Algebra & Number Theory

Volume 18
2024
No. 9

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2024 is US $525/year for the electronic version, and $770/year (+$65, 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/

ALGEBRA AND NUMBER THEORY 18:9 (2024)
https://doi.org/10.2140/ant.2024.18.1589

A bound for the exterior product of S-units

Shabnam Akhtari and Jeffrey D. Vaaler

We generalize an inequality for the determinant of a real matrix proved by A. Schinzel, to more general exterior products of vectors in Euclidean space. We apply this inequality to the logarithmic embedding of S-units contained in a number field k. This leads to a bound for the exterior product of S-units expressed as a product of heights. Using a volume formula of P. McMullen we show that our inequality is sharp up to a constant that depends only on the rank of the S-unit group but not on the field k. Our inequality is related to a conjecture of F. Rodriguez Villegas.

1. Introduction

Let k be an algebraic number field, $k^\times$ its multiplicative group of nonzero elements, and $h : k^\times \to [0, \infty)$ the absolute, logarithmic, Weil height (or simply the *height*). In [Akhtari and Vaaler 2016] we proved inequalities that compare the size of an S-regulator with the product of heights of a maximal collection of independent S-units. If $k \subseteq l$ are both number fields the results in [Akhtari and Vaaler 2022] extend inequalities of this sort to the multiplicative group of relative units. Here we prove analogous inequalities for the exterior product of a collection of independent S-units that is not a maximal collection.

At each place v of k we write k_v for the completion of k at v. We use two absolute values $\|\cdot\|_v$ and $|\cdot|_v$ from the place v. The absolute value $\|\cdot\|_v$ extends the usual archimedean or nonarchimedean absolute value on the subfield $\mathbb{Q}$. Then $|\cdot|_v$ must be a power of $\|\cdot\|_v$, and we set

$$|\cdot|_v = \|\cdot\|_v^{d_v/d}, \tag{1-1}$$

where $d_v = [k_v : \mathbb{Q}_v]$ is the local degree of the extension and $d = [k : \mathbb{Q}]$ is the global degree. With these normalizations the height of an algebraic number $\alpha \neq 0$ that belongs to k is given by

$$h(\alpha) = \sum_v \log^+ |\alpha|_v = \tfrac{1}{2} \sum_v \bigl|\log |\alpha|_v\bigr|. \tag{1-2}$$

Each sum in (1-2) is over the set of all places v of k, and the equality between the two sums follows from the product formula.

The authors are grateful to the anonymous referee. Shabnam Akhtari's research was supported by the Simons Foundation Collaboration Grant, Award Number 635880, and by the National Science Foundation Awards DMS-2001281 and DMS-2327098.
MSC2020: 05D99, 11J25, 11R27, 15A75.
Keywords: Weil height, exterior products.

Let S be a finite set of places of k such that S contains all the archimedean places. Then

$$O_S = \{\gamma \in k : \|\gamma\|_v \leq 1 \text{ for all places } v \notin S\}$$

is the ring of S-integers in k, and

$$O_S^\times = \{\gamma \in k^\times : \|\gamma\|_v = 1 \text{ for all places } v \notin S\}$$

is the multiplicative group of S-units in O_S. The abelian group $O_S^\times$ has rank r, where $|S| = r + 1$, and we assume that r is positive. We write $\boldsymbol{x} = (x_v)$ for a (column) vector in $\mathbb{R}^{r+1}$ where the coordinates of $\boldsymbol{x}$ are indexed by places v in S. We write

$$\|\boldsymbol{x}\|_1 = \sum_{v \in S} |x_v|$$

for the l^1-norm of $\boldsymbol{x}$. The *logarithmic embedding* of $O_S^\times$ into $\mathbb{R}^{r+1}$ is the homomorphism defined at each point α in $O_S^\times$ by

$$\alpha \mapsto \boldsymbol{\alpha} = (d_v \log \|\alpha\|_v), \tag{1-3}$$

where the rows of the vector $\boldsymbol{\alpha}$ on the right of (1-3) are indexed by places v in S. It follows from (1-1) and (1-2) that if α is a point in $O_S^\times$ and $\boldsymbol{\alpha}$ is the image of α in $\mathbb{R}^{r+1}$ using the logarithmic embedding (1-3), then

$$2[k : \mathbb{Q}]h(\alpha) = \sum_{v \in S} \bigl|d_v \log \|\alpha\|_v\bigr| = \|\boldsymbol{\alpha}\|_1. \tag{1-4}$$

The kernel of the logarithmic embedding (1-3) is the torsion subgroup

$$\bigl\{\alpha \in O_S^\times : (d_v \log \|\alpha\|_v) = \boldsymbol{0}\bigr\} = \mathrm{Tor}(O_S^\times) \tag{1-5}$$

of all roots of unity in $k^\times$. It is known that (1-5) is a finite, cyclic group, and from the S-unit theorem of Dirichlet, Chevalley, and Hasse (see [Narkiewicz 2004, Theorem 3.12]) we learn that the quotient

$$\mathfrak{U}_S(k) = O_S^\times / \mathrm{Tor}(O_S^\times)$$

is a free abelian group of rank r. Therefore the logarithmic embedding (1-3) induces an isomorphism from $\mathfrak{U}_S(k)$ onto the discrete subgroup

$$\Gamma_S(k) = \bigl\{(d_v \log \|\alpha\|_v) : \alpha \in O_S^\times\bigr\} \subseteq \mathbb{R}^{r+1},$$

which is a free group of rank r. It follows from the product formula

$$\sum_{v \in S} d_v \log \|\alpha\|_v = 0$$

that $\Gamma_S(k)$ is contained in the r-dimensional diagonal subspace

$$\mathcal{D}_r = \left\{\boldsymbol{x} = (x_v) : \sum_{v \in S} x_v = 0\right\} \subseteq \mathbb{R}^{r+1}.$$

The height h is constant on cosets of the quotient group $\mathfrak{U}_S(k)$ and therefore h is well defined as a map

$$h : \mathfrak{U}_S(k) \to [0, \infty).$$

Let $\eta_1, \eta_2, \dots, \eta_r$ be multiplicatively independent elements in $\mathfrak{U}_S(k)$ that form a basis for the free group $\mathfrak{U}_S(k)$. Let

$$\boldsymbol{\eta}_j = (d_v \log \|\eta_j\|_v) \quad \text{for } j = 1, 2, \dots, r$$

be the logarithmic embedding of these points in $\Gamma_S(k) \subseteq \mathcal{D}_r$. Working with the induced l^1-norm in the exterior algebra $\operatorname{Ext}(\mathbb{R}^{r+1})$ we find that

$$(r+1) \operatorname{Reg}_S(k) = \|\boldsymbol{\eta}_1 \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_r\|_1, \tag{1-6}$$

where $\operatorname{Reg}_S(k)$ is the S-regulator. More generally, let $\alpha_1, \alpha_2, \dots, \alpha_r$ be multiplicatively independent elements in $\mathfrak{U}_S(k)$, and let $\mathfrak{A} \subseteq \mathfrak{U}_S(k)$ be the multiplicative subgroup of rank r which they generate. Let

$$\boldsymbol{\alpha}_j = (d_v \log \|\alpha_j\|_v) \quad \text{for } j = 1, 2, \dots, r$$

be the image of $\alpha_1, \alpha_2, \dots, \alpha_r$ in $\Gamma_S(k)$. It follows that there exists a unique $r \times r$ nonsingular matrix $B = (b_{ij})$ with entries in $\mathbb{Z}$ such that

$$\boldsymbol{\alpha}_j = \sum_{i=1}^{r} \boldsymbol{\eta}_i b_{ij} \quad \text{for } j = 1, 2, \dots, r. \tag{1-7}$$

Then the index of the subgroup $\mathfrak{A}$ in $\mathfrak{U}_S(k)$ is

$$[\mathfrak{U}_S(k) : \mathfrak{A}] = |\det B|. \tag{1-8}$$

Combining (1-6), (1-7), and (1-8), we find that

$$(r+1) \operatorname{Reg}_S(k) \, [\mathfrak{U}_S(k) : \mathfrak{A}] = \|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_r\|_1. \tag{1-9}$$

In [Akhtari and Vaaler 2016, Theorem 1.1] we proved an upper bound for the S-regulator that is equivalent to the identity (1-9) and the inequality

$$\|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_r\|_1 \le 2^{-r}(r+1) \prod_{j=1}^{r} \|\boldsymbol{\alpha}_j\|_1. \tag{1-10}$$

The following result provides a generalization of (1-10) to an exterior product of q independent vectors in the free group $\Gamma_S(k)$, where $1 \le q \le r$.

Theorem 1.1. *Let $\alpha_1, \alpha_2, \dots, \alpha_q$ be multiplicatively independent points in $\mathfrak{U}_S(k)$, and let*

$$\boldsymbol{\alpha}_j = (d_v \log \|\alpha_j\|_v) \quad \textit{for } j = 1, 2, \dots, q$$

be the logarithmic embedding of $\alpha_1, \alpha_2, \dots, \alpha_q$ *in* $\Gamma_S(k)$. *Then we have*

$$\|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1 \le 2^{-q} C(q, r) \prod_{j=1}^{q} \|\boldsymbol{\alpha}_j\|_1, \tag{1-11}$$

where

$$C(q, r) = \min\left\{2^q, \left(\frac{r+1}{r+1-q}\right)^{r+1-q}\right\}. \tag{1-12}$$

We find that

$$C(q, r) = 2^q \quad \text{if } 2q \le r+1,$$

and

$$C(q, r) = \left(\frac{r+1}{r+1-q}\right)^{r+1-q} \quad \text{if } r+1 \le 2q.$$

In particular we have $C(r, r) = (r+1)$ so that (1-11) includes the inequality (1-10). By applying (1-4) it follows that (1-11) can be written using the Weil height as

$$\|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1 \le C(q, r) \prod_{j=1}^{q} ([k:\mathbb{Q}] h(\alpha_j)).$$

Let $\alpha_1, \alpha_2, \dots, \alpha_q$ and $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_q$ be as in the statement of Theorem 1.1, and let $\mathfrak{A}$ be the subgroup of $\Gamma_S(k)$ generated by $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_q$. Clearly $\mathfrak{A}$ is a free group of rank q. It is easy to show that the l^1-norm of the exterior product

$$\|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1 \tag{1-13}$$

depends on the subgroup $\mathfrak{A}$, but does not depend on the choice of generators. Because of (1-9) the l^1-norm of the exterior product (1-13) extends the S-regulator from the group $\Gamma_S(k)$ to subgroups of $\Gamma_S(k)$ having lower rank.

Alternatively, if $\alpha \neq 1$ belongs to $O_S^\times$ and $\boldsymbol{\alpha} \neq \boldsymbol{0}$ is the image of α with respect to the logarithmic embedding (1-3), then $\boldsymbol{\alpha}$ and $-\boldsymbol{\alpha}$ are the unique pair of generators of a subgroup of rank 1 in $\Gamma_S(k)$. In view of (1-4) we may regard $\|\boldsymbol{\alpha}\|_1$ as the height of this subgroup. Then (1-13) extends the height to more general subgroups $\mathfrak{A} \subseteq \Gamma_S(k)$ having rank q. This definition of a height on subgroups is similar to the definition stated in [Vaaler 2014, equation (6.14)].

