \delta^4 v^4; \end{aligned}$$

hence $b_1 = \lambda u + \gamma^2 \delta v$ with $\lambda^2 + \lambda = \gamma^2 a + \gamma^3$ and $\gamma^4 + \gamma = (1 + \delta)a$, as well as $b_3 = \gamma^6 u^3 + \gamma^4 \delta^2 uv^2 + \lambda v^3$.

First, assume that $\delta \neq 1$. Then, σ has order 3; hence if $(b_2, \sigma) \in G(X)$, then it fixes one of the four roots of F' . After conjugating by a suitable element of H and repeating the substitutions we used in Theorem 5.6, we may assume that (b_2, σ) fixes $[1 : 0 : 0]$. This implies that $\gamma = 0$; hence $(1 + \delta)a = 0$ implies $a = 0$. Now, we plug everything into the equation for σ^*a_6 and compare coefficients to obtain the conditions $f = g = 0$.

If $\delta = 1$, then $\gamma^4 + \gamma = 0$. Hence, if (b_2, σ) is nontrivial, then $\gamma^3 = 1$. Modifying the equation of X by an element of H , we may assume $\gamma = 1$, that is, that (b_2, σ) maps $[1 : 0 : 0]$ to $[1 : 1 : 1]$. Then, (9) implies $b = 1$. Plugging into the equation for σ^*a_6 and comparing coefficients yields $g = f + a$ and $h = 1$. The square of both lifts of (b_2, σ) to $\text{Aut}(X)$ is the Bertini involution; hence the subgroup generated by these lifts is isomorphic to C_4 .

Suppose next that $G(X)$ contains two distinct nontrivial automorphisms with $\delta = 1$. Then, we can assume that one of them acts as in the previous paragraph, so $b = h = 1$ and $g = f + a$. The other one satisfies $\gamma \neq 1$. Plugging this into the equation for σ^*a_6 and comparing coefficients yields $f = a$. As in the previous paragraph, the square of all lifts of these automorphisms is the Bertini involution; hence they generate a subgroup isomorphic to the quaternion group Q_8 .

Finally, Corollary 6.3 shows that $G(X)$ acts on the four singular points of R through A_4 , so if $G(X)$ contains a nontrivial automorphism with $\delta = 1$ and a nontrivial automorphism with $\delta \neq 1$, then $G(X) \cong A_4$. In particular, the previous two paragraphs show that $b = h = 1$ and $g = 0$ and $f = a$, while the above paragraph for $\delta \neq 1$ shows $a = f = g = 0$. In this case, $\text{Aut}(X) \cong \text{SL}_2(3)$.

(1b) and (1c) In these cases, the singularity of R over $[1 : 0 : 0]$ is not isomorphic to the other singularities of R ; hence $G(X)$ is a subgroup of C_3 acting through the subgroup of H with $\gamma = 0$. In particular, $G(X)$ fixes the points $[1 : 0 : 0]$ and $[0 : 1 : 0]$. Since the number of singular points of R that lie over points different from $[1 : 0 : 0]$ and $[0 : 1 : 0]$ is not divisible by 3, $G(X)$ fixes all of them; hence $G(X)$ is trivial.

(1d) In this case, R has singularities over $[1 : 0 : 0]$ and $[1 : c^{1/4} : c^{3/4}]$. An element of H that fixes both of these points is trivial, and the unique one that swaps the two points is of the form (b_2, σ) , where σ acts as $v \mapsto v + c^{1/4}u$ and $b_2 = c^{3/4}u^2 + c^{1/2}uv + c^{1/4}v^2$. If such an element lies in $G(X)$, then there exist polynomials b_1 and b_3 such that

$$\begin{aligned} 0 &= (ac^{1/2} + c^{3/4})u^2 + ub_1 + b_1^2 + c^{1/2}uv + c^{1/4}v^2, \\ 0 &= v^3b_1 + ub_3 + cu^2v^2 + c^{1/2}v^4; \end{aligned}$$

hence $b_1 = \lambda u + c^{1/2}v$ with $\lambda^2 + \lambda = ac^{1/2} + c^{3/4}$ and $c^4 = c$, and $b_3 = \lambda v^3 + cuv^2$. By Theorem 5.6 we have $c \neq 0$; hence we can apply an element of H to assume that $c = 1$. Plugging this into the equation for σ^*a_6 and comparing coefficients in Lemma 6.6, we obtain the conditions $h = 0$ and $g = a + f$. The square of this automorphism (b_1, b_2, b_3, σ) is the Bertini involution; hence $\text{Aut}(X) \cong C_4$ in this case.

(1e) In this case, we have $G(X) \subseteq C_3$, since $G(X)$ fixes $[1 : 0 : 0]$. Nontrivial elements of H that fix $[1 : 0 : 0]$ are of the form $(0, \sigma)$, where σ acts as $v \mapsto \delta v$ with $\delta^3 = 1$ and $\delta \neq 1$. Such an automorphism lifts to X if and only if there exist polynomials b_1 and b_3 such that

$$\begin{aligned} (1 + \delta^2)av^2 &= ub_1 + b_1^2, \\ 0 &= v^3b_1 + ub_3, \\ (1 + \delta^2)fu^4v^2 + (1 + \delta)gu^2v^4 &= v^3b_3 + b_3^2. \end{aligned}$$

The first equation implies $a = 0$ and $b_1 = \lambda u$ with $\lambda^2 + \lambda = 0$ and then the second equation implies that also $b_3 = \lambda v^3$. Finally, the third equation shows $f = g = 0$.

(2a) Here, $G(X) \subseteq H$ fixes the point $[0 : 1 : 1]$. Moreover, if $G(X)$ fixes the images of the other two singularities, then, by our description of H , $G(X)$ is trivial. Hence, $G(X) \subseteq C_2$ with equality if and only if $G(X)$ contains the involution $(0, \sigma)$, where σ acts as $v \mapsto v + e^{1/2}g^{-1/2}u$.

If this involution is in $G(X)$, then there exist polynomials b_1 and b_3 such that

$$\begin{aligned} aeg^{-1}u^2 &= ub_1 + b_1^2, \\ e^2g^{-2}u^4 &= ub_3; \end{aligned}$$

hence $b_1 = \lambda u$ with $\lambda^2 + \lambda = aeg^{-1}$, and $b_3 = e^2g^{-2}u^3$. Plugging this into the equation for σ^*a_6 and comparing coefficients in Lemma 6.6, we obtain the conditions

$$\begin{aligned} 0 &= e^4 + he^3g + fe^2g^2 + deg^3, \\ 0 &= e^{1/2}(g^{3/2} + he^{1/2}). \end{aligned}$$

Since $e \neq 0$ by Theorem 5.6, we have $h = e^{-1/2}g^{3/2}$ and $d = efg^{-1} + e^{3/2}g^{-1/2} + e^3g^{-3}$. Note that both lifts of $(0, \sigma)$ have order 2; hence $\text{Aut}(X) \cong 2^2$.

(2b) and (2c) Here, $G(X) \subseteq H$ fixes $[0 : 1 : 1]$ and $[1 : 0 : 0]$, since these are the points that lie under the singularities of the irreducible components of R , but not under the intersection of the two components R_1 and R_2 . By our description of H in Lemma 5.4, this implies that $G(X)$ is trivial.

(2d) In this case, $G(X)$ fixes $[0 : 1 : 1]$, but we get no other restrictions from the position of the singularities of R . Therefore, an element of $G(X) \subseteq H$ is of the form $(0, \sigma)$ where σ acts as $v \mapsto v + \gamma u$ for some $\gamma \in \mathbb{k}$. Such an element is in $G(X)$ if and only if there exist polynomials b_1 and b_3 such that

$$\begin{aligned} a\gamma^2u^2 &= ub_1 + b_1^2, \\ \gamma^4u^4 &= ub_3, \\ (c\gamma + d\gamma^2 + f\gamma^4 + h\gamma^6)u^6 + h\gamma^4u^4v^2 + h\gamma^2u^2v^4 &= b_3^2. \end{aligned}$$

Such b_1 and b_3 exist if and only if $h = 0$ and $\gamma^8 + h\gamma^6 + f\gamma^4 + d\gamma^2 + c\gamma = 0$, and then $b_1 = \lambda u$ with $\lambda^2 + \lambda = a\gamma^2$, and $b_3 = \gamma^4u^3$. By Theorem 5.6, we have $c \neq 0$; hence, as soon as $h = 0$, there are exactly eight choices for γ . This shows $G(X) \cong 2^3$. Every lift of every nontrivial element in $G(X)$ has order 2; hence $\text{Aut}(X) \cong 2^4$.

(2e) Here, the elements of $G(X) \subseteq H$ fix $[0 : 1 : 0]$ and preserve the pair $\{[1 : 0 : 0], [1 : 1 : 0]\}$. Using our description of γ , it is clear that an element of H that fixes all of these three points is the identity. An element that swaps $[1 : 0 : 0]$ and $[1 : 1 : 0]$ is of the form $(0, \sigma)$, where σ acts as $v \mapsto v + u$. Such an element is in $G(X)$ if and only if there exist polynomials b_1 and b_3 such that

$$\begin{aligned} au^2 &= ub_1 + b_1^2, \\ 0 &= ub_3, \\ (d + f + h)u^6 + (e + h)u^4v^2 + (e + h)u^2v^4 &= b_3^2; \end{aligned}$$

hence if and only if $h = e$ and $d = e + f$, and then $b_1 = \lambda u$ with $\lambda^2 + \lambda = a$ and $b_3 = 0$. The square of the lift of this automorphism to $\text{Aut}(X)$ is the identity; hence $\text{Aut}(X) \cong 2^2$.

(2f) In this case, $G(X) \subseteq H$ fixes $[1 : 0 : 0]$ and $[0 : 1 : 0]$. Hence, by our description of H in Lemma 5.4, every element in $G(X)$ is of the form $(0, \sigma)$, where σ acts as $v \mapsto \delta v$ for some $\delta \in \mathbb{k}^\times$. A nontrivial element of this form is in $G(X)$ if and only if there exist b_1 and b_3 such that

$$\begin{aligned} a(1 + \delta^2)v^2 &= ub_1 + b_1^2, \\ 0 &= ub_3, \\ d(1 + \delta^2)u^4v^2 + f(1 + \delta^4)u^2v^4 + (1 + \delta^5)uv^5 + h(1 + \delta^6) &= b_3^2. \end{aligned}$$

Hence, we always have $b_1 = b_3 = 0$ and $\delta^5 = 1$. Since $\delta \neq 1$ by assumption, we deduce that $(0, \sigma)$ lifts if and only if $a = d = f = h = 0$.

(3) Here, the group $G(X) \subseteq H$ fixes $[1 : 0 : 0]$, $[0 : 1 : 0]$, and $[1 : 1 : 0]$. Hence, every element of $G(X)$ is of the form $(0, \sigma)$ and σ satisfies the conditions of Lemma 5.4 (3).

First, assume that σ has even order and interchanges two components of B . Without loss of generality, we may assume that σ swaps u and v . Then, $(0, \sigma)$ lifts to X if and only if there exist b_1 and b_3 such that

$$\begin{aligned} 0 &= b_1^2, \\ (b+c)(u^3v+uv^3) &= uv(u+v)b_1, \\ (d+f)(u^5v+uv^5) &= uv(u+v)b_3 + b_3^2. \end{aligned}$$

This holds if and only if $b = c$, and then $b_1 = 0$, as well as $b_3 = \lambda uv(u+v)$ with $\lambda^2 + \lambda = 0$ and $d = f$. The square of both lifts of $(0, \sigma)$ is the identity; hence they generate a group isomorphic to 2^2 .

Next, assume that σ is nontrivial and preserves the three components of B . Then, it acts as $u \mapsto \alpha u$, $v \mapsto \alpha v$, where $\alpha^3 = 1$, $\alpha \neq 1$. This automorphism lifts to X if and only if there exist polynomials b_1 and b_3 such that

$$\begin{aligned} a(1+\alpha^{-1})uv &= b_1^2, \\ (1+\alpha)(bu^3v+(b+c)u^2v^2+cu^3v^3) &= uv(u+v)b_1, \\ 0 &= uv(u+v)b_3 + b_3^2; \end{aligned}$$

hence if and only if $a = b = c = 0$.

Finally, assume that σ has odd order and interchanges components of B . Without loss of generality, we may assume that σ acts as $u \mapsto \beta v$, $v \mapsto \beta(u+v)$ with $\beta^3 = 1$. This lifts to X if and only if there exist b_1 and b_3 such that

$$\begin{aligned} a(1+\beta^2)uv + a\beta^2v^2 &= b_1^2, \\ (b+\beta c)u^3v + (b+c+\beta b)u^2v^2 + (c+\beta(b+c))uv^3 &= uv(u+v)b_1, \\ (d+f)u^5v + fu^4v^2 + eu^2v^4 + (d+e)uv^5 + (d+e+f)v^6 &= uv(u+v)b_3 + b_3^2. \end{aligned}$$

The third equation implies $f = d$ and $d = e + e^{1/2}$ and then $b_3 = \lambda u^2v + \lambda uv^2 + e^{1/2}v^3$, where $\lambda^2 + \lambda = e + e^{1/2}$. If $\beta = 1$, the first equation implies $b_1 = a^{1/2}v$ and the second equation implies $b = c = a^{1/2}$. If $\beta \neq 1$, the first equation implies $b_1 = a = 0$ and the second equation implies $b = \beta c$.

(4) In this case, the group $G(X) \subseteq H$ fixes $[1 : 0 : 0]$ and $[0 : 1 : 0]$; hence every element of $G(X)$ is of the form (b_2, σ) , with $b_2 = \lambda uv$ for some $\lambda \in \mathbb{k}$ and where σ acts as $u \mapsto \alpha u$, $v \mapsto \alpha^{-2}v$ with $\alpha \in \mathbb{k}^\times$.

If such an automorphism lifts to X , then the condition $\sigma^*a_2 = a_2 + b_1^2 + b_2$ forces $b_2 = b_1^2$; hence $b_1 = b_2 = 0$. The other conditions of Lemma 6.6 become

$$\begin{aligned} a(1+\alpha)u^3v + b(1+\alpha^{-2})u^2v^2 + c(1+\alpha^{-5})uv^3 &= 0, \\ d(1+\alpha^3)u^5v + e(1+\alpha^{-3})u^3v^3 + f(1+\alpha^{-9})uv^5 &= u^2vb_3 + b_3^2. \end{aligned}$$

Since $d \neq 0$, the second equation implies $\alpha^3 = 1$. Hence, if σ is nontrivial, then $(0, \sigma)$ lifts to X if and only if $a = b = c = 0$.

(5) Here, $G(X) \subseteq H$ fixes $[0 : 1 : 0]$; hence every element of $G(X)$ is of the form (b_2, σ) , with $b_2 = \lambda u^2 + \mu uv$ for some $\lambda, \mu \in \mathbb{k}$ and where σ acts as $u \mapsto \alpha u$, $v \mapsto \gamma u + \delta v$, with $\alpha^3 = 1$, $\gamma \in \mathbb{k}$, $\delta \in \mathbb{k}^\times$.

If such an automorphism lifts to X , then there exists b_1 with $b_1^2 + b_2 = 0$; hence $\mu = 0$ and $b_1 = \lambda^{1/2}u$. Comparing coefficients in the equation for σ^*a_4 , we obtain

$$\lambda^2 + \lambda^{1/2} + a\gamma + b\alpha^2\gamma^2 + c\alpha\gamma^3 = 0, \quad (10)$$

$$a + a\delta + c\alpha\delta\gamma^2 = 0, \quad (11)$$

$$b + b\alpha^2\delta^2 + c\alpha\delta^2\gamma = 0, \quad (12)$$

$$c + c\alpha\delta^3 = 0. \quad (13)$$

The automorphism lifts to X if and only if, additionally, there exists a $b_3 = \lambda_0 u^3 + \lambda_1 u^2 v + \lambda_2 uv^2 + \lambda_3 v^3$ satisfying the conditions

$$\lambda_0^2 + \lambda_0 = \lambda^3 + (a\gamma + b\alpha^2\gamma^2 + c\alpha\gamma^3)\lambda + \alpha\gamma^5 + d\gamma^6,$$

$$\lambda_1 = (a\delta + c\alpha\delta\gamma^2)\lambda + \alpha\delta\gamma^4,$$

$$\lambda_1^2 + \lambda_2 = (b\alpha^2\delta^2 + c\alpha\delta^2\gamma)\lambda + d\delta^2\gamma^4,$$

$$\lambda_3 = c\alpha\delta^3\lambda, \quad (14)$$

$$\lambda_2^2 = \alpha\delta^4\gamma + d\delta^4\gamma^2,$$

$$0 = 1 + \alpha\delta^5, \quad (15)$$

$$\lambda_3^2 = d + d\delta^6. \quad (16)$$

Equation (15) shows that $\alpha = \delta^{-5}$. In particular, as $\alpha^3 = 1$, we have $\delta^{15} = 1$.

First, assume that $\delta = 1$; hence $\alpha = 1$. Then, (16) shows that $\lambda_3 = 0$. Equation (14) shows $c\lambda = 0$ and (11) shows $c\gamma = 0$. Hence, if $c \neq 0$, then (b_2, σ) is the identity, so we assume $c = 0$ in the following. Let $G_{a,b,d}$ be the group of lifts of such automorphisms to X . By the description above, these $G_{a,b,d}$ form a family $\mathcal{G}_{a,b,d}$ of finite group schemes over $\text{Spec } \mathbb{k}[a, b, d]$ cut out in $\text{Spec } \mathbb{k}[a, b, d, \lambda, \lambda_0, \gamma]$ by the equations

$$F_1 := \lambda^4 + \lambda + a^2\gamma^2 + b^2\gamma^4 = 0, \quad (17)$$

$$F_2 := a^4\lambda^4 + b^2\lambda^2 + \gamma + d\gamma^2 + d^2\gamma^8 + \gamma^{16} = 0, \quad (18)$$

$$F_3 := \lambda_0^2 + \lambda_0 + \lambda^3 + (a\gamma + b\gamma^2)\lambda + \gamma^5 + d\gamma^6 = 0.$$

In the following, we show that all geometric fibers of $\mathcal{G}_{a,b,d} \rightarrow \text{Spec } \mathbb{k}[a, b, d]$ are reduced of length 128. In particular, $\mathcal{G}_{a,b,d}$ is étale over $\text{Spec } \mathbb{k}[a, b, d]$; hence all the $G_{a,b,d}$ are isomorphic and we will show afterwards that $G_{a,b,d} \cong 2_+^{1+6}$.

• If $a \neq 0$ and $b \neq a^2$, we argue as follows: The condition $a^8 F_1^2 + F_2^2 + b^4/a^4 F_2 = 0$ yields the following expression for λ :

$$(a^{12} + b^6)\lambda^2 = b^4\gamma + (a^4 + b^4d)\gamma^2 + (a^4d^2 + b^4d^2 + a^{16})\gamma^4 + (a^{12}b^4 + b^4d^2)\gamma^8 + (a^4d^4 + b^4)\gamma^{16} + a^4\gamma^{32}.$$

By our assumptions, we can divide by $(a^{12} + b^6)$ and we obtain an expression of λ^2 in terms of γ . Plugging this back into (18), we obtain a polynomial F in γ of the form $F = \sum_{i=0}^5 c_i \gamma^{2^i}$ of degree 64 with

$$c_0 = 0, \quad c_1 = \frac{a^{12}}{a^{12} + b^6}, \quad c_5 = \frac{a^8}{(a^{12} + b^6)^2}.$$

Since $a \neq 0$, both c_1 and c_5 are nonzero, so $\partial_\gamma F = 1$ and F has only simple roots. Hence, there are exactly 64 choices for γ such that (b_2, σ) lifts and λ is uniquely determined by γ . In particular, $G_{a,b,d}$ has order 128 and it acts on the base of the associated elliptic fibration through 2^6 .

- If $a \neq 0$ and $b = a^2$, we argue as follows: The condition $a^8 F_1^2 + F_2^2 + a^4 F_2 = 0$ becomes

$$0 = a^8 \gamma + (a^4 + a^8 d) \gamma^2 + (a^4 d^2 + a^8 d^2 + a^{16}) \gamma^4 + (a^{20} + a^8 d^2) \gamma^8 + (a^4 d^4 + a^8) \gamma^{16} + a^4 \gamma^{32} =: F.$$

Note that, since $a \neq 0$, $F_1 = F_2 = 0$ holds if and only if $F_2 = F = 0$. There are 32 choices for γ with $F(\gamma) = 0$ and for each choice of γ , there are exactly two choices for λ such that $F_2(\gamma, \lambda) = 0$. As in the previous case, $G_{a,b,d}$ has order 128, but in this case, it acts on the base of the associated elliptic fibration through 2^5 .

- Next, assume that $a = 0$ and $b \neq 0$. We can immediately solve (18) for λ and obtain

$$b^2 \lambda^2 = \gamma + d \gamma^2 + d^2 \gamma^8 + \gamma^{16}.$$

Plugging this into the square of (17), we obtain a polynomial F in γ of the form $F = \sum_{i=0}^5 c_i \gamma^{2^i}$ of degree 64 with

$$c_0 = 0, \quad c_1 = b^{-2}, \quad c_5 = b^{-8}.$$

Hence, there are 64 choices for γ such that (b_2, σ) lifts and λ is uniquely determined by γ . Therefore, $G_{a,b,d}$ has order 128 and acts on the base of the associated elliptic fibration through 2^6 .

- Now, assume that $a = b = 0$. The equations simplify to

$$\begin{aligned} \lambda^4 + \lambda &= 0, & \lambda_0^2 + \lambda_0 &= \lambda^3 + \gamma^5 + d \gamma^6, & \lambda_1 &= \gamma^4, \\ \lambda_2 &= d \gamma^4 + \gamma^8, & \gamma + d \gamma^2 + d^2 \gamma^8 + \gamma^{16} &= 0. & \lambda_3 &= 0, \end{aligned}$$

Hence, there are 16 choices for γ and 4 choices for λ . Hence, $G_{a,b,d}$ has order 128 and it acts on the base of the associated elliptic fibration through 2^4 .

It remains to determine the group $G_{a,b,d}$. By the last bullet point, the subgroup of $G_{0,0,0}$ of automorphisms that act trivially on the base of the associated elliptic fibration has order 8. Thus, by Corollary 6.3, it is isomorphic to Q_8 . Hence, every $G_{a,b,d}$ contains a quaternion group Q_8 with $\beta \in Q_8$. On the other hand, in the cases where $a \neq 0$, $b \neq a^2$, we have seen that $G_{a,b,d} / \langle \beta \rangle \cong 2^6$. Hence, by Example 6.5, we have $G_{a,b,d} \cong 2_+^{1+6}$.

Next, assume that $\delta \neq 1$, $\delta^3 = 1$. Then, (11), (12), and (13) show that $a = b = c = 0$. The remaining equations become

$$\begin{aligned} \lambda^4 + \lambda &= 0, & \lambda_0^2 + \lambda_0 &= \lambda^3 + \delta\gamma^5 + d\gamma^6, & \lambda_1 &= \delta^2\gamma^4, \\ \lambda_2 &= d\delta^2\gamma^4 + \delta\gamma^8, & \delta^2\gamma^{16} + d^2\delta\gamma^8 + d\delta\gamma^2 + \delta^2\gamma &= 0. & \lambda_3 &= 0, \end{aligned}$$

We see that if $\gamma = \lambda = 0$, then (b_2, σ) admits a lift to X as an automorphism g of order 3. For a fixed γ , there are at most 128 possible choices of (γ, λ) . All of them are obtained by composing g with an element of $G_{0,0,d}$; hence all choices are realized.

Finally, assume that $\delta \neq 1$, $\delta^5 = \alpha = 1$. As in the previous paragraph, we have $a = b = c = 0$. But in this case, (16) yields the condition $d = 0$.

So, in summary, if $c \neq 0$, then $G(X)$ is trivial and if $c = 0$, then $\text{Aut}(X)$ admits a unique 2-Sylow subgroup isomorphic to 2_+^{1+6} . If a, b , or c is nonzero, this is the full automorphism group. If $a = b = c = 0$ and $d \neq 0$, then $\text{Aut}(X)/2_+^{1+6} \cong C_3$ and if $a = b = c = d = 0$, then $\text{Aut}(X)/2_+^{1+6} \cong C_{15}$. \square

Remark 6.9. The largest order of an automorphism group of a del Pezzo surface of degree 1 over the complex numbers is equal to 144 and the surface with such a group of automorphisms is unique [Dolgachev 2012]. In our case, the maximal order is equal to $1920 = 2^7 \cdot 15$ and the surface with such an automorphism group is also unique. We also see the occurrence of the group $G = 2^4$ in case (5). It is obtained as the preimage in 2_+^{1+6} of a maximal isotropic subspace of \mathbb{F}_2^6 . Since del Pezzo surfaces of degree 1 are super-rigid (see [Dolgachev and Iskovskikh 2009, Definition 7.10, Corollary 7.11]) and the corresponding G -surface is minimal, this group is not conjugate in the Cremona group of \mathbb{P}^2 to the isomorphic subgroup of the group of automorphisms of del Pezzo surfaces of degree 4 or 2 that appeared in [Dolgachev and Duncan 2019; Dolgachev and Martin 2024].

6.3. Conjugacy classes and comparison with the classification in characteristic 0. In this section, we determine the conjugacy classes in $W(E_8)$ of the elements of the groups that occur in Theorem 6.8 and, whenever possible, compare the surfaces in Theorem 6.8 with their counterparts in characteristic 0 (see [Dolgachev 2012, Table 8.14]).

For a del Pezzo surface X of degree 1, we denoted by N_X and P_X the kernel and image of the morphism $\text{Aut}(X) \rightarrow \text{Aut}(\mathbb{P}^1)$ induced by the action of $\text{Aut}(X)$ on the base of the associated elliptic pencil.

Lemma 6.10. *Let g be a nontrivial element of N_X . Then, the conjugacy class of g is either $8A_1$, $4A_2$, $2D_4(a_1)$, or $E_8(a_8)$.*

Proof. Since g acts trivially on the base of the pencil, it cannot preserve any (-1) -curve on X . Then, the lemma follows from the classification of conjugacy classes in $W(E_8)$ (see, e.g., [Dolgachev and Martin 2024, Table 3]), by checking which of them fix no (-1) -class in E_8 . \square

Corollary 6.11. *Let X be a del Pezzo surface of degree 1 in characteristic 2. Let X' be a geometric generic fiber of a lift of X to characteristic 0 and let $\text{sp} : \text{Aut}(X') \rightarrow \text{Aut}(X)$ be the specialization map.*

(1) *sp is injective and preserves conjugacy classes.*

- (2) sp induces morphisms $N_{X'} \rightarrow N_X$ and $P_{X'} \rightarrow P_X$.
- (3) The kernel H of $P_{X'} \rightarrow P_X$ is an elementary 2-group and if g is an element of $\text{Aut}(X')$ that maps to a nontrivial element of H , then the conjugacy class of g is $2D_4(a_1)$.

Proof. The proof of claim (1), including the existence of sp , is analogous to Lemma 4.7.

The existence of the morphisms in claim (2) is clear, as, for a given lift $\mathcal{X} \rightarrow S$, the groups that appear can be defined as fibers of kernel and image of the homomorphism of S -group schemes $\text{Aut}_{\mathcal{X}/S} \rightarrow \text{Aut}_{\mathbb{P}_S^1/S}$ that describes the action of $\text{Aut}_{\mathcal{X}/S}$ on the anticanonical system.

For claim (3), recall that sp preserves conjugacy classes by claim (1). Therefore, by Lemma 6.10, all nontrivial elements of H are represented by elements g of $\text{Aut}(X')$ of conjugacy class $8A_1, 4A_2, 2D_4(a_1)$, or $E_8(a_8)$. If g is of class $8A_1$, then it is the Bertini involution; hence $g \in N_{X'}$. If g is of class $4A_2$, then it has negative trace on E_8 , so, by the Lefschetz fixed-point formula, it must act trivially on the base of the elliptic pencil. Hence, $g \in N_{X'}$. If g is of class $E_8(a_8)$, then, by what we just proved, g^2 and g^3 are in $N_{X'}$; hence $g \in N_{X'}$. Thus, g must be of conjugacy class $2D_4(a_1)$. Then, g^2 is the Bertini involution, so H is 2-elementary. □

By Theorem 6.8, $|\text{Aut}(X)| \leq 36$ or $|\text{Aut}(X)| \in \{128, 384, 1920\}$, so types I and II [Dolgachev 2012, Table 8.14] do not have a reduction modulo 2 which is a del Pezzo surface.

The surfaces of type VI, VII, IX, XII, and XV from that table admit an automorphism of order $2n$ with $n > 1$ acting faithfully on \mathbb{P}^1 , which is impossible in characteristic 2, so by Corollary 6.11 they do not have good reduction mod 2.

The equation of the surfaces of type (3v) in Theorem 6.8 can be rewritten as

$$y^2 + uv(u - v)y + x^3 + a(u^2 - uv + v^2)^3 + bu^2v^2(u - v)^2$$

for certain $a, b \in \mathbb{k}$. This equation makes sense in characteristic 0, and it is stable under the \mathfrak{S}_3 -action generated by $(u, v, x, y) \mapsto (v, u, x, -y)$ and $(u, v, x, y) \mapsto (u - v, -u, x, -y)$, as well as the C_3 -action $(u, v, x, y) \mapsto (u, v, \zeta_3x, y)$, where ζ_3 is a primitive third root of unity. Hence, the automorphism group of has order at least 36; thus it is isomorphic to $6 \times \mathfrak{S}_3$. Thus, surfaces of type (3v) are reductions mod 2 of the surfaces of type III from [Dolgachev 2012, Table 8.14]. In particular, we can read off the conjugacy classes from [Dolgachev and Martin 2024, Table 8].

The equation of type (5iii) makes sense in characteristic 0, where it is isomorphic to

$$y^2 + x^3 + u(u^5 + v^5),$$

which is the equation of type IV in [Dolgachev 2012, Table 8.14].

The equation of the surfaces of type (3ii) in Theorem 6.8 can be rewritten as

$$y^2 + uv(u - v)y + x^3 + c(u^2 - uv + v^2)x^2 + a(u^2 - uv + v^2)^3 + bu^2v^2(u - v)^2$$

for certain $a, b, c \in \mathbb{k}$. Similar to the case of type (3v) above, these equations are stable under a \mathfrak{S}_3 -action, both in characteristic 0 and in characteristic 2. In characteristic 0, these equations can be simplified to the normal forms of type X from [Dolgachev 2012, Table 8.14].

The equation of the surfaces of type (3iii) makes sense in characteristic 0, where it defines a lift of X together with the action of $\text{Aut}(X)$. Both X and the lift admit an automorphism of order 6 that acts trivially on the base of the elliptic pencil. Hence, these surfaces are reductions mod 2 of the surfaces of type XI from [Dolgachev 2012, Table 8.14].

The equations of the surfaces of type (2fi) in Theorem 6.8 define a 1-dimensional family of surfaces in characteristic 0 with an action of C_{10} . These lifts must be of type XIII [Dolgachev 2012, Table 8.14].

The equations of the surfaces of type (4i) in Theorem 6.8 define a 2-dimensional family of surfaces in characteristic 0 with an action of C_6 that is trivial on the base of the elliptic pencil. Hence, these lifts are of type XVII [Dolgachev 2012, Table 8.14].

Next, consider the equations

$$y^2 + (aux + bu^3 + cv^3)y + x^3 + (du^4 + euv^3)x + fu^6 + gu^3v^3 + hv^6,$$

where a, b, c, d, e, f, g, h are parameters. In characteristic 0, we can simplify this equation to the normal form of type XVIII from [Dolgachev 2012, Table 8.14]. In characteristic 2, these equations cover three of the families of Theorem 6.8: If $a, c \neq 0$, we can simplify the equation to the normal form for type (1ai) which, in turn, specializes to type (1ei) for special values of the parameters. If $a = 0$ but $b, c \neq 0$, we can simplify the equation to

$$y^2 + (u^3 + v^3)y + x^3 + euv^3x + fu^6.$$

This is an alternative normal form for our surfaces of type (3iv).

Finally, consider the equations

$$y^2 + (a(u+v)x + b(u+v)^3 + cuv(u+v))y + x^3 + (d(u+v)^4 + euv(u+v)^2 + fu^2v^2)x \\ + (g(u+v)^6 + huv(u+v)^4 + iu^2v^2(u+v)^2 + ju^3v^3).$$

In characteristic 0, we can simplify this equation to the normal form of type XX from [Dolgachev 2012, Table 8.14]. In characteristic 2, these equations cover four of the families of Theorem 6.8:

If $a \neq 0$, we can simplify the equation to

$$y^2 + (u+v)xy + x^3 + cuv^2x + g(u+v)^6 + huv(u+v)^4 + iu^2v^2(u+v)^2 + ju^3v^3.$$

If $d, j \neq 0$, we can rescale one of them to 1 and obtain an alternative normal form for type (2ai). If $d \neq 0$ and $j = 0$, we obtain a normal form for type (2ei). If $j \neq 0$ and $d = 0$, we obtain an alternative normal form for type (2di). Note that $d = j = 0$ would lead to a singular surface. Since the family of type (2di) occurs as a reduction mod 2 of certain surfaces of type XX from [Dolgachev 2012, Table 8.14], we call them type XX'.

If $a = 0$ and $c \neq 0$, we can simplify the equation to

$$y^2 + (b(u+v)^3 + uv(u+v))y + x^3 + (euv(u+v)^2 + fu^2v^2)x + (g(u+v)^6 + ju^3v^3).$$

This defines a 4-dimensional family of surfaces with 2^2 -action (one parameter is redundant). By Theorem 6.8, the corresponding surfaces must be of type (3i).

The surfaces in the families (1ai), (1aiii), (1aiv), (1di), (5i), (5ii), and (5iii) admit an automorphism of order 4 and it turns out that writing down integral equations for such automorphisms similar to the ones above is hard. So, instead, to determine the conjugacy classes of the automorphisms of this family and to compare with the classification in characteristic 0, we will use the following observation.

Lemma 6.12. *Let g be an automorphism of a del Pezzo surface of degree 1. Let $m := \text{ord}(g)$ and let n be the order of the induced automorphism of \mathbb{P}^1 . Assume that m is even. Then, the conjugacy class Γ of g in $W(E_8)$ is one of the following:*

- (1) If $(m, n) = (2, 1)$, then $\Gamma = 8A_1$.
- (2) If $(m, n) = (2, 2)$, then $\Gamma = 4A_1$.
- (3) If $m = 4$, then $\Gamma = 2D_4(a_1)$.
- (4) If $(m, n) = (6, 1)$, then $\Gamma = E_8(a_8)$.
- (5) If $(m, n) = (6, 2)$, then $\Gamma = E_6(a_2) + A_2$.
- (6) If $(m, n) = (6, 3)$ and g^2 is of class $3A_2$, then $\Gamma = E_7(a_4) + A_1$.
- (7) If $(m, n) = (6, 3)$ and g^2 is of class $2A_2$, then $\Gamma = 2D_4$.
- (8) If $m = 10$, then $\Gamma = E_8(a_6)$.
- (9) If $m = 12$, then $\Gamma = E_8(a_3)$.
- (10) If $m = 20$, then $\Gamma = E_8(a_2)$.
- (11) If $m = 30$, then $\Gamma = E_8$.

Proof. By Theorem 6.8, we know that the only possible values for m and n are the ones in the statement.

In case (1), g is the Bertini involution; hence $\Gamma = 8A_1$. In case (2), we may assume that g acts as $u \leftrightarrow v$. Then, we proved in this section that g lifts to characteristic 0, so by [Dolgachev and Martin 2024, Table 8], $\Gamma = 4A_1$. In case (3), g^2 is the Bertini involution, because $\text{PGL}_2(\mathbb{k})$ does not contain elements of order 4, and it is known (see [loc. cit., Table 3]) that the only conjugacy class of automorphisms of order 4 whose square is the Bertini involution is $2D_4(a_1)$. In cases (8) and (11), $g^{m/2}$ is the Bertini involution and g^2 lifts to characteristic 0; hence g lifts to characteristic 0 and we can read off the conjugacy class Γ from [loc. cit., Table 8]. Then, we deduce case (10) from case (8). In case (9), g^2 must be of type $E_8(a_8)$, since $\text{PGL}_2(\mathbb{k})$ does not contain any elements of order 4 or 6. Then, from [loc. cit., Table 3], we see that $\Gamma = E_8(a_3)$. Finally, cases (4), (5), (6), and (7) follow from [loc. cit., Table 3] by comparing the conjugacy classes of g^2 and g^3 . \square

Now, we can complete Table 9 by using the description of $\text{Aut}(X)$ in Theorem 6.8. We observe that the conjugacy classes for types (1ai) and (1di) are the same as for type XIX from [Dolgachev 2012, Table 8.14], the conjugacy classes for type (1aiii) are the same as for type XIV from [loc. cit., Table 8.14], and the conjugacy classes for type (1aiv) are the same as for type V from [loc. cit., Table 8.14]. The only groups in Theorem 6.8 that contain D_8 are 2_+^{1+6} , $2_+^{1+6} : 3$, and $2_+^{1+6} : 15$, and the only group that contains an

name	Aut(X) order	id	4A ₁	8A ₁	2A ₂	3A ₂	4A ₂	2A ₃ +A ₁	2A ₄	A ₅ +A ₃ +A ₁	D ₄ (a ₁)+A ₁	2D ₄ (a ₁)	D ₈ (a ₃)	E ₆ +A ₁
I-II	(do not exist)													
III/(3v)	6 × D ₆	36	1	6	1	2	4	2				2		
IV/(5iii)	2 ¹⁺⁶ :15	1920	1	70	1		8						56	
M/(5ii)	2 ¹⁺⁶ :3	384	1	70	1		8	64					56	
V/(1aiv)	SL ₂ (3)	24	1	1	1	8							6	
VI-VII	(do not exist)													
VIII	(same as IV)													
IX	(does not exist)													
X/(3ii)	D ₁₂	12	1	6	1	2					2			
XI/(3iii)	2 × 6	12	1	2	1		2							
XII	(does not exist)													
XIII/(2fi)	10	10	1	1	1			4					6	
XIV/(1aiii)	Q ₈	8	1	1	1									
XV	(does not exist)													
XVI/(5i)	2 ¹⁺⁶	128	1	70	1								56	
XVII/(4i)	6	6	1	1	1		2							
XVIII/(1aii), (1ei), (3iv)	6	6	1	1	1									
XIX/(1ai), (1di)	4	4	1	1	1								2	
XX/(2ai), (2ei), (3i)	2 ²	4	1	2	1									
XX'/(2di)	2 ⁴	16	1	14	1									
XXI	2	2	1	1	1									

name	Aut(X) order	E ₆ (a ₂)	E ₆ (a ₂)	E ₆ (a ₂)+A ₂	E ₇ (a ₂)	E ₇ (a ₂)+A ₁	E ₈	E ₈ (a ₁)	E ₈ (a ₂)	E ₈ (a ₃)	E ₈ (a ₅)	E ₈ (a ₆)	E ₈ (a ₈)
I-II	(do not exist)												
III/(3v)	6 × D ₆	36		12		4							2
IV/(5iii)	2 ¹⁺⁶ :15	1920		80			512		384	160	512	64	8
M/(5ii)	2 ¹⁺⁶ :3	384		80					160	160			8
V/(1aiv)	SL ₂ (3)	24				8							
VI-VII	(do not exist)												
VIII	(same as IV)												
IX	(does not exist)												
X/(3ii)	D ₁₂	12											
XI/(3iii)	2 × 6	12											
XII	(does not exist)												
XIII/(2fi)	10	10											
XIV/(1aiii)	Q ₈	8										4	
XV	(does not exist)												
XVI/(5i)	2 ¹⁺⁶	128											
XVII/(4i)	6	6											
XVIII/(1aii), (1ei), (3iv)	6	6											2
XIX/(1ai), (1di)	4	4											
XX/(2ai), (2ei), (3i)	2 ²	4											
XX'/(2di)	2 ⁴	16											
XXI	2	2											

Table 9. Automorphism groups of del Pezzo surfaces of degree 1.

automorphism of order 20 is $2_+^{1+6} : 15$. Hence, if the types XVI, M, and VIII from [loc. cit., Table 8.14] and [Dolgachev and Martin 2024, Table 8] have good reduction modulo 2, then they must reduce to our types (5i) and (5ii), respectively. In each of these cases, we determine the conjugacy classes using Lemma 6.12.

We summarize the classification of automorphism groups of del Pezzo surfaces of degree 1 in Table 9. In the first column, we give the name of the corresponding family, both in the notation of Theorem 4.3 and in the notation of [Dolgachev 2012, Table 8.14]. The second and third columns give the group $\text{Aut}(X)$ and its size. In the remaining columns, we list the number of elements of a given Carter conjugacy class in $\text{Aut}(X)$.

References

- [Aschbacher 2000] M. Aschbacher, *Finite group theory*, 2nd ed., Cambridge Stud. Adv. Math. **10**, Cambridge Univ. Press, 2000. MR Zbl
- [Barth 1977] W. Barth, “Moduli of vector bundles on the projective plane”, *Invent. Math.* **42** (1977), 63–91. MR Zbl
- [Beauville 1977] A. Beauville, “Prym varieties and the Schottky problem”, *Invent. Math.* **41**:2 (1977), 149–196. MR Zbl
- [Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Math. (3) **21**, Springer, 1990. MR Zbl
- [Carter 1972] R. W. Carter, “Conjugacy classes in the Weyl group”, *Compos. Math.* **25**:1 (1972), 1–59. MR Zbl
- [Conway et al. 1985] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups: maximal subgroups and ordinary characters for simple groups*, Oxford Univ. Press, 1985. MR Zbl
- [Demazure 1980] M. Demazure, “Surfaces de del Pezzo, I–V”, pp. 21–69 in *Séminaire sur les singularités des surfaces* (Palaiseau, 1976–1977), edited by M. Demazure et al., Lecture Notes in Math. **777**, Springer, 1980. Zbl
- [Dolgachev 2012] I. V. Dolgachev, *Classical algebraic geometry: a modern view*, Cambridge Univ. Press, 2012. MR Zbl
- [Dolgachev and Duncan 2019] I. Dolgachev and A. Duncan, “Automorphisms of cubic surfaces in positive characteristic”, *Izv. Ross. Akad. Nauk Ser. Mat.* **83**:3 (2019), 15–92. MR Zbl
- [Dolgachev and Iskovskikh 2009] I. V. Dolgachev and V. A. Iskovskikh, “Finite subgroups of the plane Cremona group”, pp. 443–548 in *Algebra, arithmetic, and geometry, I*, Progr. Math. **269**, Birkhäuser, Boston, MA, 2009. MR Zbl
- [Dolgachev and Martin 2024] I. Dolgachev and G. Martin, “Automorphisms of del Pezzo surfaces in odd characteristic”, *J. Lond. Math. Soc.* (2) **109**:5 (2024), art. id. e12905. MR Zbl
- [Ekedahl 1988] T. Ekedahl, “Canonical models of surfaces of general type in positive characteristic”, *Inst. Hautes Études Sci. Publ. Math.* **67** (1988), 97–144. MR Zbl
- [Katz and Sarnak 1999] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, Amer. Math. Soc. Colloq. Publ. **45**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl
- [Lang 2000] W. E. Lang, “Configurations of singular fibres on rational elliptic surfaces in characteristic two”, *Comm. Algebra* **28**:12 (2000), 5813–5836. MR Zbl
- [Neusel and Smith 2002] M. D. Neusel and L. Smith, *Invariant theory of finite groups*, Math. Surv. Monogr. **94**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl
- [Oguiso and Shioda 1991] K. Oguiso and T. Shioda, “The Mordell–Weil lattice of a rational elliptic surface”, *Comment. Math. Univ. St. Paul.* **40**:1 (1991), 83–99. MR Zbl

Communicated by Gavril Farkas

Received 2023-05-28 Revised 2024-02-14 Accepted 2024-03-29

idolga@umich.edu

Department of Mathematics, University of Michigan, Ann Arbor, MI, United States

gmartin@math.uni-bonn.de

Mathematisches Institut, Universität Bonn, Bonn, Germany

On the \mathcal{D} -module of an isolated singularity

Thomas Bitoun

Let Z be the germ of a complex hypersurface isolated singularity of equation f , with Z at least of dimension 2. We consider the family of analytic \mathcal{D} -modules generated by the powers of $1/f$ and describe it in terms of the pole order filtration on the de Rham cohomology of the complement of $\{f = 0\}$ in the neighbourhood of the singularity.

1. Introduction

The \mathcal{D} -modules generated by powers of a polynomial (or analytic function) f have been the topic of several noted publications in the last decade, for example, [Bitoun and Schedler 2018; Mustața and Olano 2023; Saito 2021; 2022]. On the one hand, they are elementary objects accessible to beginners in \mathcal{D} -module theory. On the other hand, they relate to analytic invariants and Hodge theory in deep and subtle ways.

This note provides a new, elementary approach to describing these \mathcal{D} -modules in the general isolated singularity case in terms of the pole order filtration on the de Rham cohomology of the algebraic link of the singularity.

Our results include:

- A new approach to, and a new proof of, Vilonen's characterization of the intersection cohomology \mathcal{D} -module [Vilonen 1985, Theorem], presented in Theorem 2.2.4 and Remark 2.2.7.
- A new computation of the length of the \mathcal{D} -module of meromorphic functions (Theorem 2.2.4).
- An elementary description of the Hodge structure of the \mathcal{D} -module of meromorphic functions (Theorem 2.2.4).
- A full solution to the question of the length of the \mathcal{D} -module generated by $1/f$ and of the corresponding Poisson cohomology (see [Bitoun and Schedler 2018; Etingof and Schedler 2014, Conjecture 3.8]) in Corollary 3.0.4.
- Connections between top-forms decompositions with prescribed pole order and the \mathcal{D} -modules generated by a power of $1/f$ via generalizations of Vilonen's theorem, described in Corollary 3.0.1.
- Demonstrating the importance of \mathcal{D} -submodules generated by pieces of the Hodge or pole order filtrations, first considered in [Mustața and Olano 2023] and studied in Corollary 2.2.6 and Theorem 2.2.8.

MSC2020: primary 14F10; secondary 14B05.

Keywords: \mathcal{D} -modules, isolated singularities, Hodge filtration, de Rham cohomology, mixed Hodge structures.

© 2025 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via [Subscribe to Open](#).

- Explaining the failure of the conjecture from [Bitoun and Schedler 2018] (equivalent in terms of the Hamiltonian flow to [Etingof and Schedler 2014, Conjecture 3.8]), as first noted in [Mustață and Olano 2023] (see also [Saito 2022]). This is discussed at the end of the introduction.

Finally, we note that our approach has already led to new results; see, e.g., the updates to [Saito 2022].

We now describe the contents in more technical detail. Let f be a complex analytic function in n variables, $n \geq 3$, and assume that $Z := \{f = 0\}$ is reduced and has an isolated singularity at o . Our main tool is the product pairing in the neighbourhood of the singularity between the meromorphic functions with poles along Z and the top regular forms, with values in the n -th de Rham cohomology group H' of the complement of Z .

Let δ_o be the irreducible \mathcal{D} -module supported at o . In Theorem 2.2.4, we show that the pairing can be interpreted as a \mathcal{D} -module map $r : \mathcal{O}(*Z)_o \rightarrow \delta_o \otimes H'$, which is surjective with kernel equal to the intersection homology \mathcal{D} -module \mathcal{L}_o . The latter relies on Vilonen's characterization of that \mathcal{D} -module, of which our result can be viewed as a generalization (e.g., Corollary 3.0.1; see also Remark 2.2.7). The morphism r is especially convenient to study the \mathcal{D} -submodules of $\mathcal{O}(*Z)_o$ generated by powers of $1/f$ or by the pieces of Hodge filtration; see Corollaries 2.2.6 and 3.0.1.

Using the above, [Mustață and Olano 2023] implies that the dimension of the first piece $F_0 H'$ of the Hodge filtration is the reduced genus g of [Bitoun and Schedler 2018], while the length of $\mathcal{D}(1/f)/\mathcal{L}_o$ is $\dim P_0 H'$, where $P_0 H'$ is the set of classes generated by forms with pole order at most 1 along Z . However it is well known that the pole order filtration is, in general, strictly greater than the former; see, e.g., [Dimca 1991, 5.4 ii; Karpishpan 1991, (b) of Theorem 0.3]. This explains the failure of [Bitoun and Schedler 2018, Conjecture 1.7]. Finally, let us note that even though the natural language of the note is that of analytic \mathcal{D} -modules, we deduce results on length in the algebraic case as well; see Corollary 3.0.4.

2. A morphism of \mathcal{D} -modules

2.1. Setup and conventions. Let $n \geq 3$. Let X be a complex analytic manifold of dimension n , let Z be a hypersurface of X that has isolated singularities and let $U \subset X$ be the open complement of Z . By restricting to an open neighbourhood of a singularity we may assume that Z has a unique singularity o , which we do from now on. By a \mathcal{D} -module, we mean a left coherent, analytic D_X -module, and by a D_o -module, we mean a finitely generated left module over the stalk of D_X at the point o . For a holonomic \mathcal{D} -module M , we let $DR^l(M)$ be the l -th cohomology group of the de Rham complex of M . We use the same notation $DR^l(N)$ for N a holonomic D_o -module. We let $\mathcal{O}(*Z)$ be the \mathcal{D} -module of meromorphic functions on X with poles along Z .

2.2. Construction. Let us first recall standard facts.

Let k be a field, let B be an arbitrary k -algebra and let M be a right (resp. left) B -module. Then for an arbitrary k -vector space W , the space of k -linear maps $L(M, W)$ from M to W is a left (resp. right) B -module, where the action is given by $bf(m) := f(mb)$ (resp. $fb(m) := f(bm)$), for all $b \in B$, $m \in M$, $f \in L(M, W)$. In the following lemma, we apply this to the right D_0 -module Ω_o^n for $k = \mathbb{C}$.

Lemma 2.2.1. *Let V be a finite-dimensional complex vector space and let o be a point of X . For Ω_o^n the stalk at o of the sheaf of differential n -forms, the space of linear maps $L(\Omega_o^n, V)$ from the right D_o -module Ω_o^n to V is naturally a D_o -module. Moreover the D_o -submodule $L(\Omega_o^n, V)_o$ of linear maps annihilated by a power of the ideal of o is isomorphic to the D_o -module $\delta_o \otimes_{\mathbb{C}} V$, where δ_o is the irreducible D_o -module supported at o .*

Proof. Only the last part needs further proof. Under Kashiwara’s equivalence [Björk 1993, Lemma 2.6.18], a holonomic left D -module M supported at o corresponds to the finite-dimensional vector space $M/(\mathfrak{m}M)$, where \mathfrak{m} is the ideal of o , and $M \simeq \delta_o \otimes_{\mathbb{C}} M/(\mathfrak{m}M)$. Therefore, the D -module $L(\Omega_o^n, V)_o$ corresponds to the vector space of linear maps from $\Omega_o^n/\mathfrak{m}\Omega_o^n \simeq \mathbb{C}$ to V , which is isomorphic to V . Hence the existence of an isomorphism $L(\Omega_o^n, V)_o \simeq \delta_o \otimes_{\mathbb{C}} V$. \square

We will use the following lemma. For an element λ of $L(\Omega_o^n, V)$, we denote by $\text{Im}(\lambda)$ the image of the corresponding linear map $\Omega_o^n \rightarrow V$.

Lemma 2.2.2. *Let V be a finite-dimensional complex vector space and let N be a D_o -submodule of $L(\Omega_o^n, V)_o$. Assume that for all $v \in V$, there exists an element λ_v of N such that $v \in \text{Im}(\lambda_v)$. Then $N = L(\Omega_o^n, V)_o$.*

Proof. By Lemma 2.2.1 and Kashiwara’s equivalence [Björk 1993, Lemma 2.6.18], $N = L(\Omega_o^n, V')_o$ for a vector subspace V' of V . Thus if $v \in V \setminus V'$, then for all $\lambda \in N = L(\Omega_o^n, V')_o$, $v \notin \text{Im}(\lambda)$. This contradicts the assumption on N . Hence $N = L(\Omega_o^n, V)_o$. \square

In our setup 2.1, we denote by \mathcal{L} the D -module preimage in $\mathcal{O}(\star Z)$ of the intersection cohomology D -module $\mathcal{L}_Z \subseteq \mathcal{O}(\star Z)/\mathcal{O}$ associated with Z . We now recall Vilonen’s description of the intersection homology D -module in terms of residues.

Theorem 2.2.3 (Vilonen). *An element s of the stalk $\mathcal{O}(\star Z)_o$ is in the stalk \mathcal{L}_o if and only if $\forall \omega' \in \Omega_o^n$, $s\omega'$ is exact, i.e.,*

$$\mathcal{L}_o = \{s \in \mathcal{O}(\star Z)_o \mid \text{for all } \omega' \in \Omega_o^n, s\omega' \in d(\Omega_o^{n-1}(\star Z))\}.$$

Proof. This is a reformulation of [Vilonen 1985, Theorem]; see [Björk 1993, 5.7.21] for a textbook treatment. We include the proof below for the benefit of the reader. Let

$$V := \{s \in \mathcal{O}(\star Z)_o \mid \text{for all } \omega' \in \Omega_o^n, s\omega' \in d(\Omega_o^{n-1}(\star Z))\}.$$

It follows by a special case of the argument given in the proof of Theorem 2.2.4 below that V is a D_o -submodule. We want to prove that $V = \mathcal{L}_o$.

Let us first show that $\mathcal{L}_o \subseteq V$. Note that for $N \subseteq \mathcal{O}(\star Z)_o$ a D_o -submodule, $N \subseteq V$ if and only if the image of $DR^n(N)$ in the de Rham cohomology group $DR^n(\mathcal{O}(\star Z)_o)$ vanishes. But $DR^n(\mathcal{L}_o) = DR^n(\mathcal{L})_o = 0$ by [Björk 1993, Lemma 5.7.18]; hence $\mathcal{L}_o \subseteq V$. Let us now show that $\mathcal{L}_o = V$. The quotient V/\mathcal{L}_o is supported at the singularity since it is the case for $\mathcal{O}(\star Z)_o/\mathcal{L}_o$. Therefore $V/\mathcal{L}_o \simeq \delta_o^j$ for some $j \geq 0$. By the long exact sequence of the DR^i ’s applied to the short exact sequence $0 \rightarrow V \rightarrow \mathcal{O}(\star Z)_o \rightarrow \mathcal{O}(\star Z)_o/V \rightarrow 0$, we have that the natural map $DR^n(V) \rightarrow DR^n(\mathcal{O}(\star Z)_o)$ is an injection,

because $DR^{n-1}(\delta_o) = 0$. Hence $DR^n(V) = 0$ by the definition of V . But using the long exact sequence of the DR^i 's associated with the short exact sequence $0 \rightarrow \mathcal{L}_o \rightarrow V \rightarrow \delta_o^j \rightarrow 0$, we deduce from $DR^{n-1}(\delta_o) = DR^{n+1}(\mathcal{L}_o) = 0$ and $DR^n(\delta_o) = \mathbb{C}$ that

$$\mathbb{C}^j \simeq \frac{DR^n(V)}{DR^n(\mathcal{L}_o)}.$$

Since the latter vanishes, we must have $j = 0$ and $\mathcal{L}_o = V$. □

Let us now prove the main theorems of this note. Note that the de Rham cohomology $DR^n(\mathcal{O}(\star Z)_o)$ is endowed with a Hodge structure, which we denote by H' .

Theorem 2.2.4. *Under the hypotheses in Section 2.1, the pairing*

$$\mathcal{O}(\star Z)_o \times \Omega_o^n \xrightarrow{B} H', \quad (s, \omega') \mapsto [s\omega'],$$

where $[-]$ is the cohomology class of a form, induces a surjective homomorphism of D_o -modules

$$\mathcal{O}(\star Z)_o \xrightarrow{r} L(\Omega_o^n, H')_o, \quad s \mapsto B(s, -),$$

where o is the singularity. This homomorphism is compatible with the Hodge filtrations, where the Hodge filtration on $L(\Omega_o^n, H')_o$ is the one induced by that of H' under Kashiwara's equivalence for Hodge D -modules. The kernel of r is the D_o -module \mathcal{L}_o .

Proof. Let us first show that the map

$$\mathcal{O}(\star Z)_o \rightarrow L(\Omega_o^n, DR^n(\mathcal{O}(\star Z)_o)), \quad s \mapsto (\omega' \mapsto [s\omega']),$$

defines a morphism of D_o -modules and takes its values in $L(\Omega_o^n, DR^n(\mathcal{O}(\star Z)_o))_o$.

It follows directly from the definitions that the map is \mathcal{O} -linear. Moreover, the fact that the class of an exact form in $DR^n(\mathcal{O}(\star Z)_o)$ vanishes implies that r is compatible with the actions of derivations. We may restrict ourselves to verifying it for the action of the partials $(\partial_i)_i$ corresponding to coordinates $(x_i)_i$. Let ω be a volume form in the neighbourhood of o and let $\omega^{(i)}$ be an $(n-1)$ -form such that $dx_i \wedge \omega^{(i)} = \omega$. Then $s\omega' = sg\omega$ for some holomorphic function g and $d(sg\omega^{(i)}) = \partial_i(s)\omega + s\partial_i(g)\omega$. That is,

$$\partial_i s \mapsto (g\omega \mapsto [\partial_i(s)g\omega] = -[s\partial_i(g)\omega]).$$

Hence the map is compatible with the right D_o -module action on Ω_o^n . Therefore we have a morphism of D_o -modules $\mathcal{O}(\star Z)_o \rightarrow L(\Omega_o^n, DR^n(\mathcal{O}(\star Z)_o))$. Note that by Theorem 2.2.3, the kernel of this morphism is \mathcal{L}_o . But $\mathcal{O}(\star Z)_o/\mathcal{L}_o$ is supported at o ; hence r factors through $L(\Omega_o^n, DR^n(\mathcal{O}(\star Z)_o))_o$.

That r is surjective follows directly from Lemma 2.2.2. Finally, the compatibility of r with the Hodge filtrations is a direct consequence of the construction of the Hodge filtration on the de Rham complex. □

Remark 2.2.5. Letting Z_∞ be the Milnor fibre at o , we note that for $H := H^{n-1}(Z_\infty)_1$ the unipotent monodromy part of the cohomology group of the Milnor fibre and N the logarithm of the unipotent part of the monodromy, we have a natural identification of mixed Hodge structures $\gamma : H' \simeq H/(NH)$. This

follows from applying DR^n to the short exact sequence $0 \rightarrow M_f \simeq \mathcal{L} \rightarrow M'_f \simeq \mathcal{O}(\star Z)/\mathcal{O} \rightarrow M''_f/M_f \rightarrow 0$ of [Saito 2009, Remarks 3.2i] and using the isomorphisms [Saito 2009, 3.2.5 and 3.2.4].

As a direct corollary, we get the following.

Corollary 2.2.6. *The image by r of the \mathcal{D} -submodule $\mathcal{D}_o F_l \mathcal{O}(\star Z)_o$ generated by the l -th piece of the Hodge filtration on $\mathcal{O}(\star Z)_o$ is $L(\Omega_o^n, F_l H')_o$, where $F_l H'$ is the l -th piece of the Hodge filtration on H' . Therefore the length of $\mathcal{D}_o F_l \mathcal{O}(\star Z)_o/\mathcal{L}_o$ is $\dim F_l H'$.*

Proof. Since r is a surjective morphism of Hodge D_o -modules by Theorem 2.2.4, we have $r(F_l \mathcal{O}(\star Z)_o) = \sum_{i+j \leq l} G_i \delta_o \otimes F_j H'$, where G is the good filtration on δ_o induced by the standard generator of δ_o and the usual good filtration of \mathcal{D}_o ; see, e.g., [Saito 2009, 1.5.3]. But the submodules $\mathcal{D}_o(\sum_{i+j \leq l} G_i \delta_o \otimes F_j H')$ and $\mathcal{D}_o(G_o \otimes F_l H') = L(\Omega_o^n, F_l H')_o$ are equal. Indeed the \mathcal{D}_o -module structure on $L(\Omega_o^n, H')_o \simeq \delta_o \otimes H'$ is such that \mathcal{D}_o acts only on the left factor δ_o , which is generated by $G_o \delta_o$. The equality follows. Therefore $r(\mathcal{D}_o F_l \mathcal{O}(\star Z)_o) = \mathcal{D}_o r(F_l \mathcal{O}(\star Z)_o) = \mathcal{D}_o(\sum_{i+j \leq l} G_i \delta_o \otimes F_j H') = L(\Omega_o^n, F_l H')_o$, as stated. \square

Remark 2.2.7. We draw the reader's attention to the fact that Theorem 2.2.4 can be thought of as providing a new proof of Vilonen's theorem. Namely, we know that r is surjective. So if we accept the elementary fact that $\mathcal{O}(\star Z)_o$ is of length $1 + \dim H'$ [Björk 1993, 5.7.17], we must have that the kernel of r is \mathcal{L}_o . But the kernel of r exactly matches the description in Vilonen's theorem.

We now consider D_o -submodules generated by an \mathcal{O}_o -submodule of $\mathcal{O}(\star Z)_o$. For M an \mathcal{O}_o -submodule of $\mathcal{O}(\star Z)_o$, we set $\Omega_o^n(M) = M \otimes_{\mathcal{O}_o} \Omega_o^n$ and let $[\Omega_o^n(M)] \subseteq DR^n(\mathcal{O}(\star Z)_o)$ be the vector subspace of the corresponding classes of forms, namely the classes that can be represented as mw' , for some $m \in M$ and $w' \in \Omega_o^n$.

Theorem 2.2.8. *Let M be an \mathcal{O}_o -submodule of $\mathcal{O}(\star Z)_o$ and let $D_o M$ be the D_o -submodule of $\mathcal{O}(\star Z)_o$ generated by M . Assume that $\mathcal{O}(\star Z)_o$ and $D_o M$ agree generically on Z . Then the quotient $D_o M/\mathcal{L}_o$ is isomorphic to $L(\Omega_o^n, [\Omega_o^n(M)])_o$. Moreover,*

$$D_o M = \{s \in \mathcal{O}(\star Z)_o \mid \text{for all } \omega' \in \Omega_o^n, s\omega' \in M \otimes_{\mathcal{O}_o} \Omega_o^n + d(\Omega_o^{n-1}(\star Z))\}.$$

Proof. Since \mathcal{L}_o is the minimal extension and $D_o M$ extends $\mathcal{O}(\star Z)_o$ generically on Z , $D_o M$ contains \mathcal{L}_o . It thus makes sense to consider the quotient $D_o M/\mathcal{L}_o$. We claim that $r(D_o M) = L(\Omega_o^n, [\Omega_o^n(M)])_o$, where r is the morphism from Theorem 2.2.4. Indeed, we have $M \subseteq r^{-1}(L(\Omega_o^n, [\Omega_o^n(M)])_o)$. Therefore $D_o M \subseteq r^{-1}(L(\Omega_o^n, [\Omega_o^n(M)])_o)$, because $r^{-1}(L(\Omega_o^n, [\Omega_o^n(M)])_o)$ is a D_o -module containing M , and hence $r(D_o M) \subseteq L(\Omega_o^n, [\Omega_o^n(M)])_o$. We then note that the equality $r(D_o M) = L(\Omega_o^n, [\Omega_o^n(M)])_o$ follows immediately from Lemma 2.2.2. \square

3. Some corollaries

In what follows, let us consider the filtration P by order of the pole on the de Rham cohomology $DR^n(\mathcal{O}(\star Z)_o) \simeq H'$. Namely, we let $P_l H'$ be the subspace of the classes that can be represented, via the isomorphism above, by forms having a pole of order at most $l + 1$ along Z .

Corollary 3.0.1. *Let f be a local equation of Z , and let $l \geq 0$. We have the following description of $D_o(1/f^{l+1})$, the D_o -submodule of $\mathcal{O}(\star Z)_o$ generated by $1/f^{l+1}$:*

$$D_o \frac{1}{f^{l+1}} = \{s \in \mathcal{O}(\star Z)_o \mid \text{for all } \omega' \in \Omega_o^n, s\omega' \in \Omega_o^n((l+1)Z) + d(\Omega_o^{n-1}(\star Z))\}.$$

It follows that the D_o -module length of the quotient $D_o(1/f^{l+1})/\mathcal{L}_o$ is $\dim_{\mathbb{C}} P_l H'$.

Proof. Apply Theorem 2.2.8 to $M = \mathcal{O}_o/f^{l+1}$. It then follows that the quotient $D_o(1/f^{l+1})/\mathcal{L}_o = D_o(\mathcal{O}_o/f^{l+1})/\mathcal{L}_o$ is isomorphic to

$$L\left(\Omega_o^n, \left[\Omega_o^n\left(\frac{\mathcal{O}_o}{f^{l+1}}\right)\right]\right)_o = L(\Omega_o^n, P_l H')_o.$$

Since the latter is isomorphic to $\delta_o \otimes_{\mathbb{C}} P_l H'$ by Lemma 2.2.1, the length assertion is proved. □

Therefore we deduce the following properties, first proved in [Mustařa and Olano 2023, Theorems 1.1 and 1.3], from those of the pole order filtration.

Corollary 3.0.2. *Recall the hypotheses in Section 2.1.*

- (1) *The D_o -module length of the quotient $D_o(1/f^{l+1})/\mathcal{L}_o$ is at least $\dim_{\mathbb{C}} F_l H'$.*
- (2) *If Z is quasihomogeneous, then the inequality from 1 is an equality.*

Proof. Since the D_o -module length of the quotient $D_o(1/f^{l+1})/\mathcal{L}_o$ is $\dim_{\mathbb{C}} P_l H'$ by Corollary 3.0.1, assertions (1) and (2) follow from Theorems (b) and (a) of [Karpishpan 1991], respectively. □

Remark 3.0.3. We note that, conversely, any result on the length of $D_o(1/f^{l+1})/\mathcal{L}_o$ transfers by Corollary 3.0.1 to a statement about the pole order filtration. For example, [Saito 2022, Theorem 1] describes those lengths in terms of the Gauss–Manin connection (compare with [Karpishpan 1991, Theorem (c)]). Moreover, [Mustařa and Olano 2023, §5; Saito 2022, 3.2 Example I] provide new examples where the Hodge filtration is strictly contained in the pole order filtration.

We also obtain results for algebraic D -modules. Let $n \geq 3$ and let g be a complex polynomial in n variables defining a reduced irreducible hypersurface Y with an isolated singularity at the origin, i.e., $|Y^{\text{sing}}| = 1$. Then for all $l \geq 0$, we denote by $D^{\text{alg}}(1/g^{l+1})$ the left D^{alg} -submodule of $R[1/g]$ generated by $1/g^{l+1}$, where R is the ring of complex polynomials in n variables and D^{alg} is the n -th Weyl algebra $A_n(\mathbb{C})$. We let IC be the D^{alg} -module preimage in $R[1/g]$ of the intersection cohomology D^{alg} -module IC_Y .

Corollary 3.0.4. *The quotient D^{alg} -module $D^{\text{alg}}(1/g^{l+1})/(IC)$ is of length $\dim_{\mathbb{C}} P_l H_{dR}^n(B \setminus Y)$, where P_l is the pole order filtration of the de Rham cohomology of the complement of Y in a small analytic ball B centred at the origin.*

Proof. Using the notation of Section 2.1, the analytification functor $D_{\mathbb{C}^n} \otimes_{D^{\text{alg}}} -$ is an equivalence between the category of regular holonomic D^{alg} -modules and a full subcategory of regular $D_{\mathbb{C}^n}$ -modules [Brylinski 1986, Proposition 7.8]. It follows directly from the definition that the analytification of the regular holonomic D^{alg} -module $R[1/g]$ is the sheaf of meromorphic functions $\mathcal{O}(\star Y^{\text{an}})$, and the analytification of

$D^{\text{alg}}(1/g^{l+1})$ is $D_{\mathbb{C}^n}(1/g^{l+1})$. Moreover, because the analytification is an equivalence, the minimality of IC and \mathcal{L} force them to correspond to each other under the analytification functor. But, as $D_{\mathbb{C}^n}(1/g^{l+1})/\mathcal{L}$ is supported at the origin, its length is the same as that of its stalk at the origin. But the natural map from $H_{dR}^n(B \setminus Y)$ to $H' = DR^n(\mathcal{O}(\star Y^{\text{an}}))_o$ is an isomorphism for a small enough analytical ball B around o , and it is compatible with the pole order filtration. Therefore the assertion follows from Corollary 3.0.1. \square

Remark 3.0.5. While the corollary is presented with the constraint $|Y^{\text{sing}}| = 1$ for clarity, the assertion can be extended to polynomials g with multiple isolated singularities. This would involve introducing a summation over all singularities.

Acknowledgements

I am grateful to M. Mustařă for sharing his then-preprint with S. Olano [Mustařă and Olano 2023] and especially to C. Sabbah for his continuous feedback highlighting in particular the importance of the analytic topology in Vilonen’s theorem. I thank also A. Dimca, for sending me some useful references. Finally, I am grateful to the referees for remarks that have helped improve the quality of the presentation. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), [RGPIN-2020-06075]. Cette recherche a été financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), [RGPIN-2020-06075].

References

- [Bitoun and Schedler 2018] T. Bitoun and T. Schedler, “On \mathcal{D} -modules related to the b -function and Hamiltonian flow”, *Compos. Math.* **154**:11 (2018), 2426–2440. MR Zbl
- [Björk 1993] J.-E. Björk, *Analytic \mathcal{D} -modules and applications*, Math. Appl. **247**, Kluwer, Dordrecht, 1993. MR
- [Brylinski 1986] J.-L. Brylinski, “Transformations canoniques, dualité projective, théorie de Lefschetz, transformations de Fourier et sommes trigonométriques”, pp. 3–134 in *Géométrie et analyse microlocales*, edited by J.-L. Brylinski and T. Monteiro Fernandes, Astérisque **140-141**, Soc. Math. France, Paris, 1986. MR Zbl
- [Dimca 1991] A. Dimca, “Differential forms and hypersurface singularities”, pp. 122–153 in *Singularity theory and its applications, I* (Coventry, 1988–1989), edited by D. Mond and J. Montaldi, Lecture Notes in Math. **1462**, Springer, 1991. MR Zbl
- [Etingof and Schedler 2014] P. Etingof and T. Schedler, “Invariants of Hamiltonian flow on locally complete intersections”, *Geom. Funct. Anal.* **24**:6 (2014), 1885–1912. MR Zbl
- [Karpishpan 1991] Y. Karpishpan, “Pole order filtration on the cohomology of algebraic links”, *Compos. Math.* **78**:2 (1991), 213–226. MR Zbl
- [Mustařă and Olano 2023] M. Mustařă and S. Olano, “On a conjecture of Bitoun and Schedler”, *Int. Math. Res. Not.* **2023**:21 (2023), 18254–18272. MR Zbl
- [Saito 2009] M. Saito, “On the Hodge filtration of Hodge modules”, *Mosc. Math. J.* **9**:1 (2009), 161–191. MR Zbl
- [Saito 2021] M. Saito, “ D -modules generated by rational powers of holomorphic functions”, *Publ. Res. Inst. Math. Sci.* **57**:3-4 (2021), 867–891. MR Zbl
- [Saito 2022] M. Saito, “Length of $D_X f^{-\alpha}$ in the isolated singularity case”, preprint, 2022. arXiv 2208.08977
- [Vilonen 1985] K. Vilonen, “Intersection homology D -module on local complete intersections with isolated singularities”, *Invent. Math.* **81**:1 (1985), 107–114. MR Zbl

Communicated by H el ene Esnault

Received 2023-09-20 Revised 2024-04-16 Accepted 2024-06-15

thomas.bitoun@ucalgary.ca

University of Calgary, Calgary, AB, Canada

Ribbon Schur functors

Keller VandeBogert

We investigate a generalization of the classical notion of a Schur functor associated to a ribbon diagram. These functors are defined with respect to an arbitrary algebra, and in the case that the underlying algebra is the symmetric/exterior algebra, we recover the classical definition of Schur/Weyl functors, respectively. In general, we construct a family of 3-term complexes categorifying the classical concatenation/near-concatenation identity for symmetric functions, and one of our main results is that the exactness of these 3-term complexes is equivalent to the Koszul property of the underlying algebra A . We further generalize these ribbon Schur functors to the notion of a multi-Schur functor and construct a canonical filtration of these objects whose associated graded pieces are described explicitly; one consequence of this filtration is a complete equivariant description of the syzygies of arbitrary Segre products of Koszul modules over the Segre product of Koszul algebras. Further applications to the equivariant structure of derived invariants, symmetric function identities, and Koszulness of certain classes of modules are explored at the end, along with a characteristic-free computation of the regularity of the Schur functor \mathbb{S}^λ applied to the tautological subbundle on projective space.