In [Akhtari and Vaaler 2016, Theorem 1.2] we showed that if $\mathfrak{A} \subseteq \Gamma_S(k)$ is a subgroup with full rank r, then there exist r linearly independent points in $\mathfrak{A}$ such that the product of their heights is bounded by a number depending only on r multiplied by

$$\mathrm{Reg}_S(k)\,[\mathfrak{U}_S(k) : \mathfrak{A}]. \tag{1-14}$$

The following result generalizes [Akhtari and Vaaler 2016, Theorem 1.2] to arbitrary subgroups $\mathfrak{A} \subseteq \Gamma_S(k)$ having positive rank q where $1 \le q \le r$. In this result the S-regulator (1-14) is replaced by the l^1-norm (1-13) of the exterior product of a set of generators for the subgroup $\mathfrak{A}$.

Theorem 1.2. *Let $\mathfrak{A} \subseteq \Gamma_S(k)$ be a subgroup of positive rank q, and let the points*

$$\boldsymbol{\alpha}_j = (d_v \log \|\alpha_j\|_v), \quad \text{where } j = 1, 2, \ldots, q,$$

generate the subgroup $\mathfrak{A}$. Then there exists a subgroup $\mathfrak{B} \subseteq \mathfrak{A}$ of rank q and a set of generators

$$\boldsymbol{\beta}_j = (d_v \log \|\beta_j\|_v), \quad \text{where } j = 1, 2, \ldots, q,$$

for $\mathfrak{B}$ such that

$$\|\boldsymbol{\beta}_1 \wedge \boldsymbol{\beta}_2 \wedge \cdots \wedge \boldsymbol{\beta}_q\|_1 = [\mathfrak{A} : \mathfrak{B}]\, \|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1 \tag{1-15}$$

and

$$\prod_{j=1}^{q} \|\boldsymbol{\beta}_j\|_1 \le q!\, \|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1. \tag{1-16}$$

We have $[\mathfrak{A} : \mathfrak{B}] \le q!$.

By applying (1-4) we find that the product on the left of (1-16) can be written using the Weil height as

$$\prod_{i=1}^{q} \|\boldsymbol{\beta}_i\|_1 = 2^q \prod_{j=1}^{q} ([k : \mathbb{Q}]\, h(\beta_j)).$$

Because the subgroups $\mathfrak{B} \subseteq \mathfrak{A}$ both have rank q, the identity (1-15) follows as in our derivation of (1-8) from (1-7).

It would be of interest to know if there exist absolute constants $b_0 > 0$ and $b_1 > 1$ such that the factor $q!$ on the right of (1-16) could be replaced by $b_0 b_1^q$. This could have implications for a conjecture of F. Rodriguez Villegas which we discuss in Section 2.

2. A conjecture of F. Rodriguez Villegas

In a well-known paper D. H. Lehmer [1933] proposed an important problem about the roots of irreducible polynomials in $\mathbb{Z}[x]$. An equivalent form of Lehmer's problem stated using the absolute, logarithmic, Weil height (1-2) is this: does there exist an absolute constant $c > 0$ such that

$$c \le [\mathbb{Q}(\alpha) : \mathbb{Q}]\, h(\alpha)$$

whenever $\alpha \ne 0$ is an algebraic number and not a root of unity? If $\alpha \ne 0$ and α is not a unit, the lower bound

$$\log 2 \le [\mathbb{Q}(\alpha) : \mathbb{Q}]\, h(\alpha)$$

follows easily. Therefore when considering Lehmer's problem we may restrict our attention to algebraic units α which are not roots of unity. Further information about Lehmer's problem can be found in [Bombieri and Gubler 2006, Section 1.6.15; Smyth 2008; Waldschmidt 2000, Section 3.6].

Let S_∞ be the set of archimedean places of k and assume that $|S_\infty| \ge 2$. We continue to write $|S_\infty| = r + 1$ so that the logarithmic embedding (1-3) is an isomorphism from the free group

$$\mathfrak{U}_{S_\infty}(k) = O_{S_\infty} / \operatorname{Tor}(O_{S_\infty}^\times)$$

onto the discrete subgroup $\Gamma_{S_\infty}(k)$ of rank r contained in the diagonal subspace $\mathcal{D}_r \subseteq \mathbb{R}^{r+1}$. Then Lehmer's problem asks if there exists an absolute constant $c > 0$ such that the inequality

$$c \le 2[k:\mathbb{Q}]h(\alpha) = \|\boldsymbol{\alpha}\|_1 \tag{2-1}$$

holds at all points $\boldsymbol{\alpha} \neq \mathbf{0}$ in $\Gamma_{S_\infty}(k)$. A generalization of this conjecture to independent subsets $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_q$ in $\Gamma_{S_\infty}(k)$ with $2 \le q \le r$ was proposed by Bertrand [1997]. More precisely, Bertrand asked if for each integer $2 \le q$ there exists a constant $c_q > 0$ such that

$$c_q \le \|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_2, \tag{2-2}$$

where the l^2-norm of the wedge product on the right of (2-2) is the covolume of the subgroup of $\Gamma_{S_\infty}(k)$ generated by $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_q$. Examples found by Siegel [1969] show that the inequality (2-2) cannot hold for $q = 1$. However, a positive answer for $q \ge 3$ was established by Amoroso and David [1999].

An alternative generalization of Lehmer's problem to subgroups of rank q has been proposed in a conjecture of F. Rodriguez Villegas stated in [Chinburg et al. 2022, Appendix], and also discussed in [Amoroso and David 2021]. We state a special case of this conjecture for pure wedges.

Conjecture 2.1 (F. Rodriguez Villegas). *There exist two absolute constants $c_0 > 0$ and $c_1 > 1$ with the following property. If q is an integer such that*

$$1 \le q \le r = \operatorname{rank} \Gamma_{S_\infty}(k),$$

and if $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_q$ are linearly independent points in $\Gamma_{S_\infty}(k)$, then

$$c_0 c_1^q \le \|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1. \tag{2-3}$$

If $q = 1$ then the truth of (2-3) would solve the problem originally proposed by Lehmer, and if $q = r$ then (2-3) follows from a known lower bound for the regulator proved by R. Zimmert [1981]. Thus the conjecture of Rodriguez Villegas interpolates between the unsolved problem of Lehmer and Zimmert's result. It follows from earlier work of Pohst [1978] and Schinzel [1973] that Conjecture 2.1 holds for the collection of totally real algebraic number fields k.

Let $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_q$ be linearly independent points in $\Gamma_{S_\infty}(k)$ and let $\mathfrak{A} \subseteq \Gamma_{S_\infty}(k)$ be the subgroup of rank q that they generate. We have already observed in connection with (1-13) that the l^1-norm

$$\|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1$$

depends on the subgroup $\mathfrak{A}$, but does *not* depend on the choice of generators. Thus Conjecture 2.1 can be regarded as a generalization of Lehmer's problem (reformulated as a conjecture) from subgroups of rank 1 to more general subgroups of rank q where $1 \le q \le r$.

Here is a related conjecture.

Conjecture 2.2. *There exist two absolute constants $d_0 > 0$ and $d_1 > 1$ with the following property. If q is an integer such that*

$$1 \le q \le r = \operatorname{rank} \Gamma_{S_\infty}(k),$$

and if $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_q$ *are linearly independent points in* $\Gamma_{S_\infty}(k)$, *then*

$$d_0 d_1^q \le \|\boldsymbol{\alpha}_1\|_1 \|\boldsymbol{\alpha}_2\|_1 \cdots \|\boldsymbol{\alpha}_q\|_1.$$

It follows from (1-12) that the constant on the right of (1-11) satisfies

$$2^{-q} C(q, r) \le 1.$$

Therefore if the conjectured inequality (2-3) is correct, then from Theorem 1.1 we also get

$$c_0 c_1^q \le \|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1 \le \prod_{j=1}^{q} \|\boldsymbol{\alpha}_j\|_1.$$

Thus Conjecture 2.1 implies Conjecture 2.2 with $d_0 = c_0$ and $d_1 = c_1$.

Now assume that Conjecture 2.2 is correct. Let $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_q$ be linearly independent points in the logarithmic embedding $\Gamma_{S_\infty}(k)$, and let $\mathfrak{A}$ be the subgroup of rank q that they generate. By Theorem 1.2 there exist linearly independent points $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_q$ in $\mathfrak{A}$ such that

$$d_0 d_1^q \le \|\boldsymbol{\beta}_1\|_1 \|\boldsymbol{\beta}_2\|_1 \cdots \|\boldsymbol{\beta}_q\|_1 \le q!\, \|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1, \tag{2-4}$$

where the inequality on the left of (2-4) follows from Conjecture 2.2, and the inequality on the right of (2-4) follows from (1-16). However, as $q!$ grows faster than an exponential function of q, at present we are unable to conclude that Conjecture 2.2 implies Conjecture 2.1. This could change if the factor $q!$ in the inequality (1-16) could be replaced by a factor of the form $b_0 b_1^q$, where $b_0 > 0$ and $b_1 > 1$ are absolute constants.

3. Generalization of Schinzel's inequality, I

For a real number x we write

$$x^+ = \max\{0, x\} \quad \text{and} \quad x^- = \max\{0, -x\},$$

so that $x = x^+ - x^-$ and $|x| = x^+ + x^-$. Let $\boldsymbol{x} = (x_n)$ be a (column) vector in $\mathbb{R}^N$. As in [Akhtari and Vaaler 2016, equation (4.3)], the *Schinzel norm* is the function

$$\delta : \mathbb{R}^N \to [0, \infty)$$

defined by

$$\delta(\boldsymbol{x}) = \max\left\{ \sum_{m=1}^{N} x_m^+, \ \sum_{n=1}^{N} x_n^- \right\} = \tfrac{1}{2}\left| \sum_{n=1}^{N} x_n \right| + \tfrac{1}{2} \sum_{n=1}^{N} |x_n|.$$

It is clear that δ is in fact a norm on $\mathbb{R}^N$, and we write

$$K_N = \{\boldsymbol{x} \in \mathbb{R}^N : \delta(\boldsymbol{x}) \le 1\}$$

for the corresponding closed unit ball. Then K_N is a compact, convex, symmetric subset of $\mathbb{R}^N$ with a nonempty interior. The N-dimensional volume of K_N was computed in [Akhtari and Vaaler 2016, Lemma 4.1]. The connection between the Schinzel norm and the Weil height follows from (1-4) and (5-2) (see also [Akhtari and Vaaler 2016, Lemma 5.1]).

In Lemma 3.2 we will determine the finite collection of extreme points of K_N. Then a combinatorial argument in Section 4 applied to the extreme points of K_N will lead to a proof of the following inequalities.

Theorem 3.1. *Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_L$ be linearly independent vectors in $\mathbb{R}^N$. If $L = N$ then*

$$|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_N| \le \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\cdots\delta(\boldsymbol{x}_N), \tag{3-1}$$

if $L < N \le 2L$ then

$$\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1 \le \left(\frac{N}{N-L}\right)^{N-L} \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\cdots\delta(\boldsymbol{x}_L), \tag{3-2}$$

and if $2L \le N$ then

$$\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1 \le 2^L\, \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\cdots\delta(\boldsymbol{x}_L). \tag{3-3}$$

Alternatively, for $L < N$ we have

$$\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1 \le \min\left\{2^L, \left(\frac{N}{N-L}\right)^{N-L}\right\} \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\cdots\delta(\boldsymbol{x}_L). \tag{3-4}$$

If $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N$, are (column) vectors in $\mathbb{R}^N$, then Schinzel [1978] proved the inequality

$$|\det(\boldsymbol{x}_1\ \ \boldsymbol{x}_2\ \ \cdots\ \ \boldsymbol{x}_N)| \le \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\cdots\delta(\boldsymbol{x}_N), \tag{3-5}$$

which is equivalent to (3-1). It can be shown that there exist nontrivial cases of equality in the inequality (3-2) whenever the integer $N - L$ is a divisor of N. And it can be shown that there always exist nontrivial cases of equality in the inequality (3-3). It is instructive to define the function

$$g_L : [L, \infty] \to [1, e^L]$$

by

$$g_L(x) = \begin{cases} 1 & \text{if } x = L, \\ \left(\dfrac{x}{x-L}\right)^{x-L} & \text{if } L < x < \infty, \\ e^L & \text{if } x = \infty. \end{cases}$$

It follows that $x \mapsto g_L(x)$ is continuous, and has a continuous, positive derivative on (L, ∞). Then $x \mapsto g_L(x)$ is strictly increasing on $[L, \infty]$. We have $g_L(2L) = 2^L$, and this clarifies the behavior of the function

$$x \mapsto \min\{2^L, g_L(x)\}$$

which occurs on the right of (3-4).