1. Introduction

1.1. Ribbon Schur functors. Schur functors are fundamental objects that lie at the intersection of representation theory, algebraic geometry, combinatorics, and commutative algebra. Representation theoretically, the Schur functors \mathbb{S}^λ corresponding to a partition λ are irreducible $GL(V)$ -representations in characteristic 0, and (up to twists by the determinant representation) these constitute all irreducible representations of the general linear group; see, for instance, [Weyman 2003, Chapter 2]. From the algebro-geometric perspective, Schur functors corresponding to partitions may be constructed via the famous *Borel–Weil theorem* (see [Serre 1959; Bott 1957], or [Kempf 1976]), which realizes these objects as the global sections of canonical line bundles on the complete flag variety. Combinatorially, Schur functors may be identified with their multigraded characters to obtain *Schur polynomials*, which are a fundamental basis for the ring of symmetric functions; see, for instance, [Grinberg and Reiner 2014] for more on this perspective. Finally, for a commutative algebraist, Schur modules over arbitrary commutative rings were constructed in the foundational work of Akin, Buchsbaum and Weyman [Akin et al. 1982], where these objects may be described as the image of a map constructed using the (graded) commutative Hopf algebra structure on the symmetric/exterior powers.

MSC2020: 05E40, 13D02, 14F06.

Keywords: Schur functors, Koszul algebras, free resolutions, sheaf cohomology.



Figure 1. The skew shape on the left is *not* a ribbon, since the bottom two rows have overlap size 2. The shape on the right *is* a ribbon.

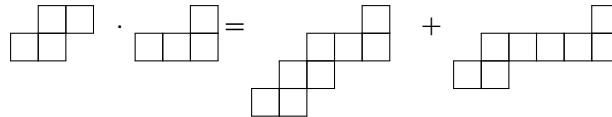


Figure 2. The concatenation/near-concatenation identity (here we are identifying the ribbon diagrams with their corresponding Schur polynomials).

We take the perspective that Schur functors corresponding to *ribbon diagrams* are “more natural” than Schur functors corresponding to partitions. A *ribbon diagram* is a skew shape for which consecutive rows overlap by exactly one block.

From a representation-theoretic point of view, ribbon Schur functors may seem quite *unnatural*: they are highly reducible, even in characteristic 0. However, experts in the theory of symmetric functions have long known that ribbon Schur polynomials also generate the ring of symmetric functions and are in many ways much more well-behaved than Schur polynomials corresponding to partitions; see [Lascoux and Pragacz 1988; Billera et al. 2006; Reiner et al. 2007; Huang 2016]. For example, ribbon Schur polynomials satisfy a much simpler Pieri-type formula (the *concatenation/near-concatenation identity*) than Schur polynomials corresponding to partitions, which require the Littlewood–Richardson rule to expand in general.

One of the most fundamental reasons ribbon diagrams are desirable is that Schur functors corresponding to ribbons can be defined solely using the algebra structure on the symmetric algebra $S(V)$, whereas Schur functors for partitions almost always need the full Hopf algebra structure; this is immediate from the work of Akin, Buchsbaum and Weyman [1982], which establishes a canonical, characteristic-free presentation of Schur modules associated to skew partitions. This means that there is an evident way to generalize ribbon Schur functors to arbitrary algebras, and the main theme of this paper is that in order for this theory to work “as it should”, the underlying algebra needs to be Koszul.

Example 1.1. In this example, identify the skew shapes with their corresponding Schur functors. Then

$$\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} = \text{Ker} \left(\begin{array}{c} S_2(V) \\ \oplus \\ S_3(V) \otimes_k V \end{array} \right) \quad \text{and} \quad \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array} = \text{Ker} (S_2(V) \otimes_k S_2(V) \rightarrow S_4(V)).$$

In the first equality, the map $S_2(V) \otimes_k S_2(V) \rightarrow S_3(V) \otimes_k V$ is the composition of *both* comultiplication and multiplication on the symmetric algebra, whereas the second ribbon diagram is simply the kernel of the canonical multiplication map $S_2(V) \otimes_k S_2(V) \rightarrow S_4(V)$.

1.2. Koszul algebras and Backelin’s theorem. A k -algebra A is *Koszul* if its residue field k has a linear homogeneous minimal free resolution over A . Koszul algebras are one method of generalizing standard graded polynomial rings and make their appearance in a wide range of seemingly disconnected settings. Topologically, Koszul duality for quadratic algebras can be used to translate between facts about equivariant and standard cohomology; see [Goresky et al. 1998]. In the context of number theory, Positselski [2014] has shown that certain classes of Milnor rings are Koszul algebras and relates the Milnor–Bloch–Kato conjecture to Koszulness of certain quotient rings. Combinatorially, early work of Backelin [1981] showed that there was an equivalence between Koszul algebras and distributivity of certain associated subspace lattices, and since then Koszulness of rings associated to combinatorial objects has been an active and fruitful area of research; see for instance [Yuzvinskiĭ 2001; Mastroeni and McCullough 2023].

Suppose for the time being that V is a vector space over a field k ; all tensor products will be taken over k . Recall that a quadratic algebra A is any quotient of the tensor algebra $T(V)$ by a quadratic ideal $(Q_2) \subset T_{\geq 2}(V)$, where $Q_2 \subset V \otimes V$. In particular, each graded piece of A is the quotient

$$A_n = \frac{V^{\otimes n}}{Q_2 \otimes V^{\otimes n-2} + V \otimes Q_2 \otimes V^{\otimes n-3} + \dots + V^{\otimes n-2} \otimes Q_2}.$$

Although Koszulness is often defined in terms of a homological property of the residue field of A , a well-known result of Backelin establishes the elegant fact that Koszulness has an equivalent formulation in terms of a purely combinatorial property of the collections $Q_2 \otimes V^{\otimes n-2}, V \otimes Q_2 \otimes V^{\otimes n-3}, \dots, V^{\otimes n-2} \otimes Q_2$.

Theorem 1.2 ([Backelin 1981], see also [Beilinson et al. 1996]). *Let A be a quadratic algebra. Then A is Koszul if and only if the collection $Q_2 \otimes V^{\otimes n-2}, V \otimes Q_2 \otimes V^{\otimes n-3}, \dots, V^{\otimes n-2} \otimes Q_2$ generates a distributive subspace lattice for all $n \geq 2$.*

This alternative (equivalent) formulation of Koszulness is surprising because it implies that the Koszulness assumption gives you “more than you bargained for”; the backwards implication of Backelin’s theorem is evident, but a priori distributivity seems like a much stronger property than Koszulness. In this paper, we reinterpret this distributivity property as being equivalent to the exactness of a family of 3-term sequences; see Proposition 3.14. A key observation here is that these 3-term sequences yield a defining identity for a generalization of ribbon Schur functors to any Koszul algebra.

1.3. Generalized ribbon Schur functors and the concatenation/near-concatenation criterion for Koszulness. Let A be any standard graded quadratic R -algebra and $\alpha = (\alpha_1, \dots, \alpha_n)$ any sequence of positive integers.¹ Then the *ribbon Schur module* \mathbb{S}_A^α is defined as the kernel of the map

$$A_{\alpha_1} \otimes_R A_{\alpha_2} \otimes_R \dots \otimes_R A_{\alpha_n} \rightarrow \bigoplus_{i=1}^{n-1} A_{\alpha_1} \otimes_R \dots \otimes_R A_{\alpha_i + \alpha_{i+1}} \otimes_R \dots \otimes_R A_{\alpha_n},$$

¹The quadratic assumption is unnecessary, but since these objects are most well-behaved when the algebra A is assumed to be Koszul, we will often be in this setting anyway.

where each component of the map is induced by the natural multiplication maps $A_{\alpha_i} \otimes_R A_{\alpha_{i+1}} \rightarrow A_{\alpha_i + \alpha_{i+1}}$. As mentioned previously, when $A = S^*(V)$ is the symmetric algebra, this definition recovers the Akin, Buchsbaum and Weyman [1982] definition of the Schur functor associated to the ribbon diagram induced by α . The importance of using ribbon Schur functors to canonically describe syzygies of modules over Veronese subalgebras of the polynomial ring was discovered in [Almoussa et al. 2024], and the goal of this paper is to fully expand upon this perspective in a much more general setting.

One of our first main results is the following (for notation, see Definition 4.1). This statement may be interpreted as saying that the concatenation/near-concatenation identity observed in the classical theory of Schur polynomials is the shadow of a canonical short exact sequence *of functors*, and that the concatenation/near-concatenation identity is actually a defining property of Koszulness.

Theorem 1.3. *Let A be any Koszul R -algebra (where R is any commutative ring) and α, β be any two compositions:*

- (1) *There is a canonical isomorphism of R -modules*

$$(\mathbb{S}_A^\alpha)^* \cong \mathbb{S}_{A^t}^{\alpha^t},$$

where $(-)^!$ denotes the quadratic dual, α^t denotes the transposed ribbon diagram, and $(-)^* := \text{Hom}_R(-, R)$.

- (2) *There is a canonical short exact sequence*

$$0 \rightarrow \mathbb{S}_A^{\alpha \cdot \beta} \rightarrow \mathbb{S}_A^\alpha \otimes_R \mathbb{S}_A^\beta \rightarrow \mathbb{S}_A^{\alpha \odot \beta} \rightarrow 0.$$

*Conversely, this sequence is exact for **all** compositions **if and only if** A is Koszul.*

This is surprising since, intuitively, the classical concatenation/near-concatenation identity for symmetric functions may seem like a consequence of the combinatorial/representation-theoretic structure on the symmetric algebra $S^*(V)$. Theorem 1.3 tells us that this additional structure is more of a red herring, and the classical identity is actually a consequence of a much more fundamental algebraic property.

One could also ask if this equivalent formulation of Koszulness is actually useful for proving classes of modules are Koszul, and in Section 6.4 we use this criterion to give quick proofs of the existence of large classes of Koszul modules over arbitrary Koszul algebras A . In the case of the symmetric/exterior algebras, this criterion for Koszulness gives a very simple proof that a general class of modules parametrized by arbitrary skew-partitions are Koszul. Previous proofs that these modules were Koszul only worked for diagrams corresponding to partitions, and resorted to geometric arguments that realized these objects as arising from taking cohomology of line bundles on flag varieties; see, for instance, [Gao and Raicu 2024, Theorem 2.2]. Our proof is totally characteristic-independent, allows for the algebra to be over any commutative ring, and requires no machinery coming from algebraic geometry.

We also generalize this definition of Schur functors to allow for module inputs M and N , denoted by $\mathbb{S}_{M,A,N}^\alpha$; this is a generalization even in the classical case of Schur functors, and allows for elegant descriptions of the higher derived invariants associated to pairs of Koszul modules.

Theorem 1.4. *Let A be a Koszul algebra and N (resp. M) a left (resp. right) Koszul A -module. Then there is a canonical isomorphism of A -modules*

$$\mathrm{Tor}_i^A(M, N) = \mathbb{S}_{M, A, N}^{(1^i)}.$$

If M is instead a left A -module, then there is a canonical isomorphism of A -modules

$$\mathrm{Ext}_A^i(M, N) = \mathbb{S}_{M^!, A^!, (N^*)^!}^{(i)}.$$

If A , M , and N have any ambient group action, then the above isomorphisms are equivariant.

Remark 1.5. Note that there are other places in the literature where generalizations of Schur functors have been considered, such as the work of Sam and Snowden [2017; 2019] which approaches the problem from a much more representation-theoretic perspective. These constructions are done in the standard type ABCD framework and do not apply to arbitrary Koszul algebras like the constructions in this paper do.

1.4. Multi-Schur functors. Let us first motivate the construction of a multi-Schur functor. Segre subalgebras of the tensor product of Koszul algebras are well known to be Koszul algebras. Given Koszul algebras A and B , let $A \circ B$ denote the Segre product; the homogeneous components of the (dual) quadratic dual may be computed as the kernel of a map that applies the multiplication on A and B “diagonally”. For example, there is an equality

$$((A \circ B)^!)^*_3 = \mathrm{Ker} \left(A_1^{\otimes 3} \otimes B_1^{\otimes 3} \rightarrow \begin{array}{c} A_2 \otimes A_1 \otimes B_2 \otimes B_1 \\ \oplus \\ A_1 \otimes A_2 \otimes B_1 \otimes B_2. \end{array} \right).$$

This is the same thing as applying the defining relations for the Schur modules $\mathbb{S}_A^{(1^3)}$ and $\mathbb{S}_B^{(1^3)}$ diagonally, and leads us directly to the notion of a *multi-Schur functor*.

A multi-Schur functor \mathbb{S}_A^α takes as inputs *tuples* of compositions and algebras

$$\underline{\alpha} = (\alpha^1, \dots, \alpha^n), \quad \underline{A} = (A^1, \dots, A^n),$$

and is defined by taking the kernel of the defining relations of each of the ribbon Schur modules $\mathbb{S}_A^{\alpha^i}$ applied diagonally, exactly as above. Surprisingly, multi-Schur modules satisfy an identical concatenation/near-concatenation sequence (appropriately generalized; see Lemma 5.8), and we can further extend this definition to allow for tuples of modules. This gives us a similar clean description of Tor modules over Segre products.

Theorem 1.6. *Consider tuples of the form*

$$\underline{A} = (A^1, \dots, A^n), \quad \underline{N} = (N^1, \dots, N^n), \quad \underline{M} = (M^1, \dots, M^n),$$

where each A^i is a Koszul R -algebra and N^i (resp. M^i) is a left (resp. right) A^i -module. Then there is a canonical isomorphism of $A^1 \circ \dots \circ A^n$ -modules

$$\mathrm{Tor}_i^{A^1 \circ \dots \circ A^n}(M^1 \circ \dots \circ M^n, N^1 \circ \dots \circ N^n) = \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{(1^i)}.$$

This isomorphism is natural with respect to morphisms of algebras; in particular it is equivariant if each of the Koszul algebras has any ambient group action.

Canonical filtrations. Although Theorem 1.6 is simple to state, it still gives us very little information about the relationship between a multi-Schur functor (which is a priori a purely formal construction) and objects that we may better understand. A standard method of trying to understand an object is to construct a filtration whose associated graded objects are “simple” in the appropriate sense.

As it turns out, multi-Schur functors $\mathbb{S}_{M,A,N}^\alpha$ admit a canonical (that is, totally functorial) filtration whose associated graded pieces are tensor products of the Schur functors $\mathbb{S}_{M^i,A^i,N^i}^{\beta^i}$, where $1 \leq i \leq n$. The difficult part is determining the poset that parametrizes this filtration, and our main result related to multi-Schur functors is an explicit description of this parametrization.

Before stating the result, we need to introduce one piece of notation: given a composition $\alpha = (\alpha_1, \dots, \alpha_\ell)$ and a subset $I \subset [\ell - 1]$, the notation $\sigma_I(\alpha)$ denotes the composition obtained by adding α_i and α_{i+1} for all $i \in I$. For example

$$\sigma_{\{1,2,4,6,7\}}(2, 1, 3, 5, 3, 6, 5, 3) = (6, 8, 14).$$

With this notation in hand, we have:

Theorem 1.7. *Consider the tuple*

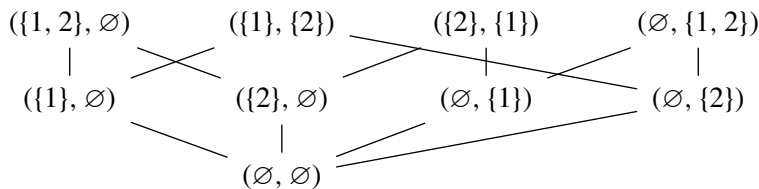
$$\underline{\alpha} = (\alpha^1, \dots, \alpha^n),$$

where each composition α^i has a fixed length ℓ . With notation and hypotheses as in Theorem 1.6, the multi-Schur module $\mathbb{S}_{M,A,N}^\alpha$ admits a canonical filtration with associated graded pieces of the form

$$\mathbb{S}_{M^1,A^1,N^1}^{\sigma_{I_1}(\alpha^1)} \otimes_R \mathbb{S}_{M^2,A^2,N^2}^{\sigma_{I_2}(\alpha^2)} \otimes_R \cdots \otimes_R \mathbb{S}_{M^n,A^n,N^n}^{\sigma_{I_n}(\alpha^n)},$$

where the subsets $I_1, \dots, I_n \subset [\ell - 1]$ range over all choices such that $I_1 \cap I_2 \cap \cdots \cap I_n = \emptyset$. This filtration is equivariant with respect to any kind of ambient group action.

Example 1.8. If in the statement of Theorem 1.7, one has $n = 2$ and each composition has length 3, the poset parametrizing the filtration factors has Hasse diagram:



For a concrete example of the explicit filtration factors, see Example 5.25.

1.5. Organization of paper. This paper is organized as follows. In Section 2 we recall some background on augmented bar complexes and their homogeneous strands. We also establish conventions and notation that will be used throughout the paper. In Section 3 we recall the notion of distributivity for collections of arbitrary R -submodules; much of this material is essentially contained in [Polishchuk and Positselski 2005], but since we do not assume the ambient ring is a field, there are some additional details that need to be checked. We also introduce a collection of modules L_{M_1, \dots, M_n}^I associated to a distributive collection of R -submodules that will end up being an equivalent way to define ribbon Schur functors in the setting

of arbitrary quadratic algebras. Under appropriate assumptions, we also show that distributivity behaves well with respect to “dualization”.

Sections 4 and 5 develop the theory of ribbon Schur functors associated to arbitrary quadratic algebras/modules. In Section 4, we define ribbon Schur functors associated to arbitrary quadratic algebras/modules and prove the concatenation/near-concatenation Koszulness criterion. Section 4.2 is devoted to giving explicit examples and illustrating the statements for concrete choices of Koszul algebras, and the main proofs of the results in full generality are given in Section 4.3. We will see that many duality properties of Schur and Weyl functors in the classical setting as a consequence of Koszul duality for these more general ribbon Schur functors.

In Section 5 we define multi-Schur functors as described earlier. The beginning of this section explains how to upgrade all of the statements of Section 4 to this level of generality; it turns out that there is a large amount of additional bookkeeping needed when doing this. Sections 5.1 and 5.2 prove all of the results necessary to construct the filtration as stated in Theorem 1.7.

Section 6 is where we get to see the utility of the theory developed in Sections 4 and 5 for giving canonical descriptions/filtrations to many of the invariants associated to Koszul algebras/modules. Section 6.1 proves the isomorphisms of Tor and Ext between Koszul modules as stated in Theorem 1.4, and Sections 6.2 and 6.3 give canonical descriptions of the derived invariants over Veronese/Segre subalgebras in terms of the original algebra(s), and even deduce an interesting symmetric function identity as a result taking a character count on the minimal free resolution over a Segre product. In Section 6.4, we construct a large class of Koszul modules over an arbitrary Koszul algebra and, in the case of symmetric/exterior algebras, we generalize a construction of Koszul modules associated to Schur functors corresponding to arbitrary skew partitions. We use this to give a characteristic free computation of the sheaf cohomology of $\mathbb{S}^\lambda(\mathcal{R})$ on $\mathbb{P}(V)$, where \mathcal{R} denotes the tautological subbundle on projective space.

The Appendix recalls the equivalence between Koszulness and distributivity (that is, Backelin’s theorem) and interprets distributivity in terms of exactness properties of so-called *refinement complexes*. Again, much of this material follows from straightforward generalizations of the material in [Polishchuk and Positselski 2005], but we check the details for sake of completeness. In Section A.1 we first discuss some generalities and definitions related to general quadratic algebras. The bulk of this appendix is dedicated to the material of Section A.2, which defines *refinement complexes* and establishes some important conventions for these refinement complexes. In Sections A.3 and A.4, we recall Backelin’s theorem along with the appropriate analogs for Koszul modules and translate these results into statements about the family of refinement complexes. Finally, in Section A.5 we recall the Priddy complex and some of its properties in the generality in which we are working.

2. Augmented bar constructions

In this section, we establish some conventions and notation to be used for the remainder of the paper. We also define one of the most important elements of the paper: the augmented bar complex associated

to an augmented R -algebra. Unsurprisingly, understanding homogeneous strands of the augmented bar complex is equivalent to understanding Koszulness of the algebra A .

Definition 2.1. Let R be a commutative Noetherian ring and A any associative unital R -algebra:

- (1) The algebra A is (\mathbb{Z}) -graded if $A = \bigoplus_{i \in \mathbb{Z}} A_i$ with $A_i \cdot A_j \subset A_{i+j}$ for all $i, j \in \mathbb{Z}$, and $1_A \in A_0$.
- (2) The algebra A is *augmented* if $A = R \oplus A_+$, where $R = R \cdot 1_A$ and A_+ is a two-sided ideal in A .
- (3) The algebra A is *quadratic* if $A = T(A_1)/(Q_2^A)$, where $T(-)$ denotes the tensor algebra functor, $Q_2^A \subset A_1 \otimes_R A_1$, and (Q_2) denotes the two-sided ideal generated by Q_2 .
- (4) A left A -module M is *graded* if $M = \bigoplus_{i \in \mathbb{Z}} M_i$ and $A_i M_j \subset M_{i+j}$ for all $i, j \in \mathbb{Z}$.

The notation A^{op} denotes the *opposite algebra* of A ; that is, the underlying set of A^{op} is the same as A but with multiplication defined by $a \cdot b := b \cdot a$. Any right A -module is equivalently a left A^{op} -module.

Remark 2.2. For quadratic R -algebras, we will always assume that the grading is induced by the standard grading on the tensor algebra. In particular, all quadratic algebras are nonnegatively graded.

The following convention is extremely important, and without it most of the arguments given later in this paper cannot even get off the ground.

Convention 2.3. Throughout this paper, all graded R -algebras A will be assumed to be finitely-generated flat R -modules in each degree. Similarly, all graded A -modules will be assumed to be finitely-generated and flat in each degree. Sometimes we will assume that A or M is even R -projective in each degree, but it is *always* true that they are at least finitely-generated and flat in each degree.

There are two fundamental operations on algebras that will be particularly interesting for us.

Definition 2.4. Let A and B be graded R -algebras. The d -th Veronese power $A^{(d)}$ of A is defined to be the subalgebra

$$A^{(d)} := \bigoplus_{i \equiv 0 \pmod d} A_i \subset A.$$

Let M be a graded (left) A -module of initial degree t . Then the d -th Veronese power $M^{(d)}$ is defined to be the $A^{(d)}$ -module

$$M^{(d)} := \bigoplus_{i \equiv t \pmod d} M_i \subset M.$$

The *Segre product* $A \circ B$ of A and B is defined to be the subalgebra

$$A \circ B := \bigoplus_{i \geq 0} A_i \otimes_R B_i \subset A \otimes_R B.$$

The d -th Segre power $A^{[d]}$ of A is defined to be the iterated Segre product

$$\underbrace{A \circ A \circ \dots \circ A}_{d \text{ times}} \subset A^{\otimes d}.$$

Given a graded left A (resp. B)-module M (resp. N) of initial degree s (resp. t), the Segre product $M \circ N$ is the left $A \circ B$ module defined as

$$M \circ N := \bigoplus_{i \geq 0} M_{i+s} \otimes_R N_{i+t}.$$

The Segre product of Koszul right modules is defined identically.

The d -th Segre power $M^{[d]}$ of M is defined to be the Segre product

$$\underbrace{M \circ \cdots \circ M}_{d \text{ times}}.$$

Remark 2.5. Often one uses the convention that the generators of the Veronese/Segre subalgebras have been rescaled to have degree 1. We will actually have no need for this convention in the paper as long as we define an algebra to be Koszul if the associated Priddy complex (see Theorem A.25) is a resolution; this is equivalent to the distributivity of a certain set of submodules (by Backelin’s theorem), which is again a condition that still makes sense regardless of the generators having degree 1. That being said, many of the results stated in this paper will be written as if the generators are in degree 1, but again this condition is only required “up to rescaling”.

Likewise, the graded dual of an algebra will be useful for describing certain Ext modules. All duals in this paper are understood to be *graded* duals; that is

$$\text{Hom}_A(M, R) := \bigoplus_{i \in \mathbb{Z}} \text{Hom}_R(M_i, R).$$

Definition 2.6. Let M be a left A -module, where A is any graded R -algebra. Then the graded dual $M^* := \text{Hom}_A(M, R)$ is canonically a right A -module via the action

$$(m^* a)(n) := m^*(an), \quad m^* \in M^*, \quad a \in A, \quad n \in M.$$

Moreover, there is a natural isomorphism of A^{op} -modules $(M^{\text{op}})^* \cong (M^*)^{\text{op}}$.

Recall that modules over (positively) graded algebras satisfy a strong form of Nakayama’s lemma:

Lemma 2.7 (Nakayama’s lemma). *Let A be any nonnegatively graded R -algebra and M a graded left A -module with $M_i = 0$ for $i \ll 0$. If $A_+ M = M$, then $M = 0$.*

Now we define the augmented bar complex.

Definition 2.8. Let A be any augmented R -algebra and let $A_+ := A/A_0$. Given any left A -module M , the *augmented Bar complex* $\text{Bar}^A(M)$ is the complex of A -modules with

$$\text{Bar}_i^A(M) := A \otimes_R A_+^{\otimes i} \otimes_R M,$$

with differential

$$d^B(a_0 \otimes a_1 \otimes \cdots \otimes a_i \otimes m) := \sum_{j=0}^{i-1} (-1)^j a_0 \otimes \cdots \otimes a_j \cdot a_{j+1} \otimes \cdots \otimes a_i \otimes m + (-1)^i a_0 \otimes \cdots \otimes a_{i-1} \otimes a_i m.$$

The *augmented cobar complex* $\text{Cobar}^A(M)$ on M is the graded R -dual of $\text{Bar}^A(M)$. The bar complex and cobar complex are both bigraded by the homological and internal degree. In other words

$$\text{Bar}_i^A(M)_j = \bigoplus_{k_0+\dots+k_i+\ell=j} A_{j_1} \otimes_R \dots \otimes_R A_{j_k} \otimes_R M_\ell.$$

With respect to the internal grading, the differential of $\text{Bar}^A(M)$ has degree 0.

The notation $\text{Bar}^A(M)_n$ will denote the (internal) *degree n strand* of the bar complex.

Remark 2.9. For right A -modules M , there is a similar bar complex construction $\text{Bar}^A(M)^{\text{op}}$, where M appears as the leftmost tensor factor and the differential is defined analogously. This construction is compatible with taking the opposite algebra in the sense that there is an isomorphism of complexes

$$\text{Bar}^{A^{\text{op}}}(M) \cong \text{Bar}^A(M)^{\text{op}},$$

where M is viewed as a left A^{op} -module on the left-hand side of this isomorphism. This means that throughout this section, it is of no loss of generality to assume that M is a left A -module.

Example 2.10. The degree-4 strand $(R \otimes_A \text{Bar}^A(M))_4$ is the following complex of projective R -modules:

$$A_1^{\otimes 4} \rightarrow \begin{array}{c} A_2 \otimes_R A_1^{\otimes 2} \\ \oplus \\ A_1 \otimes_R A_2 \otimes_R A_1 \\ \oplus \\ A_1^{\otimes 2} \otimes_R A_2 \end{array} \rightarrow \begin{array}{c} A_3 \otimes_R A_1 \\ \oplus \\ A_2 \otimes_R A_2 \\ \oplus \\ A_1 \otimes_R A_3 \end{array} \rightarrow A_4.$$

The following are some fundamental properties of (co)bar constructions that will be useful in later sections.

Proposition 2.11. *Let A be an augmented graded R -algebra and M any left A -module:*

- (1) *The augmented bar complex $\text{Bar}^A(M)$ is a flat resolution of M .*
- (2) *The augmented cobar complex $\text{Cobar}^A(A)$ is an associative DG-algebra via the standard tensor algebra product. Likewise, the cobar construction $\text{Cobar}^A(M)$ is a right DG-module over $\text{Bar}^A(A)$.*
- (3) *There are isomorphisms*

$$H_i(R \otimes_A \text{Bar}^A(M)) = \text{Tor}_i^A(R, M), \quad H^i(\text{Bar}^A(M)^*) = \text{Ext}_A^i(M, R).$$

- (4) *For any right A -module N , there are isomorphisms*

$$H_i(N \otimes_A \text{Bar}^A(M)) = \text{Tor}_i^A(N, M), \quad \text{Tor}_i^A(N, M)^* \cong \text{Ext}_A^i(M, N^*),$$

where N^* is a left A -module via the convention of Definition 2.6. If N is instead a left A -module, there is an isomorphism

$$H^i(\text{Hom}_A(\text{Bar}^A(M), N)) = \text{Ext}_A^i(M, N).$$

Proof. Both (1) and (2) are well known. The statements (3) and (4) follow from the fact that Tor and Ext may be computed using flat resolutions instead of projective resolutions. \square

We conclude this section with a straightforward observation that will be useful to write explicitly, since it will be cited many times in later sections.

Observation 2.12. Let M be any R -module and

$$0 \rightarrow M \rightarrow F_0 \rightarrow F_1 \rightarrow \dots \rightarrow F_n \rightarrow 0$$

any exact complex such that F_i is a flat (resp. projective) R -module for each $i = 0, \dots, n$. Then M is flat (resp. projective).

Proof. Proceed by induction on n , where the case $n = 0$ implies $M = F_0$ is evidently flat (resp. projective). If $n > 0$, let $C := \text{im}(F_0 \rightarrow F_1)$. By the inductive hypothesis, the module C is flat (resp. projective) and there is a short exact sequence

$$0 \rightarrow M \rightarrow F_0 \rightarrow C \rightarrow 0.$$

If C is projective, this sequence splits and hence M is also projective. If C is instead flat, apply the functor $-\otimes_R N$ for every R -module N and employ the long exact sequence of homology to deduce that M is flat. \square

3. Distributivity and submodule lattices

In this section, we recall a general family of complexes that may be associated to any collection of R -submodules $M_1, \dots, M_n \subset M$; see [Polishchuk and Positselski 2005, Chapter 2] for the case over a field, though the theory is essentially identical here. The exactness properties of these complexes will be used to define distributivity, and under some additional hypotheses on the collection M_1, \dots, M_n we will see that distributivity satisfies a simple duality. The purpose of this section is to introduce the modules L_{M_1, \dots, M_n}^I of Definition 3.11 and study their flatness/projectivity/duality properties in the generality established in the previous section.

All modules in this section are assumed to be finitely generated (including all submodules). For convenience, we recall the definition of a lattice.

Definition 3.1. A poset is a set S equipped with a partial order \leq . A poset is a *lattice* if any two pairs of elements $a, b \in S$ have a well-defined meet and join, denoted $a \wedge b$ and $a \vee b$, respectively.

Notation 3.2. Let $[n] := \{1, \dots, n\}$ for some integer n . The j -th degeneracy map $s_j : [n] \rightarrow [n - 1]$ is defined to be the surjection

$$s_j(i) := \begin{cases} i & \text{if } i \leq j, \\ i - 1 & \text{if } i > j. \end{cases}$$

Likewise, the j -th face map $d_j : [n - 1] \rightarrow [n]$ is defined to be the map

$$d_j(i) := \begin{cases} i & \text{if } i < j, \\ i + 1 & \text{if } i \geq j. \end{cases}$$

Likewise, given any set $I \subset [n]$, the notation $\text{sgn}(j, I)$ for any $j \notin I$ denotes the sign of the permutation that reorders the set (I, j) into ascending order.

Remark 3.3. The terminology “face” and “degeneracy” maps is borrowed from the terminology used in the simplicial category Δ .

The following definition associates a lattice to any collection of R -submodules:

Definition 3.4. Let M be any R -module and $M_1, \dots, M_n \subset M$ a collection of R -submodules. The collection of submodules M_1, \dots, M_n generates a lattice with operations

$$M_i \wedge M_j := M_i \cap M_j, \quad M_i \vee M_j := M_i + M_j.$$

The following construction introduces a fundamental family of complexes that can be associated to any collection of submodules. The terms of these complexes are built by taking “intervals” above elements in the associated submodule lattices, and the exactness properties of these complexes will be very important in later sections.

Construction 3.5. Let M be any R -module and $M_1, \dots, M_n \subset M$ a collection of R -submodules. Given any subset $I \subset J$, let $\rho_{I,J}$ denote the canonical surjection

$$\rho_{I,J} : \frac{M}{\bigvee_{j \in I} M_j} \rightarrow \frac{M}{\bigvee_{j \in J} M_j}.$$

Define the (cochain) complex $C^\bullet(M; M_1, \dots, M_n)$ via

$$C^i(M; M_1, \dots, M_n) := \bigoplus_{|I|=i} \frac{M}{\bigvee_{j \in I} M_j},$$

with differential

$$d^{C^\bullet} \Big|_{\frac{M}{\bigvee_{i \in I} M_i}} := \sum_{j \notin I} \text{sgn}(j, I) \rho_{I \cup j, I}.$$

In the above, we use the convention that $C^0(M; M_1, \dots, M_n) := M$.

Likewise, given any subset $I \subset J$, let $\iota_{J,I}$ denote the natural inclusion

$$\iota_{J,I} : \bigwedge_{i \in J} M_i \rightarrow \bigwedge_{i \in I} M_i.$$

Define the (chain) complex $C_\bullet(M; M_1, \dots, M_n)$ via

$$C_i(M; M_1, \dots, M_n) := \bigoplus_{|I|=i} \bigwedge_{j \in I} M_j,$$

with differential

$$d^{C_\bullet} \Big|_{\bigwedge_{i \in I} M_i} := \sum_{j \notin I} \text{sgn}(j, I) \iota_{I \cup j, I}.$$

If $I \subset [n]$ is any subset, define the complex $C_I^\bullet(M; M_1, \dots, M_n)$ by restricting to all direct summands of the form

$$C_I^i(M; M_1, \dots, M_n) = \bigoplus_{\substack{I \subset J, \\ |J|-|I|=i}} \frac{M}{\bigvee_{j \in J} M_j}.$$

Likewise, define the complex $C_\bullet^I(M; M_1, \dots, M_n)$ by restricting to all direct summands of the form

$$C_\bullet^I(M; M_1, \dots, M_n) := \bigoplus_{\substack{I \subset J, \\ |J|-|I|=i}} \bigwedge_{j \in J} M_j.$$

Remark 3.6. Notice that by definition there are equalities

$$\begin{aligned} C_\bullet(M; M_1, \dots, M_n) &= C_\bullet^\emptyset(M; M_1, \dots, M_n), \\ C^\bullet(M; M_1, \dots, M_n) &= C_\bullet^\emptyset(M; M_1, \dots, M_n). \end{aligned}$$

Example 3.7. If M_1, M_2, M_3 is a length 3 collection of R -submodules of M , the associated complexes of Construction 3.5 are explicitly given by

$$\begin{aligned} C_\bullet(M; M_1, M_2, M_3) : & \quad M_1 \cap M_2 \cap M_3 \rightarrow \begin{array}{c} M_1 \cap M_2 \\ \oplus \\ M_1 \cap M_3 \\ \oplus \\ M_2 \cap M_3 \end{array} \rightarrow \begin{array}{c} M_1 \\ \oplus \\ M_2 \\ \oplus \\ M_3 \end{array} \rightarrow M, \\ \\ C^\bullet(M; M_1, M_2, M_3) : & \quad M \rightarrow \begin{array}{c} \frac{M}{M_1} \\ \oplus \\ \frac{M}{M_2} \\ \oplus \\ \frac{M}{M_3} \end{array} \rightarrow \begin{array}{c} \frac{M}{M_1+M_2} \\ \oplus \\ \frac{M}{M_1+M_3} \\ \oplus \\ \frac{M}{M_2+M_3} \end{array} \rightarrow \frac{M}{M_1 + M_2 + M_3}. \end{aligned}$$

In the above, note that $C_\bullet(M; M_1, M_2, M_3)$ is homologically indexed, whereas $C^\bullet(M; M_1, M_2, M_3)$ is cohomologically indexed. By construction, both complexes have terms parametrized by the Boolean poset on $\{1, 2, 3\}$.

We can now define the notion of distributivity using the complexes of Construction 3.5.

Definition 3.8. Let M be an R -module and M_1, \dots, M_n a collection of R -submodules of M . The collection M_1, \dots, M_n is called *distributive* if the complexes

$$C_\bullet^I(M; M_1, \dots, M_n) \quad \text{and} \quad C_I^\bullet(M; M_1, \dots, M_n)$$

are exact in positive (co)homological degrees for all subsets $I \subset [n]$.

Remark 3.9. At first glance, this definition of distributive might seem quite different from the definition in terms of the submodule lattice generated by M_1, \dots, M_n , but it turns out that these definitions are in

fact equivalent by, for instance, [Polishchuk and Positselski 2005, Chapter 1, Proposition 7.2] (the result here is stated over a field, but it holds over any commutative ring).

The following is a trivial consequence of the inductive structure of the complexes C_i^\bullet and C_i^I :

Observation 3.10. If $M_1, \dots, M_n \subset M$ is a distributive collection, then every subcollection of M_1, \dots, M_n is distributive.

Next, we introduce a collection of subquotients that may be associated to any collection of R -submodules. In later sections, we will see that all ribbon Schur functors arise as modules of this form for some appropriate collection of submodules:

Definition 3.11. Let M be an R -module and M_1, \dots, M_n a collection of R -submodules of M . Given an indexing set $I \subset [n]$, define the R -module

$$L_{M_1, \dots, M_n}^I := \frac{\bigwedge_{i \notin I} M_i}{\bigvee_{i \in I} M_i}.$$

Remark 3.12. By convention, we set

$$L_{M_1, \dots, M_n}^{[n]} := \frac{M}{M_1 + \dots + M_n} \quad \text{and} \quad L_{M_1, \dots, M_n}^\emptyset := M_1 \cap \dots \cap M_n.$$

Moreover, notice that for the sake of conciseness of notation, the notation N/L for two R -submodules $N, L \subset M$ is understood to mean $N/(L \cap N)$.

The following observation is proved just by the definition of the complexes of Construction 3.5 and the equivalence between the exactness of these complexes and distributivity:

Observation 3.13. Let M be an R -module and M_1, \dots, M_n any collection of R -submodules of M . For every $I \subset [n]$ the complexes

$$C_i^\bullet(M; M_1, \dots, M_n) \quad \text{and} \quad C_i^I(M; M_1, \dots, M_n)$$

satisfy

$$H_0(C_i^I(M; M_1, \dots, M_n)) \twoheadrightarrow L_{M_1, \dots, M_n}^{[n] \setminus I} \quad \text{and} \quad L_{M_1, \dots, M_n}^I \hookrightarrow H^0(C_i^\bullet(M; M_1, \dots, M_n)).$$

If the collection M_1, \dots, M_n is distributive, then the above surjection/inclusion are equalities.

Proof. By definition, the zeroth homology of $C_i^I(M; M_1, \dots, M_n)$ is the cokernel of the map

$$\bigoplus_{j \notin I} \bigwedge_{\ell \in I \cup j} M_\ell \rightarrow \bigwedge_{\ell \in I} M_\ell,$$

which is precisely the quotient

$$\frac{\bigwedge_{\ell \in I} M_\ell}{\sum_{j \notin I} \bigwedge_{\ell \in I \cup j} M_\ell}.$$

There is always a containment

$$\sum_{j \notin I} \bigwedge_{\ell \in I \cup j} M_\ell \subset \left(\sum_{j \notin I} M_j \right) \cap \bigwedge_{\ell \in I} M_\ell,$$

and thus there is a natural surjection $H_0(C_\bullet^I(M; M_1, \dots, M_n)) \twoheadrightarrow L_{M_1, \dots, M_n}^{[n] \setminus I}$. If the collection M_1, \dots, M_n is distributive, then the above containment is an equality and thus the surjection is also an equality.

Similarly, the zeroth cohomology of $C_I^\bullet(M; M_1, \dots, M_n)$ is the kernel of the map

$$\frac{M}{\bigvee_{\ell \in I} M_\ell} \rightarrow \bigoplus_{j \notin I} \frac{M}{\bigvee_{\ell \in I \cup j} M_\ell}.$$

Again, there is always a containment

$$\bigwedge_{j \notin I} \left(\bigvee_{\ell \in I \cup j} M_\ell \right) \supset \bigwedge_{j \notin I} M_j + \bigvee_{\ell \in I} M_\ell$$

and thus a natural inclusion $L_{M_1, \dots, M_n}^I \hookrightarrow H^0(C_I^\bullet(M; M_1, \dots, M_n))$. When the collection M_1, \dots, M_n is distributive the above containment is an equality in which case the inclusion is also an equality. \square

The following proposition gives 3 different equivalent conditions equivalent to distributivity. In particular, the exactness of C_\bullet^I for all $I \subset [n]$ is equivalent to the exactness of C_I^\bullet , and this exactness is in turn equivalent to a family of short exact sequences. These short exact sequences will end up modeling the concatenation/near-concatenation sequences in later sections.

Recall the notation d_j and s_j for face and degeneracy maps, respectively, of Notation 3.2 in the statement of the following result:

Proposition 3.14. *Let M be any R -module and $M_1, \dots, M_n \subset M$ a collection of R -submodules. Then the following are equivalent:*

- (1) *For all nonempty $I \subset [n]$, the sequence*

$$0 \rightarrow L_{M_1, \dots, M_n}^{I \setminus j} \rightarrow L_{M_1, \dots, \widehat{M}_j, \dots, M_n}^{s_j(I \setminus j)} \rightarrow L_{M_1, \dots, M_n}^I \rightarrow 0$$

is exact.

- (2) *For all $I \subset [n]$, the cochain complex $C_I^\bullet(M; M_1, \dots, M_n)$ is exact in positive cohomological degrees.*
- (3) *For all $I \subset [n]$, the chain complex $C_\bullet^I(M; M_1, \dots, M_n)$ is exact in positive homological degrees.*

In other words, to check distributivity it suffices to check exactness of either the complex $C_\bullet^I(M; M_1, \dots, M_n)$ or $C_I^\bullet(M; M_1, \dots, M_n)$ for all $I \subset [n]$.

Remark 3.15. When $|I| = 1$ is a singleton set, the sequence above simply reads

$$0 \rightarrow M_1 \rightarrow M \rightarrow M/M_1 \rightarrow 0.$$

In general, the sequence

$$0 \rightarrow L_{M_1, \dots, M_n}^{I \setminus j} \rightarrow L_{M_1, \dots, \widehat{M}_j, \dots, M_n}^{s_j(I \setminus j)} \rightarrow L_{M_1, \dots, M_n}^I \rightarrow 0$$

is a complex that is exact at the left and rightmost nontrivial terms by definition. This means that Proposition 3.14 gives equivalent conditions for this sequence to be exact at the middle term.

Proof. Notice first that there are canonical short exact sequences of complexes (where $\widehat{}$ denotes omission), for $j \in I$

$$0 \rightarrow C_I^\bullet(M; M_1, \dots, M_n)[-1] \rightarrow C_{I \setminus j}^\bullet(M; M_1, \dots, M_n) \rightarrow C_{s_j(I \setminus j)}^\bullet(M; M_1, \dots, \widehat{M}_j, \dots, M_n) \rightarrow 0, \quad (3.15.1)$$

and, for $j \notin I$,

$$0 \rightarrow C_{\bullet}^{s_j(I)}(M; M_1, \dots, M_n) \rightarrow C_{\bullet}^I(M; M_1, \dots, M_n) \rightarrow C_{\bullet}^{I \cup j}(M; M_1, \dots, M_n)[-1] \rightarrow 0. \quad (3.15.2)$$

(1) \iff (2): By induction on n , we may assume that the collection $M_1, \dots, \widehat{M}_j, \dots, M_n$ is distributive; in particular, the complex $C_{s_j(I \setminus j)}^\bullet(M; M_1, \dots, \widehat{M}_j, \dots, M_n)$ is exact in positive cohomological degrees, and moreover by a downward induction on $|I|$ (with base case $I = [n]$) we may also assume that $C_I^\bullet(M; M_1, \dots, M_n)$ is exact in positive cohomological degrees.

Employing the long exact sequence of cohomology on the short exact sequence (3.15.1) along with Observation 3.13 yields

$$0 \rightarrow H^0(C_{I \setminus j}^\bullet(M; M_1, \dots, M_n)) \rightarrow L_{M_1, \dots, \widehat{M}_j, \dots, M_n}^{s_j(I \setminus j)} \rightarrow L_{M_1, \dots, M_n}^I \rightarrow H^1(C_I^\bullet(M; M_1, \dots, M_n)) \rightarrow H^1(C_{I \setminus j}^\bullet(M; M_1, \dots, M_n)) \rightarrow H^1(C_{s_j(I \setminus j)}^\bullet(M; M_1, \dots, \widehat{M}_j, \dots, M_n)) \rightarrow \dots$$

The inductive hypothesis immediately implies that

$$H^i(C_I^\bullet(M; M_1, \dots, M_n)) = 0 \quad \text{for all } i > 1,$$

and there is an exact sequence

$$0 \rightarrow H^0(C_{I \setminus j}^\bullet(M; M_1, \dots, M_n)) \rightarrow L_{M_1, \dots, \widehat{M}_j, \dots, M_n}^{s_j(I \setminus j)} \rightarrow L_{M_1, \dots, M_n}^I \rightarrow H^1(C_I^\bullet(M; M_1, \dots, M_n)) \rightarrow 0.$$

Assuming (2), we may employ Observation 3.13 to deduce the short exact sequence of (1). On the other hand, assuming (1) there is a commutative diagram with exact rows:

$$\begin{array}{ccccccc} 0 & \longrightarrow & L_{M_1, \dots, M_n}^{I \setminus j} & \longrightarrow & L_{M_1, \dots, \widehat{M}_j, \dots, M_n}^{s_j(I \setminus j)} & \longrightarrow & L_{M_1, \dots, M_n}^I & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & H^0(C_I^\bullet(M; M_1, \dots, M_n)) & \longrightarrow & L_{M_1, \dots, \widehat{M}_j, \dots, M_n}^{s_j(I \setminus j)} & \longrightarrow & L_{M_1, \dots, M_n}^I & \longrightarrow & H^1(C_I^\bullet(M; M_1, \dots, M_n)) \end{array}$$

Since the inner two maps are isomorphisms, both of the outer two inclusions are isomorphisms and thus $C_{I \setminus j}^\bullet(M; M_1, \dots, M_n)$ is also exact in positive degree for all $j \in I$. It follows that $C_I^\bullet(M; M_1, \dots, M_n)$ is exact in positive cohomological degrees for all I if and only if the sequence of (1) is exact for all I .

(1) \iff (3): This proof proceeds similarly: by induction on n , we may assume that the collection $M_1, \dots, \widehat{M}_j, \dots, M_n$ is distributive and $C_{\bullet}^{s_j(I)}(M; M_1, \dots, M_n)$ is exact in positive homological degrees, and likewise by a downward induction on $|I|$ we may assume that $H_1(C_{\bullet}^{I \cup j}(M; M_1, \dots, M_n))$ is exact in positive homological degrees.

Employing the long exact sequence of homology on the second short exact sequence (3.15.2) along with Observation 3.13 yields

$$\begin{aligned} \dots \rightarrow H_1(C_{\bullet}^{I \cup j}(M; M_1, \dots, M_n)) &\rightarrow H_1(C_{\bullet}^{s_j(I)}(M; M_1, \dots, M_n)) \\ &\rightarrow H_1(C_{\bullet}^I(M; M_1, \dots, M_n)) \rightarrow L_{M_1, \dots, M_n}^{[n] \setminus (I \cup j)} \rightarrow L_{M_1, \dots, \widehat{M}_j, \dots, M_n}^{[n-1] \setminus s_j(I)} \rightarrow L_{M_1, \dots, M_n}^{[n] \setminus I} \rightarrow 0. \end{aligned}$$

Set $I' := [n] \setminus I$ and notice that

$$[n] \setminus (I \cup j) = I' \setminus j \quad \text{and} \quad [n-1] \setminus s_j(I) = s_j([n] \setminus (I \cup j)) = s_j(I' \setminus j).$$

Again, this at least implies that $H_i(C_{\bullet}^I(M; M_1, \dots, M_n)) = 0$ for $i > 1$ and there is an exact sequence

$$0 \rightarrow H_1(C_{\bullet}^I(M; M_1, \dots, M_n)) \rightarrow L_{M_1, \dots, M_n}^{I' \setminus j} \rightarrow L_{M_1, \dots, \widehat{M}_j, \dots, M_n}^{s_j(I' \setminus j)} \rightarrow H_0(C_{\bullet}^I(M; M_1, \dots, M_n)) \rightarrow 0.$$

The result thus follows by identical reasoning as in (1) \iff (2). □

The following lemma shows that the property of distributivity “dualizes well”. This will be essential to prove the appropriate analogs of Koszul duality.

Lemma 3.16. *Let M be a flat R -module and $M_1, \dots, M_n \subset M$ a sequence of submodules with M/M_i a flat R -module for all $1 \leq i \leq n$. For a submodule $N \subset M$, use the notation $N^\vee := (M/N)^* := \text{Hom}_R(M/N, R)$. Then:*

(1) *Assume that $M/\sum_{i \in I} M_i$ is a flat R -module for all $I \subset [n]$. Then*

the collection $M_1, \dots, M_n \subset M$ is distributive \iff the collection $M_1^\vee, \dots, M_n^\vee \subset M^$ is distributive*

(notice that each M_i^\vee may be viewed as a submodule of M^ by dualizing the surjection $M \rightarrow M/M_i$). In this case, each of the modules L_{M_1, \dots, M_n}^I are R -flat.*

(2) *Assume that M_1, \dots, M_n is a distributive collection. Then there are isomorphisms*

$$\left(\frac{M}{\bigvee_{i \in I} M_i} \right)^* \cong \bigwedge_{i \in I} M_i^\vee \quad \text{and} \quad \left(\bigwedge_{i \in I} M_i \right)^* \cong \frac{M^*}{\bigvee_{i \in I} M_i^\vee},$$

where $I \subset [n]$.

(3) *Assume that M_1, \dots, M_n is a distributive collection. Then there are isomorphisms of complexes*

$$C_{\bullet}^I(M; M_1, \dots, M_n)^* \cong C_{\bullet}^I(M^*; M_1^\vee, \dots, M_n^\vee) \quad \text{and} \quad C_{\bullet}^I(M; M_1, \dots, M_n)^* \cong C_{\bullet}^I(M^*; M_1^\vee, \dots, M_n^\vee).$$

In particular, for all $I \subset [n]$ there is an isomorphism

$$(L_{M_1, \dots, M_n}^I)^* \cong L_{M_1^\vee, \dots, M_n^\vee}^{[n] \setminus I}.$$

Proof. The progression of the proof actually follows by first proving (2) and (3), then noting that (1) is an immediate consequence of (3).

Proof of (2). Proceed by induction on n , where the base case is for $n = 1$. In this case, the short exact sequence

$$0 \rightarrow M_i \rightarrow M \rightarrow M/M_i \rightarrow 0$$

implies that M_i is a flat R -module for all i , and hence dualizing yields the short exact sequence

$$0 \rightarrow (M/M_i)^* := M_i^\vee \rightarrow M^* \rightarrow M_i^* \rightarrow 0.$$

This implies that $M_i^* = M^*/M_i^\vee$, yielding the base case.

Let $n > 1$ and consider the complexes $C^\bullet(M; M_1, \dots, M_n)$ and $C_\bullet(M; M_1, \dots, M_n)$. Each term of these complexes falls within the inductive hypothesis, in which case we may dualize to obtain the isomorphisms

$$C^\bullet(M; M_1, \dots, M_n)^* \cong C_\bullet(M^*; M_1^\vee, \dots, M_n^\vee) \quad \text{and} \quad C_\bullet(M; M_1, \dots, M_n)^* \cong C^\bullet(M^*; M_1^\vee, \dots, M_n^\vee).$$

Taking zeroth (co)homology of each of the above complexes and employing Proposition 3.14 yields the isomorphisms

$$\left(\frac{M}{\bigvee_{i \in [n]} M_i} \right)^* \cong \bigwedge_{i \in [n]} M_i^\vee \quad \text{and} \quad \left(\bigwedge_{i \in [n]} M_i \right)^* \cong \frac{M^*}{\bigvee_{i \in [n]} M_i^\vee}.$$

Proof of (3). Since each term of the complex $C_i^\bullet(M; M_1, \dots, M_n)$ is of the form $M/\bigvee_{i \in J} M_i$ for some subset $J \subset [n]$, dualizing and using part (2) implies that $C_i^\bullet(M; M_1, \dots, M_n)^*$ has terms $\bigwedge_{i \in J} M_i^\vee$, and it is clear that the dualized differentials are the same as those of $C_i^I(M^*; M_1^\vee, \dots, M_n^\vee)$. The proof for $C_i^I(M; M_1, \dots, M_n)$ is identical and the isomorphism $(L_{M_1, \dots, M_n}^I)^* \cong L_{M_1^\vee, \dots, M_n^\vee}^{[n]I}$ follows upon taking zeroth (co)homology and using Proposition 3.14.

Proof of (1). Assume that M_1, \dots, M_n is distributive. The assumption that each quotient $M/\sum_{i \in I} M_i$ is flat implies that $\bigwedge_{i \in I} M_i$ is flat for all $I = (i_1, \dots, i_k) \subset [n]$, since each of the augmented complexes

$$\bigwedge_{i \in I} M_i \rightarrow C^\bullet(M; M_{i_1}, \dots, M_{i_k})$$

is exact and by assumption $C^\bullet(M; M_{i_1}, \dots, M_{i_k})$ is a complex of flat R -modules. Thus $\bigwedge_{i \in I} M_i$ is flat by Observation 2.12.

Using this, it follows that for every $I \subset [n]$ the chain complex $C_i^I(M; M_1, \dots, M_n)$ is exact in positive homological degrees and the zeroth homology is a flat R -module. This means that the dual $C_i^I(M; M_1, \dots, M_n)^*$ is a cochain complex with cohomology concentrated in degree 0, and by the isomorphism of (3) combined with Proposition 3.14, the collection $M_1^\vee, \dots, M_n^\vee$ is distributive. This argument is inherently symmetric in the roles of M_i and M_i^\vee , whence the result follows. \square

Corollary 3.17. *Let M be any projective R -module and $M_1, \dots, M_n \subset M$ a distributive collection of R -submodules. Assume that*

$$M / \sum_{i \in I} M_i$$

is a projective R -module for all $I \subset [n]$. Then for all subsets $I \subset [n]$, the R -module

$$L_{M_1, \dots, M_n}^I$$

is a projective R -submodule.

Proof. By (3) of Lemma 3.16, each of the modules L_{M_1, \dots, M_n}^I has a right resolution by projective R -modules given by the complex $C_j^*(M; M_1, \dots, M_n)$. By Observation 2.12, the R -module L_{M_1, \dots, M_n}^I must itself be projective. \square

4. A generalization of Schur modules for ribbon diagrams

In this section, we introduce ribbon Schur functors associated to arbitrary Koszul algebra (and module) inputs. We show that the exactness of the concatenation/near-concatenation sequence is actually equivalent to Koszulness and establish a set of properties generalizing many well-known properties for classically defined Schur functors corresponding to ribbon diagrams.

4.1. Standard operations between compositions. Before defining ribbon Schur functors, it will be helpful to establish multiple conventions and define certain natural operations on the Boolean/refinement posets. These operations are standard in the combinatorial literature, but for the sake of establishing conventions, we define things explicitly with examples here.

Definition 4.1 (operations on compositions). Given any integer $d > 0$, a *composition* of d is a tuple $\alpha = (\alpha_1, \dots, \alpha_k)$ with $\alpha_i > 0$ for each $1 \leq i \leq k$ and $|\alpha| := \alpha_1 + \dots + \alpha_k = d$. The integer k is the *length* of α , denoted $\ell(\alpha)$.

Given two compositions α and β , the *concatenation* of α and β , denoted $\alpha \cdot \beta$, is defined as the composition

$$\alpha \cdot \beta := (\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_j).$$

The *near-concatenation* of α and β , denoted $\alpha \odot \beta$, is defined as the composition

$$\alpha \odot \beta := (\alpha_1, \dots, \alpha_{k-1}, \alpha_k + \beta_1, \beta_2, \dots, \beta_j).$$

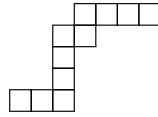
Finally, given any integer $d > 0$, the notation $\alpha^{(d)}$ is defined to be the composition

$$\alpha^{(d)} := (d\alpha_1, \dots, d\alpha_k).$$

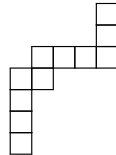
Definition 4.2 (ribbon diagrams associated to compositions). Given any composition α , one can associate the *ribbon diagram* by building a diagram whose row lengths (read from bottom to top) are given by $\alpha_1, \dots, \alpha_n$, and such that consecutive rows have overlap size precisely 1.

The *transpose* of a composition, denoted α^t , is the composition obtained by transposing the ribbon diagram associated to α .

Example 4.3. The ribbon diagram associated to the composition $(3, 1, 1, 2, 4)$ is the diagram



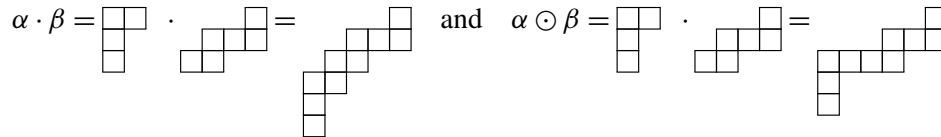
The transpose of this shape is given by the shape



and hence, upon reading the row lengths from bottom to top we find

$$(3, 1, 1, 2, 4)^t = (1, 1, 1, 2, 4, 1, 1).$$

Likewise, if $\alpha = (1, 1, 2)$ and $\beta = (2, 3, 1)$ then:



With the ribbon diagrams, we see that concatenation corresponds to “stacking” the diagrams, and near concatenation corresponds to merging the diagrams along their rows.

Definition 4.4. Let $n \in \mathbb{N}$ be any nonnegative integer and $C(n)$ denote the set of compositions of n into positive parts. Let $[n] := \{1, \dots, n\}$ with the convention that $[0] := \emptyset$, and let $[a, b] := \{a, a + 1, \dots, b\}$. The set $2^{[n-1]}$ is a poset with the standard Boolean poset structure, and $C(n)$ is also a poset with the standard refinement poset structure. Whenever the notation $\alpha \leq \beta$ is used for two compositions α and β , the partial order \leq is understood to be the refinement order.

There is moreover a standard isomorphism of posets

$$\phi : 2^{[n-1]} \rightarrow C(n).$$

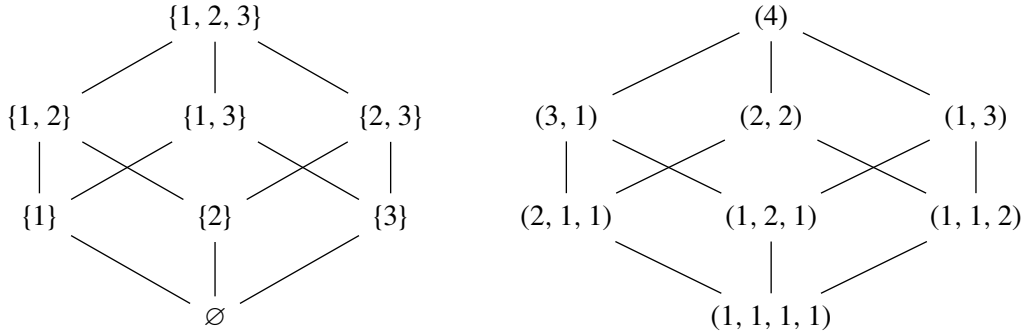
This isomorphism may be given explicitly as follows: given $I \subset [n - 1]$, write

$$I = [a_1, b_1] \cup [a_1, b_2] \cup \dots \cup [a_k, b_k]$$

for $a_i < b_i$ and $b_i < a_{i+1} - 1$ for each i . Then

$$\phi(I) = (1^{a_1-1}, b_1 - a_1 + 1, 1^{a_2-b_1-1}, b_2 - a_2 + 1, \dots, b_k - a_k + 1, n - b_k).$$

Example 4.5. The following example shows the Boolean poset $2^{\{1,2,3\}}$ and the refinement poset $C(4)$ side by side. The map ϕ is even a morphism of distributive lattices:



Observation 4.6. Let α and β be two compositions and $I \subset [n]$. Then

$$(\alpha \cdot \beta)^t = \beta^t \odot \alpha^t, \quad (\alpha \odot \beta)^t = \beta^t \cdot \alpha^t \quad \text{and} \quad \text{rev}(\phi(I)^t) = \phi([n] \setminus I),$$

where rev denotes the *reversal* operator, which simply reverses the order of the entries of a composition.

Definition 4.7 (partitioned compositions). Let $\alpha = (\alpha_1, \dots, \alpha_n)$ be a composition of some integer d . A *partition* of α is any choice of decomposition of α as a concatenation of subcompositions of α . The composition α is ℓ -partitioned if α is endowed with a partition that decomposes α into ℓ parts.

Given an ℓ -partitioned composition α , the notation $p_i(\alpha)$ denotes the i -th piece of partition of α . In other words, every ℓ -partitioned composition α may be written as the concatenation

$$\alpha = p_1(\alpha) \cdot p_2(\alpha) \cdots p_\ell(\alpha).$$

Remark 4.8. We will employ a minor abuse of notation when dealing with partitioned compositions, since the chosen partition will often not be specified. More precisely, the data of a partitioned composition includes both the data of a composition α and a chosen partition P , which can be encoded by any chosen subset of $[\ell(\alpha) - 1]$. The word “composition” without any adjective will only refer to a composition as defined in Definition 4.1.

Convention 4.9. By convention, any ℓ -partitioned composition α may be viewed as a j -partitioned composition for any $j < \ell$ by concatenating the last $\ell - j$ pieces of α . In other words, if

$$\alpha = p_1(\alpha) \cdots p_\ell(\alpha),$$

then α may be viewed as the j -partitioned composition

$$\alpha = p_1(\alpha) \cdots p_{j-1}(\alpha) \cdot (p_j(\alpha) \cdot p_{j+1}(\alpha) \cdots p_\ell(\alpha)).$$

This convention will become essential when we deal with multi-Schur modules, as defined later.

Example 4.10. Let $\alpha = (2, 2, 1, 4, 3)$ and consider the 4-partition of α induced by the subset $\{1, 3, 4\} \subset [4]$. This decomposes α as the concatenation

$$(2) \cdot (2, 1) \cdot (4) \cdot (3).$$

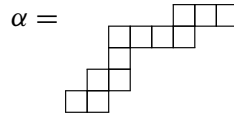
Via Convention 4.14, α may also be viewed as the 3-partitioned composition

$$(2) \cdot (2, 1) \cdot (4, 3).$$

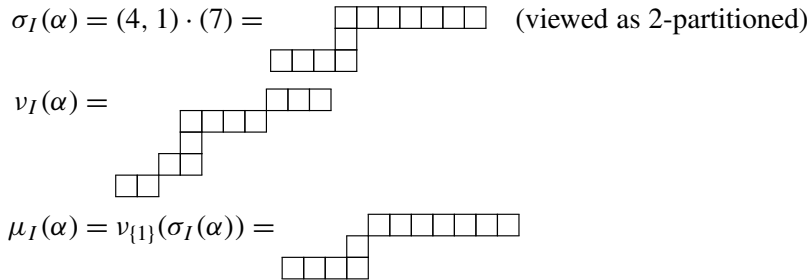
Definition 4.11. Let α be an ℓ -partitioned composition and let $I = \{i_1 < \dots < i_k\} \subset [\ell - 1]$ be any subset:

- (1) The notation $\sigma_I(\alpha)$ denotes the composition obtained by near-concatenating $p_j(\alpha)$ and $p_{j+1}(\alpha)$ for every $j \in I$.
- (2) The notation $\nu_I(\alpha)$ denotes the ribbon diagram obtained by disconnecting $p_j(\alpha)$ and $p_{j+1}(\alpha)$ for every $j \in I$.
- (3) The notation $\mu_I(\alpha)$ denotes the composition $\nu_{[\ell-1-|I|]} \circ \sigma_I(\alpha)$.

Example 4.12. Let $\alpha := (2) \cdot (2, 1) \cdot (4) \cdot (3)$ be the 4-partitioned shape of Example 4.10. As a ribbon diagram:



Let $I := \{1, 3\}$. Then:



Remark 4.13. The operation μ_I can be reformulated in words as the operation that near-concatenates all elements $p_i(\alpha)$ and $p_{i+1}(\alpha)$ for $i \in I$, then disconnects everything else.

Convention 4.14. Let α and β be j and k -partitioned compositions, respectively. By convention, the concatenation $\alpha \cdot \beta$ will be viewed as a $j + k$ -partitioned composition with

$$p_s(\alpha \cdot \beta) = \begin{cases} p_s(\alpha) & \text{if } s \leq j, \\ p_{s-j}(\beta) & \text{if } s > j. \end{cases}$$

Likewise, the near-concatenation $\alpha \odot \beta$ will be viewed as a $j + k - 1$ -partitioned composition with

$$p_s(\alpha \odot \beta) = \begin{cases} p_s(\alpha) & \text{if } s < k, \\ p_k(\alpha) \odot p_1(\beta) & \text{if } s = k, \\ p_{s-j+1}(\beta) & \text{if } s > k. \end{cases}$$

Definition 4.15. Let A be any quadratic R -algebra. The notation A_d for any integer $d \in \mathbb{Z}$ denotes the degree d component of A . Given a tuple $\alpha = (\alpha_1, \dots, \alpha_n)$, use the notation

$$A_\alpha := A_{\alpha_1} \otimes_R \cdots \otimes_R A_{\alpha_n}.$$

Likewise, given a quadratic left (resp. right) A -module N (resp. M), use the notation

$$(M \otimes_R A)_\alpha := M_{\alpha_1} \otimes_R A_{\alpha_2} \otimes_R \cdots \otimes_R A_{\alpha_n} \quad \text{and} \quad (A \otimes_R N)_\alpha := A_{\alpha_1} \otimes_R \cdots \otimes_R A_{\alpha_{n-1}} \otimes_R N_{\alpha_n}.$$

We use the convention that $(M \otimes_R A)_\alpha = M_{\alpha_1}$ and $(A \otimes_R N)_\alpha := N_{\alpha_1}$ if $\ell(\alpha) = 1$.

Similarly, for $\ell(\alpha) \geq 2$, use the notation

$$(M \otimes_R A \otimes_R N)_\alpha := M_{\alpha_1} \otimes_R A_{\alpha_2} \otimes_R \cdots \otimes_R A_{\alpha_{n-1}} \otimes_R N_{\alpha_n},$$

where we use the convention that $(M \otimes_R A \otimes_R N)_\alpha = M_{\alpha_1} \otimes_R N_{\alpha_2}$ if $\ell(\alpha) = 2$.

4.2. Ribbon Schur modules for Koszul algebras. In this subsection, we will hold off on proving the statements until the next subsection (where the statements will be proved in more generality). This subsection will mainly be used to illustrate the construction of ribbon Schur functors and their various properties with examples.

Definition 4.16. Let A be a Koszul A -module and α any composition. Define the *ribbon Schur module* \mathbb{S}_A^α as the kernel of the natural map

$$\mathbb{S}_A^\alpha = \text{Ker} \left(A_\alpha \rightarrow \bigoplus_{\alpha < \beta} A_\beta \right).$$

where $<$ denotes the covering relation in the refinement poset.

Remark 4.17. Since the ribbon Schur module \mathbb{S}_A^α is defined only using the algebra structure on A , this definition is canonical and is well-defined for any Koszul algebra A (in fact, it is well-defined for any algebra, but the Koszul property will ensure uniformity in the properties of these objects).

Example 4.18. Let $\alpha = (3, 2, 1, 3)$. Then in the refinement poset, α is covered by the compositions

$$(5, 1, 3), \quad (3, 3, 3) \quad \text{and} \quad (3, 2, 4).$$

Thus for any Koszul algebra $\mathbb{S}_A^{(3,2,1,3)}$ is defined to be the kernel of the map:

$$\begin{array}{c} A_5 \otimes_R A_1 \otimes_R A_3 \\ \oplus \\ A_3 \otimes_R A_2 \otimes_R A_1 \otimes_R A_3 \rightarrow A_3 \otimes_R A_3 \otimes_R A_3 \\ \oplus \\ A_3 \otimes_R A_2 \otimes_R A_4 \end{array}$$

The following proposition justifies usage of the terminology ‘‘ribbon Schur functor’’.

Proposition 4.19. *Formation of the ribbon Schur module \mathbb{S}_A^α is functorial in the ring argument; that is, given a morphism of R -algebras $\psi : A \rightarrow B$, there is an induced morphism*

$$\mathbb{S}_\psi^\alpha : \mathbb{S}_A^\alpha \rightarrow \mathbb{S}_B^\alpha.$$

Example 4.20. When $A = S(V)$, the symmetric algebra on a free R -module V , the ribbon Schur module coincides with the classical Schur module definition corresponding to the skew ribbon shape defined by α .

Likewise, for $A = \bigwedge^\bullet V$ the exterior algebra on some free R -module, the ribbon Schur module coincides with the Weyl module associated to the ribbon diagram of α .

The functoriality of the classical Schur and Weyl functors is precisely that of Proposition 4.19, and there are isomorphisms of functors

$$\mathbb{S}_{S(-)}^\alpha = \mathbb{S}^\alpha(-), \quad \mathbb{S}_{\bigwedge^\bullet(-)}^\alpha = \mathbb{W}^\alpha(-).$$

Example 4.21. When $A = T(V)$, the tensor algebra on some flat R -module V , the ribbon Schur modules are particularly simple:

$$\mathbb{S}_{T(V)}^\alpha = \begin{cases} V^{\otimes \alpha_1} & \text{if } \ell(\alpha) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Example 4.22. If $\alpha = (1^i)$ for some integer $i \geq 1$, then by definition

$$\mathbb{S}_A^{(1^i)} = (A^!)_i^*.$$

Thus the Priddy complex may be reformulated as the complex

$$\cdots \rightarrow A \otimes_R \mathbb{S}_A^{(1^i)} \rightarrow A \otimes_R \mathbb{S}_A^{(1^{i-1})} \rightarrow \cdots \rightarrow A \otimes_R \mathbb{S}_A^{(1)} \rightarrow A \rightarrow 0.$$

Lemma 4.23. *Let A be a Koszul R -algebra. Then:*

- (1) *For all compositions α , the ribbon Schur module \mathbb{S}_A^α is a flat R -module.*
- (2) *For any two compositions α and β there is a canonical short exact sequence of R -modules*

$$0 \rightarrow \mathbb{S}_A^{\alpha \cdot \beta} \rightarrow \mathbb{S}_A^\alpha \otimes_R \mathbb{S}_A^\beta \rightarrow \mathbb{S}_A^{\alpha \odot \beta} \rightarrow 0.$$

Moreover, this sequence is exact for all compositions if and only if A is a Koszul algebra.

- (3) *There is a canonical isomorphism of R -modules*

$$(\mathbb{S}_A^\alpha)^* \cong \mathbb{S}_{A^!}^{\alpha^!}.$$

If A is assumed R -projective, then the ribbon Schur module \mathbb{S}_A^α is R -projective for all compositions α .

Remark 4.24. As mentioned in the introduction, in the theory of symmetric functions the short exact sequence of (2) is an explicit realization of the so-called *concatenation/near-concatenation* identity.

Remark 4.25. Notice that Observation 4.6 implies that the sequences

$$0 \rightarrow \mathbb{S}_A^{\alpha \cdot \beta} \rightarrow \mathbb{S}_A^\alpha \otimes_R \mathbb{S}_A^\beta \rightarrow \mathbb{S}_A^{\alpha \odot \beta} \rightarrow 0 \quad \text{and} \quad 0 \rightarrow \mathbb{S}_{A^!}^{\beta^! \cdot \alpha^!} \rightarrow \mathbb{S}_{A^!}^{\beta^!} \otimes_R \mathbb{S}_{A^!}^{\alpha^!} \rightarrow \mathbb{S}_{A^!}^{\beta^! \odot \alpha^!} \rightarrow 0$$

are naturally dual to each other.