We recall that a point $\boldsymbol{k}$ in K_N is an *extreme point* of K_N if $\boldsymbol{k}$ cannot be written as a proper convex combination of two distinct points in K_N. Obviously all extreme points of K_N occur on the boundary of K_N. Let

$$\varphi : \mathbb{R}^N \to \mathbb{R}$$

be a continuous linear functional, and write

$$\delta^*(\varphi) = \sup\{\varphi(\boldsymbol{x}) : \delta(\boldsymbol{x}) \le 1\}$$

for the dual norm of φ. As K_N is compact there exists a point $\boldsymbol{\eta}$ in K_N such that

$$\delta^*(\varphi) = \varphi(\boldsymbol{\eta}).$$

If there exists a linear functional φ such that

$$\{\boldsymbol{\eta} \in K_N : \delta^*(\varphi) = \varphi(\boldsymbol{\eta})\} = \{\boldsymbol{k}\},$$

then $\boldsymbol{k}$ is an *exposed point* of K_N. It is known (see [Eggleston 1958, section 1.8, exercise 3]) that an exposed point of K_N is also an extreme point of K_N.

We define two finite, disjoint subsets of $\mathbb{R}^N$ by

$$E_N = \{\pm \boldsymbol{e}_m : 1 \le m \le N\} \quad \text{and} \quad F_N = \{\boldsymbol{e}_m - \boldsymbol{e}_n : m \ne n\}, \tag{3-6}$$

where $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N$ are the standard basis vectors in $\mathbb{R}^N$. Clearly we have

$$|E_N| = 2N \quad \text{and} \quad |F_N| = N^2 - N.$$

It follows easily that each point of $E_N \cup F_N$ is on the boundary of K_N.

Lemma 3.2. *The subset $E_N \cup F_N$ is the collection of all extreme points of K_N.*

Proof. For $1 \le m \le N$ let $\varphi_m : \mathbb{R}^N \to \mathbb{R}$ be the linear functional defined by

$$\varphi_m(\boldsymbol{x}) = \tfrac{1}{2} \sum_{n=1}^{N} x_n + \tfrac{1}{2} x_m.$$

Then we have

$$\varphi_m(\boldsymbol{x}) \le \tfrac{1}{2} \left| \sum_{n=1}^{N} x_n \right| + \tfrac{1}{2} |x_m|, \tag{3-7}$$

and there is equality in the inequality (3-7) if and only if

$$0 \le \sum_{n=1}^{N} x_n \quad \text{and} \quad 0 \le x_m.$$

We also have

$$\tfrac{1}{2} \left| \sum_{n=1}^{N} x_n \right| + \tfrac{1}{2} |x_m| \le \delta(\boldsymbol{x}), \tag{3-8}$$

and there is equality in the inequality (3-8) if and only if

$$x_n = 0 \quad \text{for each } n \ne m.$$

Combining (3-7) and (3-8) we find that

$$\varphi_m(\boldsymbol{x}) \le \delta(\boldsymbol{x}) \tag{3-9}$$

for all $\boldsymbol{x}$ in $\mathbb{R}^N$, and there is equality in the inequality (3-9) if and only if $\boldsymbol{x} = t\boldsymbol{e}_m$ with $0 \le t$. Therefore

$$\delta^*(\varphi_m) = \sup\{\varphi_m(\boldsymbol{x}) : \delta(\boldsymbol{x}) \le 1\} = \varphi_m(\boldsymbol{e}_m) = 1$$

and

$$\{\boldsymbol{\eta} \in K_N : \delta^*(\varphi_m) = \varphi_m(\boldsymbol{\eta})\} = \{\boldsymbol{e}_m\}.$$

This shows that $\boldsymbol{e}_m$ is an exposed point of K_N, and therefore $\boldsymbol{e}_m$ is an extreme point of K_N. As K_N is symmetric, we find that $-\boldsymbol{e}_m$ is also an extreme point.

Next we suppose that $m \ne n$, and we define the linear functional $\psi_{mn} : \mathbb{R}^N \to \mathbb{R}$ by

$$\psi_{mn}(\boldsymbol{x}) = \tfrac{1}{2}(x_m - x_n).$$

Then we have

$$\psi_{mn}(\boldsymbol{x}) \le \tfrac{1}{2}\left|\sum_{\ell=1}^{N} x_\ell\right| + \tfrac{1}{2}|x_m| + \tfrac{1}{2}|x_n|, \tag{3-10}$$

and there is equality in the inequality (3-10) if and only if

$$\sum_{\ell=1}^{N} x_\ell = 0, \quad 0 \le x_m \text{ and } x_n \le 0.$$

We get

$$\tfrac{1}{2}\left|\sum_{\ell=1}^{N} x_\ell\right| + \tfrac{1}{2}|x_m| + \tfrac{1}{2}|x_n| \le \delta(\boldsymbol{x}), \tag{3-11}$$

with equality in the inequality (3-11) if and only if

$$x_\ell = 0 \quad \text{for all } \ell \ne m \text{ and } \ell \ne n.$$

By combining (3-10) and (3-11) we find that

$$\psi_{mn}(\boldsymbol{x}) \le \delta(\boldsymbol{x}), \tag{3-12}$$

and there is equality in the inequality (3-12) if and only if $\boldsymbol{x} = t(\boldsymbol{e}_m - \boldsymbol{e}_n)$ with $0 \le t$. As in the previous case we conclude that

$$\delta^*(\psi_{mn}) = \sup\{\psi_{mn}(\boldsymbol{x}) : \delta(\boldsymbol{x}) \le 1\} = \psi_{mn}(\boldsymbol{e}_m - \boldsymbol{e}_n) = 1$$

and

$$\{\boldsymbol{\eta} \in K : \delta^*(\psi_{mn}) = \psi_{mn}(\boldsymbol{\eta})\} = \{\boldsymbol{e}_m - \boldsymbol{e}_n\}.$$

This shows that $\boldsymbol{e}_m - \boldsymbol{e}_n$ is an exposed point of K_N, and therefore $\boldsymbol{e}_m - \boldsymbol{e}_n$ is an extreme point of K_N.

We have now shown that each point in $E_N \cup F_N$ is an extreme point of K_N. To complete the proof we will show that if $\boldsymbol{x}$ is a point on the boundary of K_N, then $\boldsymbol{x}$ can be written as a convex combination of points in $E_N \cup F_N$. Thus we assume that

$$\delta(\boldsymbol{x}) = \max\left\{ \sum_{m=1}^{N} x_m^+, \sum_{n=1}^{N} x_n^- \right\} = 1, \tag{3-13}$$

and we write

$$\sigma^+ = \sum_{m=1}^{N} x_m^+ \quad \text{and} \quad \sigma^- = \sum_{n=1}^{N} x_n^-.$$

Then we have

$$\begin{aligned}
\sum_{\substack{m=1 \\ m\neq n}}^{N} \sum_{n=1}^{N} x_m^+ x_n^- (\boldsymbol{e}_m - \boldsymbol{e}_n) &= \left(\sum_{n=1}^{N} x_n^- \right) \sum_{m=1}^{N} x_m^+ \boldsymbol{e}_m - \left(\sum_{m=1}^{N} x_m^+ \right) \sum_{n=1}^{N} x_n^- \boldsymbol{e}_n \\
&= \sigma^- \sum_{m=1}^{N} x_m^+ \boldsymbol{e}_m - \sigma^+ \sum_{n=1}^{N} x_n^- \boldsymbol{e}_n \\
&= \sum_{m=1}^{N} x_m^+ \boldsymbol{e}_m - \sum_{n=1}^{N} x_n^- \boldsymbol{e}_n - (1-\sigma^-) \sum_{m=1}^{N} x_m^+ \boldsymbol{e}_m + (1-\sigma^+) \sum_{n=1}^{N} x_n^- \boldsymbol{e}_n \\
&= \boldsymbol{x} - (1-\sigma^-) \sum_{m=1}^{N} x_m^+ \boldsymbol{e}_m - (1-\sigma^+) \sum_{n=1}^{N} x_n^- (-\boldsymbol{e}_n),
\end{aligned}$$

and therefore

$$\boldsymbol{x} = (1-\sigma^-) \sum_{m=1}^{N} x_m^+ \boldsymbol{e}_m + (1-\sigma^+) \sum_{n=1}^{N} x_n^- (-\boldsymbol{e}_n) + \sum_{\substack{m=1 \\ m\neq n}}^{N} \sum_{n=1}^{N} x_m^+ x_n^- (\boldsymbol{e}_m - \boldsymbol{e}_n). \tag{3-14}$$

The identity (3-14) shows that $\boldsymbol{x}$ is a linear combination of points in $E_N \cup F_N$ with nonnegative coefficients. Using (3-13), the sum of the coefficients in (3-14) is

$$\begin{aligned}
(1-\sigma^-) \sum_{m=1}^{N} x_m^+ + (1-\sigma^+) \sum_{n=1}^{N} x_n^- + \sum_{\substack{m=1 \\ m\neq n}}^{N} \sum_{n=1}^{N} x_m^+ x_n^- &= (1-\sigma^-)\sigma^+ + (1-\sigma^+)\sigma^- + \sigma^+\sigma^- \\
&= 1 - (1-\sigma^+)(1-\sigma^-) \\
&= 1.
\end{aligned}$$

It follows that $\boldsymbol{x}$ is a convex combination of points in $E_N \cup F_N$. We have shown that if $\boldsymbol{x}$ is on the boundary of K_N, then $\boldsymbol{x}$ is a convex combination of points in $E_N \cup F_N$. Therefore the only extreme points of K_N are the points in $E_N \cup F_N$. □

Let

$$I = \{i_1 < i_2 < \cdots < i_L\} \subseteq \{1, 2, \ldots, N\}$$

be a subset of positive cardinality L. If $\boldsymbol{x} = (x_n)$ is a point in $\mathbb{R}^N$ we write $\boldsymbol{x}_I$ for the point in $\mathbb{R}^L$ given by $\boldsymbol{x}_I = (x_{i_\ell})$. Alternatively, $\boldsymbol{x}_I$ is the $L \times 1$ submatrix of $\boldsymbol{x}$ having rows indexed by the integers in the subset I. The following result is now an immediate consequence of Lemma 3.2.

Corollary 3.3. *Let $\boldsymbol{\xi}$ be an element in the set of extreme points $E_N \cup F_N$, and let*

$$I \subseteq \{1, 2, \ldots, N\}$$

be a subset of positive cardinality L. Then either $\boldsymbol{\xi}_I = \boldsymbol{0}$ in $\mathbb{Z}^L$, or $\boldsymbol{\xi}_I$ belongs to the set of extreme points $E_L \cup F_L$.

Let

$$\Phi_{L,N} : \mathbb{R}^N \times \mathbb{R}^N \times \cdots \times \mathbb{R}^N \to \mathbb{R}^M, \quad \text{where } M = \binom{N}{L},$$

be the continuous, alternating, multilinear function taking values in $\mathbb{R}^M$ and defined by

$$\Phi_{L,N}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L) = \boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L.$$

By compactness the continuous, nonnegative function

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L) \mapsto \|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1$$

assumes its maximum value on the L-fold product

$$K_N \times K_N \times \cdots \times K_N.$$

We write

$$\mu_{L,N} = \max\{\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1 : \boldsymbol{x}_\ell \in K_N \text{ for } \ell = 1, 2, \ldots, L\} \tag{3-15}$$

for this maximum value. We show that $\mu_{L,N}$ can be determined by restricting each variable $\boldsymbol{x}_\ell$ to the set $E_N \cup F_N$ of extreme points in K_N.

Lemma 3.4. *There exist points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ in the set of extreme points $E_N \cup F_N$ such that*

$$\mu_{L,N} = \|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1. \tag{3-16}$$

If $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L$ are vectors in $\mathbb{R}^N$ then

$$\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1 \le \mu_{L,N}\, \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2) \cdots \delta(\boldsymbol{x}_L). \tag{3-17}$$

Proof. Let $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_L$ be points in K_N such that

$$\mu_{L,N} = \|\boldsymbol{\eta}_1 \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_L\|_1. \tag{3-18}$$

Because $\Phi_{L,N}$ is linear in each variable, it is easy to show that $\delta(\boldsymbol{\eta}_\ell) = 1$ for each $\ell = 1, 2, \ldots, L$. Also, among all the collections of L points from the boundary of K_N that satisfy (3-18), we may assume that the collection $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_L$ contains the maximum number of extreme points. If this maximum number is L then we are done. Therefore we may assume that the maximum number of extreme points is less than L.

If, for example, $\boldsymbol{\eta}_1$ is not an extreme point, then there exist extreme points $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_J$ in K_N, and positive numbers $\theta_1, \theta_2, \dots, \theta_J$, such that

$$\boldsymbol{\eta}_1 = \sum_{j=1}^{J} \theta_j \boldsymbol{u}_j \quad \text{and} \quad \sum_{j=1}^{J} \theta_j = 1.$$

It follows that

$$\mu_{L,N} = \Bigl\| \sum_{j=1}^{J} \theta_j (\boldsymbol{u}_j \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_L) \Bigr\|_1 \le \sum_{j=1}^{J} \theta_j \|\boldsymbol{u}_j \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_L\|_1 \le \mu_{L,N} \sum_{j=1}^{J} \theta_j = \mu_{L,N} \tag{3-19}$$

Hence there is equality throughout the inequality (3-19), and we conclude that

$$\mu_{L,N} = \|\boldsymbol{u}_j \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_L\|_1$$

for each $j = 1, 2, \dots, J$. But each collection of points $\boldsymbol{u}_j, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_L$ plainly contains one more extreme point than the collection $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_L$. The contradiction shows that there exists a collection of points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$ from the boundary of K_N such that (3-16) holds and each $\boldsymbol{\xi}_\ell$ is an extreme point of K_N.

Next we verify the inequality (3-17). If one of the vectors in the collection $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_L$ is the zero vector, then both sides of (3-17) are zero. Thus we may assume that $\boldsymbol{x}_\ell \ne \boldsymbol{0}$ for each $\ell = 1, 2, \dots, L$. Let

$$\boldsymbol{y}_\ell = \delta(\boldsymbol{x}_\ell)^{-1} \boldsymbol{x}_\ell, \tag{3-20}$$

so that $\delta(\boldsymbol{y}_\ell) = 1$ for each $\ell = 1, 2, \dots, L$. Then we certainly have

$$\|\boldsymbol{y}_1 \wedge \boldsymbol{y}_2 \wedge \cdots \wedge \boldsymbol{y}_L\|_1 \le \mu_{L,N} \tag{3-21}$$

by the definition of $\mu_{L,N}$. Then (3-17) follows using (3-20), (3-21), and the multilinearity of the exterior product. □

The extreme points $E_N \cup F_N$ for the δ-unit ball K_N have the following useful property.

Lemma 3.5. *Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$ be extreme points in the set $E_N \cup F_N$, and let*

$$\Xi = (\boldsymbol{\xi}_1 \;\; \boldsymbol{\xi}_2 \;\; \cdots \;\; \boldsymbol{\xi}_L)$$

be the $N \times L$ matrix having $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$ as columns. If

$$I \subseteq \{1, 2, \dots, N\}$$

is a subset of cardinality $|I| = L$, and Ξ_I is the $L \times L$ submatrix having rows indexed by I, then the integer $\det \Xi_I$ *belongs to the set* $\{-1, 0, 1\}$.

Proof. Clearly the columns of the $L \times L$ submatrix Ξ_I are the $L \times 1$ column vectors $(\boldsymbol{\xi}_1)_I, (\boldsymbol{\xi}_2)_I, \dots, (\boldsymbol{\xi}_L)_I$. If a column of Ξ_I is $\boldsymbol{0}$, then $\det \Xi_I = 0$ is obvious. If each column of Ξ_I is not $\boldsymbol{0}$, then it follows from

Corollary 3.3 that each column of Ξ_I belongs to the set of extreme points $E_L \cup F_L$. Applying Schinzel's determinant inequality (3-5) to the matrix Ξ_I, we get

$$|\det \Xi_I| \leq \delta((\boldsymbol{\xi}_1)_I)\delta((\boldsymbol{\xi}_2)_I)\cdots\delta((\boldsymbol{\xi}_L)_I) = 1.$$

As $\det \Xi_I$ is an integer, the lemma is proved. □

If $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ are extreme points in $E_N \cup F_N$, then it follows from Lemma 3.5 that

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 = \sum_{\substack{I \subseteq \{1,2,\ldots,N\} \\ |I|=L}} |\det \Xi_I| \leq \binom{N}{L}. \tag{3-22}$$

Using (3-16) we get the simple upper bound

$$\mu_{L,N} \leq \binom{N}{L} \quad \text{for } 1 \leq L \leq N. \tag{3-23}$$

It follows from (3-5) that there is equality in (3-23) when $L = N$. There is also equality in (3-23) when $L + 1 = N$; this follows from the example

$$\Xi = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -1 & -1 & -1 & \cdots & -1 & -1 \end{pmatrix}.$$

By squaring each of the subdeterminants in the sum (3-22) we can determine the value of $\mu_{L,N}$ for $2L \leq N$.

Lemma 3.6. *If $1 \leq L < N$ then*

$$\mu_{L,N} \leq 2^L. \tag{3-24}$$

If $2L \leq N$ then there is equality in the inequality (3-24).

Proof. Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ be extreme points in $E_N \cup F_N$, and let

$$\Xi = (\boldsymbol{\xi}_1 \ \ \boldsymbol{\xi}_2 \ \ \cdots \ \ \boldsymbol{\xi}_L)$$

be the $N \times L$ matrix having $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ as columns. It follows from Lemma 3.5 that

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 = \sum_{\substack{I \subseteq \{1,2,\ldots,N\} \\ |I|=L}} |\det \Xi_I| = \sum_{\substack{I \subseteq \{1,2,\ldots,N\} \\ |I|=L}} (\det \Xi_I)^2.$$

Then from the Cauchy–Binet identity we get

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 = \sum_{\substack{I \subseteq \{1,2,\ldots,N\} \\ |I|=L}} (\det \Xi_I)^2 = \det(\Xi^T \Xi). \tag{3-25}$$

The $L \times L$ matrix in the determinant on the right of (3-25) is

$$\Xi^T \Xi = (\boldsymbol{\xi}_k^T \boldsymbol{\xi}_\ell),$$

where $k = 1, 2, \ldots, L$ indexes rows and $\ell = 1, 2, \ldots, L$ indexes columns. As $\Xi^T \Xi$ is an $L \times L$ real, symmetric matrix, we can apply Hadamard's inequality to estimate its determinant. We find that

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 = \det(\Xi^T \Xi) \le \prod_{\ell=1}^{L} \|\boldsymbol{\xi}_\ell\|_2^2 \le 2^L. \tag{3-26}$$

This proves the inequality (3-24).

If the columns of the matrix Ξ are orthogonal, then there is equality in Hadamard's inequality. Therefore, if $2L \le N$ we select $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ in F_N so that

$$\Xi = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

For this choice of Ξ the columns of Ξ are orthogonal. Hence for this choice of Ξ there is equality in (3-26), and equality in (3-24). □

If $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L$ belong to $\mathbb{R}^N$ and $2L \le N$, then it follows from (3-17) and the case of equality in (3-24) that

$$\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1 \le 2^L\, \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\cdots\delta(\boldsymbol{x}_L). \tag{3-27}$$

This proves the inequality (3-3) in the statement of Theorem 3.1.

The following lemma, together with combinatorial arguments in Section 4, will be used in the proof of the inequality (3-2).

Lemma 3.7. *Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$ be linearly independent extreme points in the set $E_N \cup F_N$. Assume that exactly K of the points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$ belong to the subset E_N, where $1 \le K < L$. Then there exist linearly independent extreme points $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_{L-K}$ in the set $E_{N-K} \cup F_{N-K}$ such that*

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 = \|\boldsymbol{\eta}_1 \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_{L-K}\|_1.$$

Proof. By using a suitable permutation of the points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$, we may assume that

$$\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_K\} \subseteq E_N \quad \text{and} \quad \{\boldsymbol{\xi}_{K+1}, \boldsymbol{\xi}_{K+2}, \dots, \boldsymbol{\xi}_L\} \subseteq F_N.$$

We may further assume that for $k = 1, 2, \dots, K$ we have

$$\boldsymbol{\xi}_k = \pm \boldsymbol{e}_{m_k}, \quad \text{where } 1 \le m_1 < m_2 < \cdots < m_K \le N.$$

It will be convenient to write

$$M = \{m_1, m_2, \dots, m_K\}.$$

Now let

$$\Xi = (\boldsymbol{\xi}_1 \;\; \boldsymbol{\xi}_2 \;\; \cdots \;\; \boldsymbol{\xi}_L)$$

be the $N \times L$ matrix having $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$ as columns. We partition Ξ into submatrices

$$\Xi = (U \;\; V),$$

where

$$U = (\boldsymbol{\xi}_1 \;\; \boldsymbol{\xi}_2 \;\; \cdots \;\; \boldsymbol{\xi}_K) \quad \text{and} \quad V = (\boldsymbol{\xi}_{K+1} \;\; \boldsymbol{\xi}_{K+2} \;\; \cdots \;\; \boldsymbol{\xi}_L)$$

are $N \times K$ and $N \times (L-K)$, respectively. We suppose that $I \subseteq \{1, 2, \dots, N\}$ is a subset of cardinality $|I| = L$ such that

$$\det \Xi_I = \det(U_I \;\; V_I) \ne 0. \tag{3-28}$$

On the right of (3-28) the submatrix U_I is $L \times K$ and the submatrix V_I is $L \times (L-K)$. If the integer m_k, which occurs in M, does not belong to I, then the k-th column of Ξ_I is identically zero and (3-28) cannot hold. Therefore (3-28) implies that

$$M \subseteq I.$$

Next we apply the Laplace expansion of the determinant to Ξ_I partitioned as in (3-28). In view of our previous remarks we find that

$$\det \Xi_I = \sum_{\substack{J \subseteq I \\ |J| = K}} (-1)^{\varepsilon(J)} (\det U_J)(\det V_{\widetilde{J}}), \tag{3-29}$$

where

$$\widetilde{J} = I \setminus J$$

is the complement of J in I, and $\varepsilon(J)$ is an integer that depends on J. As before, if the integer m_k which occurs in M does not belong to the subset J, then the k-th column of U_J is identically zero and therefore