Example 4.26. Let $\mathbb{S}^\alpha(-)$ and $\mathbb{W}^\beta(-)$ denote the classical Schur and Weyl functors corresponding to the skew ribbon shapes defined by the compositions α and β , respectively. Then by definition, for any free R -module V there are equalities

$$\mathbb{S}^\alpha(V) = \mathbb{S}_{S(V)}^\alpha, \quad \mathbb{W}^{\alpha'}(V^*) = \mathbb{S}_{\bigwedge^\bullet V^*}^\alpha.$$

The duality mentioned in the statement of Lemma 4.23 is thus a generalization of the well-known isomorphism

$$\mathbb{S}^\alpha(V)^* \cong \mathbb{W}^{\alpha'}(V^*).$$

Next, we introduce a class of complexes whose terms are tensor products of ribbon Schur functors; this complex generalizes the short exact sequence of Lemma 4.23(2); more defining this complex we make a simple observation:

Observation 4.27. Recall the operator μ_I as in Definition 4.11. Let A be a Koszul algebra and α any ℓ -partitioned composition. Then for all $I \subset J$ there is a natural surjection

$$\rho_{I,J} : \mathbb{S}_A^{\mu_I(\alpha)} \rightarrow \mathbb{S}_A^{\mu_J(\alpha)}.$$

This surjection is defined by taking the tensor products of the surjections as in the right map of the short exact sequence of Lemma 4.23 (as dictated by the subsets I and J).

Definition 4.28. Let α be any ℓ -partitioned composition. Define the cochain complex $(\mathcal{H}_A(\alpha), \delta)$ whose i -th term is

$$\mathcal{H}_A^i(\alpha) := \bigoplus_{\substack{I \subseteq [\ell-1]: \\ |I|=i}} \mathbb{S}_A^{\mu_I(\alpha)},$$

with differential

$$d^{\mathcal{H}_A} \big|_{\mathbb{S}^{\mu_I(\alpha)}} := \sum_{j \notin I} \text{sgn}(j, I) \rho_{I, I \cup j}.$$

Theorem 4.29. For all ℓ -partitioned compositions α , the complex $\mathcal{H}_A(\alpha)$ is a cochain complex with

$$H^0(\mathcal{H}_A^i(\alpha)) \cong \mathbb{S}_A^\alpha,$$

and which is exact in positive cohomological degrees. If A also has a compatible action by a group G , then the complex $\mathcal{H}(\alpha)$ is also G -equivariant.

Remark 4.30. The complexes $\mathcal{H}_A^\bullet(\alpha)$ for an ℓ -partitioned composition α may be viewed as a categorification of the Hamel–Goulden identities (see [Hamel and Goulden 1995]) associated to the (horizontal) ribbon decomposition induced by the partitioning data of α .

Example 4.31. Let $\alpha = (3, 1, 2, 2, 4, 1, 5)$, viewed as a 4-partitioned composition with $p_1(\alpha) = (3, 1)$, $p_2(\alpha) = (2, 2)$, $p_3(\alpha) = (4)$, and $p_4(\alpha) = (1, 5)$. Then the complex $\mathcal{H}_A(\alpha)$ takes the following form:

$$\begin{array}{c}
 \mathbb{S}_A^{(3,1)} \otimes_R \mathbb{S}_A^{(2,2)} \otimes_R \mathbb{S}_A^{(4)} \otimes_R \mathbb{S}_A^{(1,5)} \\
 \downarrow \\
 \mathbb{S}_A^{(3,3,2)} \otimes_R \mathbb{S}_A^{(4)} \otimes_R \mathbb{S}_A^{(1,5)} \oplus \mathbb{S}_A^{(3,1)} \otimes_R \mathbb{S}_A^{(2,6)} \otimes_R \mathbb{S}_A^{(1,5)} \oplus \mathbb{S}_A^{(3,1)} \otimes_R \mathbb{S}_A^{(2,2)} \otimes_R \mathbb{S}_A^{(5,5)} \\
 \downarrow \\
 \mathbb{S}_A^{(3,3,6)} \otimes_R \mathbb{S}_A^{(1,5)} \oplus \mathbb{S}_A^{(3,3,2)} \otimes_R \mathbb{S}_A^{(5,5)} \oplus \mathbb{S}_A^{(3,1)} \otimes_R \mathbb{S}_A^{(2,7,5)} \\
 \downarrow \\
 \mathbb{S}_A^{(3,3,7,5)}
 \end{array}$$

4.3. Ribbon Schur modules with Koszul module inputs. In this section, we generalize the ribbon Schur functors to allow for Koszul module inputs as well. The same type of concatenation/near-concatenation sequence may be used to detect Koszulness of modules. In the following, recall the notation established in Definition 4.15.

Definition 4.32. Let α be any composition and recall the conventions of Definition 4.15. Given a Koszul left (resp. right) A -module M with initial degree t , define the ribbon Schur module $\mathbb{S}_{A,M}^\alpha$ (resp. $\mathbb{S}_{M,A}^\alpha$) as the kernel of the natural map

$$\begin{aligned}
 \mathbb{S}_{A,M}^\alpha &:= \text{Ker} \left((A \otimes_R M)_{\alpha \cdot (t)} \rightarrow \bigoplus_{\alpha \cdot (t) \leq \beta} (A \otimes_R M)_\beta \right), \\
 \text{resp. } \mathbb{S}_{M,A}^\alpha &:= \text{Ker} \left((M \otimes_R A)_{(t) \cdot \alpha} \rightarrow \bigoplus_{(t) \cdot \alpha \leq \beta} (M \otimes_R A)_\beta \right).
 \end{aligned}$$

Given a Koszul left (resp. right) A -module N (resp. M) of initial degree s (resp. t), define the ribbon Schur module $\mathbb{S}_{M,A,N}^\alpha$ as the kernel of the natural map

$$\mathbb{S}_{M,A,N}^\alpha := \text{Ker} \left((M \otimes_R A \otimes_R N)_{(t) \cdot \alpha \cdot (s)} \rightarrow \bigoplus_{(t) \cdot \alpha \cdot (s) \leq \beta} (M \otimes_R A \otimes_R N)_\beta \right).$$

Remark 4.33. Notice that if α is the empty partition, we use the convention that

$$\mathbb{S}_{A,M}^0 = M_t, \quad \mathbb{S}_{M,A,N}^0 = M_t \otimes_R N_s.$$

Notation 4.34. The definition of $\mathbb{S}_{M,A,N}^\alpha$ may be extended to disconnected ribbon diagrams by using the convention that disconnected portions of the diagram correspond to tensor products of the respective Schur modules. We will use this convention for the remainder of the paper, since it drastically simplifies notational issues in the following proofs/constructions.

Example 4.35. Let $\alpha := (3, 1, 1)$ and $\beta = (2, 2)$. Then by the convention of Notation 4.34, there is an equality

$$\mathbb{S}_A^{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}} = \mathbb{S}_A^{\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}} \otimes_R \mathbb{S}_A^{\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}} = \mathbb{S}_A^{(3,1,1)} \otimes_R \mathbb{S}_A^{(2,2)}.$$

Example 4.36. Let A be any Koszul algebra and α a composition. Recall that every truncation of A is Koszul by Corollary A.21, whence A_+^r is a Koszul (left and right) A -module for all $r \geq 0$. By definition, there are equalities

$$\mathbb{S}_{A_+^r, A, A_+^{r'}}^\alpha = \mathbb{S}_A^{(r) \cdot \alpha \cdot (r')}.$$

The following definition is an evident extension of Definition A.10 to allow for both left and right A -modules:

Definition 4.37. Given a Koszul left (resp. right) A -module N (resp. M) with initial degrees s and t , respectively, define the collection $S_{M,A,N,1}^n, \dots, S_{M,A,N,n-1}^n \subset M_t \otimes_R A_1^{\otimes n-1} \otimes_R N_s$ for $n \geq 2$ via

$$S_{M,A,N,i}^n := \begin{cases} Q_{t+1}^M \otimes_R A_1^{\otimes n-2} \otimes_R N_s & \text{if } i = 1, \\ Q_t^M \otimes_R A_1^{\otimes i-3} \otimes_R Q_2^A \otimes_R A_1^{\otimes n-i+1} \otimes_R N_s & \text{if } 2 \leq i \leq n-2, \\ M_t \otimes_R A_1^{\otimes n-2} \otimes_R Q_{s+1}^N & \text{if } i = n-1. \end{cases}$$

If $n = 1$, use the convention that $S_{M,A,N,0}^1 := 0$.

Recall the notation of μ_I for a subset I as established in Definition 4.11 in the following:

Definition 4.38. Let A be a Koszul algebra and M any left A -module of initial degree s . Given any ℓ -partitioned composition α , view the concatenation $\alpha \cdot (s)$ as $(\ell + 1)$ -partitioned via Convention 4.14. Then, for any subset $I \subset [\ell]$ write $\mu_I(\alpha \cdot (s)) = \alpha^1 \cup \dots \cup \alpha^{\ell-|I|}$ as a disjoint union of compositions $\alpha^1, \dots, \alpha^{\ell-|I|}$, where $\alpha^{\ell-|I|}$ has length g . Then we define

$$\mathbb{S}_{A,M}^{\mu_I(\alpha)} := \mathbb{S}_A^{\alpha^1} \otimes_R \mathbb{S}_A^{\alpha^2} \otimes_R \dots \otimes_R \mathbb{S}_{A, M_{\geq \alpha_g^{\ell-|I|} + s}}^{\alpha^{\ell-|I|}}.$$

As a quick example let $\alpha = (2, 1) \cdot (3, 4) \cdot (4, 3, 3)$ be a 3-partitioned composition and $I = \{1, 3\} \subset [3]$. Then for any Koszul algebra A and left A -module M of initial degree s , there is an equality

$$\mathbb{S}_{A,M}^{\mu_I(\alpha)} := \mathbb{S}_A^{(2,4,4)} \otimes_R \mathbb{S}_{A, M_{\geq 3+s}}^{(4,3)}.$$

Remark 4.39. The slight technicality in defining $\mathbb{S}_{A,M}^{\mu_I(\alpha)}$ when taking account of a module input stems from the fact that M has its own initial degree, in which case we should be defining the corresponding ribbons with respect to the composition $\alpha \cdot (s)$.

The following observation is an immediate consequence of the identification of the ribbon Schur functors with the modules of Definition 3.11, noted in Observation 4.42.

Observation 4.40. Let A be a Koszul algebra and M any left A -module. Given an ℓ -partitioned composition α and subsets $I \subset J \subset [\ell]$, there is a canonical morphism of R -modules

$$\mathbb{S}_A^{\mu_I(\alpha)} \rightarrow \mathbb{S}_A^{\mu_J(\alpha)}.$$

The following observation is a direct consequence of the definition of a ribbon Schur functor, but will be very useful later on:

Observation 4.41. Let $\alpha, \beta,$ and γ be any (possibly empty) compositions. Given a Koszul algebra A and a Koszul left (resp. right) A -module N (resp. M), there is an equality

$$\mathbb{S}_{M,A,N}^{\alpha \cdot \beta \cdot \gamma} = (\mathbb{S}_{M,A}^{\alpha \cdot \beta} \otimes_R \mathbb{S}_{A,N}^\gamma) \cap (\mathbb{S}_{M,A}^\alpha \otimes_R \mathbb{S}_{A,N}^{\beta \cdot \gamma}),$$

where the intersection is being viewed as taking place in $(M \otimes_R A \otimes_R N)_{\alpha \cdot \beta \cdot \gamma}$.

Finally, we are able to tie ribbon Schur functors to the modules introduced in Definition 3.11. This reformulation combined with the equivalence of Proposition 3.14 will yield quick proofs of the concatenation/near-concatenation formulation of Koszulness.

Observation 4.42. Let α be any composition of length ℓ . Given a Koszul algebra A and a Koszul left A -module M , there is an isomorphism of R -modules

$$\mathbb{S}_{A,M}^\alpha = L_{S_{A,M,1}^n, \dots, S_{A,M,n-1}^n}^{\phi^{-1}(\alpha)},$$

where the map ϕ is the isomorphism of posets of Definition 4.4 and the module L_{M_1, \dots, M_n}^I is from Definition 3.11. The analogous statement for right Koszul modules also holds.

Likewise, given a Koszul left (resp. right) A -module N (resp. M), there is an equality

$$\mathbb{S}_{M,A,N}^\alpha = L_{S_{M,A,N,1}^n, \dots, S_{M,A,N,n-1}^n}^{\phi^{-1}(\alpha)}.$$

We arrive at the statement and proof of the relevant properties for ribbon Schur functors associated to Koszul algebras/modules; again, the proof here is deceptively short, but the proof combines all of the machinery developed thus far. Recall the notation for the reversal operator rev from Observation 4.6.

Lemma 4.43. *Let A be a Koszul R -algebra and M (resp. N) a Koszul right (resp. left) A -module. Then:*

(1) *For all compositions α , the ribbon Schur module $\mathbb{S}_{A,M}^\alpha$ (resp. $\mathbb{S}_{N,A}^\alpha$) is a flat R -module. If A and M are R -projective, then $\mathbb{S}_{A,M}^\alpha$ (resp. $\mathbb{S}_{N,A}^\alpha$) is also R -projective.*

(2) *For any two compositions α and β there is a short exact sequence of R -modules*

$$0 \rightarrow \mathbb{S}_{A,M}^{\alpha \cdot \beta} \rightarrow \mathbb{S}_A^\alpha \otimes_R \mathbb{S}_{A,M}^\beta \rightarrow \mathbb{S}_{A,M}^{\alpha \circ \beta} \rightarrow 0 \quad \text{resp.} \quad 0 \rightarrow \mathbb{S}_{N,A}^{\alpha \cdot \beta} \rightarrow \mathbb{S}_{N,A}^\alpha \otimes_R \mathbb{S}_A^\beta \rightarrow \mathbb{S}_{N,A}^{\alpha \circ \beta} \rightarrow 0.$$

Conversely, if the sequences

$$0 \rightarrow \mathbb{S}_{A,M_{>d}}^{\alpha \cdot \beta} \rightarrow \mathbb{S}_A^\alpha \otimes_R \mathbb{S}_{A,M_{>d}}^\beta \rightarrow \mathbb{S}_{A,M_{>d}}^{\alpha \circ \beta} \rightarrow 0 \quad \text{resp.} \quad 0 \rightarrow \mathbb{S}_{N_{>d},A}^{\alpha \cdot \beta} \rightarrow \mathbb{S}_{N_{>d},A}^\alpha \otimes_R \mathbb{S}_A^\beta \rightarrow \mathbb{S}_{N_{>d},A}^{\alpha \circ \beta} \rightarrow 0.$$

are exact for all integers $d \geq 0$ and compositions α, β , then M (resp. N) is a Koszul left (resp. right) A -module.

(3) *There is a canonical isomorphism of R -modules*

$$(\mathbb{S}_{A,M}^\alpha)^* \cong \mathbb{S}_{M',A'}^{\text{rev}(\alpha')} \quad \text{resp.} \quad (\mathbb{S}_{N,A}^\alpha)^* \cong \mathbb{S}_{A',N'}^{\text{rev}(\alpha')}.$$

Remark 4.44. There is a convention here that is important to take note of in the short exact sequence of (2): if the composition β is the empty composition, write $\alpha = \alpha' \cdot (\alpha_n)$; the short exact sequence (2) then reads

$$0 \rightarrow \mathbb{S}_{A,M}^\alpha \rightarrow \mathbb{S}_A^\alpha \otimes_R M_t \rightarrow \mathbb{S}_{A,M_{\geq t+\alpha_n}}^{\alpha'} \rightarrow 0.$$

There is a similar convention in the right module case: if α is empty, write $\beta = (\beta_1) \cdot \beta'$. Then the short exact sequence reads

$$0 \rightarrow \mathbb{S}_{N,A}^\beta \rightarrow N_t \otimes_R \mathbb{S}_A^\beta \rightarrow \mathbb{S}_{N_{\geq t+\beta_1},A}^{\beta'} \rightarrow 0.$$

Notice moreover that the reason Lemma 4.23(3) does not have the reversal $\text{rev}(\alpha')$ is because there is by construction a canonical isomorphism $\mathbb{S}_A^{\alpha'} \cong \mathbb{S}_A^{\text{rev}(\alpha')}$ for any composition α .

Proof. Proof of (1). This just combines Lemma 3.16(1), Theorem A.19, and Observation 4.42.

Proof of (2). This is precisely the short exact sequence of Proposition 3.14, so again the result follows from Theorem A.19 combined with Observation 4.42.

Proof of (3). This just combines Lemma 3.16(3), Theorem A.19, and Observation 4.42. □

The case of having double module inputs in the associated Schur functor has to be treated separately, since the distributivity criterion cannot be used directly:

Corollary 4.45. *Let A be a Koszul R -algebra and N (resp. M) any left (resp. right) Koszul A -module. Then:*

- (1) *For all compositions α , the ribbon Schur module $\mathbb{S}_{M,A,N}^\alpha$ is a flat R -module. If A , M , and N are R -projective, then $\mathbb{S}_{M,A,N}^\alpha$ is R -projective.*
- (2) *For any two compositions α and β there is a short exact sequence of R -modules*

$$0 \rightarrow \mathbb{S}_{M,A,N}^{\alpha \cdot \beta} \rightarrow \mathbb{S}_{M,A}^\alpha \otimes_R \mathbb{S}_{A,N}^\beta \rightarrow \mathbb{S}_{M,A,N}^{\alpha \odot \beta} \rightarrow 0.$$

(3) *There is a canonical isomorphism of R -modules*

$$(\mathbb{S}_{M,A,N}^\alpha)^* \cong \mathbb{S}_{N',A',M'}^{\text{rev}(\alpha')}.$$

Proof. Notice that (3) follows from Lemma 3.16(3) combined with Observation 4.42.

Proof of (2). As noted in Remark 3.15, the sequence

$$0 \rightarrow \mathbb{S}_{M,A,N}^{\alpha \cdot \beta} \rightarrow \mathbb{S}_{M,A}^\alpha \otimes_R \mathbb{S}_{A,N}^\beta \rightarrow \mathbb{S}_{M,A,N}^{\alpha \odot \beta} \rightarrow 0$$

is exact at the left and rightmost terms, so it suffices to prove exactness for the middle term. Assume first that $\ell(\alpha) = 1$ and $\ell(\beta) = 0$. Then there is a commutative diagram:

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 0 & \longrightarrow & \mathbb{S}_{M,A,N}^{(\alpha_1)} & \longrightarrow & \mathbb{S}_{M,A}^{(\alpha_1)} \otimes_R N_s & \longrightarrow & M_t \otimes_R N_{s+\alpha_1} \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \parallel \\
 0 & \longrightarrow & M_t \otimes_R \mathbb{S}_{A,N}^{(\alpha_1)} & \longrightarrow & M_t \otimes_R A_{\alpha_1} \otimes_R N_s & \longrightarrow & M_t \otimes_R N_{s+\alpha_1} \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \\
 0 & \longrightarrow & M_{t+\alpha_1} \otimes_R N_s & \xlongequal{\quad} & M_{t+\alpha_1} \otimes_R N_s & & \\
 & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & &
 \end{array}$$

The middle row and column is exact by Lemma 4.43(2), and the last row and column are evidently exact. Let Θ denote the surjection

$$\Theta : \mathbb{S}_{M,A}^{(\alpha_1)} \otimes_R N_s \rightarrow M_t \otimes_R N_{s+\alpha_1}.$$

A quick diagram chase shows that

$$\text{Ker}(\Theta) \subset (\mathbb{S}_{M,A}^{(\alpha_1)} \otimes_R N_s) \cap (M_t \otimes_R \mathbb{S}_{A,N}^{(\alpha_1)}),$$

and by Observation 4.41 this intersection is precisely the Schur module $\mathbb{S}_{M,A,N}^{(\alpha_1)}$. Since the reverse inclusion evidently holds, the top row is exact.

Assume now that $\ell(\alpha), \ell(\beta) > 0$. Write $\beta = \beta' \cdot (\beta_k)$ for some composition β' with $\ell(\beta') = \ell(\beta) - 1$. Then there is a commutative diagram:

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 0 & \longrightarrow & \mathbb{S}_{M,A,N}^{\alpha \cdot \beta} & \longrightarrow & \mathbb{S}_{M,A}^{\alpha} \otimes_R \mathbb{S}_{A,N}^{\beta} & \longrightarrow & \mathbb{S}_{M,A,N}^{\alpha \odot \beta} \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 0 & \longrightarrow & \mathbb{S}_{M,A}^{\alpha \cdot \beta'} \otimes_R \mathbb{S}_{A,N}^{(\beta_k)} & \longrightarrow & \mathbb{S}_{M,A}^{\alpha} \otimes_R \mathbb{S}_A^{\beta'} \otimes_R \mathbb{S}_{A,N}^{(\beta_k)} & \longrightarrow & \mathbb{S}_{M,A}^{\alpha \odot \beta'} \otimes_R \mathbb{S}_{A,N}^{(\beta_k)} \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 0 & \longrightarrow & \mathbb{S}_{M,A,N}^{\alpha \cdot \beta' \odot (\beta_k)} & \longrightarrow & \mathbb{S}_{M,A}^{\alpha} \otimes_R \mathbb{S}_{A,N}^{\beta' \odot (\beta_k)} & \longrightarrow & \mathbb{S}_{M,A,N}^{\alpha \odot \beta' \odot (\beta_k)} \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 & & 0 & & 0 & & 0
 \end{array}$$

The bottom two rows and the rightmost two columns are exact by induction (combined with the flatness proved in Lemma 4.43), and an identical diagram chase as above shows that the top row must also be exact.

Proof of (1). Proceed by induction on the length $\ell(\alpha)$, where the base cases $\ell(\alpha) = 0$ or 1 are evident since $\mathbb{S}_{M,A,N}^\emptyset := M_t \otimes_R N_s$ and there is a short exact sequence

$$0 \rightarrow \mathbb{S}_{M,A,N}^{(\alpha_1)} \rightarrow M_t \otimes_R \mathbb{S}_{A,N}^{(\alpha_1)} \rightarrow M_{t+\alpha_1} \otimes_R N_s \rightarrow 0.$$

The latter two terms in this sequence are flat (resp. projective) by Lemma 4.43, so $\mathbb{S}_{M,A,N}^{(\alpha_1)}$ is flat (resp. projective) by Observation 2.12.

Assuming now that $\ell(\alpha) \geq 2$, write $\alpha = \beta \cdot \gamma$ for two compositions β, γ with $\ell(\beta), \ell(\gamma) > 0$. Then by (2) there is a short exact sequence

$$0 \rightarrow \mathbb{S}_{M,A,N}^\alpha \rightarrow \mathbb{S}_{M,A}^\beta \otimes_R \mathbb{S}_{A,M}^\gamma \rightarrow \mathbb{S}_{M,A,N}^{\beta \circ \gamma} \rightarrow 0.$$

The latter two terms are flat (resp. projective) by the induction hypothesis, whence Observation 2.12 again implies that $\mathbb{S}_{M,A,N}^\alpha$ is flat (resp. projective). □

Definition 4.46. Let α be any ℓ -partitioned composition. Given a Koszul algebra A and a Koszul left A -module M , define the cochain complex $(\mathcal{H}_{A,M}(\alpha), \delta)$ whose i -th term is

$$\mathcal{H}_A^i(\alpha) := \bigoplus_{\substack{I \subseteq [\ell]: \\ |I|=i}} \mathbb{S}_{A,M}^{\mu_I(\alpha)},$$

with differential

$$d^{\mathcal{H}_{A,M}}|_{\mathbb{S}^{\mu_I(\alpha)}} := \sum_{j \notin I} \text{sgn}(j, I) \rho_{I, I \cup j}.$$

Remark 4.47. Note the difference between the components of the modules for the complex of Definition 4.46 versus Definition 4.28: in the first case, the direct sums are parametrized by subsets of $[\ell]$ because of the additional component coming from the module M , whereas the components of Definition 4.28 only range over subsets of $[\ell - 1]$ since there is no additional module M .

Finally, we conclude this section with the analog of the Hamel–Goulden type complexes from Definition 4.28 for ribbon Schur functors admitting a module input. Recall the notation of Definition 4.38 in the theorem below:

Theorem 4.48. *For all ℓ -partitioned composition α , the complex $\mathcal{H}_{A,M}(\alpha)$ is a cochain complex with*

$$H^0(\mathcal{H}_{A,M}(\alpha)) \cong \mathbb{S}_{A,M}^\alpha,$$

and which is exact in positive cohomological degrees.

Proof. First, observe that the statement that $H^0(\mathcal{H}_{A,M}(\alpha)) \cong \mathbb{S}_{A,M}^\alpha$ is evidently true. Proceed by induction on ℓ , with the base case $\ell = 2$. In this case, write $\alpha = \beta \cdot \gamma$; the complex $\mathcal{H}_{A,M}(\alpha)$ simply becomes the length 1 complex

$$\mathcal{H}_{A,M}(\alpha) : \mathbb{S}_A^\beta \otimes_R \mathbb{S}_{A,M}^\gamma \rightarrow \mathbb{S}_{A,M}^{\beta \odot \gamma} \rightarrow 0.$$

These are the last two terms of the short exact sequence from Lemma 4.43(2), so the statement holds.

Assume now that $\ell(\alpha) > 2$. Write $\alpha = \alpha' \cdot p_\ell(\alpha)$. Then there is a short exact sequence of complexes:

$$0 \rightarrow \mathcal{H}_{A,M}(\alpha' \odot p_\ell(\alpha))[-1] \rightarrow \mathcal{H}_{A,M}(\alpha) \rightarrow \mathcal{H}_A(\alpha') \otimes_R \mathbb{S}_{A,M}^{p_\ell(\alpha)} \rightarrow 0.$$

By the inductive hypothesis, both of the complexes $\mathcal{H}_{A,M}(\alpha' \odot p_\ell(\alpha))$ and $\mathcal{H}_A(\alpha') \otimes_R \mathbb{S}_{A,M}^{p_\ell(\alpha)}$ are exact in positive cohomological degrees; taking the long exact sequence of cohomology yields

$$0 \rightarrow \mathbb{S}_{A,M}^\alpha \rightarrow \mathbb{S}_A^{\alpha'} \otimes_R \mathbb{S}_{A,M}^{p_\ell(\alpha)} \rightarrow \mathbb{S}_{A,M}^{\alpha' \odot p_\ell(\alpha)} \rightarrow H^1(\mathcal{H}_{A,M}(\alpha)) \rightarrow 0.$$

The map $\mathbb{S}_A^{\alpha'} \otimes_R \mathbb{S}_{A,M}^{p_\ell(\alpha)} \rightarrow \mathbb{S}_{A,M}^{\alpha' \odot p_\ell(\alpha)}$ is precisely the map of Lemma 4.43(2), which is of course surjective. Thus $H^1(\mathcal{H}_{A,M}(\alpha)) = 0$ and the result follows. \square

Recall that Example 4.31 gives an explicit example of the complex of Theorem 4.48 when just using the Schur modules \mathbb{S}_A^α .

Example 4.49. As a more concrete example of Theorem 4.48, let $A := S(V) \cong k[x_1, \dots, x_n]$ be the symmetric algebra (viewed as a polynomial ring) and $\mathfrak{m} := (x_1, \dots, x_n)$ the homogeneous maximal ideal. Given any composition α and integer $d \geq 1$, notice that by definition there is an equality

$$\mathbb{S}_{A,\mathfrak{m}^d}^\alpha = \mathbb{S}_A^{\alpha \cdot (d)}.$$

Consider the composition $\alpha := (1^4)$ for any integer. Then we can see the difference in the complex of Theorem 4.48 based on how we view the composition α . Viewed as a 3-partitioned composition, the complex of Theorem 4.48 yields:

$$A_1 \otimes_R A_1 \otimes A_1 \otimes_R A_d \rightarrow \begin{array}{c} A_2 \otimes_R A_1 \otimes_R A_d \\ \oplus \\ A_1 \otimes_R A_2 \otimes_R A_d \\ \oplus \\ A_1 \otimes_R A_1 \otimes_R A_{d+1} \end{array} \rightarrow \begin{array}{c} A_3 \otimes_R A_d \\ \oplus \\ A_2 \otimes_R A_{d+1} \\ \oplus \\ A_1 \otimes_R A_{d+2} \end{array} \rightarrow A_{d+3}$$

This is simply the degree $d + 3$ homogeneous strand of the Bar complex on \mathfrak{m}^d . If we instead view α as 2-partitioned via $(1^3) = (1^2) \cdot (1)$, we obtain something distinct from a strand of the Bar complex:

$$\mathbb{S}_A^{(1^2)} \otimes_R A_1 \otimes_R A_d \rightarrow \begin{array}{c} \mathbb{S}_A^{(1^2)} \otimes_R A_{d+1} \\ \oplus \\ \mathbb{S}_A^{(1,2)} \otimes_R A_d \end{array} \rightarrow \mathbb{S}_A^{(1,d+2)}$$

Finally, we can also view (1^3) as 1-partitioned; recalling that $\mathbb{S}_A^{(1^3)} = \bigwedge^3 V$ in this setting, Theorem 4.48 implies that there is a short exact sequence

$$0 \rightarrow \mathbb{S}^{(1^3,d)} \rightarrow \bigwedge^3 V \otimes_R \mathcal{S}^d(V) \rightarrow \mathbb{S}_A^{(1^2,d+1)} \rightarrow 0.$$

Note that $\mathbb{S}_A^{(1^3,d)}$ and $\mathbb{S}_A^{(1^2,d+1)}$ are equal to the classical defined Schur modules $\mathbb{S}^{(d,1^3)}(V)$ and $\mathbb{S}^{(d+1,1^2)}(V)$, respectively, in which case the above short exact sequence recovers the well-known description of hook Schur modules as arising from homogeneous strands of the Koszul complex; see [Buchsbaum and Eisenbud 1975]. Thus Theorem 4.48 yields a family of complexes that interpolates between the full strands of the Bar complex and homogeneous strands of the Koszul complex.

5. Multi-Schur functors

In this section, we define multi-Schur functors. As mentioned in the introduction, the intuition behind these objects is that they are obtained by taking kernels of the defining ribbon Schur module relations diagonally; they will be particularly helpful for describing canonical equivariant decompositions of the derived invariants over Segre subalgebras.

Definition 5.1. Let $\alpha^1, \dots, \alpha^n$ be a sequence of compositions of fixed length ℓ and A^1, \dots, A^n be a sequence of Koszul R -algebras. The *multi-Schur module* $\mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n}$ is defined to be the kernel of the natural map

$$(A^1 \otimes_R \cdots \otimes_R A^n)_{(\alpha^1, \dots, \alpha^n)} \rightarrow \bigoplus_{i=1}^{\ell-1} (A^1 \otimes_R \cdots \otimes_R A^n)_{(\sigma_i(\alpha^1), \dots, \sigma_i(\alpha^n))}.$$

Given a sequence of Koszul left A^i -modules M^i of initial degree t^i for $1 \leq i \leq n$, the *multi-Schur module* $\mathbb{S}_{(A^1, M^1), \dots, (A^n, M^n)}^{\alpha^1, \dots, \alpha^n}$ is defined to be the kernel of the natural map

$$\begin{aligned} & ((A^1 \otimes_R M^1) \otimes_R \cdots \otimes_R (A^n \otimes_R M^n))_{(\alpha^1 \cdot (t^1), \dots, \alpha^n \cdot (t^n))} \\ & \rightarrow \bigoplus_{i=1}^{\ell-1} ((A^1 \otimes_R M^1) \otimes_R \cdots \otimes_R (A^n \otimes_R M^n))_{(\sigma_i(\alpha^1 \cdot (t^1)), \dots, \sigma_i(\alpha^n \cdot (t^n)))}. \end{aligned}$$

The definition for a right A -module is analogous. Assume now that $\ell(\alpha^1) \geq 2$. Given Koszul right (resp. left) A -modules M^1, \dots, M^n (resp. N^1, \dots, N^n) of initial degrees t^i (resp. s^i), the multi-Schur module $\mathbb{S}_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\alpha^1, \dots, \alpha^n}$ is defined to be the kernel of the natural map

$$\begin{aligned} & ((M^1 \otimes_R A^1 \otimes_R N^1) \otimes_R \cdots \otimes_R (M^n \otimes_R A^n \otimes_R N^n))_{((t^1) \cdot \alpha^1 \cdot (s^1), \dots, (t^n) \cdot \alpha^n \cdot (s^n))} \\ & \rightarrow \bigoplus_{i=1}^{\ell-1} ((M^1 \otimes_R A^1 \otimes_R N^1) \otimes_R \cdots \otimes_R (M^n \otimes_R A^n \otimes_R N^n))_{(\sigma_i((t^1) \cdot \alpha^1 \cdot (s^1)), \dots, \sigma_i((t^n) \cdot \alpha^n \cdot (s^n)))}. \end{aligned}$$

In a similar way to the ribbon Schur functors for single inputs, the multi-Schur modules can also be described as modules of the form $L^I_{M_1, \dots, M_n}$ for some collection of submodules, but the translation is a little more subtle.

Construction 5.2. Adopt notation and hypotheses as in Definition 5.1, and recall the notation for the R -submodules $S^j_{A,N,i}$ and $S^j_{M,A,N,i}$ as introduced in Definitions A.18 and 4.37, respectively. For each of the compositions α^i , there are isomorphisms of R -modules

$$(A^i \otimes_R N^i)_{\alpha^i \cdot (s^i)} \cong \frac{A_1^{i \otimes |\alpha^i|} \otimes_R N_{s^i}^i}{\bigvee_{j \in \phi^{-1}(\alpha^i)} S_{A,N,j}^{|\alpha^i|}} \quad \text{and} \quad (M^i \otimes_R A^i \otimes_R N^i)_{(t^i) \cdot \alpha^i \cdot (s^i)} \cong \frac{M_{t^i} \otimes_R A_1^{i \otimes |\alpha^i|} \otimes_R N_{s^i}^i}{\bigvee_{j \in \phi^{-1}(\alpha^i)} S_{M,A,N,j}^{|\alpha^i|}}.$$

By the assumption that each α^i has length ℓ , each of the sets

$$[|\alpha^i| - 1] \setminus \phi^{-1}(\alpha^i)$$

has length $\ell - 1$.² Let $\psi_{\alpha^i} : [|\alpha^i| - 1] \setminus \phi^{-1}(\alpha^i) \rightarrow [\ell - 1]$ denote the unique order-preserving isomorphism between these two sets and consider the induced collections

$$\begin{aligned} S_{A,N,j}^\alpha &:= S_{A^1, N^1, \psi_{\alpha^1}^{-1}(j)}^{|\alpha^1|} + S_{A^2, N^2, \psi_{\alpha^2}^{-1}(j)}^{|\alpha^2|} + \dots + S_{A^n, N^n, \psi_{\alpha^n}^{-1}(j)}^{|\alpha^n|} \\ &\subset (A_1^{1 \otimes |\alpha^1|} \otimes_R N_{s_1}^1) \otimes_R (A_1^{2 \otimes |\alpha^2|} \otimes_R N_{s_2}^2) \otimes_R \dots \otimes_R (A_1^{n \otimes |\alpha^n|} \otimes_R N_{s_n}^n), \\ S_{M,A,N,j}^\alpha &:= S_{M^1, A^1, N^1, \psi_{\alpha^1}^{-1}(j)}^{|\alpha^1|} + S_{M^2, A^2, N^2, \psi_{\alpha^2}^{-1}(j)}^{|\alpha^2|} + \dots + S_{M^n, A^n, N^n, \psi_{\alpha^n}^{-1}(j)}^{|\alpha^n|} \\ &\subset (M_{t_1}^1 \otimes_R A_1^{1 \otimes |\alpha^1|} \otimes_R N_{s_1}^1) \otimes_R (M_{t_2}^2 \otimes_R A_1^{2 \otimes |\alpha^2|} \otimes_R N_{s_2}^2) \otimes_R \dots \otimes_R (M_{t_n}^n \otimes_R A_1^{n \otimes |\alpha^n|} \otimes_R N_{s_n}^n). \end{aligned}$$

In the above, we are abusing notation for sake of clarity: the module $S_{A^i, N^i, j}^{|\alpha^i|}$ is defined as an R -submodule of $A_1^{i \otimes |\alpha^i|} \otimes_R N_{s^i}^i$, but in the above expressions we are viewing each of these submodules as the i -th tensor factor of

$$(A_1^{1 \otimes |\alpha^1|} \otimes_R N_{s_1}^1) \otimes_R (A_1^{2 \otimes |\alpha^2|} \otimes_R N_{s_2}^2) \otimes_R \dots \otimes_R (A_1^{n \otimes |\alpha^n|} \otimes_R N_{s_n}^n).$$

The running flatness assumption implies that inclusion into the i -th tensor factor is actually a well-defined injection. The same abuse of notation is used for $S_{M^i, A^i, N^i, j}^{|\alpha^i|}$.