$\det U_J = 0$. As $|J| = |M| = K$, we conclude that there is exactly one nonzero term in the sum on the right of (3-29), and the nonzero term occurs when $J = M$. From these observations we conclude that the Laplace expansion (3-29) is simply

$$\det \Xi_I = (-1)^{\varepsilon(M)} (\det U_M)(\det V_{I\setminus M}). \tag{3-30}$$

It is obvious that $\det U_M = \pm 1$, and therefore (3-30) leads to the identity

$$|\det \Xi_I| = |\det V_{I\setminus M}|.$$

Let

$$V' = (\boldsymbol{\xi}'_{K+1} \;\; \boldsymbol{\xi}'_{K+2} \;\; \cdots \;\; \boldsymbol{\xi}'_L)$$

be the $(N-K)\times(L-K)$ submatrix of V obtained by removing the rows of V that are indexed by the integers m_k in the subset M. It follows from Lemma 3.4 that the columns of V' belong to the set of extreme points $E_{N-K} \cup F_{N-K}$. Moreover, we have

$$|\det \Xi_I| = |\det V_{I\setminus M}| = |\det V'_J|, \tag{3-31}$$

where

$$J = I \setminus M \subseteq \{1, 2, \ldots, N\} \setminus M \quad \text{and} \quad |J| = L - K.$$

We note that

$$I \mapsto J = I \setminus \{m_1, m_2, \ldots, m_K\}$$

is a bijection from the set of subsets I that contain M onto the set of subsets of $\{1, 2, \ldots, N\} \setminus M$ that have cardinality $L - K$. Using (3-31) we find that

$$\sum_{\substack{I\subseteq\{1,2,\ldots,N\}\\ M\subseteq I}} |\det \Xi_I| = \sum_{\substack{J\subseteq\{1,2,\ldots,N\}\setminus M\\ |J|=L-K}} |\det V'_J|. \tag{3-32}$$

Because the rows of V' are indexed by the elements of the set $\{1, 2, \ldots, N\}\setminus M$, it follows from (3-32) that

$$\|\boldsymbol{\xi}_1\wedge\boldsymbol{\xi}_2\wedge\cdots\wedge\boldsymbol{\xi}_L\|_1 = \sum_{\substack{I\subseteq\{1,2,\ldots,N\}\\ M\subseteq I}} |\det \Xi_I| = \sum_{\substack{J\subseteq\{1,2,\ldots,N\}\setminus M\\ |J|=L-K}} |\det V'_J| = \|\boldsymbol{\xi}'_{K+1}\wedge\boldsymbol{\xi}'_{K+2}\wedge\cdots\wedge\boldsymbol{\xi}'_L\|_1. \tag{3-33}$$

As the columns of V' belong to $E_{N-K} \cup F_{N-K}$ and satisfy (3-33), they are linearly independent. Therefore we set

$$\boldsymbol{\eta}_\ell = \boldsymbol{\xi}'_{K+\ell} \quad \text{for } \ell = 1, 2, \ldots, L-K,$$

and the lemma is proved. □

4. Generalization of Schinzel's inequality, II

We develop a combinatorial method which leads to an asymptotically sharp upper bound for the quantity $\mu_{L,N}$ defined in (3-15). The bound we prove here applies when $L < N \le 2L$, and will be used to verify the inequality (3-2) in the statement of Theorem 3.1.

We suppose throughout this section that

$$\{S(1), S(2), S(3), \dots, S(L)\} \tag{4-1}$$

is a collection of L distinct subsets of $\{1, 2, \dots, N\}$ such that

$$|S(\ell)| = 2 \quad \text{for each } \ell = 1, 2, \dots, L \tag{4-2}$$

and

$$\bigcup_{\ell=1}^{L} S(\ell) = \{1, 2, \dots, N\}. \tag{4-3}$$

It follows from (4-2) and (4-3) that

$$N \le 2L \le N(N-1),$$

but for our later applications we will make the more restrictive assumption that

$$L < N \le 2L. \tag{4-4}$$

Let $\mathcal{A}$ be the collection of *all* subsets $A \subseteq \{1, 2, \dots, N\}$. We define a map $\eta : \mathcal{A} \to \mathcal{A}$ by

$$\eta(A) = \bigcup_{\substack{\ell=1 \\ S(\ell)\cap A \neq \varnothing}}^{L} S(\ell). \tag{4-5}$$

Then it follows from (4-3) that

$$A \subseteq \eta(A) \quad \text{for each subset } A \in \mathcal{A}. \tag{4-6}$$

We are interested in subsets A in $\mathcal{A}$ that satisfy $\eta(A) = A$. Obviously $\varnothing$ and $\{1, 2, \dots, N\}$ have this property. More generally we define

$$\mathcal{P} = \{A \in \mathcal{A} : \eta(A) = A\}. \tag{4-7}$$

If A belongs to the collection $\mathcal{P}$ and $S(\ell) \cap A \neq \varnothing$, then $S(\ell) \subseteq A$. Thus a nonempty subset A in $\mathcal{P}$ must have $2 \le |A|$. We show that the collection $\mathcal{P}$ forms an algebra of subsets.

Lemma 4.1. *Let $\mathcal{P} \subseteq \mathcal{A}$ be the collection of subsets defined by (4-7).*

(i) *If A_1 belongs to $\mathcal{P}$ then its complement*

$$A_2 = \{1, 2, \dots, N\} \setminus A_1$$

also belongs to $\mathcal{P}$.

(ii) *If* A_3 *and* A_4 *belong to* $\mathcal{P}$ *then* $A_3 \cup A_4$ *belongs to* $\mathcal{P}$.

(iii) *If* A_5 *and* A_6 *belong to* $\mathcal{P}$ *then* $A_5 \cap A_6$ *belongs to* $\mathcal{P}$.

Proof. Assume that $S(\ell) \cap A_2 \neq \varnothing$. Then $S(\ell) \cap A_1 \neq \varnothing$ is impossible. Hence we have $S(\ell) \subseteq A_2$, and this implies that A_2 belongs to $\mathcal{P}$.

Let $S(\ell) \cap (A_3 \cup A_4) \neq \varnothing$. Then either $S(\ell) \cap A_3 \neq \varnothing$ or $S(\ell) \cap A_4 \neq \varnothing$. Hence either $S(\ell) \subseteq A_3$ or $S(\ell) \subseteq A_4$, and therefore $S(\ell) \subseteq A_3 \cup A_4$. It follows that $A_3 \cup A_4$ belongs to $\mathcal{P}$.

By what we have already proved the sets

$$A_7 = \{1, 2, \dots, N\} \setminus A_5 \quad \text{and} \quad A_8 = \{1, 2, \dots, N\} \setminus A_6$$

both belong to $\mathcal{P}$, and therefore the set

$$A_5 \cap A_6 = \{1, 2, \dots, N\} \setminus (A_7 \cup A_8)$$

belongs to $\mathcal{P}$. □

Lemma 4.2. *Let* A_1 *be a nonempty subset in* $\mathcal{A}$, *and let* B *be a subset in* $\mathcal{P}$. *Assume that* $A_1 \subseteq B$. *Define an increasing sequence of subsets*

$$A_1, A_2, A_3, \dots$$

from $\mathcal{A}$ *inductively by*

$$A_{n+1} = \eta(A_n) \quad \textit{for each } n = 1, 2, 3, \dots.$$

Then

$$A_n \subseteq B \quad \textit{for each } n = 1, 2, 3, \dots.$$

Proof. We argue by induction on n. If $n = 1$ then $A_1 \subseteq B$ by hypothesis. Now assume that $2 \leq n$ and $A_{n-1} \subseteq B$. Then we have

$$A_n = \eta(A_{n-1}) = \bigcup_{\substack{\ell=1 \\ S(\ell) \cap A_{n-1} \neq \varnothing}}^{L} S(\ell). \tag{4-8}$$

If $S(\ell) \cap A_{n-1} \neq \varnothing$ then $S(\ell)$ contains a point of B, and therefore $S(\ell) \subseteq B$. It follows from (4-8) that $A_n \subseteq B$. This proves the lemma. □

We say that a subset A in $\mathcal{A}$ is *minimal* if A is not empty and belongs to $\mathcal{P}$, but no proper subset of A belongs to $\mathcal{P}$. That is, a nonempty set A in $\mathcal{P}$ is *minimal* if for every nonempty subset $B \subseteq A$ such that $B \neq A$, we have $\eta(B) \neq B$. We will show that each element of $\{1, 2, \dots, N\}$ is contained in a minimal subset in $\mathcal{P}$.

Lemma 4.3. *Let* A_1 *in* $\mathcal{A}$ *have cardinality* 1. *Define an increasing sequence of subsets*

$$A_1, A_2, A_3, \dots$$

from $\mathcal{A}$ *inductively by*

$$A_{n+1} = \eta(A_n) \quad \textit{for } n = 1, 2, 3, \dots. \tag{4-9}$$

Let K be the smallest positive integer such that

$$A_K = \eta(A_K) = A_{K+1}. \tag{4-10}$$

Then K exists, $2 \le K$, and the subset A_K is minimal.

Proof. From (4-6) we get

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots \subseteq A_n \subseteq \cdots.$$

As $|A_n| \le N$ for each $n = 1, 2, \ldots,$ it is obvious that K exists.

Let $A_1 = \{k_1\}$ where $1 \le k_1 \le N$. It follows from (4-3) that there exists a subset $S(\ell_1)$ that contains k_1. Write $S(\ell_1) = \{k_1, k_2\}$ where $k_1 \ne k_2$. From (4-5) we conclude that

$$S(\ell_1) = \{k_1, k_2\} \subseteq \eta(A_1) = A_2,$$

and therefore $A_1 = \{k_1\}$ is a proper subset of $\eta(A_1) = A_2$. Hence we have $2 \le K$.

If A_K is not minimal there exists a proper subset $B \subseteq A_K$ such that $\eta(B) = B$, and therefore B belongs to $\mathcal{P}$. Let

$$C = A_K \setminus B = A_K \cap (\{1, 2, \ldots, N\} \setminus B) \tag{4-11}$$

be the complement of B in A_K. It follows from Lemma 4.1, and the representation on the right of (4-11), that C is a proper subset of A_K and C belongs to $\mathcal{P}$. Thus we have the disjoint union of proper subsets

$$A_K = B \cup C, \quad \text{where } B \in \mathcal{P} \text{ and } C \in \mathcal{P}. \tag{4-12}$$

Plainly $A_1 = \{k_1\}$ is a subset of either B or C, and by renaming these sets if necessary we may assume that $A_1 = \{k_1\}$ is contained in B. Then it follows from Lemma 4.2 that

$$A_n \subseteq B \quad \text{for each } n = 1, 2, 3, \ldots.$$

But this is inconsistent with the representation of A_K as the disjoint union (4-12). We conclude that B and C do not exist, and therefore A_K is minimal. □

It follows from Lemma 4.3 that each element of $\{1, 2, \ldots, N\}$ is contained in a minimal subset. This minimal subset is unique, and leads to a partition of $\{1, 2, \ldots, N\}$ into a disjoint union of minimal subsets.

Lemma 4.4. *Let B and C be nonempty, minimal subsets in $\mathcal{P}$. Then either*

$$B = C \quad \textit{or} \quad B \cap C = \varnothing.$$

Proof. If $B \cap C = \varnothing$ we are done. Therefore we assume that k_1 is a point in $B \cap C$. Let $A_1 = \{k_1\}$, and let $A_1, A_2, A_3, \ldots$ be the sequence of subsets defined by (4-9). Let K be the smallest positive integer such that (4-10) holds. By Lemma 4.3 the subset A_K is minimal, and by Lemma 4.2 we have both $A_K \subseteq B$ and $A_K \subseteq C$. But A_K is minimal and therefore A_K cannot be a proper subset of the minimal subset B. Similarly, A_K cannot be a proper subset of the minimal subset C. We conclude that

$$B = A_K = C.$$

□

Lemma 4.5. *Let (4-1) be a collection of distinct subsets of* $\{1, 2, \dots, N\}$ *such that*

$$|S(\ell)| = 2 \quad \textit{for each } \ell = 1, 2, \dots, L$$

and

$$\bigcup_{\ell=1}^{L} S(\ell) = \{1, 2, \dots, N\}.$$

Let $\mathcal{P} \subseteq \mathcal{A}$ *be the collection of subsets of* $\{1, 2, \dots, N\}$ *defined by (4-7), and let* $A_1, A_2, \dots, A_r$ *be the collection of all distinct, minimal subsets in* $\mathcal{P}$. *Then the subsets* $A_1, A_2, \dots, A_r$ *are disjoint and*

$$A_1 \cup A_2 \cup \cdots \cup A_r = \{1, 2, \dots, N\}.$$

Proof. The subsets $A_1, A_2, \dots, A_r$ exist by Lemma 4.3. Then it follows from Lemma 4.4 that the subsets $A_1, A_2, \dots, A_r$ are disjoint. Therefore we get

$$A_1 \cup A_2 \cup \cdots \cup A_r \subseteq \{1, 2, \dots, N\}. \tag{4-13}$$

It follows from Lemma 4.3 that each point in $\{1, 2, \dots, N\}$ is contained in a minimal subset, hence there is equality in (4-13). □

We continue to assume that L and N are positive integers that satisfy (4-4). Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$ be vectors from the set of extreme points F_N, and write

$$\Xi = (\boldsymbol{\xi}_1 \ \ \boldsymbol{\xi}_2 \ \ \cdots \ \ \boldsymbol{\xi}_L)$$

for the $N \times L$ matrix having $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_L$ as columns. We assume that no row of the matrix Ξ is identically zero, and we assume that rank $\Xi = L$. We write $\boldsymbol{\xi}_\ell = (\xi_{n\ell})$ and use the vectors $\boldsymbol{\xi}_\ell$ to define a collection of subsets

$$S(\ell) \subseteq \{1, 2, \dots, N\} \quad \text{for each } \ell = 1, 2, \dots, L. \tag{4-14}$$

More precisely, we define

$$S(\ell) = \{n : 1 \le n \le N \text{ and } \xi_{n\ell} \ne 0\} \quad \text{for each } \ell = 1, 2, \dots, L. \tag{4-15}$$

As each column vector $\boldsymbol{\xi}_\ell$ belongs to the set of extreme points F_N, it follows that each subset $S(\ell)$ has cardinality 2 and

$$\sum_{n=1}^{N} \xi_{n\ell} = \sum_{n \in S(\ell)} \xi_{n\ell} = 0.$$

Because no row of the matrix Ξ is identically zero, we find that

$$\bigcup_{\ell=1}^{L} S(\ell) = \{1, 2, \dots, N\}.$$

Therefore the subsets $S(\ell)$ defined by (4-15) satisfy the conditions (4-2) and (4-3) that were assumed in the previous lemmas. We continue to write $\mathcal{A}$ for the collection of all subsets of $\{1, 2, \dots, N\}$, and we write $\mathcal{P}$ for the collection of subsets defined by (4-7).

Next we suppose that $A_1, A_2, \dots, A_r$ is the collection of distinct, nonempty, minimal subsets in $\mathcal{P}$. Then it follows from Lemma 4.5 that

$$A_1 \cup A_2 \cup \cdots \cup A_r = \{1, 2, \dots, N\} \tag{4-16}$$

is a disjoint union of nonempty sets. Because each subset A_j is minimal we have

$$A_j = \bigcup_{\substack{\ell=1 \\ S(\ell)\subseteq A_j}}^{L} S(\ell) = \bigcup_{\substack{\ell=1 \\ S(\ell)\cap A_j \neq \varnothing}}^{L} S(\ell). \tag{4-17}$$

We use each subset A_j to define a subset $D_j \subseteq \{1, 2, \dots, L\}$ by

$$D_j = \{\ell : 1 \le \ell \le L \text{ and } S(\ell) \subseteq A_j\} \quad \text{for } j = 1, 2, \dots, r. \tag{4-18}$$

Then it follows from (4-16), (4-17), and (4-18), that

$$D_1 \cup D_2 \cup \cdots \cup D_r = \{1, 2, \dots, L\} \tag{4-19}$$

is a disjoint union of nonempty sets. For each $j = 1, 2, \dots, r$ we write Y_j for the $N \times |D_j|$ submatrix of Ξ having columns indexed by the integers in D_j. That is, we define

$$Y_j = (\boldsymbol{\xi}_\ell), \quad \text{where } \ell \in D_j \text{ indexes columns.} \tag{4-20}$$

We assemble the matrices $Y_1, Y_2, \dots, Y_r$ as $N \times |D_j|$ blocks so as to define the $N \times L$ matrix

$$Z = (Y_1 \;\; Y_2 \;\; \cdots \;\; Y_r). \tag{4-21}$$

Because of the disjoint union (4-19), the columns of the matrix Z can also be obtained by permuting the columns of the matrix Ξ. That is, there exists an $L \times L$ permutation matrix P such that

$$\Xi = ZP.$$

As $\det P = \pm 1$ and the columns of Ξ are linearly independent, it follows that the matrix Y_j has rank $|D_j|$ for each $j = 1, 2, \dots, r$. We also find that

$$\det(\Xi^T \Xi) = \det(P^T Z^T Z P) = \det(Z^T Z)$$

is a positive integer.

Now suppose that $1 \le i \le r$, that $1 \le j \le r$, and $i \ne j$. It follows from (4-14), (4-18), and (4-19), that each nonzero row of the matrix Y_i is indexed by an integer in the set A_i, and each nonzero row of the matrix Y_j is indexed by an integer in the set A_j. As A_i and A_j are disjoint we conclude that $Y_i^T Y_j$ is a zero matrix. Because we have organized Z into blocks as in (4-21), we find that

$$\det(\Xi^T \Xi) = \det(Z^T Z) = \prod_{j=1}^{r} \det(Y_j^T Y_j). \tag{4-22}$$

Since the extreme points $\boldsymbol{\xi}_l$ that form the columns of Ξ belong to F_N, it follows that

$$\sum_{n=1}^{N} \xi_{n\ell} = 0 \quad \text{for each } \ell = 1, 2, \ldots, L.$$

For each $j = 1, 2, \ldots, r$ the nonzero rows of Y_j are indexed by the elements of A_j, and so we get

$$\sum_{n \in A_j} \xi_{n\ell} = 0 \quad \text{for each } \ell \in D_j. \tag{4-23}$$

As Y_j has rank $|D_j|$ we find that

$$|D_j| + 1 \le |A_j|. \tag{4-24}$$

Next we will show that there is equality in the inequality (4-24). Each subset A_j is minimal in $\mathcal{P}$ and therefore no proper subset of A_j belongs to $\mathcal{P}$. It follows from (4-23) that the $|A_j|$ distinct (row) vectors

$$\{(\xi_{n\ell}) : n \in A_j\} \tag{4-25}$$

are linearly dependent. Let $f : A_j \to \mathbb{Z}$ be a function that is supported on the subset

$$B = \{n \in A_j : f(n) \neq 0\},$$

where B is a proper subset of A_j. As B does not belong to $\mathcal{P}$ it follows that there exists ℓ_1 in D_j such that

$$|S(\ell_1) \cap B| = 1.$$

We conclude that

$$\sum_{n \in A_j} f(n)\xi_{n\ell_1} = \sum_{n \in B} f(n)\xi_{n\ell_1} \neq 0,$$

because this sum contains exactly one nonzero term. This shows that no proper subset of the collection of (row) vectors (4-25) is linearly dependent. In particular, each subset of the (row) vectors in (4-25) with cardinality $|A_j| - 1$ is linearly independent. As the rank of the matrix Y_j is $|D_j|$ we conclude by (4-24) that

$$|D_j| + 1 = |A_j| \quad \text{for each } j = 1, 2, \ldots, r. \tag{4-26}$$

We also get the identity

$$L + r = \sum_{j=1}^{r} (|D_j| + 1) = \sum_{j=1}^{r} |A_j| = N, \tag{4-27}$$

which determines the value of r.

Lemma 4.6. *Let the columns of the $N \times L$ matrix*

$$\Xi = (\xi_1 \ \ \xi_2 \ \ \cdots \ \ \xi_L)$$

be vectors from the set of extreme points F_N defined in (3-6). If $L < N \le 2L$ then

$$\|\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_L\|_1 \le \left(\frac{N}{N-L}\right)^{N-L}. \tag{4-28}$$

Proof. Clearly we may assume that rank $\Xi = L$. We assume to begin with that no row of the matrix Ξ is identically zero. As in our proof of Lemma 3.6 we have

$$\|\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_L\|_1 = \sum_{\substack{I \subseteq \{1,2,\ldots,N\} \\ |I|=L}} (\det \Xi_I)^2 = \det(\Xi^T \Xi) \tag{4-29}$$

by the Cauchy–Binet identity. By combining (4-22) and (4-29) we find that

$$\|\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_L\|_1 = \prod_{j=1}^{r} \det(Y_j^T Y_j),$$

where each $N \times |D_j|$ matrix Y_j is defined as in (4-20). Let W_j be the $|A_j| \times |D_j|$ submatrix of Y_j obtained by removing all rows which are identically zero. Because there is equality in the inequality (4-24) the submatrix W_j is also $(|D_j|+1) \times |D_j|$. That is, W_j is an $(M+1) \times M$ matrix with columns in the set of extreme points F_M, where $M = |D_j|$. Then it follows from the inequality (3-23) and (4-26) that

$$\prod_{j=1}^{r} \det(Y_j^T Y_j) = \prod_{j=1}^{r} \det(W_j^T W_j) \le \prod_{j=1}^{r} (|D_j| + 1) = \prod_{j=1}^{r} |A_j|. \tag{4-30}$$

We estimate the product on the right of (4-30) by applying the arithmetic/geometric mean inequality and using the identity (4-27). In this way we arrive at the inequality

$$\prod_{j=1}^{r} \det(Y_j^T Y_j) \le \left(r^{-1} \sum_{j=1}^{r} |A_j|\right)^r = (r^{-1} N)^r = \left(\frac{N}{N-L}\right)^{N-L}.$$

This proves (4-28) under the assumption that no row of Ξ is identically zero.

Next we suppose that $L < N \le 2L$, that

$$\Xi = (\xi_1 \ \ \xi_2 \ \ \cdots \ \ \xi_L)$$

is an $N \times L$ matrix with columns $\xi_1, \xi_2, \ldots, \xi_L$ from F_N, that rank $\Xi = L$, and that Ξ has exactly $N - M > 0$ rows that are identically zero. Because rank $\Xi = L$, we find that $L \le M < N \le 2L$. We write

$$\Xi' = (\xi'_1 \ \ \xi'_2 \ \ \cdots \ \ \xi'_L)$$

for the $M \times L$ matrix obtained from Ξ by removing the rows of Ξ that are identically zero. It follows from Lemma 3.4 that each column $\xi'_1, \xi'_2, \ldots, \xi'_L$ belongs to F_M. Clearly each $L \times L$ submatrix of Ξ

with a row that is identically zero has a zero determinant. Thus we have

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 = \|\boldsymbol{\xi}'_1 \wedge \boldsymbol{\xi}'_2 \wedge \cdots \wedge \boldsymbol{\xi}'_L\|_1.$$

If $L = M$ then Ξ' is $L \times L$, and it follows from Lemma 3.5 that

$$\|\boldsymbol{\xi}'_1 \wedge \boldsymbol{\xi}'_2 \wedge \cdots \wedge \boldsymbol{\xi}'_L\|_1 = 1 \le \left(\frac{N}{N-L}\right)^{N-L}.$$

If $L < M < N \le 2L$ then by the case already considered we get

$$\|\boldsymbol{\xi}'_1 \wedge \boldsymbol{\xi}'_2 \wedge \cdots \wedge \boldsymbol{\xi}'_L\|_1 \le \left(\frac{M}{M-L}\right)^{M-L} < \left(\frac{N}{N-L}\right)^{N-L}.$$

This verifies the bound (4-28) in general. □

We now combine Lemma 3.7 and Lemma 4.6 to obtain the inequality (4-28) in full generality.