Observation 5.3. Adopt notation and hypotheses as in Construction 5.2. Then there are isomorphisms of R -modules

$$\begin{aligned} S_{(A^1, N^1), \dots, (A^n, N^n)}^{\alpha^1, \dots, \alpha^n} &\cong \frac{\bigwedge_{j \in [\ell-1]} S_{A,M}^\alpha}{\bigvee_{j \notin \phi^{-1}(\alpha^1)} S_{A^1, N^1, j}^{|\alpha^1|} + \dots + \bigvee_{j \notin \phi^{-1}(\alpha^n)} S_{A^n, N^n, j}^{|\alpha^n|}}, \\ S_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\alpha^1, \dots, \alpha^n} &\cong \frac{\bigwedge_{j \in [\ell-1]} S_{M,A,M}^\alpha}{\bigvee_{j \notin \phi^{-1}(\alpha^1)} S_{M^1, A^1, N^1, j}^{|\alpha^1|} + \dots + \bigvee_{j \notin \phi^{-1}(\alpha^n)} S_{M^n, A^n, N^n, j}^{|\alpha^n|}}. \end{aligned}$$

²Put more informally, quotienting by the module $S_{A,i}^n$ has the effect of deleting the i -th comma in the composition $(1^{|\alpha|})$ and replacing it with addition; if α is ℓ -partitioned, that means there are $\ell - 1$ commas *not* deleted from $(1^{|\alpha|})$.

An interesting aspect of multi-Schur modules is that one can define the multi-Schur modules when the input compositions have different sizes; to make sense of this, the inputs need to be ℓ -partitioned into the same number of parts, instead:

Definition 5.4. Let $\alpha^1, \dots, \alpha^n$ be a sequence of ℓ -partitioned compositions and A^1, \dots, A^n be a sequence of Koszul R -algebras. The *multi-Schur module* $\mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n}$ is defined as the kernel of the natural map

$$\mathbb{S}_{A^1}^{\mu_\emptyset(\alpha^1)} \otimes_R \mathbb{S}_{A^2}^{\mu_\emptyset(\alpha^2)} \otimes_R \cdots \otimes_R \mathbb{S}_{A^n}^{\mu_\emptyset(\alpha^n)} \rightarrow \bigoplus_{i=1}^{\ell-1} \mathbb{S}_{A^1}^{\mu_i(\alpha^1)} \otimes_R \mathbb{S}_{A^2}^{\mu_i(\alpha^2)} \otimes_R \cdots \otimes_R \mathbb{S}_{A^n}^{\mu_i(\alpha^n)}.$$

If $\alpha := \alpha^1 = \alpha^2 = \dots = \alpha^n$ then the more concise notation $\mathbb{S}_{A^1, \dots, A^n}^\alpha$ will be used to denote $\mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n}$.

Likewise, given a Koszul left A -module M , define the multi-Schur module $\mathbb{S}_{(A^1, M^1), \dots, (A^n, M^n)}^{\alpha^1, \dots, \alpha^n}$ as the kernel of the natural map

$$\mathbb{S}_{(A^1, M^1)}^{\mu_\emptyset(\alpha^1)} \otimes_R \mathbb{S}_{(A^2, M^2)}^{\mu_\emptyset(\alpha^2)} \otimes_R \cdots \otimes_R \mathbb{S}_{(A^n, M^n)}^{\mu_\emptyset(\alpha^n)} \rightarrow \bigoplus_{i=1}^{\ell} \mathbb{S}_{(A^1, M^1)}^{\mu_i(\alpha^1)} \otimes_R \mathbb{S}_{(A^2, M^2)}^{\mu_i(\alpha^2)} \otimes_R \cdots \otimes_R \mathbb{S}_{(A^n, M^n)}^{\mu_i(\alpha^n)}.$$

Finally, given a Koszul left (resp. right) A -module N (resp. M), define the multi-Schur module

$$\mathbb{S}_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\alpha^1, \dots, \alpha^n}$$

as the kernel of the natural map

$$\mathbb{S}_{(M^1, A^1, N^1)}^{\mu_\emptyset(\alpha^1)} \otimes_R \cdots \otimes_R \mathbb{S}_{(M^n, A^n, N^n)}^{\mu_\emptyset(\alpha^n)} \rightarrow \bigoplus_{i=1}^{\ell} \mathbb{S}_{(M^1, A^1, N^1)}^{\mu_i(\alpha^1)} \otimes_R \cdots \otimes_R \mathbb{S}_{(M^n, A^n, N^n)}^{\mu_i(\alpha^n)}.$$

The following is the evident analog of Observation 4.41 for multi-Schur modules:

Observation 5.5. Let α^i, β^i , and γ^i be ℓ (resp. j, k)-partitioned partitions. Given a collection of Koszul algebras A^i and left (resp. right) Koszul A^i -modules N_i (resp. M_i) for $1 \leq i \leq n$, there is an equality

$$\mathbb{S}_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\alpha^1 \cdot \beta^1 \cdot \gamma^1, \dots, \alpha^n \cdot \beta^n \cdot \gamma^n} = (\mathbb{S}_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\alpha^1 \cdot \beta^1, \dots, \alpha^n \cdot \beta^n} \otimes_R \mathbb{S}_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\gamma^1, \dots, \gamma^n}) \cap (\mathbb{S}_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\alpha^1, \dots, \alpha^n} \otimes_R \mathbb{S}_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\beta^1 \cdot \gamma^1, \dots, \beta^n \cdot \gamma^n}).$$

The following properties are immediate from the definition of multi-Schur modules:

Proposition 5.6. Let $\alpha^1, \dots, \alpha^n$ be a sequence of ℓ -partitioned compositions and A^1, \dots, A^n be a sequence of Koszul R -algebras. Then:

(1) Let $\tau \in \Sigma_n$ be any permutation of $[n]$. Then there is a natural isomorphism

$$\mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n} \cong \mathbb{S}_{A^{\tau(1)}, \dots, A^{\tau(n)}}^{\alpha^{\tau(1)}, \dots, \alpha^{\tau(n)}}.$$

(2) If $\ell = 1$, there is an isomorphism

$$\mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n} \cong \mathbb{S}_{A^1}^{\alpha^1} \otimes_R \cdots \otimes_R \mathbb{S}_{A^n}^{\alpha^n}.$$

(3) If each A^i admits the structure of a $R[G^i]$ -module for some group G^i (where $1 \leq i \leq n$), then the multi-Schur module $\mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n}$ admits the structure of a $R[G^1 \times \dots \times G^n]$ -module, and all of the above isomorphisms are $G^1 \times \dots \times G^n$ -equivariant.

In view of Observation 5.5, the notation for multi-Schur modules can quickly become overwhelming. For this reason, we introduce the following shorthand notation:

Notation 5.7. Let $\alpha^1, \dots, \alpha^n$ be a collection of ℓ -partitioned compositions. Given a collection of Koszul algebras A^i and left (resp. right) Koszul A^i -modules N_i (resp. M_i) for $1 \leq i \leq n$, use the more concise notation

$$\mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^\alpha := \mathbb{S}_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^{\alpha^1, \dots, \alpha^n}.$$

For given tuples of compositions $\underline{\alpha}$ and $\underline{\beta}$, extend the operations of Definitions 4.4 and 4.1 by applying them coordinatewise to the tuples. With this identification, we have the equalities

$$\underline{\alpha} := \{\alpha^1, \dots, \alpha^n\}, \quad \underline{M} := \{M^1, \dots, M^n\}, \quad \underline{A} := \{A^1, \dots, A^n\}, \quad \underline{N} := \{N^1, \dots, N^n\}.$$

The following lemma is the multi-Schur analog of Lemma 4.43, but the proof is actually a little more subtle due to the added difficulty of allowing compositions of different sizes.

Lemma 5.8. Let $\underline{\alpha}$ and $\underline{\beta}$ be sequences of k -partitioned and $\ell - k$ -partitioned compositions, respectively, for some fixed $1 \leq k \leq \ell$. Let $\underline{A} = \{A^1, \dots, A^n\}$ be a sequence of Koszul R -algebras and \underline{M} a sequence of Koszul left \underline{A} -modules. Then:

- (1) Every multi-Schur module is R -flat.
- (2) There is a canonical short exact sequence

$$0 \rightarrow \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\alpha} \cdot \underline{\beta}} \rightarrow \mathbb{S}_{\underline{A}}^\alpha \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^\beta \rightarrow \mathbb{S}_{\underline{A}, \underline{M}}^{\alpha \odot \beta} \rightarrow 0.$$

If all Koszul algebras/modules are R -projective, then every multi-Schur module $\mathbb{S}_{\underline{A}, \underline{M}}^\alpha$ is also R -projective. The analogous statement for right modules holds as well.

Proof. It suffices to prove (2), since (1) is a consequence of (2) combined with Observation 2.12.

Proof of (2). Let $p := \max\{\ell(p_i(\alpha^j)) \mid 1 \leq i \leq k, 1 \leq j \leq n\}$ and define $q := |\{i \mid \ell(p_i(\alpha^j)) = p\}|$. The proof is by a double induction on the values p and q . Notice that when $p = 1$ and q is arbitrary, recall that $A^1 \otimes_R \dots \otimes_R A^n$ is a Koszul algebra and the tensor product $M^1 \otimes_R \dots \otimes_R M^n$ is a left Koszul module over $A^1 \otimes_R \dots \otimes_R A^n$ by Corollary A.23. By Observation 5.3 combined with Proposition 3.14, the sequence of (2) is exact.

Assume now that $p > 1$ and $q = 1$. This means that there is some element of one of the compositions $\underline{\alpha}$ or $\underline{\beta}$ whose largest part is strictly greater than 1. Reversing the order of α and β , it is of no loss of generality to assume that $\underline{\alpha}$ contains a part of size p . Moreover, for simplicity of notation, let us assume that the first element of $\underline{\alpha}$ has size p (the general case is identical but notationally more cumbersome).

Write $\underline{\alpha} = (\underline{a}) \cdot \underline{\alpha}'$, so that by construction $\underline{\alpha}'$ has all parts of size $< p$. There is a commutative diagram:

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 0 & \longrightarrow & \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\alpha} \cdot \underline{\beta}} & \longrightarrow & \mathbb{S}_{\underline{A}}^{\underline{\alpha}} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\beta}} & \longrightarrow & \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\alpha} \circ \underline{\beta}} \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 0 & \longrightarrow & \mathbb{S}_{\underline{A}}^{(\underline{a})} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\alpha}' \cdot \underline{\beta}} & \longrightarrow & \mathbb{S}_{\underline{A}}^{(\underline{a})} \otimes_R \mathbb{S}_{\underline{A}}^{\underline{\alpha}'} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\beta}} & \longrightarrow & \mathbb{S}_{\underline{A}}^{(\underline{a})} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\alpha}' \circ \underline{\beta}} \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 0 & \longrightarrow & \mathbb{S}_{\underline{A}, \underline{M}}^{(\underline{a}) \circ \underline{\alpha}' \cdot \underline{\beta}} & \longrightarrow & \mathbb{S}_{\underline{A}}^{(\underline{a}) \circ \underline{\alpha}'} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\beta}} & \longrightarrow & \mathbb{S}_{\underline{A}, \underline{M}}^{(\underline{a}) \circ \underline{\alpha}' \circ \underline{\beta}} \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 & & 0 & & 0 & & 0
 \end{array}$$

In the above diagram, notice that the bottom two rows and the last two columns are exact by the inductive hypothesis. We may also assume by induction on p that the map

$$\Theta : \mathbb{S}_{\underline{A}}^{(\underline{a})} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\alpha}' \cdot \underline{\beta}} \rightarrow \mathbb{S}_{\underline{A}, \underline{M}}^{(\underline{a}) \circ \underline{\alpha}' \cdot \underline{\beta}}$$

is a surjection. We claim that with this information the first column of the above diagram is exact. It is evident just by definition that the map

$$\mathbb{S}_{\underline{A}, \underline{M}}^{(\underline{a}) \cdot \underline{\alpha}' \cdot \underline{\beta}} \rightarrow \mathbb{S}_{\underline{A}}^{\underline{\alpha}} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\beta}}$$

is always an injection, so it remains to prove exactness at the middle term $\mathbb{S}_{\underline{A}}^{\underline{\alpha}} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\beta}}$. However, a diagram chase employing exactness of the middle column shows that

$$\text{Ker } \Theta \subset (\mathbb{S}_{\underline{A}}^{\underline{\alpha}} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\beta}}) \cap (\mathbb{S}_{\underline{A}}^{(\underline{a})} \otimes_R \mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\alpha}' \cdot \underline{\beta}}),$$

and this latter intersection is precisely equal to $\mathbb{S}_{\underline{A}, \underline{M}}^{\underline{\alpha} \cdot \underline{\beta}}$ by Observation 5.5. Since the reverse containment evidently holds, exactness of the first column follows, and hence all columns of the above diagram are exact. Employing the long exact sequence of homology, it follows that the first row is exact.

For the inductive step on q , the argument is actually identical. After choosing a similar decomposition of α , there is an identical commutative diagram. By construction, the bottom two rows and rightmost two columns are exact by induction on q . A verbatim argument works for showing that the first column is exact, and hence the long exact sequence of homology yields the statement in general. \square

Corollary 5.9. *Let $\underline{\alpha}$ and $\underline{\beta}$ be a collection of k and $\ell - k$ -partitioned compositions, respectively. Let A^i be a collection of Koszul algebras and N_i (resp. M_i) a collection of left (resp. right) Koszul A^i -modules for $1 \leq i \leq n$. Then:*

- (1) Every multi-Schur module $\mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^\alpha$ is R -flat.
- (2) There is a canonical short exact sequence

$$0 \rightarrow \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\alpha \cdot \beta} \rightarrow \mathbb{S}_{\underline{M}, \underline{A}}^\alpha \otimes_R \mathbb{S}_{\underline{A}, \underline{N}}^\beta \rightarrow \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\alpha \odot \beta} \rightarrow 0.$$

If all the Koszul algebras/modules are R -projective, then every multi-Schur module $\mathbb{S}_{\underline{A}, \underline{M}}^\alpha$ is also R -projective.

Proof. This proof is formally identical to the proof of Corollary 4.45 but with the multi-index notation of Notation 5.7 used instead. □

Finally, we have the generalization of the complexes $\mathcal{H}(\alpha)$ of Definition 4.46 for the multi-Schur setting.

Definition 5.10. Let $\alpha^1, \dots, \alpha^n$ be a sequence of ℓ -partitioned compositions. Given Koszul algebras A^i and left Koszul modules M^i for $1 \leq i \leq n$, define the cochain complex $\mathcal{H}_{(A^1, M^1), \dots, (A^n, M^n)}^\bullet(\alpha^1, \dots, \alpha^n)$ via

$$\mathcal{H}_{(A^1, M^1), \dots, (A^n, M^n)}^i(\alpha^1, \dots, \alpha^n) := \bigoplus_{|I|=i} \mathbb{S}_{(A^1, M^1), \dots, (A^n, M^n)}^{\mu_I(\alpha^1), \dots, \mu_I(\alpha^n)}$$

with differential

$$d_{\mathcal{H}_{(A^1, M^1), \dots, (A^n, M^n)}^\bullet} \Big|_{\mathbb{S}_{A^1, \dots, A^n}^{\mu_I(\alpha^1), \dots, \mu_I(\alpha^n)}} := \sum_{j \notin I} \text{sgn}(j, I) \rho_{I, I \cup j}.$$

Corollary 5.11. Let $\alpha^1, \dots, \alpha^n$ be a sequence of ℓ -partitioned compositions. Given Koszul algebras A^i and left Koszul modules M^i for $1 \leq i \leq n$, the cochain complex $\mathcal{H}_{(A^1, M^1), \dots, (A^n, M^n)}^\bullet(\alpha^1, \dots, \alpha^n)$ is exact in positive cohomological degrees and

$$H^0(\mathcal{H}_{(A^1, M^1), \dots, (A^n, M^n)}^\bullet(\alpha^1, \dots, \alpha^n)) = \mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n}.$$

5.1. Some generalities on filtrations. We now turn our attention to the task of filtering the multi-Schur modules $\mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^\alpha$ for given choices of $\underline{\alpha}$. This section collects a few general results on “splicing” filtrations together; these results are essentially trivial, but it will be useful to refer to them explicitly. First, let us recall the definition of a filtration.

Definition 5.12. Let M be an R -module. A (ascending) *filtration* of M is a chain of R -submodules

$$0 = F_0 \subset F_1 \subset \dots \subset F_{n-1} \subset F_n = M.$$

The successive quotients F_i/F_{i-1} are called the *associated graded pieces*.

The following observation is likely a common first exercise on properties of filtrations, but we state and prove it here for completeness. The intuition here is that filtrations whose graded pieces admit further filtrations may be refined to a single filtration of the larger object with the same graded pieces.

Observation 5.13. Let M be an R -module equipped with a finite filtration

$$F : \quad 0 = F_0 \subset F_1 \subset \cdots \subset F_{n-1} \subset F_n = M.$$

Suppose that for all $i \geq 1$, the i -th graded piece $G_i := F_i/F_{i-1}$ admits a finite filtration

$$F^i : \quad 0 = F_0^i \subset F_1^i \subset \cdots \subset F_{n_i-1}^i \subset F_{n_i}^i = G_i$$

with graded pieces $G_j^i := F_j^i/F_{j-1}^i$ for each $j \geq 1$. Then the filtration F of M may be refined to a filtration F' of M with associated graded pieces G_j^i for $1 \leq i \leq n, 1 \leq j \leq n_i$.

The above observation may be used to understand how to filter tensor products of modules, each equipped separately with their own filtrations; this is the content of the following corollary.

Corollary 5.14. Let M^1, M^2, \dots, M^ℓ be a collection of R -modules equipped with filtrations

$$F^i : \quad 0 = F_0^i \subset F_1^i \subset \cdots \subset F_{n_i-1}^i \subset F_{n_i}^i = M^i,$$

and assume that each associated graded piece $G_j^i := F_j^i/F_{j-1}^i$ is a flat R -module. Then each of the R -modules M^i is flat and there is a filtration of $M^1 \otimes_R M^2 \otimes_R \cdots \otimes_R M^\ell$ with associated graded pieces of the form

$$G_{i_1}^1 \otimes_R G_{i_2}^2 \otimes_R \cdots \otimes_R G_{i_\ell}^\ell,$$

where $1 \leq i_k \leq n_k$ for each $1 \leq k \leq \ell$.

Proof. Proceed by induction on ℓ , with the base case $\ell = 1$ being vacuous. Assume $\ell > 1$ and let F' be a filtration of $M^1 \otimes_R M^2 \otimes_R \cdots \otimes_R M^{\ell-1}$ with associated graded pieces of the form

$$G_{i_1}^1 \otimes_R \cdots \otimes_R G_{i_{\ell-1}}^{\ell-1}.$$

By the flatness assumption, the induced filtration $F' \otimes_R M^\ell$ has associated graded pieces of the form

$$G_{i_1}^1 \otimes_R \cdots \otimes_R G_{i_{\ell-1}}^{\ell-1} \otimes_R M^\ell.$$

Further filtering each graded piece by the filtration F^ℓ of M^ℓ , we may employ Observation 5.13 to deduce the result. □

5.2. A canonical filtration of multi-Schur functors. The following short exact sequence, when combined with some of the other filtration results proved in this subsection, will be the essential ingredient for proving Theorem 5.20. It will be used to place an object we want to understand in the middle of a short exact sequence whose outer terms have a well-understood filtration.

Lemma 5.15. Let $\alpha^1, \dots, \alpha^n$ and β^1, \dots, β^n be sequences of k -partitioned and $\ell - k$ -partitioned compositions, respectively, for some fixed $0 \leq k < \ell$. Let A^1, \dots, A^n be a sequence of Koszul R -algebras and N^i (resp. M^i) be Koszul left (resp. right) A^i -modules for $1 \leq i \leq n$. Then there is a canonical short exact sequence

$$0 \rightarrow \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\alpha^1 \cdot \beta^1, \dots, \alpha^{n-1} \cdot \beta^{n-1}, \alpha^n} \otimes_R \mathbb{S}_{M^n, A^n, N^n}^{\beta^n} \rightarrow \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\alpha^1 \cdot \beta^1, \dots, \alpha^n \cdot \beta^n} \rightarrow \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\alpha^1 \circ \beta^1, \dots, \alpha^{n-1} \circ \beta^{n-1}, \alpha^n \cdot \beta^n} \rightarrow 0.$$

In the above, the sequences of $\alpha^1 \cdot \beta^1, \dots, \alpha^{n-1} \cdot \beta^{n-1}, \alpha^n$ and $\alpha^1 \odot \beta^1, \dots, \alpha^{n-1} \odot \beta^{n-1}, \alpha^n \cdot \beta^n$ are being viewed as k and $\ell - 1$ -partitioned compositions via the convention of Convention 4.14.

Example 5.16. Let A and B be two Koszul R -algebras and $\alpha^1 = \beta^1 = (1^3)$, viewed as a 3-partitioned composition. Then the short exact sequence of Lemma 5.15 takes the form

$$\mathbb{S}_{A,B}^{(1^3),(1^2)} \otimes_R B_1 \rightarrow \mathbb{S}_{A,B}^{(1^3),(1^3)} \rightarrow \mathbb{S}_{A,B}^{(1,2),(1^3)} \rightarrow 0.$$

In the above, (1^3) is being viewed as the 2-partitioned composition $(1) \cdot (1^2)$ in the first term of the sequence, and likewise for the last term in the sequence.

Proof. For simplicity of notation, we will prove the theorem only for Koszul algebra inputs (otherwise the relevant diagrams are too large to display). The proof follows by examining the commutative diagram of Figure 3; the base case and the inductive step are outlined in the caption. \square

The next lemma can be seen as a “first approximation” of the filtration in Theorem 5.20.

Lemma 5.17. Let $\alpha^1, \dots, \alpha^n$ be a sequence of ℓ -partitioned compositions and A^1, \dots, A^n be a sequence of Koszul R -algebras and N^i (resp. M^i) be Koszul left (resp. right) A^i -modules for $1 \leq i \leq n$. Then the multi-Schur module $\mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^\alpha$ admits a canonical filtration with graded pieces of the form

$$\mathbb{S}_{(M^1, A^1, N^1), \dots, (M^{n-1}, A^{n-1}, N^{n-1})}^{\sigma_I(\alpha^1), \dots, \sigma_I(\alpha^{n-1})} \otimes_R \mathbb{S}_{(M^n, A^n, N^n)}^{v_{[\ell] \setminus I}(\alpha^n)},$$

where I ranges over all subsets of $[\ell - 1]$.

Proof. Proceed by induction on ℓ , where the base case $\ell = 1$ is trivial. Assume now that $\ell > 1$ and write each $\alpha^i = \beta^i \cdot \gamma^i$, where $\gamma^i := p_\ell(\alpha^i)$. By Lemma 5.15 there is a canonical short exact sequence

$$0 \rightarrow \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\beta^1 \cdot \gamma^1, \dots, \beta^{n-1} \cdot \gamma^{n-1}, \beta^n} \otimes_R \mathbb{S}_{M^n, A^n, N^n}^{\gamma^n} \rightarrow \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\beta^1 \cdot \gamma^1, \dots, \beta^n \cdot \gamma^n} \rightarrow \mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\beta^1 \odot \gamma^1, \dots, \beta^{n-1} \odot \gamma^{n-1}, \beta^n \cdot \gamma^n} \rightarrow 0.$$

By the inductive hypothesis, the multi-Schur module $\mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\beta^1 \cdot \gamma^1, \dots, \beta^{n-1} \cdot \gamma^{n-1}, \beta^n}$ admits a filtration with graded pieces of the form

$$\mathbb{S}_{(M^1, A_1, N^1), \dots, (M^{n-1}, A^{n-1}, N^{n-1})}^{\sigma_I(\alpha^1), \dots, \sigma_I(\alpha^{n-1})} \otimes_R \mathbb{S}_{(M^n, A^n, N^n)}^{v_{[\ell-1] \setminus I}(\beta^n)},$$

where $I \subset [\ell - 2]$. Likewise, the multi-Schur module $\mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^{\beta^1 \odot \gamma^1, \dots, \beta^{n-1} \odot \gamma^{n-1}, \beta^n \cdot \gamma^n}$ admits a filtration with graded pieces of the form

$$\mathbb{S}_{(M^1, A_1, N^1), \dots, (M^{n-1}, A^{n-1}, N^{n-1})}^{\sigma_J(\beta^1 \odot \gamma^1), \dots, \sigma_J(\beta^{n-1} \odot \gamma^{n-1})} \otimes_R \mathbb{S}_{(M^n, A^n, N^n)}^{v_{[\ell-1] \setminus J}(\beta^n \cdot \gamma^n)},$$

where $J \subset [\ell - 2]$. Note that ranging over all $I \subset [\ell - 2]$ as in the first case is the same as ranging over all $I \subset [\ell - 1]$ with $\ell - 1 \notin I$, and ranging over all J as in the second case is equivalent to ranging over all $J \subset [\ell - 1]$ with $\ell - 1 \in J$. Combining both of these filtrations with Corollary 5.14 yields the result. \square

Lemma 5.18. *Let α be any ℓ -partitioned composition. Let A be a Koszul R -algebra and N (resp. N) any left (resp. right) Koszul A -module. For any $I \subset [\ell - 1]$, the module $\mathbb{S}_{M,A,N}^{\nu_I(\alpha)}$ admits a canonical filtration with associated graded pieces of the form*

$$\mathbb{S}_{M,A,N}^{\sigma_J(\alpha)},$$

where $J \subset [\ell - 1]$ ranges over all subsets with $J \subseteq I$.

Proof. Proceed by induction on $|I|$, with base case $I = \emptyset$ being vacuous since there is an equality $\mathbb{S}_{M,A,N}^{\nu_\emptyset(\alpha)} = \mathbb{S}_{M,A,N}^{\sigma_\emptyset(\alpha)}$. For $|I| > 0$, let $j \in I$ be the largest element of I . Write $\alpha = p_{<j}(\alpha) \cdot p_{\geq j}(\alpha)$ and consider the short exact sequence

$$0 \rightarrow \mathbb{S}_{M,A,N}^{\nu_{I \setminus j}(\alpha)} \rightarrow \mathbb{S}_{M,A,N}^{\nu_I(\alpha)} \rightarrow \mathbb{S}_{M,A,N}^{\nu_{I \setminus j}(p_{<j}(\alpha) \odot p_{\geq j}(\alpha))} \rightarrow 0.$$

Let us consider the filtrations of the outer two terms: by the inductive hypothesis, $\mathbb{S}_{M,A,N}^{\nu_{I \setminus j}(\alpha)}$ has a filtration with graded pieces of the form

$$\mathbb{S}_A^{\sigma_K(\alpha)},$$

where $K \subset [\ell - 1]$ ranges over all subsets with $K \subset I \setminus j$. Likewise, the term $\mathbb{S}_{M,A,N}^{\nu_{I \setminus j}(p_{<j}(\alpha) \odot p_{\geq j}(\alpha))}$ has a filtration with graded pieces of the form

$$\mathbb{S}_{M,A,N}^{\sigma_L(p_{<j}(\alpha) \odot p_{\geq j}(\alpha))},$$

where $L \subset [\ell - 2]$ ranges over all subsets $L \subset s_j(I \setminus j)$. Notice that $p_{<j}(\alpha) \odot p_{\geq j}(\alpha)$ is the same as $\sigma_j(\alpha)$, in which case ranging over all K and L as above is the same as just ranging over all subsets $J \subset [\ell - 1]$ with $J \subset I$. The result thus follows from Corollary 5.14. \square

Example 5.19. Let $A = S(V)$, the symmetric algebra on some free R -module V . Then A is a $GL(V)$ -representation and Lemma 5.18 implies that there is a $GL(V)$ -equivariant filtration of the tensor power $V^{\otimes d}$ with graded pieces of the form \mathbb{S}_A^α , where α ranges over all compositions of d (since $V^{\otimes d} = \mathbb{S}_{S(V)}^{\nu_{[d-1]}(1^d)}$).

Assuming R is a field of characteristic 0, the ring $R[GL(V)]$ is semisimple and hence this yields a $GL(V)$ -equivariant direct sum decompositions

$$V^{\otimes d} = \bigoplus_{|\alpha|=d} \mathbb{S}_A^\alpha.$$

Finally, we arrive at the main result of this section; this result furnishes the multi-Schur functors with a canonical filtration whose associated graded pieces are easily described as tensor products of the ribbon Schur functors of Section 4.

Theorem 5.20. *Let $\alpha^1, \dots, \alpha^n$ be a sequence of ℓ -partitioned compositions and A^1, \dots, A^n be a sequence of Koszul R -algebras and N^i (resp. M^i) be Koszul left (resp. right) A^i -modules for $1 \leq i \leq n$. Then the multi-Schur module $\mathbb{S}_{\underline{M}, \underline{A}, \underline{N}}^\alpha$ admits a canonical filtration with graded pieces of the form*

$$\mathbb{S}_{(M^1, A^1, N^1)}^{\sigma_{I_1}(\alpha^1)} \otimes_R \mathbb{S}_{(M^2, A^2, N^2)}^{\sigma_{I_2}(\alpha^2)} \otimes_R \cdots \otimes_R \mathbb{S}_{(M^n, A^n, N^n)}^{\sigma_{I_n}(\alpha^n)},$$

where the subsets $I_1, \dots, I_n \subset [\ell - 1]$ range over all choices such that $I_1 \cap I_2 \cap \cdots \cap I_n = \emptyset$.

Remark 5.21. Notice that the parametrizing set for the associated graded pieces is inherently symmetric under permuting the tensor factors, which is expected by the invariance of the multi-Schur module on the ordering of the inputs.

In general, the number of ways to choose n subsets of $[\ell - 1]$ with no common intersection is $(2^n - 1)^{\ell - 1}$, so the number of associated graded pieces grows exponentially with respect to both n and ℓ .

Remark 5.22. The adjective “canonical” used in Theorem 5.20 means that this filtration is really true at the level of functors; in other words, the functor \mathbb{S}^α which takes as inputs n -tuples of Koszul R -algebras and outputs the associated multi-Schur module admits a canonical filtration whose associated graded pieces are given by the functors of the form

$$\mathbb{S}^{\sigma_{I_1}(\alpha^1)} \otimes_R \mathbb{S}^{\sigma_{I_2}(\alpha^2)} \otimes_R \cdots \otimes_R \mathbb{S}^{\sigma_{I_n}(\alpha^n)}.$$

Proof of Theorem 5.20. Proceed by induction on n (the number of Koszul algebras), with base case $n = 2$. To prove the base case, we induct on ℓ (where the case $\ell = 1$ is vacuous). Let $\ell > 1$ and write $\alpha = \alpha' \cdot p_\ell(\alpha)$ and $\beta = \beta' \cdot p_\ell(\beta)$. By Lemma 5.15 there is a short exact sequence

$$0 \rightarrow \mathbb{S}_{A,B}^{\alpha,\beta'} \otimes_R \mathbb{S}_B^{p_\ell(\beta)} \rightarrow \mathbb{S}_{A,B}^{\alpha,\beta} \rightarrow \mathbb{S}_{A,B}^{\alpha' \odot p_\ell(\alpha),\beta} \rightarrow 0.$$

By the inductive hypothesis, the outer 2 terms admit filtrations with graded pieces of the correct form. This establishes the base case.

Assume now that $n > 2$. By Lemma 5.17 the multi-Schur module $\mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n}$ admits a canonical filtration with associated graded pieces of the form

$$\mathbb{S}_{A^1, \dots, A^{n-1}}^{\sigma_I(\alpha^1), \dots, \sigma_I(\alpha^{n-1})} \otimes_R \mathbb{S}_{A^n}^{v_{[\ell] \setminus I}(\alpha^n)},$$

where I ranges over all subsets of $[\ell - 1]$. By the inductive hypothesis, each of the modules $\mathbb{S}_{A^1, \dots, A^{n-1}}^{\sigma_I(\alpha^1), \dots, \sigma_I(\alpha^{n-1})}$ admits a canonical filtration with associated graded pieces of the form

$$\mathbb{S}_{A^1}^{\sigma_{I_1}(\alpha^1)} \otimes_R \mathbb{S}_{A^2}^{\sigma_{I_2}(\alpha^2)} \otimes_R \cdots \otimes_R \mathbb{S}_{A^{n-1}}^{\sigma_{I_{n-1}}(\alpha^{n-1})},$$

where the above ranges over all choices of I_1, \dots, I_{n-1} with $I_1 \cap \cdots \cap I_{n-1} = I$. On the other hand, the module $\mathbb{S}_{A^n}^{v_{[\ell] \setminus I}(\alpha^n)}$ admits a canonical filtration with associated graded pieces of the form $\mathbb{S}_{A^n}^{\sigma_{I_n}(\alpha^n)}$, where I_n ranges over all subsets $I_n \subset [\ell - 1]$ such that $I_n \cap I = \emptyset$. This is overall the same thing as ranging over all choices $I_1, \dots, I_n \subset [\ell - 1]$ such that $I_1 \cap \cdots \cap I_n = \emptyset$. \square

Example 5.23. Let A and B be Koszul algebras. Let us use the argument of Theorem 5.20 to filter the multi-Schur module $\mathbb{S}_{A,B}^{(1^3), (1^3)}$, helping to illustrate the idea of the proof in a concrete setting. Recall by Example 5.16 that we have the short exact sequence

$$\mathbb{S}_{A,B}^{(1^3), (1^2)} \otimes_R B_1 \rightarrow \mathbb{S}_{A,B}^{(1^3), (1^3)} \rightarrow \mathbb{S}_{A,B}^{(1,2), (1^3)} \rightarrow 0.$$

Iteratively applying this sequence to the outer two terms yields

$$\begin{aligned}
 0 \rightarrow A_3 \otimes_R (B_1)^{\otimes 3} &\rightarrow \mathbb{S}_{A,B}^{(1^3),(1^2)} \rightarrow \mathbb{S}_A^{(2,1)} \otimes_R \mathbb{S}_B^{(1^2)} \otimes_R B_1 \rightarrow 0, \\
 0 \rightarrow \mathbb{S}_A^{(1,2)} \otimes_R B_1 \otimes_R \mathbb{S}_A^{(1^2)} &\rightarrow \mathbb{S}_{A,B}^{(1,2),(1^3)} \rightarrow A_3 \otimes_R \mathbb{S}_B^{(1^3)} \rightarrow 0.
 \end{aligned}$$

The terms involving B appearing on the ends of these short exact sequences are precisely the filtration factors of Lemma 5.17, and are filtered further by Lemma 5.18, yielding the filtration factors of Theorem 5.20.

Remark 5.24. Let $2^{[\ell-1]}$ denote the Boolean poset on n elements. Then the product $(2^{[\ell-1]})^{\times n}$ is naturally a ranked poset (in fact, isomorphic to the Boolean poset $2^{[\ell-1] \times [n]}$), and the set \mathcal{S} of all tuples (I_1, \dots, I_n) with $I_1 \cap \dots \cap I_n = \emptyset$ is a ranked subposet. Choosing any total order $<$ refining the partial order on this subposet, the filtration of the multi-Schur module $\mathbb{S}_{A^1, \dots, A^n}^{\alpha^1, \dots, \alpha^n}$ is parametrized by $<$. In other words, the filtration is of the form

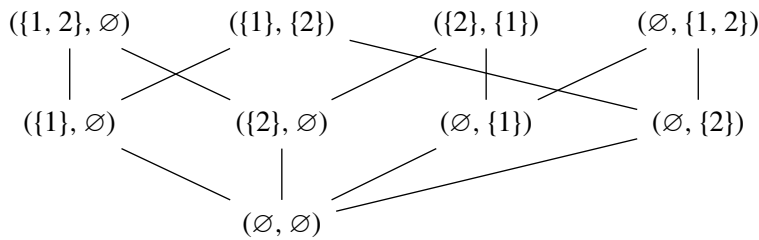
$$\{F_{(I_1, \dots, I_n)}\}_{(I_1, \dots, I_n) \in \mathcal{S}},$$

with

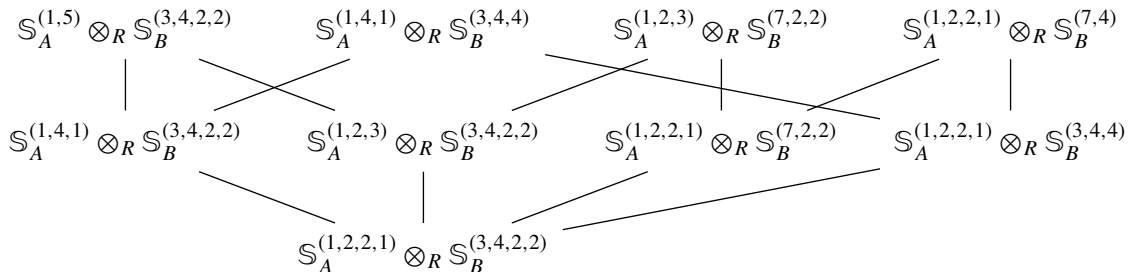
$$F_{(I_1, \dots, I_n)} / F_{\text{pred}(I_1, \dots, I_n)} \cong \mathbb{S}_{A^1}^{\sigma_{I_1}(\alpha^1)} \otimes_R \dots \otimes_R \mathbb{S}_{A^n}^{\sigma_{I_n}(\alpha^n)}.$$

In the above, pred denotes the predecessor function (i.e., the largest element strictly smaller with respect to $<$).