Theorem 4.7. *Let the columns of the $N \times L$ matrix*

$$\Xi = (\boldsymbol{\xi}_1 \;\; \boldsymbol{\xi}_2 \;\; \cdots \;\; \boldsymbol{\xi}_L)$$

be vectors in the set of extreme points $E_N \cup F_N$. If $L < N \le 2L$ then

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 \le \left(\frac{N}{N-L}\right)^{N-L}. \tag{4-31}$$

Proof. We argue by induction on the positive integer L. If $L = 1$ then $N = 2$ and the result is trivial to check. Next we assume that $2 \le L$, and we assume that (4-31) holds for all pairs (L', N') such that $L' < N' \le 2L'$ and $1 \le L' < L$.

If the extreme points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ all belong to the set of extreme points E_N, then

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 = 1$$

and the inequality (4-31) is trivial. If the extreme points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ all belong to the set of extreme points F_N, then the inequality (4-31) follows from Lemma 4.6. To complete the proof we assume that K of the extreme points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ belong to E_N and $L-K$ extreme points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L$ belong to F_N, where $1 \le K < L$. In this case the set of extreme points satisfies the hypotheses of Lemma 3.7. It follows from the conclusion of Lemma 3.7 that there exist linearly independent extreme points $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_{L-K}$ in the set $E_{N-K} \cup F_{N-K}$ such that

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 = \|\boldsymbol{\eta}_1 \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_{L-K}\|_1. \tag{4-32}$$

We write $L' = L - K$, $N' = N - K$, and we consider two cases. First we suppose that

$$N' \le 2L'.$$

In this case we apply the inductive hypothesis and conclude that

$$\|\boldsymbol{\eta}_1 \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_{L-K}\|_1 \le \left(\frac{N'}{N'-L'}\right)^{N'-L'} = \left(\frac{N-K}{N-L}\right)^{N-L} < \left(\frac{N}{N-L}\right)^{N-L}. \tag{4-33}$$

Next we suppose that

$$2L' \le N'.$$

In this case we appeal to the inequality (3-27) which we have already proved. By that result we have

$$\begin{aligned}\|\boldsymbol{\eta}_1 \wedge \boldsymbol{\eta}_2 \wedge \cdots \wedge \boldsymbol{\eta}_{L-K}\|_1 \le 2^{L'} &= \min\left\{2^{L'}, \left(\frac{N'}{N'-L'}\right)^{N'-L'}\right\} \\ &\le \left(\frac{N'}{N'-L'}\right)^{N'-L'} = \left(\frac{N-K}{N-L}\right)^{N-L} < \left(\frac{N}{N-L}\right)^{N-L}.\end{aligned} \tag{4-34}$$

Combining (4-32), (4-33), and (4-34), establishes the inequality

$$\|\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\|_1 \le \left(\frac{N}{N-L}\right)^{N-L}$$

whenever $L < N \le 2L$. This proves the lemma. □

If $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L$ belong to $\mathbb{R}^N$ and $L < N \le 2L$, then it follows from (4-31) that

$$\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1 \le \left(\frac{N}{N-L}\right)^{N-L} \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\cdots\delta(\boldsymbol{x}_L).$$

This proves the inequality (3-2), and so completes the proof of Theorem 3.1.

5. Proof of Theorem 1.1

We apply Theorem 3.1 with $N = r+1$ and $L = q$, and we apply the theorem to the collection of linearly independent points $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_q$ in

$$\Gamma_S(k) \subseteq \mathcal{D}_r \subseteq \mathbb{R}^{r+1}.$$

From (3-4) we find that

$$\begin{aligned}\|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1 &\le \min\left\{2^q, \left(\frac{r+1}{r+1-q}\right)^{r+1-q}\right\} \delta(\boldsymbol{\alpha}_1)\delta(\boldsymbol{\alpha}_2)\cdots\delta(\boldsymbol{\alpha}_q) \\ &= C(r,q)\,\delta(\boldsymbol{\alpha}_1)\delta(\boldsymbol{\alpha}_2)\cdots\delta(\boldsymbol{\alpha}_q).\end{aligned} \tag{5-1}$$

By the product formula the points $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_q$ belong to the diagonal subspace $\mathcal{D}_r$. Therefore we get

$$\delta(\boldsymbol{\alpha}_j) = \tfrac{1}{2}\|\boldsymbol{\alpha}_j\|_1 \quad \text{for each } j = 1, 2, \ldots, q. \tag{5-2}$$

Combining (5-1) and (5-2) establishes the inequality (1-11).

6. Proof of Theorem 1.2

Let $1 \le L < N$ and let

$$X = (\boldsymbol{x}_1 \ \ \boldsymbol{x}_2 \ \ \cdots \ \ \boldsymbol{x}_L)$$

be an $N \times L$ real matrix with columns $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L$. We assume that the columns of X are $\mathbb{R}$-linearly independent so that rank $X = L$ and

$$\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L \neq \boldsymbol{0}.$$

We use the matrix X to define a norm on $\mathbb{R}^L$ by

$$\boldsymbol{y} \mapsto \|X\boldsymbol{y}\|_1. \tag{6-1}$$

The unit ball associated to the norm (6-1) is obviously the set

$$B_X = \{\boldsymbol{y} \in \mathbb{R}^L : \|X\boldsymbol{y}\|_1 \le 1\}.$$

It is not difficult to show that the dual unit ball is

$$B_X^* = \{X^T\boldsymbol{w} : \boldsymbol{w} \in \mathbb{R}^N \text{ and } \|\boldsymbol{w}\|_\infty \le 1\}.$$

It can be shown (see [Bolker 1969; Schneider and Weil 1983] or, for a more general result, [Vaaler 2014, Lemma 2]) that the dual unit ball B_X^* is an example of a zonoid. Therefore by an inequality of S. Reisner [1985, Theorem 2], we have

$$\frac{4^L}{L!} \le \mathrm{Vol}_L(B_X)\,\mathrm{Vol}_L(B_X^*). \tag{6-2}$$

An identity for the L-dimensional volume of B_X^* was established by P. McMullen [1984] and C. G. Shephard [1974, equation (57)]. These results assert that

$$\mathrm{Vol}_L(B_X^*) = 2^L \sum_{|I|=L} |\det X_I| = 2^L \|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1. \tag{6-3}$$

By combining Reisner's inequality (6-2) and the volume formula (6-3), we obtain the lower bound

$$\frac{2^L}{L!} \le \mathrm{Vol}_L(B_X)\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1. \tag{6-4}$$

Now let

$$0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_L < \infty$$

be the successive minima for the convex symmetric set B_X and the integer lattice $\mathbb{Z}^L$. By Minkowski's theorem on successive minima (see [Cassels 1959, Section VIII.4.3]) we have

$$\mathrm{Vol}_L(B_X)\lambda_1\lambda_2\cdots\lambda_L \le 2^L. \tag{6-5}$$

We combine the lower bound (6-4) and the upper bound (6-5), and obtain the inequality

$$\lambda_1\lambda_2\cdots\lambda_L \le L!\,\|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1. \tag{6-6}$$

This leads to the following general result.

Theorem 6.1. *Let $\mathcal{X} \subseteq \mathbb{R}^N$ be the free group of rank L generated by the linearly independent vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L$. Then there exist linearly independent points $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_L$ in $\mathcal{X}$ such that*

$$\|\boldsymbol{y}_1\|_1 \|\boldsymbol{y}_2\|_1 \cdots \|\boldsymbol{y}_L\|_1 \leq L! \, \|\boldsymbol{x}_1 \wedge \boldsymbol{x}_2 \wedge \cdots \wedge \boldsymbol{x}_L\|_1. \tag{6-7}$$

If $\mathcal{Y} \subseteq \mathcal{X}$ is the subgroup generated by the points $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_L$, then $[\mathcal{X} : \mathcal{Y}] \leq L!$.

Proof. By Minkowski's theorem on successive minima there exist linearly independent points $\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots, \boldsymbol{m}_L$ in the integer lattice $\mathbb{Z}^L$ such that

$$\|X\boldsymbol{m}_\ell\|_1 = \lambda_\ell \quad \text{for } \ell = 1, 2, \ldots, L. \tag{6-8}$$

As rank $X = L$ the points

$$\{X\boldsymbol{m}_\ell : \ell = 1, 2, \ldots, L\}$$

are linearly independent points in the free abelian group $\mathcal{X}$. We write $\boldsymbol{y}_\ell = X\boldsymbol{m}_\ell$ for each $\ell = 1, 2, \ldots, L$. Then (6-7) follows from (6-6) and (6-8). The bound $[\mathcal{X} : \mathcal{Y}] \leq L!$ also follows from Minkowski's theorem. □

Now let $L = q$, $N = r+1$ and let $\mathfrak{A} \subseteq \mathbb{R}^{r+1}$ be the subgroup of rank q generated by the linearly independent vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_q$. By Theorem 6.1 there exist linearly independent vectors $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_q$ in $\mathfrak{A}$ such that

$$\|\boldsymbol{\beta}_1\|_1 \|\boldsymbol{\beta}_2\|_1 \cdots \|\boldsymbol{\beta}_q\|_1 \leq q! \, \|\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge \cdots \wedge \boldsymbol{\alpha}_q\|_1.$$

Moreover, the free group $\mathfrak{B} \subseteq \mathfrak{A}$ generated by the vectors $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_q$ has rank q and index

$$[\mathfrak{A} : \mathfrak{B}] \leq q!.$$

This proves Theorem 1.2.