Example 5.25. Let $\alpha = (1, 2) \cdot (2) \cdot (1)$ and $\beta = (3) \cdot (4, 2) \cdot (2)$ be two 3-partitioned compositions and let us compute the composition factors of the multi-Schur module $\mathbb{S}_{A,B}^{\alpha,\beta}$ for any two Koszul algebras A and B . The subposet (in fact, meet semilattice) of $(2^{[2]})^{\times 2}$ that parametrizes the filtration factors has Hasse diagram:



This translates to filtration factors of the following form:



If for instance $A = B = S(V)$, the symmetric algebra, then the above filtration is also $GL(V) \times GL(V)$ equivariant (and in characteristic 0 yields a direct sum decomposition).

6. Applications

The following section is the reward for enduring the technical details of Sections 4 and 5; we are able to arrive at the other end with a rather robust theory that allows us to give elegant and simple closed form descriptions of higher derived invariants associated to (Veronese/Segre subalgebras of) Koszul algebras; see also [Bărcănescu and Manolache 1981]. In Section 6.4 we show how to use this theory to build a large class of Koszul modules over an arbitrary Koszul algebra A , and specialize to the case of a polynomial ring to prove a uniform, characteristic-free regularity result for certain classes of vector bundles on projective space.

6.1. Tor and Ext. In this section we prove the following theorem, which gives some concise descriptions of Tor and Ext between pairs of Koszul modules in terms of ribbon Schur functors.

Theorem 6.1. *Let A be a Koszul algebra and M (resp. N) a Koszul right (resp. left) A -module of initial degree t (resp. s). Then there is a canonical isomorphism of A -modules*

$$\text{Tor}_i^A(M, N) \cong \mathbb{S}_{M,A,N}^{(1^i)} \quad \text{for all } i > 0.$$

If M is instead a Koszul left A -module, then there is a canonical isomorphism

$$\text{Ext}_A^i(M, N) \cong \mathbb{S}_{M^!, A^!, (N^*)^!}^{(i)} \quad \text{for all } i > 0.$$

In particular, both of the above modules are flat R -modules annihilated by A_+ .

Proof. By the definition of Tor combined with Theorem A.25, the module $\text{Tor}_i^A(M, N)$ may be computed by looking at the homology of the complex

$$\dots \rightarrow M \otimes_R \mathbb{S}_{A,N}^{(1^{i+1})} \rightarrow M \otimes_R \mathbb{S}_{A,N}^{(1^i)} \rightarrow M \otimes_R \mathbb{S}_{A,N}^{(1^{i-1})} \rightarrow \dots$$

Splitting this complex into graded pieces, there is a commutative diagram:

$$\begin{array}{ccc}
 M_j \otimes_R \mathbb{S}_{A,N}^{(1^i)} & \xrightarrow{(d_i)_{j+i+s}} & M_{j+1} \otimes_R \mathbb{S}_{A,N}^{(1^{i-1})} \\
 \searrow & & \nearrow \\
 & M_j \otimes_R A_1 \otimes_R \mathbb{S}_{A,N}^{(1^{i-1})} & \\
 \nearrow & & \\
 \mathbb{S}_{M_{\geq j}, A}^{(1)} \otimes_R \mathbb{S}_{A,N}^{(1^{i-1})} & &
 \end{array}$$

This implies that there is an equality

$$\text{Ker}(d_i)_{i+j+s} = (M_j \otimes_R \mathbb{S}_{A,N}^{(1^i)}) \cap (\mathbb{S}_{M_{\geq j}, A}^{(1)} \otimes_R \mathbb{S}_{A,N}^{(1^{i-1})}).$$

By Observation 4.41, this latter intersection is precisely $\mathbb{S}_{(M_{\geq j}, A, N)}^{(1^i)}$. On the other hand, by definition of the Priddy differential there is also a commutative diagram for all $j > s$:

$$\begin{array}{ccc} M_{j-1} \otimes_R \mathbb{S}_{A, N}^{(1^{i+1})} & \xrightarrow{(d_{i+1})_{j+i+s}} & M_j \otimes_R \mathbb{S}_{A, N}^{(1^i)} \\ \downarrow & & \downarrow \\ \mathbb{S}_{M_{\geq j}, A, N}^{(1^i)} & \hookrightarrow & M_j \otimes_R A_1 \otimes_R \mathbb{S}_{A, N}^{(1^{i-1})} \end{array}$$

This implies that there is also an equality

$$\text{im}(d_{i+1})_{j+i+s} = \mathbb{S}_{M_{\geq j}, A, N}^{(1^i)}.$$

Putting both of the above equalities together, it follows that

$$\text{Tor}_i^A(M, N)_{i+j+s} = \begin{cases} \mathbb{S}_{M, A, N}^{(1^i)} & \text{if } j = s, \\ 0 & \text{otherwise.} \end{cases}$$

To prove the isomorphism for Ext, recall first that there is a canonical isomorphism

$$\text{Ext}_A^i(M, N) \cong (\text{Tor}_i^A(N^*, M))^*.$$

By Observation A.22, the graded dual N^* is a Koszul right A -module, and by the isomorphism just proved for Tor there is an isomorphism

$$\text{Tor}_i^A(N^*, M) \cong \mathbb{S}_{N^*, A, M}^{(1^i)}.$$

Dualizing and using the isomorphism of Corollary 4.45(3), the result follows immediately. □

Example 6.2. Assume that A is any Koszul algebra and recall that

$$\mathbb{S}_{A, A_+^d}^{(1^i)} = \mathbb{S}_A^{(1^i, d)}.$$

By Theorem A.25, the minimal free resolution of A_+^d thus has the form

$$\cdots \rightarrow A \otimes_R \mathbb{S}_A^{(1^i, d)} \rightarrow A \otimes_R \mathbb{S}_A^{(1^{i-1}, d)} \rightarrow \cdots \rightarrow A \otimes_R A_d \rightarrow A_+^d.$$

The ribbon Schur module $\mathbb{S}_A^{(1^i, d)}$ may be presented as the cokernel of the composition

$$(A^1)_{i+2}^* \otimes_R A_{d-2} \rightarrow (A^1)_{i+1}^* \otimes_R A_1 \otimes_R A_{d-2} \rightarrow (A^1)_{i+1}^* \otimes_R A_{d-1},$$

in which case we see that there is an isomorphism with the ribbon Schur functor

$$\mathbb{S}_A^{(1^i, d)} \cong L_{i-1, d-1}^A,$$

where the module $L_{i+1, d-1}^A$ is defined as in [Faber et al. 2021]. Thus Theorem 6.1 at least recovers the minimal free resolution of powers of the maximal ideal of a Koszul algebra constructed in [loc. cit.].

Corollary 6.3. *Let A be any Koszul algebra such that $\mathbb{S}_A^{(1^i)} = 0$ for all $i > 1$. Then every Koszul module over A is a flat R -module. If A is R -projective, then every Koszul module over A is R -projective.*

6.2. Veronese subalgebras. Recall the definition of the Veronese subalgebra as in Definition 2.4. The following observation shows that the operation $(-)^{(d)}$ on compositions as defined in Definition 4.1 interacts functorially with the formation of the ribbon Schur functor:

Observation 6.4. Let A be any Koszul algebra and M (resp. N) any Koszul right (resp. left) A -module. For any integer $d > 0$ and integers $i, j \in \mathbb{Z}$ there are isomorphisms

$$\mathbb{S}_{A^{(d)}}^\alpha = \mathbb{S}_A^{\alpha^{(d)}} \quad \text{and} \quad \mathbb{S}_{M^{(d)}, A^{(d)}, N^{(d)}}^\alpha \cong \mathbb{S}_{M, A, N}^{\alpha^{(d)}}.$$

Combining Theorem 6.1 with Observation 6.4 immediately yields:

Corollary 6.5. Let A be any Koszul algebra and M (resp. N) any Koszul right (resp. left) A -module. For any integer $d > 0$ there is an isomorphism

$$\text{Tor}_i^{A^{(d)}}(M^{(d)}, N^{(d)}) = \mathbb{S}_{M, A, N}^{(d^i)},$$

where t (resp. s) is the initial degree of M (resp. N). In particular, there are canonical isomorphisms

$$\text{Tor}_i^{A^{(d)}}(A^{\geq r, d}, A^{\geq r', d}) \cong \mathbb{S}_A^{(r, d^i, r')}.$$

Remark 6.6. The isomorphism

$$\text{Tor}_i^{A^{(d)}}(A^{\geq r, d}, A^{\geq r', d}) \cong \mathbb{S}_A^{(r, d^i, r')}.$$

was originally proved in the case $A = S(V)$ (the symmetric algebra) in the work [Almoussa et al. 2024]. However, one notices that the original proof of this isomorphism does not invoke anything more than the Koszulness properties of the symmetric algebra (and its truncations), which leads to the generalization presented in Corollary 6.5.

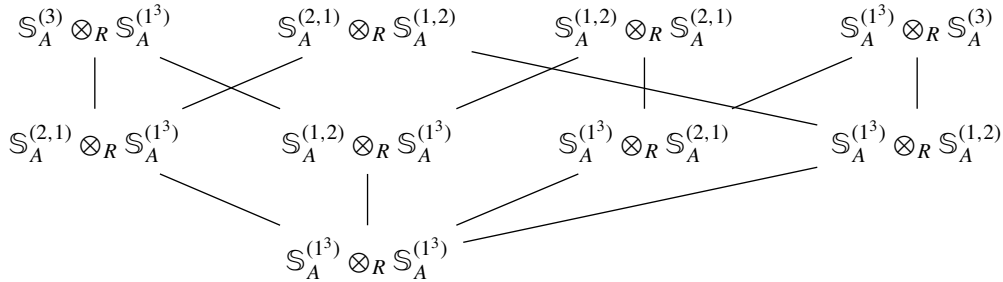
6.3. Segre subalgebras. In this subsection, we apply the construction of multi-Schur modules and their filtrations to study Segre products of Koszul algebras. The following observation is a straightforward translation of the quadratic dual for a Koszul R -algebra:

Observation 6.7. Let A^1, \dots, A^n be a collection of Koszul algebras and M^i a Koszul left A^i -module for $1 \leq i \leq n$. Then there are isomorphisms of R -modules

$$((A^1 \circ \dots \circ A^n)^!)^*_i = \mathbb{S}_{A^1, \dots, A^n}^{(1^i)} \quad \text{and} \quad ((M^1 \circ \dots \circ M^n)^!)^*_i = \mathbb{S}_{(A^1, M^1), \dots, (A^n, M^n)}^{(1^i)}.$$

Example 6.8. Let A be any Koszul algebra and consider the filtration of Theorem 5.20 applied to computing $(A^{[2]})^*_3 = \mathbb{S}_{A, A}^{(1^3)}$. The poset parametrizing the filtration terms are the same as those of

Example 5.25, which yields the filtration factors:



Suppose now that $A = S(V)$, the symmetric algebra on a vector space V . If R is a field of characteristic 0, this induces a $GL(V) \times GL(V)$ -equivariant direct sum decomposition of $(A^{[2] \ 1})_3^*$ into irreducibles

$$(A^{[2] \ 1})_3^* = \left(\bigwedge^3 V \otimes \bigwedge^3 V \right) \oplus \left(\bigwedge^3 V \otimes_R S^{(2,1)}(V) \right)^{\oplus 4} \oplus \left(\bigwedge^3 V \otimes S_3(V) \right)^{\oplus 2} \oplus (S^{(2,1)}(V) \otimes S^{(2,1)}(V))^{\oplus 2}.$$

Observation 6.9. Let $\alpha \in C(d)$ be any composition of some integer $d > 0$ and A^1, \dots, A^n a sequence of Koszul R -algebras admitting a compatible G^i -action for each $1 \leq i \leq n$. Assume that M^i (resp. N^i) is a Koszul right (resp. left) A^i -module admitting a compatible G^i -action. Then the multi-Schur module

$$S_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^\alpha$$

admits a $G^1 \times \dots \times G^n$ -equivariant filtration with associated graded pieces of the form

$$S_{(M^1, A^1, N^1)}^{\alpha^1} \otimes_R \dots \otimes_R S_{(M^n, A^n, N^n)}^{\alpha^n},$$

where $\alpha^1, \dots, \alpha^n$ range over all compositions of $|\alpha|$ satisfying $\alpha^1 \wedge \dots \wedge \alpha^n = \alpha$.

In particular, if each of the group rings $R[G^i]$ is semisimple, then there is a $G^1 \times \dots \times G^n$ -equivariant decomposition

$$S_{(M^1, A^1, N^1), \dots, (M^n, A^n, N^n)}^\alpha \cong \bigoplus_{\substack{(\alpha^1, \dots, \alpha^n) \in C(d)^{\times n} \\ \alpha^1 \wedge \dots \wedge \alpha^n = \alpha}} S_{(M^1, A^1, N^1)}^{\alpha^1} \otimes_R \dots \otimes_R S_{(M^n, A^n, N^n)}^{\alpha^n}.$$

Proof. This is just a retranslation of Theorem 5.20 combined with the fact that the refinement poset on α is isomorphic to the Boolean poset on $[\ell(\alpha) - 1]$, and the meet operation corresponds to intersection under this isomorphism. □

Remark 6.10. The appearance of compositions ranging over the refinement poset as filtration factors is likely related to the connection between Segre products and internal cohomomorphisms as discovered by Manin [1987; 1991]. Indeed, another perspective on the filtration given in Theorem 5.20 is as a canonical filtration of the graded pieces of the cohomomorphism algebra.

Next, we use Theorem 5.20 to prove a symmetric function identity; we first need to recall some notation related to Schur polynomials and establish some multi-index conventions.

Notation 6.11. Let $n \geq 1$ be any integer and consider sets of indeterminates $\mathbf{x}^1, \dots, \mathbf{x}^n$. Recall that the Schur polynomial associated to a skew shape λ/μ is the polynomial

$$s_{\lambda/\mu}(\mathbf{x}) := \sum_{T \in \text{SST}(\lambda/\mu)} \mathbf{x}_T,$$

where \mathbf{x}_T denotes the multigraded character of the semistandard tableau T . If α is a composition, the notation $s_\alpha(\mathbf{x})$ denotes the Schur polynomial corresponding to the ribbon shape determined by α . The complete symmetric polynomial $h_d(\mathbf{x})$ is defined to be $s_{(d)}(\mathbf{x})$.

Given a tuple of compositions $\underline{\alpha} = (\alpha^1, \dots, \alpha^n)$, use the notation

$$s_{\underline{\alpha}}(\underline{\mathbf{x}}) := s_{\alpha^1}(\mathbf{x}^1) \cdot s_{\alpha^2}(\mathbf{x}^2) \cdots s_{\alpha^n}(\mathbf{x}^n).$$

In particular, for a single composition $\alpha = (\alpha_1, \dots, \alpha_n)$ there is the equality

$$h_\alpha(\underline{\mathbf{x}}) = h_{\alpha_1}(\mathbf{x}^1) \cdot h_{\alpha_2}(\mathbf{x}^2) \cdots h_{\alpha_n}(\mathbf{x}^n).$$

Corollary 6.12. With notation as in Notation 6.11, there is an equality of symmetric polynomials

$$h_{d^n+\alpha}(\underline{\mathbf{x}}) = \sum_{i=1}^d \sum_{\substack{\beta^1, \dots, \beta^n \\ |\beta^1| = \dots = |\beta^n| = i \\ \beta^1 \wedge \dots \wedge \beta^n = 1^i}} (-1)^{i+1} h_{(d-i)^n}(\underline{\mathbf{x}}) \cdot s_{\underline{\beta \cdot \alpha}}(\underline{\mathbf{x}}).$$

Remark 6.13. In the statement of Corollary 6.12, if $\underline{\beta} = (\beta^1, \dots, \beta^n)$ is a tuple of compositions and $\alpha = (\alpha_1, \dots, \alpha_n)$ is a composition, then we use the convention

$$\underline{\beta} \cdot \alpha := (\beta^1 \cdot (\alpha_1), \beta^2 \cdot (\alpha_2), \dots, \beta^n \cdot (\alpha_n)).$$

Proof. Assume $R = k$ is a field of characteristic 0 and let $A = S(V)$ for some vector space V . Note that A and $A_{\geq d} = A_+^d$ are polynomial functors for all $d \geq 1$. Taking Segre products, this means that $A^{[n]}$ and $A_+^\alpha := A_+^{\alpha_1} \circ A_+^{\alpha_2} \circ \dots \circ A_+^{\alpha_n}$ are also polynomial functors, and hence there is an equality

$$\text{Ch}(A_+^\alpha) = \text{Ch}(A^{[n]}) \cdot \sum_{i \geq 0} (-1)^i \text{Ch}(\text{Tor}_i^{A^{[n]}}(A_+^\alpha, k)), \tag{6.13.1}$$

where $\text{Ch}(-)$ denotes the multigraded character. By definition there are equalities

$$\text{Ch}(A_+^\alpha) = \sum_{d \geq 1} h_{d^n+\alpha}(\underline{\mathbf{x}}), \quad \text{Ch}(A^{[n]}) = \sum_{d \geq 0} h_{d^n}(\underline{\mathbf{x}}).$$

On the other hand, by Theorem 5.20 there is an equality

$$\text{Ch}(\text{Tor}_i^{A^{[n]}}(A_+^\alpha, k)) = \sum_{\substack{\beta^1, \dots, \beta^n \\ |\beta^1| = \dots = |\beta^n| = i \\ \beta^1 \wedge \dots \wedge \beta^n = 1^i}} s_{\underline{\beta \cdot \alpha}}(\underline{\mathbf{x}}).$$

Combining all of these expressions and comparing degrees on each side of the equality (6.13.1) yields the result. □

Remark 6.14. When $n = 1$, this reduces to the well-known classical character identity

$$h_{d+a}(\mathbf{x}) = \sum_{i=1}^d (-1)^{i+1} h_{d-i}(\mathbf{x}) s_{(1^i, a)}(\mathbf{x}).$$

Moreover, taking s -th Veronese powers of the polynomial ring and performing an identical argument yields the same identity, but every composition is replaced with its s -th rescaling (i.e., apply the operation $(-)^{(s)}$ to all compositions).

Example 6.15. Let $n = 2, d = 2$, and $\alpha = (\alpha_1, \alpha_2)$ be any composition. Then the identity of Corollary 6.12 reads

$$\begin{aligned} h_{2+\alpha_1}(\mathbf{x}^1) h_{2+\alpha_2}(\mathbf{x}^2) &= h_{d-1}(\mathbf{x}^1) h_{d-1}(\mathbf{x}^2) s_{(1, \alpha_1)}(\mathbf{x}^1) s_{(1, \alpha_2)}(\mathbf{x}^2) \\ &\quad - s_{(1^2, \alpha_1)}(\mathbf{x}^1) s_{(1^2, \alpha_2)}(\mathbf{x}^2) - s_{(1^2, \alpha_1)}(\mathbf{x}^1) s_{(2, \alpha_2)}(\mathbf{x}^2) - s_{(2, \alpha_1)}(\mathbf{x}^1) s_{(1^2, \alpha_2)}(\mathbf{x}^2). \end{aligned}$$

Our next corollary is an evident consequence of Theorem 6.1 combined with Observation 6.9 (by setting $\alpha = (1^i)$).

Corollary 6.16. Let A^1, \dots, A^n a sequence of Koszul R -algebras admitting a compatible G^i -action for each $1 \leq i \leq n$. Assume that M^i (resp. N^i) is a Koszul right (resp. left) A^i -module admitting a compatible G^i -action. Then the module

$$\text{Tor}_i^{A^1 \circ \dots \circ A^n} (M^1 \circ \dots \circ M^n, N^1 \circ \dots \circ N^n)$$

admits a $G^1 \times \dots \times G^n$ -equivariant filtration with associated graded pieces of the form

$$\mathbb{S}_{M^1, A^1, N^1}^{\alpha^1} \otimes_R \dots \otimes_R \mathbb{S}_{M^n, A^n, N^n}^{\alpha^n},$$

where the compositions range over all tuples $(\alpha^1, \dots, \alpha^n) \in C(i)^{\times n}$ with $\alpha^1 \wedge \dots \wedge \alpha^n = (1^i)$.

Likewise, the module

$$\text{Tor}_i^{(A^1 \circ \dots \circ A^n)^{(d)}} ((M^1 \circ \dots \circ M^n)^{(d)}, (N^1 \circ \dots \circ N^n)^{(d)})$$

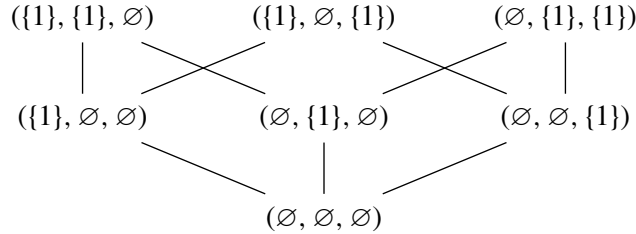
admits a $G^1 \times \dots \times G^n$ -equivariant filtration with associated graded pieces of the form

$$\mathbb{S}_{M^1, A^1, N^1}^{\alpha^1} \otimes_R \dots \otimes_R \mathbb{S}_{M^n, A^n, N^n}^{\alpha^n},$$

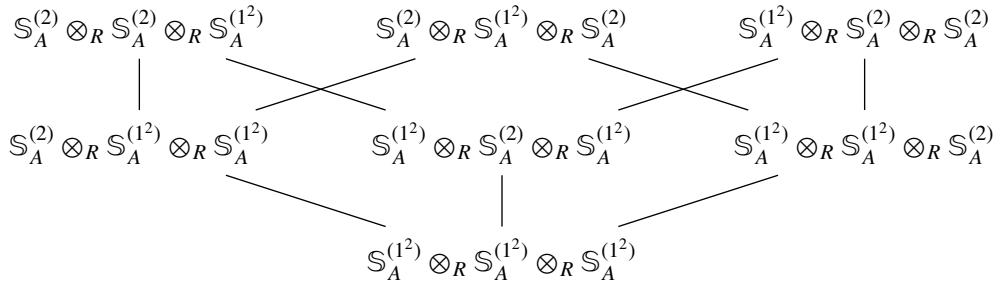
where the compositions range over all tuples $(\alpha^1, \dots, \alpha^n) \in C(di)^{\times n}$ with $\alpha^1 \wedge \dots \wedge \alpha^n = (d^i)$.

Example 6.17. Let A be any Koszul algebra and let us compute the filtration factors of $(A^{[3]})_2^* = \mathbb{S}_{A^{[3]}}^{(1^2)}$ (unfortunately, trying this for $\mathbb{S}_{A^{[3]}}^{(1^3)}$ yields 49 filtration factors, which is too big of an example). The poset

parametrizing the filtration factors is:



This yields filtration factors:



Example 6.18. Let A be any Koszul algebra and let $T(V)$ denote the tensor algebra on some projective R -module V . Then by definition the R -module $S_{A,T(V)}^{(1^i)}$ admits a filtration with graded pieces of the form

$$S_A^{\alpha^1} \otimes_R S_{T(V)}^{\alpha^2},$$

with $\alpha^1 \wedge \alpha^2 = (1^i)$. Note that the only choice of α^2 for which $S_{T(V)}^{\alpha^2}$ is nonzero is for $\alpha^2 = (i)$. Thus $\alpha^1 = (1^i)$ and we find

$$S_{A,T(V)}^{(1^i)} = S_A^{(1^i)} \otimes_R S_{T(V)}^{(i)} = (A^1)_i^* \otimes_R V^{\otimes i}.$$

In particular, after dualizing and collecting graded pieces, there is an isomorphism of R -algebras

$$(A \circ T(V))^! = A^! \circ T(V^*).$$

Of course, this could be verified using more direct methods, but the point is to demonstrate the utility of Theorem 5.20.

Example 6.19. Let $A = S(V)$, $B = S(W)$ be symmetric algebras on free R -modules V and W both of rank 2. Let $M = A'_+ \circ B$ and $N = A \circ B'_+$, viewed as modules over the Segre product $A \circ B$. By Corollary 6.16 there is an equality

$$\mathrm{Tor}_i^{A \circ B}(M, N) = S_{(A,A'_+) \circ (B,B'_+)}^{(1^i)}.$$

The module $S_{(A,A'_+) \circ (B,B'_+)}^{(1^i)}$ has a $\mathrm{GL}(V) \times \mathrm{GL}(W)$ -equivariant filtration with associated graded pieces of the form

$$S_{A,A'_+}^\alpha \otimes_R S_{B,B'_+}^\beta,$$

where α and β range over all partitions with $\alpha \wedge \beta = (1^i)$. Notice that since V and W have rank 2, we are really only ranging over all partitions α and β such that all parts of α^t and β^t are at most 2, and such that $\alpha^t \vee \beta^t = (i)$. Retranslating this in terms of subsets of the Boolean poset, this is asking for all ways to partition the set $[i - 1]$ into the union of two totally disconnected sets $I_\alpha \cup I_\beta$, where $1 \in I_\beta$ and $i - 1 \in I_\alpha$. One quickly sees that there is no such decomposition if $i - 1$ is odd and only 1 such decomposition when $i - 1$ is even:

$$[i - 1] = \{2, 4, \dots, i - 1\} \cup \{1, \dots, i - 2\}.$$

Retranslating this in terms of compositions, we find there is a $\text{GL}(V) \times \text{GL}(W)$ -equivariant isomorphism

$$\text{Tor}_i^{A \circ B}(M, N) = \begin{cases} S_{r-1}(V) \otimes_R \det(V)^{(i+1)/2} \otimes_R S_{r'-1}(W) \otimes_R \det(W)^{(i+1)/2} & \text{if } i > 0 \text{ is odd,} \\ 0 & \text{if } i > 0 \text{ is even.} \end{cases}$$

6.4. Koszul modules built from ribbons and general skew shapes. In this section we apply the Koszulness criterion of Lemma 4.43 to deduce that a large class of modules parametrized by ribbons are Koszul modules. This immediately yields interesting Koszul modules over any Koszul algebra (generalizing powers of A_+), and in the case of the symmetric algebra we are able to give a quick and much more general proof of the Koszulness of certain classes of modules formed by attaching rows to a fixed Schur functor associated to a skew-partition. We conclude with an application of these results that allows us to compute the regularity of the sheaf $\mathbb{S}^\lambda(\mathcal{R})$ in arbitrary characteristic.

Notation 6.20. Let A be any Koszul algebra and α any fixed composition. Define the right A -module $\mathbb{S}_A^{\alpha \circ \bullet}$ via

$$\mathbb{S}_A^{\alpha \circ \bullet} := \bigoplus_{d \geq 0} \mathbb{S}_A^{\alpha \circ (d)},$$

with right A -module action induced by the canonical surjections $\mathbb{S}_A^\alpha \otimes_R A_d \twoheadrightarrow \mathbb{S}_A^{\alpha \circ (d)}$. Similarly, the right $(A^!)^*$ -comodule $\mathbb{S}_A^{\alpha \cdot (1^\bullet)}$ is defined via

$$\mathbb{S}_A^{\alpha \cdot (1^\bullet)} := \bigoplus_{d \geq 0} \mathbb{S}_A^{\alpha \cdot (1^d)},$$

with comodule action induced by the canonical injections $\mathbb{S}_A^{\alpha \cdot (1^d)} \hookrightarrow \mathbb{S}_A^\alpha \otimes_R (A_d^!)^*$. The left A -modules $\mathbb{S}_A^{\bullet \circ \alpha}$ and left $(A^!)^*$ -comodules $\mathbb{S}_A^{(1^\bullet) \cdot \alpha}$ are defined identically, but with the appropriate concatenation/near-concatenation appearing on the left.

Example 6.21. If $\alpha = (e)$ is a single integer, then the module $\mathbb{S}_A^{(e) \circ \bullet}$ is simply $(A_+)^e$, viewed as a right A -module. More generally, let $\Omega_{\geq j}^i(R)$ denote the i -th syzygy of the A -module R ,³ truncated past degree j . Then

$$\Omega_{\geq j}^i(R) = \mathbb{S}_A^{(1^{i-1}, j) \circ \bullet}.$$

³That is, the image of the i -th differential in any A -projective resolution of R .

Theorem 6.22. *Let A be any Koszul algebra and α any composition. Then the right A -module $\mathbb{S}_A^{\alpha \odot \bullet}$ is a Koszul A -module, and the minimal free resolution over A has the form*

$$\dots \rightarrow \mathbb{S}_A^{\alpha \cdot (1^i)} \otimes_R A(-i) \rightarrow \mathbb{S}_A^{\alpha \cdot (1^{i-1})} \otimes_R A(-i+1) \rightarrow \dots \rightarrow \mathbb{S}_A^{\alpha \cdot (1)} \otimes_R A(-1) \rightarrow \mathbb{S}_A^\alpha \otimes_R A.$$

In particular, the quadratic dual of $\mathbb{S}_A^{\alpha \odot \bullet}$ is precisely the left A^1 -module $\mathbb{S}_A^{\bullet \odot \text{rev}(\alpha^t)}$.

Proof. Let α be any fixed partition and set $M := \mathbb{S}_A^{\alpha \odot \bullet}$. We use the criterion of Lemma 4.43(2). Let $d \geq 0$ be any integer and observe first that

$$\mathbb{S}_{M \geq d, A}^\beta = \mathbb{S}_A^{\alpha \odot d \cdot \beta}.$$

Thus for any compositions β, γ the exactness of the sequence

$$0 \rightarrow \mathbb{S}_{M \geq d, A}^{\beta \cdot \gamma} \rightarrow \mathbb{S}_{M \geq d, A}^\beta \otimes_R \mathbb{S}_A^\gamma \rightarrow \mathbb{S}_{M \geq d, A}^{\beta \odot \gamma} \rightarrow 0$$

is equivalent to the exactness of the sequence

$$0 \rightarrow \mathbb{S}_A^{\delta \cdot \gamma} \rightarrow \mathbb{S}_A^\delta \otimes_R \mathbb{S}_A^\gamma \rightarrow \mathbb{S}_A^{\delta \odot \gamma} \rightarrow 0,$$

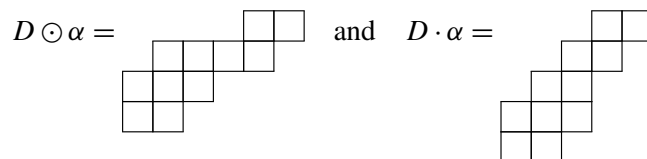
where $\delta = \alpha \odot (d) \cdot \beta$. This latter sequence is evidently exact, since the algebra A is assumed to be Koszul, whence the module M is Koszul. The latter statements are trivial consequences of the Priddy complex associated to a Koszul module (see Theorem A.25). □

The modules $\mathbb{S}_A^{\alpha \odot \bullet}$ are constructed in such a way that they are totally compatible with any kind of ambient group actions, and the naturality of this construction leads one to wonder if there are classes of Koszul modules in the literature that are “secretly” of the form $\mathbb{S}_A^{\alpha \odot \bullet}$ for some α . We pose this question formally.

Question 6.23. Are there interesting examples of Koszul modules in the literature of the form $\mathbb{S}_A^{\alpha \odot \bullet}$ for some composition α ? (One such class of examples arises as in Example 6.21.)

For the remainder of this subsection, we assume that $A = S^\bullet(V)$ (the symmetric algebra) or $\bigwedge^\bullet V$ (the exterior algebra), where V is any free R -module (R is still assumed to be a commutative ring). In this setting, we have access to the classically defined Schur functors $\mathbb{S}^D(V)$ of Akin, Buchsbaum and Weyman [1982], where $D = \lambda/\mu$ is a skew partition. For a composition α , the notation $D \odot \alpha$ will denote the skew partition obtained by attaching the bottom row of α to the top row of the diagram D . Likewise, the notation $D \cdot \alpha$ is defined to be the skew partition obtained by concatenating the ribbon diagram associated with α to the top row of D .

Example 6.24. Let $D = (3, 3, 2)/(1)$ and $\alpha = (2, 2)$. Then:



Remark 6.25. The concatenation/near-concatenation of arbitrary diagrams D and D' is defined in [Almoussa et al. 2024, Definition 3.4], but we will not need this level of generality here.

Definition 6.26. Let $D = \lambda/\mu$ be a skew-partition. The notation $\mathbb{S}^{D \circ \bullet}(V)$ will denote the $S^\bullet(V)$ -module with

$$\mathbb{S}^{D \circ \bullet}(V) := \bigoplus_{d \geq 0} \mathbb{S}^{D \circ (d)}(V),$$

with multiplication induced by the canonical surjections

$$\mathbb{S}^D(V) \otimes_R S^d(V) \rightarrow \mathbb{S}^{D \circ d}(V).$$

Likewise, the notation $\mathbb{S}^{D \cdot (1^\bullet)}(V)$ denotes the $\bigwedge^\bullet V$ -comodule with

$$\mathbb{S}^{D \cdot (1^\bullet)}(V) := \bigoplus_{d \geq 0} \mathbb{S}^{D \cdot (1^d)}(V)$$

with comultiplication induced by the natural inclusions

$$\mathbb{S}^{D \cdot (1^d)}(V) \hookrightarrow \mathbb{S}^D(V) \otimes_R \bigwedge^d V.$$

The analogous definitions for \mathbb{S}^D replaced by the Weyl functors \mathbb{W}^D will be used, with the tacit knowledge that the module $\mathbb{W}^{D \circ \bullet}(V)$ is instead a $\bigwedge^\bullet V$ -module. Likewise, the modules $\mathbb{S}^{\bullet \circ D}(V)$ (resp. $\mathbb{W}^{\bullet \circ D}(V)$) and $\mathbb{S}^{(1^\bullet) \cdot D}(V)$ (resp. $\mathbb{W}^{(1^\bullet) \cdot D}(V)$) are defined analogously.

Remark 6.27. A simple way to define the module structure on $\mathbb{S}^{D \circ \bullet}(V)$ is to take advantage of the short exact sequence, see [Almoussa et al. 2024, Proposition 3.6],

$$0 \rightarrow \mathbb{S}^{D \cdot (1)}(V) \rightarrow \mathbb{S}^D \otimes_R V \rightarrow \mathbb{S}^{D \circ (1)}(V) \rightarrow 0,$$

and then define $\mathbb{S}^{D \circ \bullet}(V)$ to be the quadratic $S^\bullet(V)$ -module induced by the above short exact sequence, there is always a standard way to do this; see, for instance, [Polishchuk and Positselski 2005].

Theorem 6.28. Let $A := S^\bullet(V)$, the symmetric algebra on a free R -module V . For any skew partition $D := \lambda/\mu$, the $S^\bullet(V)$ -module $\mathbb{S}^{D \circ \bullet}(V)$ is Koszul, and the minimal free resolution over A has the form

$$\dots \rightarrow \mathbb{S}^{D \cdot (1^i)} \otimes_R A(-i) \rightarrow \mathbb{S}^{D \cdot (1^{i-1})}(V) \otimes_R A(-i+1) \rightarrow \dots \rightarrow \mathbb{S}^{D \cdot (1)}(V) \otimes_R A(-1) \rightarrow \mathbb{S}^D(V) \otimes_R A.$$

In particular, the quadratic dual of $\mathbb{S}^{D \circ \bullet}(V)$ is precisely the $\bigwedge^\bullet V^*$ -module $\mathbb{W}^{\bullet \circ D^t}(V)$. The analogous statement for the $A = \bigwedge^\bullet V$ -module $\mathbb{W}^{D \circ \bullet}(V)$ also holds.

Proof. Define $M := \mathbb{S}^{D \circ \bullet}(V)$ and notice that by identical reasoning to the proof of Theorem 6.22, for any integer $d \geq 0$ and composition α there is an equality

$$\mathbb{S}_{M \geq d, A}^\alpha = \mathbb{S}^{D \circ (d) \cdot \alpha}(V),$$

whence the sequences of Lemma 4.43 read

$$0 \rightarrow \mathbb{S}^{D \circ (d) \cdot \alpha \cdot \beta}(V) \rightarrow \mathbb{S}^{D \circ (d) \cdot \alpha}(V) \otimes_R \mathbb{S}_A^\beta \rightarrow \mathbb{S}^{D \circ (d) \cdot \alpha \circ \beta}(V) \rightarrow 0.$$

By [Almoussa et al. 2024, Proposition 3.6], this sequence is exact; in the notation of [Almoussa et al. 2024], the diagram denoted D is the diagram $D \odot (d) \cdot \alpha$ in our notation and D' is the ribbon diagram associated to β . By Lemma 4.43(2), the module M is Koszul, and the latter statements are again an immediate consequence of the Priddy complex (see Theorem A.25). \square

As a further application, we conclude this section with a characteristic-free computation of the regularity of a certain class of vector bundles on projective space by using the resolution of Theorem 6.28.

Notation 6.29. Let V be any k -vector space and $\mathbb{P}(V)$ denote projective space on V . The *tautological subbundle* \mathcal{R} on $\mathbb{P}(V)$ is the twisted sheaf $\Omega(1)$, where Ω denotes the cotangent bundle. More concretely, \mathcal{R} is defined via a twist of the Euler sequence

$$0 \rightarrow \mathcal{R} \rightarrow V \otimes_k \mathcal{O}_{\mathbb{P}(V)} \rightarrow \mathcal{O}_{\mathbb{P}(V)}(1) \rightarrow 0.$$

For convenience, recall that a sheaf \mathcal{F} on some variety X is *r-regular* if

$$H^i(X, \mathcal{F}(r - i)) = 0$$

for all $i > 0$. The *regularity* of a sheaf \mathcal{F} is defined to be the minimal integer r such that \mathcal{F} is r -regular. We conclude our applications with a characteristic-free regularity computation for a canonical class of vector bundles on projective space.

Theorem 6.30. *Let V be a k -vector space and \mathcal{R} the tautological subbundle on $\mathbb{P}(V)$. Given any partition $\lambda = (\lambda_1, \dots, \lambda_n)$ (where $n = \dim V$), there is an exact sequence of vector bundles of $\mathbb{P}(V)$*

$$\begin{aligned} \dots \rightarrow \mathbb{S}^{(\lambda_1, \lambda) \cdot 1^i}(V) \otimes_k \mathcal{O}_{\mathbb{P}(V)}(-\lambda_1 - i) \rightarrow \dots \rightarrow \mathbb{S}^{(\lambda_1, \lambda) \cdot 1}(V) \otimes_k \mathcal{O}_{\mathbb{P}(V)}(-\lambda_1 - 1) \\ \rightarrow \mathbb{S}^{(\lambda_1, \lambda)}(V) \otimes_k \mathcal{O}_{\mathbb{P}(V)}(-\lambda_1) \rightarrow \mathbb{S}^\lambda(\mathcal{R}) \rightarrow 0. \end{aligned} \quad (6.30.1)$$

Moreover, writing $\lambda^t = (\lambda_1^t, \dots, \lambda_m^t)$ (where $\lambda_m^t > 0$),⁴ there is an equality

$$H^{\lambda_m^t}(\mathbb{P}(V), \mathbb{S}^\lambda(\mathcal{R})(\lambda_1 - \lambda_m^t - 1)) = \mathbb{S}^{((\lambda_1 - 1)^{\lambda_m^t + 1}, \lambda_2, \dots, \lambda_n)}(V).$$

In particular, the sheaf $\mathbb{S}^\lambda(\mathcal{R})$ has regularity λ_1 .

Remark 6.31. Theorem 6.30 generalizes a theorem of Gao and Raicu [2024, Theorem 2.2], where the authors used the Kempf–Weyman geometric technique [Weyman 2003] to construct the above resolution in the case $\lambda = (a)$; they used this to prove that the regularity of the sheaf $\mathbb{S}^a(\mathcal{R})$ is precisely a .

The subtlety of needing to concatenate the partition λ on the rightmost column was not present in [Gao and Raicu 2024], since concatenating a column of the left/right of the diagram for the partition (a, a) yields isomorphic representations.

⁴In other words, λ_m^t is the length of the rightmost column of the tableau associated to λ .

Example 6.32. Consider the partition

$$\lambda := (3, 3, 2, 1) = \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \\ \hline \square & & \\ \hline \end{array}$$

so that $\lambda^t = (4, 3, 2)$. Then Theorem 6.30 implies that there is a resolution of the form

$$\cdots \rightarrow \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & & \\ \hline \end{array} \otimes_{\mathbf{k}} \mathcal{O}_{\mathbb{P}(V)}(-5) \rightarrow \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & & \\ \hline \end{array} \otimes_{\mathbf{k}} \mathcal{O}_{\mathbb{P}(V)}(-4) \rightarrow \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & & \\ \hline \end{array} \otimes_{\mathbf{k}} \mathcal{O}_{\mathbb{P}(V)}(-3) \rightarrow \mathbb{S}^{(3,3,2,1)}(\mathcal{R}),$$

where in the above complex a Young diagram for shape λ/μ corresponds to the Schur module $\mathbb{S}^{\lambda/\mu}(V)$. By stripping off the rightmost column of the shapes appearing in the above complex, Theorem 6.30 also implies that there is an isomorphism

$$H^2(\mathbb{P}(V), \mathbb{S}^{(3,3,2,1)}(\mathcal{R})) = \mathbb{S}^{(2,2,2,2,1)}(V).$$

This is particularly evident if the ambient vector space V has dimension 4, since the above resolution becomes the short exact sequence

$$0 \rightarrow \mathbb{S}^{(2,2,2,2,1)}(V) \otimes_{\mathbf{k}} \det(V) \otimes_{\mathbf{k}} \mathcal{O}_{\mathbb{P}(V)}(-4) \rightarrow \mathbb{S}^{(3,3,3,2,1)}(V) \otimes_{\mathbf{k}} \mathcal{O}_{\mathbb{P}(V)}(-3) \rightarrow \mathbb{S}^{(3,3,2,1)}(\mathcal{R}) \rightarrow 0.$$

Proof of Theorem 6.30. The complex (6.30.1) arises by taking sheaves associated to the resolutions of Theorem 6.28 for the partition λ . The sheaf associated to the module $\mathbb{S}^{\lambda \odot \bullet}(V)$ is precisely $\mathbb{S}^{\lambda}(\mathcal{R})$, whence the sequence (6.30.1) is indeed an exact sequence of vector bundles.

Observe that it is of no loss of generality to assume that $\lambda_n = 0$, since if $\lambda_n > 0$ we may write

$$\mathbb{S}^{\lambda}(\mathcal{R}) = \det(\mathcal{R})^{\lambda_n} \otimes_{\mathcal{O}_{\mathbb{P}(V)}} \mathbb{S}^{(\lambda_1 - \lambda_n, \lambda_2 - \lambda_n, \dots, \lambda_{n-1} - \lambda_n, 0)}(\mathcal{R}).$$

Using the fact that $\det(\mathcal{R}) = \mathcal{O}_{\mathbb{P}(V)}(-1)$, we see that it indeed suffices to prove the statement of Theorem 6.30 with $\lambda_n = 0$.

Twist the complex (6.30.1) by $\lambda_1 - \lambda_m^t - 1$. The cohomology of each of the terms

$$\mathbb{S}^{(\lambda_1, \lambda) \cdot 1^i}(V) \otimes_{\mathbf{k}} \mathcal{O}_{\mathbb{P}(V)}(\underbrace{-i - \lambda_m^t - 1}_{= -\lambda_1 - i + (\lambda_1 - \lambda_m^t - 1)})$$

is 0 unless $i = n - \lambda_m^t - 1$, since if $i < n - \lambda_m^t - 1$ then $0 > -i - \lambda_m^t - 1 > -n$ and hence the twists $\mathcal{O}_{\mathbb{P}(V)}(-i - \lambda_m^t - 1)$ have 0 cohomology identically. If $i > n - \lambda_m^t - 1$, then the Schur module $\mathbb{S}^{(\lambda_1, \lambda) \cdot (1^{n - \lambda_m^t - 1})}(V)$ is identically 0, since the rightmost column has length strictly greater than n (which is the rank of V). It follows that

$$H^j(\mathbb{P}(V), \mathbb{S}^{(\lambda_1, \lambda) \cdot (1^{n - \lambda_m^t - 1})}(V) \otimes_{\mathbf{k}} \mathcal{O}_{\mathbb{P}(V)}(-n)) = \begin{cases} \mathbb{S}^{(\lambda_1, \lambda) \cdot 1^{n - \lambda_m^t - 1}}(V) \otimes_{\mathbf{k}} \det(V^*) & \text{if } j = n - 1, \\ 0 & \text{otherwise.} \end{cases}$$

Combining the fact that $\mathbb{S}^{(\lambda_1, \lambda) \cdot 1^{n-\lambda_m^t-1}}(V) = \mathbb{S}^{((\lambda_1-1)^{\lambda_m^t+1}, \lambda_2, \dots, \lambda_n)}(V) \otimes_k \det(V)$ with the above equality, the hypercohomology spectral sequence implies that

$$H^j(\mathbb{P}(V), \mathbb{S}^\lambda(\mathcal{R})(\lambda_1 - \lambda_m^t - 1)) = \begin{cases} \mathbb{S}^{((\lambda_1-1)^{\lambda_m^t+1}, \lambda_1-1, \lambda_2, \dots, \lambda_n)}(V) & \text{if } j = \lambda_m^t, \\ 0 & \text{otherwise.} \end{cases}$$

The fact that $H^i(\mathbb{P}(V), \mathbb{S}^\lambda(\mathcal{R})(r - i)) = 0$ for all $r \geq \lambda_1$ is an immediate consequence of the complex (6.30.1), since twisting by any $s > \lambda_1 - \lambda_m^t - 1$ will yield a complex of vector bundles whose terms have at most global sections. \square

Appendix: Koszul algebras and modules over commutative rings

The purpose of this appendix is to define Koszul algebras/modules and their quadratic duals and recall Backelin’s theorem in the generality established in Section 2. After developing the machinery of refinement complexes, we relate Backelin’s theorem to the exactness properties of these complexes. Much of the material in this section follows from straightforward extensions of the material of [Polishchuk and Positselski 2005], but since we assume that R is an arbitrary commutative ring and our algebras are only flat R -modules in each homogeneous component, there are some additional details/technicalities to be verified.

A.1. Generalities on quadratic algebras and modules.

Definition A.1. Let A be any quadratic algebra and M any graded (left) A -module M of initial degree t . There is a canonical multiplication map $A_1^{\otimes d} \otimes_R M_t \rightarrow M_{t+d}$ for every $d \geq 0$; the kernel of this map will be denoted Q_{t+d}^M .

The module M is called *quadratic* if

- (1) the canonical map $A_1^{\otimes d} \otimes_R M_t \rightarrow M_{t+d}$ is surjective for all $d \geq 0$, and
- (2) for every $d \geq 0$, there is an equality

$$Q_{d+t}^M = Q_2^A \otimes_R A_1^{\otimes d-2} \otimes_R M_t + \dots + \underbrace{A_1^{\otimes i} \otimes_R Q_2^A \otimes_R A_1^{\otimes d-i-2} \otimes_R M_t + \dots + A_1^{\otimes d-1} \otimes_R Q_{t+1}^M}_{(i+1)\text{-th position}}$$

Definition A.2 (quadratic duals). Let A be a quadratic R -algebra and M any quadratic left A -module of initial degree t . The *quadratic dual* $A^! \subset \text{Ext}_A^*(R, R)$ is defined to be the subalgebra

$$\bigoplus_{i \in \mathbb{Z}} \text{Ext}_A^i(R, R)_i \subset \text{Ext}_A^*(R, R).$$

Notice that this is indeed a well-defined subalgebra, since the Yoneda product respects both the cohomological and internal grading. Viewing $\text{Ext}_A^i(M, R)$ as a right $A^!$ -module (via Yoneda composition), define the *quadratic dual* $M^! \subset \text{Ext}_A^i(M, R)$ to be the $A^!$ -submodule

$$\bigoplus_{i \in \mathbb{Z}} \text{Ext}_A^i(M, R)_{i+t} \subset \text{Ext}_A^*(M, R).$$

The quadratic dual $M^!$ of a right A -module M is defined analogously and is a left $A^!$ -module.

Remark A.3. The above definition is indeed well-defined for right A -modules, since a right A -module is equivalently a left A^{op} -module, and it is evident that there is an isomorphism of algebras

$$\text{Ext}_{A^{\text{op}}}^{\bullet}(R, R) \cong \text{Ext}_A^{\bullet}(R, R)^{\text{op}}.$$

Thus $\text{Ext}_A^{\bullet}(M, R)$ is a right $\text{Ext}_A^{\bullet}(R, R)^{\text{op}}$ -module, and hence a left $\text{Ext}_A^{\bullet}(R, R)$ -module.

Remark A.4. The notion of a *quadratic dual* is typically only reserved for Koszul algebras. The modules A^{\dagger} and M^{\dagger} as defined in Definition A.2 are sometimes referred to as diagonal subalgebras and diagonal submodules of the Ext algebra/module, but in view of the observation below it seems appropriate to use the name quadratic dual for the general construction.

Observation A.5. Let A be a quadratic R -algebra and M any quadratic left A -module of initial degree t . Then the quadratic dual A^{\dagger} is a quadratic algebra, and likewise the quadratic dual M^{\dagger} is a quadratic right A^{\dagger} -module.

Proof. Dualizing the Bar complex $\text{Bar}^A(A)$, the n -th graded piece of the quadratic dual A^{\dagger} is by definition defined to be the cokernel of the map

$$\bigoplus_{i=1}^{n-1} A_1^{*\otimes i-1} \otimes_R A_2^* \otimes_R A_1^{*\otimes n-i-1} \rightarrow A_1^{*\otimes n}.$$

Moreover, the Yoneda product is induced by the tensor algebra product on the cobar construction, in which case the algebra A^{\dagger} is by definition a quadratic algebra. Similarly, dualizing the bar complex $\text{Bar}^A(M)$ implies that

$$M_n^{\dagger} := \text{Coker} \left(\begin{array}{c} M_{t+1}^* \otimes_R A_1^{*\otimes n-t-1} \\ \oplus \\ \bigoplus_{i=1}^{n-t-1} M_t^* \otimes_R A_1^{*\otimes i-1} \otimes_R A_2^* \otimes_R A_1^{*\otimes n-t-i-1} \end{array} \rightarrow M_t^* \otimes_R A_1^{*\otimes n-t} \right).$$

Again, the right Yoneda module structure is induced by the right tensor algebra structure on the cobar complex, in which case M^{\dagger} is a quadratic right A^{\dagger} -module. □

Finally, we conclude this section by defining a Koszul algebra/module:

Definition A.6. Let A be a quadratic R -algebra. The algebra A is *Koszul* if there is an isomorphism of R -algebras

$$A^{\dagger} \cong \text{Ext}_A^{\bullet}(R, R).$$

Likewise, let M be any left A -module. Then the module M is *Koszul* if there is an isomorphism of right A^{\dagger} -modules

$$M^{\dagger} \cong \text{Ext}_A^{\bullet}(M, R).$$

In other words, the inclusions of Definition A.2 are actually equalities.

A.2. Refinement complexes. In this section, we recall the notion of *refinement complexes*; it should be noted here that the terminology is new, but such complexes (often unnamed) have been studied before; see for instance [Polishchuk and Positselski 2005, Chapter 2.8]. These complexes may be understood as (subquotients of) homogeneous strands of the augmented bar complex associated to a quadratic algebra A .

Throughout this section, we will use the notation for compositions and the standard operations between them established in Section 4.1.

Definition A.7 (refinement complexes). Let $\alpha = (\alpha_1, \dots, \alpha_n)$ be a composition of some integer d . Define the *refinement (chain) complex* $R_{\bullet}^{A,M}(\alpha)$ to be the chain complex with

$$R_i^{A,M}(\alpha) = \bigoplus_{\substack{\beta \geq \alpha \\ \ell(\beta) - \ell(\alpha) = i}} ((A^1)^* \otimes_R (M^1)^*)_{\beta},$$

and differential induced by the cobar differential on $\text{Cobar}^{A^1}(M^1)$. Likewise, define the *refinement (cochain) complex* $R_{A,M}^{\bullet}(\alpha)$ to be the cochain complex with

$$R_{A,M}^i(\alpha) = \bigoplus_{\substack{\beta \geq \alpha \\ \ell(\beta) - \ell(\alpha) = i}} (A \otimes_R M)_{\beta},$$

and differential induced by the bar differential on $\text{Bar}^A(M)$. The notation $R_i^A(\alpha)$ and $R_i^A(\alpha)$ will be shorthand for $R_i^{A,A^+}(\alpha)$ and $R_{A,A^+}^i(\alpha)$, respectively.

Example A.8. If $A = S(V)$ is the symmetric algebra on some free R -module V , then

$$R_{\bullet}^{S(V)}(3, 2, 4) : \bigwedge^9 V \rightarrow \begin{matrix} \bigwedge^5 V \otimes \bigwedge^4 V \\ \oplus \\ \bigwedge^3 V \otimes \bigwedge^6 V \end{matrix} \rightarrow \bigwedge^3 V \otimes \bigwedge^2 V \otimes \bigwedge^4 V.$$

Likewise, using the notation of Definition 4.15,

$$R_A^{\bullet}(3, 2, 4, 3) : A_{(3,2,4,3)} \rightarrow \begin{matrix} A_{(5,4,3)} \\ \oplus \\ A_{(3,6,3)} \\ \oplus \\ A_{3,2,7} \end{matrix} \rightarrow \begin{matrix} A_{(9,3)} \\ \oplus \\ A_{(5,7)} \\ \oplus \\ A_{(3,9)} \end{matrix} \rightarrow A_{12}.$$

A.3. Koszulness and distributivity. In this section, we recall Backelin’s theorem for Koszul algebras. For convenience, we state explicitly the following equivalent conditions for Koszulness, which are trivial retranslations of the definition.

Observation A.9. Let A be any quadratic algebra. Then the following are equivalent:

- (1) The algebra A is Koszul.
- (2) For all $j > i$, one has $\text{Ext}_A^i(R, R)_j = 0$.
- (3) For all $j > i$, one has $\text{Tor}_i^A(R, R)_j = 0$.

Definition A.10. Let A be a quadratic R -algebra. Given positive integers $n, i > 0$, use the notation

$$S_{A,i}^n := A_1^{\otimes i-1} \otimes_R Q_2^A \otimes_R A_1^{\otimes n-i-1} \subset A_1^{\otimes n}.$$

The following crucial observation ties all the machinery introduced in Section A.2 to the theory of distributivity developed in Section 3.

Observation A.11. Let A be a Koszul R -algebra with $n, i > 0$ positive integers. Then there is an isomorphism of R -modules

$$\frac{A_1^{\otimes n}}{\bigvee_{i \in I} S_{A,i}^n} \cong A_{\phi(I)}.$$

In particular, with notation as in Definition 4.4 and Construction 3.5 there are isomorphisms of complexes

$$R_{\bullet}^A(\alpha) = C_{\bullet}^{\phi^{-1}(\alpha)}(A_1^{\otimes|\alpha|}; S_{A,1}^n, \dots, S_{A,n-1}^n) \quad \text{and} \quad R_A^{\bullet}(\alpha) = C_{\phi^{-1}(\alpha)}^{\bullet}(A_1^{\otimes|\alpha|}; S_{A,1}^n, \dots, S_{A,n-1}^n).$$

Finally, we state and prove the generalization of Backelin’s theorem. Notice that the proof is deceptively short, but relies on the entirety of the material developed thus far in the paper.

Theorem A.12. *Let A be any quadratic R -algebra. Then*

$$A \text{ is Koszul} \iff \text{the collection } S_{A,1}^n, \dots, S_{A,n-1}^n \subset A_1^{\otimes n} \text{ is distributive for all } n > 0.$$

In particular, the refinement complexes $R_A^{\bullet}(\alpha)$ and $R_{\bullet}^A(\alpha)$ are exact in positive (co)homological degrees for all compositions α .

Remark A.13. Notice that it is clear that distributivity implies that A is Koszul, since this means that the complex $R_A^{\bullet}(1^d)$ is in particular exact in positive cohomological degrees. It is not obvious at all that Koszulness should be sufficient to imply exactness of the refinement complexes for *all* choices of compositions.

Proof. This is an immediate consequence of Backelin’s theorem combined with Observation A.11. □

Remark A.14. Notice that with the distributivity perspective of Koszul algebras,

$$A_n \longleftrightarrow A_n^! \quad \text{corresponds to} \quad \frac{A_1^{\otimes n}}{S_{1,A}^n + \dots + S_{n-1,A}^n} \longleftrightarrow \frac{A_1^{*\otimes n}}{S_{1,A}^{n \vee} + \dots + S_{n-1,A}^{n \vee}}.$$

This is another quick way to see that $A \cong (A^!)^!$ as R -algebras.

Corollary A.15. *Let A be a quadratic R -algebra. Then A is Koszul if and only if $A^{(d)}$ is Koszul for every $d > 0$.*

Remark A.16. Of course, the nontrivial direction of Corollary A.15 is the fact that A is Koszul implies that $A^{(d)}$ is Koszul for all $d > 0$. The distributivity criterion makes this a trivial consequence; it is worth mentioning that this was originally proved by Barcanescu and Manolache [1981].

A.4. Koszul modules over Koszul algebras. The following section makes some additional observations about Koszul modules that will be useful to reference explicitly in earlier sections. We start with an analogous observation on equivalent conditions for Koszulness of a module:

Observation A.17. Let A be a Koszul algebra and M any quadratic left A -module of initial degree t . Then the following are equivalent:

- (1) The module M is Koszul.
- (2) For all $j > i$, one has $\text{Ext}_A^i(M, R)_{t+j} = 0$.
- (3) For all $j > i$, one has $\text{Tor}_i^A(M, R)_{t+j} = 0$.

The following submodule collections are the evident analog of the collection in Definition A.10 for quadratic modules:

Definition A.18. Let A be a quadratic algebra and M any quadratic left A -module of initial degree t . Given positive integers $n, i > 0$, use the notation

$$S_{A,M,i}^n := \begin{cases} A_1^{\otimes i-1} \otimes_R Q_2 \otimes_R A_1^{n-i-2} \otimes_R M_t & \text{if } i < n-1, \\ A_1^{\otimes n-1} \otimes_R Q_{t+1}^M & \text{if } i = n-1 \end{cases} \subset A_1^{\otimes n-1} \otimes_R M_t.$$

The submodules $S_{M,A,i}^n \subset M_t \otimes_R A_1^{\otimes n-1}$ for a right A -module M are defined analogously.

Theorem A.19. Let M be a left (resp. right) A -module of initial degree t , where A is a Koszul R -algebra. Then

$$M \text{ is Koszul} \iff \text{the collection } S_{A,M,1}^n, \dots, S_{A,M,n-1}^n \subset A_1^{\otimes n} \otimes_R M_t \text{ is distributive for all } n > 0.$$

The analogous statement for right A -modules holds as well. In particular, the refinement complexes $R_{A,M}^\bullet(\alpha)$ and $R_{\bullet,M}^A(\alpha)$ are exact in positive (co)homological degrees for all compositions α .

Proof. The proof is identical to the proof of Theorem A.12. □

The following observation is immediate upon applying $-\otimes_R M$ to the augmented bar complex $\text{Bar}^A(A)$:

Observation A.20. Let A be a Koszul R -algebra. Then any flat R -module is a Koszul A -module by viewing the R -module as a left (or right) A -module concentrated in some fixed degree.

Corollary A.21. If M is a Koszul left (resp. right) A -module, then the truncation $M_{\geq d}$ is a Koszul left (resp. right) A -module for all $d \in \mathbb{Z}$.

Proof. Let t denote the initial degree of M , and proceed by induction on the difference $d - t$. When $d - t = 0$, it is by assumption that $M_{\geq t} = M$ is Koszul.

Let $d - t > 0$. There is a short exact sequence

$$0 \rightarrow M_{\geq d-t} \rightarrow M_{\geq d-t-1} \rightarrow M_{d-t-1} \rightarrow 0,$$

where M_{d-t-1} is a flat R -module viewed as being concentrated in degree $d - t - 1$. By the inductive hypothesis, the truncation $M_{\geq d-t-1}$ is Koszul and by Observation A.20 the A -module M_{d-t-1} is Koszul. By the long exact sequence of cohomology, the truncation $M_{\geq d-t}$ must also be Koszul. □

Observation A.22. If M is a Koszul left A -module over a Koszul algebra A , then the graded dual M^* is a Koszul right A -module.

Proof. It is clear by definition that each graded component of M^* is a flat R -module, so it remains to prove that $\text{Ext}_{A^{\text{op}}}^{\bullet}(M^*, R)$ is generated in minimal internal degree. There is a string of isomorphisms of algebras

$$\text{Ext}_A^{\bullet}(M, R)^{\text{op}} \cong \text{Ext}_{A^{\text{op}}}^{\bullet}(R, M) \cong \text{Tor}_{\bullet}^{A^{\text{op}}}(R, M^*)^*.$$

Since $\text{Ext}_A^{\bullet}(M, R)$ is generated in minimal internal degrees, so is $\text{Tor}_{\bullet}^{A^{\text{op}}}(R, M^*)$. This implies that M^* is Koszul. \square

We conclude this subsection with a statement about the Koszulness of tensor products of Koszul algebras. This will be most useful when dealing with multi-Schur functors.

Corollary A.23. Let A and B be two Koszul algebras and let M (resp. N) a left A (resp. B)-module. Then the tensor product $A \otimes_R B$ is a Koszul algebra and $M \otimes_R N$ is a Koszul left $A \otimes_R B$ -module.

Proof. Recall that for any (left) A -module M and B -module N , the shuffle product

$$\nabla : \text{Bar}^A(M) \otimes_R \text{Bar}^B(N) \rightarrow \text{Bar}^{A \otimes_R B}(M \otimes_R N)$$

is a well-defined morphism of complexes. The complex $\text{Bar}^{A \otimes_R B}(M \otimes_R N)$ is a flat resolution of $M \otimes_R N$, and by the assumption that all the modules A, B, M , and N are flat as R -modules the tensor product $\text{Bar}^A(M) \otimes_R \text{Bar}^B(N)$ is also a flat resolution of $M \otimes_R N$. Since $R \otimes_A \text{Bar}^A(M)$ and $R \otimes_B \text{Bar}^B(N)$ both have homology concentrated in minimal degrees, so does $R \otimes_{A \otimes_R B} \text{Bar}^{A \otimes_R B}(M \otimes_R N)$. \square

A.5. The Priddy complex. In this section, we recall the well-known construction of the Priddy complex (originally proved in the work of Priddy [1970]) for our setting. The definition is identical to that of the original definition over fields, but we include the definition along with the relevant properties for completeness.

Construction A.24. Let A be a Koszul R -algebra and M a left Koszul module. For each $i \geq 0$, there is a canonical inclusion of left A -modules

$$A \otimes_R (M^{\dagger})_{i+i}^* \hookrightarrow A \otimes_R A_1^{\otimes i} \otimes_R M_i = \text{Bar}_i^A(M)_i,$$

where the left A -module structure on $A \otimes_R (M^{\dagger})^*$ comes from only allowing A to act on the leftmost tensor factor. This inclusion thus lifts to an inclusion of complexes

$$A \otimes_R (M^{\dagger})_{\bullet}^* \hookrightarrow \text{Bar}^A(M).$$

The differential on the complex $A \otimes_R (M^{\dagger})_{\bullet}^*$ is induced by the Bar complex differential (this is indeed well-defined).

Theorem A.25. Let A be a quadratic algebra and M any left A -module. Then the following are equivalent:

- (1) The left A -module M is Koszul.
- (2) The complex $A \otimes_R (M^{\dagger})^*$ is a flat resolution of M over A .

Proof. (2) \implies (1): This implication is clear, since if $A \otimes_R (M^1)^*$ is a flat resolution of M over A , tensoring with R over A implies that $\text{Tor}_i^A(R, M)$ is concentrated in degree $i + t$ (recall that Tor may be computed using flat resolutions).

(1) \implies (2): Let $\iota : A \otimes_R (M^1)^*$ be the inclusion of Construction A.24 and consider the mapping cone $\text{Cone}(\iota)$. There is the tautological short exact sequence of complexes

$$0 \rightarrow R \otimes_A \text{Bar}^A(M) \rightarrow \text{Cone}(R \otimes_A \iota) \rightarrow (M^1)^*[-1] \rightarrow 0.$$

The assumption that M is Koszul implies that $R \otimes_A \iota$ is a quasiisomorphism, so that $\text{Cone}(R \otimes_A \iota)$ is an exact complex of flat R -modules. However, this implies by Nakayama’s lemma $\text{Cone}(\iota)$ must be an exact complex, whence $A \otimes_R (M^1)^*$ is an A -flat resolution of M . \square

Acknowledgments

The author thanks Ayah Almousa, Vic Reiner, Steven Sam, and Jerzy Weyman for helpful conversations related to this material or comments/corrections on earlier drafts of this paper. The author also thanks the anonymous referee for a close reading with many helpful corrections/suggestions. The author was supported by NSF grant DMS-2202871.

References

- [Akin et al. 1982] K. Akin, D. A. Buchsbaum, and J. Weyman, “Schur functors and Schur complexes”, *Adv. Math.* **44**:3 (1982), 207–278. MR Zbl
- [Almousa et al. 2024] A. Almousa, M. Perlman, A. Pevzner, V. Reiner, and K. VandeBogert, “Equivariant resolutions over Veronese rings”, *J. Lond. Math. Soc.* (2) **109**:1 (2024), art. id. e12848. MR
- [Backelin 1981] J. Backelin, *A distributiveness property of augmented algebras and some related homological results*, Ph.D. thesis, Stockholm University, 1981.
- [Beilinson et al. 1996] A. Beilinson, V. Ginzburg, and W. Soergel, “Koszul duality patterns in representation theory”, *J. Amer. Math. Soc.* **9**:2 (1996), 473–527. MR Zbl
- [Billera et al. 2006] L. J. Billera, H. Thomas, and S. van Willigenburg, “Decomposable compositions, symmetric quasisymmetric functions and equality of ribbon Schur functions”, *Adv. Math.* **204**:1 (2006), 204–240. MR Zbl
- [Bott 1957] R. Bott, “Homogeneous vector bundles”, *Ann. of Math.* (2) **66** (1957), 203–248. MR Zbl
- [Bărcănescu and Manolache 1981] Ş. Bărcănescu and N. Manolache, “Betti numbers of Segre–Veronese singularities”, *Rev. Roumaine Math. Pures Appl.* **26**:4 (1981), 549–565. MR Zbl
- [Buchsbaum and Eisenbud 1975] D. A. Buchsbaum and D. Eisenbud, “Generic free resolutions and a family of generically perfect ideals”, *Adv. Math.* **18**:3 (1975), 245–301. MR Zbl
- [Faber et al. 2021] E. Faber, M. Juhnke-Kubitzke, H. Lindo, C. Miller, R. R. G., and A. Seceleanu, “Canonical resolutions over Koszul algebras”, pp. 281–301 in *Women in commutative algebra* (Banff, AB, 2019), edited by C. Miller et al., Assoc. Women Math. Ser. **29**, Springer, 2021. MR Zbl
- [Gao and Raicu 2024] Z. Gao and C. Raicu, “Cohomology of line bundles on the incidence correspondence”, *Trans. Amer. Math. Soc. Ser. B* **11** (2024), 64–97. MR Zbl
- [Goresky et al. 1998] M. Goresky, R. Kottwitz, and R. MacPherson, “Equivariant cohomology, Koszul duality, and the localization theorem”, *Invent. Math.* **131**:1 (1998), 25–83. MR Zbl
- [Grinberg and Reiner 2014] D. Grinberg and V. Reiner, “Hopf algebras in combinatorics”, preprint, 2014. arXiv 1409.8356

- [Hamel and Goulden 1995] A. M. Hamel and I. P. Goulden, “Planar decompositions of tableaux and Schur function determinants”, *European J. Combin.* **16**:5 (1995), 461–477. MR Zbl
- [Huang 2016] J. Huang, “A tableau approach to the representation theory of 0-Hecke algebras”, *Ann. Comb.* **20**:4 (2016), 831–868. MR Zbl
- [Kempf 1976] G. R. Kempf, “On the collapsing of homogeneous bundles”, *Invent. Math.* **37**:3 (1976), 229–239. MR Zbl
- [Lascoux and Pragacz 1988] A. Lascoux and P. Pragacz, “Ribbon Schur functions”, *European J. Combin.* **9**:6 (1988), 561–574. MR Zbl
- [Manin 1987] Y. I. Manin, “Some remarks on Koszul algebras and quantum groups”, *Ann. Inst. Fourier (Grenoble)* **37**:4 (1987), 191–205. MR Zbl
- [Manin 1991] Y. I. Manin, *Topics in noncommutative geometry*, Princeton Univ. Press, 1991. MR Zbl
- [Mastroeni and McCullough 2023] M. Mastroeni and J. McCullough, “Chow rings of matroids are Koszul”, *Math. Ann.* **387**:3–4 (2023), 1819–1851. MR Zbl
- [Polishchuk and Positselski 2005] A. Polishchuk and L. Positselski, *Quadratic algebras*, Univ. Lect. Ser. **37**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl
- [Positselski 2014] L. Positselski, “Galois cohomology of a number field is Koszul”, *J. Number Theory* **145** (2014), 126–152. MR Zbl
- [Priddy 1970] S. B. Priddy, “Koszul resolutions”, *Trans. Amer. Math. Soc.* **152** (1970), 39–60. MR Zbl
- [Reiner et al. 2007] V. Reiner, K. M. Shaw, and S. van Willigenburg, “Coincidences among skew Schur functions”, *Adv. Math.* **216**:1 (2007), 118–152. MR Zbl
- [Sam and Snowden 2017] S. V. Sam and A. Snowden, “Infinite rank spinor and oscillator representations”, *J. Comb. Algebra* **1**:2 (2017), 145–183. MR Zbl
- [Sam and Snowden 2019] S. V. Sam and A. Snowden, “Some generalizations of Schur functors”, *Proc. Amer. Math. Soc.* **147**:1 (2019), 77–90. MR Zbl
- [Serre 1959] J.-P. Serre, “Représentations linéaires et espaces homogènes kählériens des groupes de Lie compacts (d’après Armand Borel et André Weil)”, exposé 100 in *Séminaire Bourbaki*, 1953/1954, Benjamin, Amsterdam, 1959. Reprinted as pp. 447–454 in *Séminaire Bourbaki* **2**, Soc. Math. France, Paris, 1995. Zbl
- [Weyman 2003] J. Weyman, *Cohomology of vector bundles and syzygies*, Cambridge Tracts in Math. **149**, Cambridge Univ. Press, 2003. MR Zbl
- [Yuzvinskii 2001] S. Yuzvinskii, “Orlik–Solomon algebras in algebra and topology”, *Uspekhi Mat. Nauk* **56**:2(338) (2001), 87–166. In Russian; translated in *Russian Math. Surv.* **56**:2 (2001), 293–364. MR Zbl

Communicated by Victor Reiner

Received 2023-12-12 Revised 2024-04-06 Accepted 2024-05-23

kvandeb@nd.edu

*Department of Mathematics, University of Notre Dame, Notre Dame, IN,
United States*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT 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 *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

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, Corel Draw, 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.

Algebra & Number Theory

Volume 19 No. 4 2025

Odd moments in the distribution of primes	617
VIVIAN KUPERBERG	
Efficient resolution of Thue–Mahler equations	667
ADELA GHERGA and SAMIR SIKSEK	
Automorphisms of del Pezzo surfaces in characteristic 2	715
IGOR DOLGACHEV and GEBHARD MARTIN	
On the D-module of an isolated singularity	763
THOMAS BITOUN	
Ribbon Schur functors	771
KELLER VANDEBOGERT	