References

[Akhtari and Vaaler 2016] S. Akhtari and J. D. Vaaler, "Heights, regulators and Schinzel's determinant inequality", *Acta Arith.* **172**:3 (2016), 285–298. MR Zbl

[Akhtari and Vaaler 2022] S. Akhtari and J. D. Vaaler, "Independent relative units of low height", *Acta Arith.* **202**:4 (2022), 389–401. MR Zbl

[Amoroso and David 1999] F. Amoroso and S. David, "Le problème de Lehmer en dimension supérieure", *J. Reine Angew. Math.* **513** (1999), 145–179. MR Zbl

[Amoroso and David 2021] F. Amoroso and S. David, "Covolumes, unités, régulateur: conjectures de D. Bertrand et F. Rodriguez-Villegas", *Ann. Math. Qué.* **45**:1 (2021), 1–18. MR Zbl

[Bertrand 1997] D. Bertrand, "Duality on tori and multiplicative dependence relations", *J. Austral. Math. Soc. Ser. A* **62**:2 (1997), 198–216. MR Zbl

[Bolker 1969] E. D. Bolker, "A class of convex bodies", *Trans. Amer. Math. Soc.* **145** (1969), 323–345. MR Zbl

[Bombieri and Gubler 2006] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Math. Monogr. **4**, Cambridge Univ. Press, 2006. MR Zbl

[Cassels 1959] J. W. S. Cassels, *An introduction to the geometry of numbers*, Grundl. Math. Wissen. **99**, Springer, 1959. MR Zbl

[Chinburg et al. 2022] T. Chinburg, E. Friedman, and J. Sundstrom, "On Bertrand's and Rodriguez Villegas' higher-dimensional Lehmer conjecture", *Pacific J. Math.* **321**:1 (2022), 119–165. MR Zbl

[Eggleston 1958] H. G. Eggleston, *Convexity*, Cambridge Tracts in Math. Math. Phys. **47**, Cambridge Univ. Press, 1958. MR Zbl

[Lehmer 1933] D. H. Lehmer, "Factorization of certain cyclotomic functions", *Ann. of Math.* (2) **34**:3 (1933), 461–479. MR Zbl

[McMullen 1984] P. McMullen, "Volumes of projections of unit cubes", *Bull. Lond. Math. Soc.* **16**:3 (1984), 278–280. MR Zbl

[Narkiewicz 2004] W. Narkiewicz, *Elementary and analytic theory of algebraic numbers*, 3rd ed., Springer, 2004. MR Zbl

[Pohst 1978] M. Pohst, "Eine Regulatorabschätzung", *Abh. Math. Sem. Univ. Hamburg* **47** (1978), 95–106. MR Zbl

[Reisner 1985] S. Reisner, "Random polytopes and the volume-product of symmetric convex bodies", *Math. Scand.* **57**:2 (1985), 386–392. MR Zbl

[Schinzel 1973] A. Schinzel, "On the product of the conjugates outside the unit circle of an algebraic number", *Acta Arith.* **24** (1973), 385–399. MR Zbl

[Schinzel 1978] A. Schinzel, "An inequality for determinants with real entries", *Colloq. Math.* **38**:2 (1978), 319–321. MR Zbl

[Schneider and Weil 1983] R. Schneider and W. Weil, "Zonoids and related topics", pp. 296–317 in *Convexity and its applications* (Vienna, 1981/Siegen, Germany, 1982), edited by P. M. Gruber and J. M. Wills, Birkhäuser, Basel, 1983. MR Zbl

[Shephard 1974] G. C. Shephard, "Combinatorial properties of associated zonotopes", *Canadian J. Math.* **26** (1974), 302–321. MR Zbl

[Siegel 1969] C. L. Siegel, "Abschätzung von Einheiten", *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1969** (1969), 71–86. MR Zbl

[Smyth 2008] C. Smyth, "The Mahler measure of algebraic numbers: a survey", pp. 322–349 in *Number theory and polynomials* (Bristol, 2006), edited by J. McKee and C. Smyth, Lond. Math. Soc. Lect. Note Ser. **352**, Cambridge Univ. Press, 2008. MR Zbl

[Vaaler 2014] J. D. Vaaler, "Heights on groups and small multiplicative dependencies", *Trans. Amer. Math. Soc.* **366**:6 (2014), 3295–3323. MR Zbl

[Waldschmidt 2000] M. Waldschmidt, *Diophantine approximation on linear algebraic groups*, Grundl. Math. Wissen. **326**, Springer, 2000. MR Zbl

[Zimmert 1981] R. Zimmert, "Ideale kleiner Norm in Idealklassen und eine Regulatorabschätzung", *Invent. Math.* **62**:3 (1981), 367–380. MR Zbl

Communicated by Andrew Granville
Received 2022-01-22 Revised 2023-08-24 Accepted 2023-10-12

akhtari@psu.edu Department of Mathematics, Pennsylvania State University, University Park, PA, United States

vaaler@math.utexas.edu Department of Mathematics, University of Texas at Austin, Austin, TX, United States

ALGEBRA AND NUMBER THEORY 18:9 (2024)
https://doi.org/10.2140/ant.2024.18.1619

Prime values of $f(a,b^2)$ and $f(a,p^2)$, f quadratic

Stanley Yao Xiao

Dedicated to the occasion of John Friedlander's 80th birthday

We prove an asymptotic formula for primes of the shape $f(a, b^2)$ with a, b integers and of the shape $f(a, p^2)$ with p prime. Here f is a binary quadratic form with integer coefficients, irreducible over $\mathbb{Q}$ and has no local obstructions. This refines the seminal work of Friedlander and Iwaniec on primes of the form $x^2 + y^4$ and of Heath-Brown and Li on primes of the form $a^2 + p^4$, as well as earlier work of the author with Lam and Schindler on primes of the form $f(a, p)$ with f a positive definite form.

1. Introduction

Two of the most stunning results in prime number theory in the last thirty years are the seminal works of Friedlander and Iwaniec [5] and Heath-Brown [9], demonstrating that the polynomials $x^2 + y^4$ and $x^3 + 2y^3$, respectively, take on infinitely many prime values. In particular, Friedlander and Iwaniec obtained the beautiful asymptotic formula

$$\sum\sum_{a^2+b^4\leq X} \Lambda(a^2+b^4) = \frac{2\Gamma(1/4)^2}{3\pi\sqrt{2\pi}} X^{3/4}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right), \tag{1-1}$$

where $\Lambda(\cdot)$ is the von Mangoldt function and Γ is the Gamma function.

Heath-Brown's result on $x^3 + 2y^3$ was quickly generalized by Heath-Brown and Moroz in [11], which demonstrated that any admissible binary cubic form takes on infinitely many prime values. More recently, X. Li has proved that the cubic form $x^3 + 2y^3$ takes on infinitely many prime values with y restricted to a short interval [14]. One also notes the stunning work of J. Maynard on representation of primes by incomplete norm forms, a substantial generalization of Heath-Brown's work [15].

Despite the passage of more than two decades, a generalization akin to that of Heath-Brown and Moroz [11] has yet to materialize for the main result of [5], despite the authors of that paper claiming that such a result should be readily obtainable from their arguments.[1] That is, there has yet to be a proof that $f(x, y^2)$ takes on infinitely many prime values for any binary quadratic form f other than $f(x, y) = x^2 + y^2$.

More precisely, it can be seen from [9; 11] that the work needed to go from prime values of $x^3 + 2y^3$ to prime values of $F(x, y)$ for arbitrary admissible cubic forms F is purely algebraic. In particular, the

MSC2020: primary 11N32; secondary 11N35, 11N36, 11R45.
Keywords: primes, prime values of quadratic forms, Friedlander–Iwaniec theorem.

[1] "We expect, but did not check, that the methods carry over to the prime values of $\phi(a, b^2)$ for ϕ a quite general binary quadratic form." [5, p. 947].

analytic machinery established by Heath-Brown in [9] essentially only depends on the $\mathbb{Z}$-module structure of $\mathbb{Z}[\sqrt[3]{2}]$, which means it can be easily adapted to work with sets of ideal numbers.

Such is not the case with a^2+b^4. In fact the analytic machinery in [5] is far more delicate, as they needed to work around the lack of homogeneity of the polynomial. Much of this machinery is quite subtle. Therefore, in addition to establishing an appropriate algebraic framework akin to the work of Heath-Brown and Moroz, it is necessary to generalize some of the analytic machinery in [5] as well. For the algebraic framework, we use language established by Heath-Brown and Moroz, but in principle we can use the same type of explicit language used by Lam, Schindler, and the author in [12].

Fortuitously, we are able to salvage a significant portion of the analytic machinery established by Friedlander and Iwaniec [5] and Heath-Brown and Li [10]. There is one notable exception, which can be viewed as the most novel contribution of this paper: the so-called Jacobi–Kubota symbol. In fact this symbol, introduced in [5], works well *only* in the ring of Gaussian integers $\mathbb{Z}[i]$. This symbol is subtle because unlike much of the other pieces of analytic machinery, it relies also on the *arithmetic structure* of the sets of ideal numbers of the quadratic field associated to f. Obtaining a generalization of the multiplicativity of the Jacobi–Kubota symbol workable in the general setting is crucial to our arguments.

In another direction, one might ask whether *reducible polynomials* take on infinitely many *semiprime values*, with the order of the semiprime being equal to the number of irreducible factors. A first example of this type of result is due to Fouvry and Iwaniec [3], who showed that the binary cubic form $y(x^2+y^2)$ takes on infinitely many values with exactly two prime factors. This work paved the way for the later work of Friedlander and Iwaniec [5]. Heath-Brown and Li then combined the result of Fouvry and Iwaniec and Friedlander and Iwaniec in [10], showing that the polynomial $y(x^2+y^4)$ takes on infinitely many values with exactly two prime factors. In particular they obtained the asymptotic formula

$$\sum\sum_{a^2+b^4\le X} \lambda(b)\lambda(a^2+b^4) = \frac{2\Gamma(1/4)^2}{3\pi\sqrt{2\pi}}\frac{X^{3/4}}{(\log X)^2}\left(1+O_\varepsilon\left(\frac{1}{(\log X)^{1-\varepsilon}}\right)\right), \tag{1-2}$$

where λ is the prime indicator function.

Lam, Schindler and the author generalized the work of Fouvry and Iwaniec in another direction, proving that for any admissible positive definite binary quadratic form f the cubic form $yf(x,y)$ takes on infinitely many values with exactly two prime factors. Our main result implies

$$\sum\sum_{f(m,\ell)\le X} \Lambda(\ell)\Lambda(f(m,\ell)) = \nu_f \mathfrak{S}'_f X + O_A(X(\log X))^{-A}), \tag{1-3}$$

where ν_f is a product of local densities given by

$$\nu_f = \prod_{p\nmid\Delta(f)}\left(1-\frac{\rho_f(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-1}\prod_{p\mid\Delta(f)}\left(1-\frac{1}{p}\right)^{-1}, \tag{1-4}$$

$\mathfrak{S}'_f$ is given by (1-9), and $\rho_f(m) = \#\{x \pmod m : f(x,1)\equiv 0 \pmod m\}$.

We simultaneously generalize the main results of Friedlander and Iwaniec [5] and Heath-Brown and Li [10]. If f is definite put

$$\mathfrak{S}_f = \text{Area}\{(x, y) \in \mathbb{R}^2 : f(x, y^2) \le 1\}$$

and for f indefinite we define

$$\mathfrak{S}_f = \lim_{X\to\infty} \frac{\text{Area}\{(x, y) \in \mathbb{R}^2 : 0 < f(x, y^2) < X,\ 0 < y \le X^{1/4}\}}{X^{3/4}}.$$

Our first main result is:

Theorem 1.1. *Let $f(x, y) = f_2x^2 + f_1xy + f_0y^2 \in \mathbb{Z}[x, y]$ be an irreducible and primitive binary quadratic form, with the property that $f(x, 1) \not\equiv x(x+1) \pmod 2$. Then for f positive definite we have*

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ f(m,\ell^2)\le X}} \lambda(f(m, \ell^2)) = \frac{\nu_f \mathfrak{S}_f X^{3/4}}{\log X}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right) \tag{1-5}$$

and for f indefinite we have

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ 0< f(m,\ell^2)\le X\\ 0<\ell\le X^{1/4}}} \lambda(f(m, \ell^2)) = \frac{\nu_f \mathfrak{S}_f X^{3/4}}{\log X}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right). \tag{1-6}$$

The condition that $f(x, 1) \not\equiv x(x+1) \pmod 2$ is necessary, as otherwise $f(x, k)$ is divisible by 2 whenever k is odd, precluding the possibility that it could be a prime square unless $k = 4$. Theorem 1.1 recovers Theorem 1 of [5] upon setting $f(x, y) = x^2 + y^2$. It also implies, for example, that the polynomials $x^2 + xy^2 + y^4$ and $x^2 - 2y^4$ represent infinitely many primes.

The choice of cutting off the y-variable at $X^{1/4}$ is somewhat arbitrary, and is mostly done for aesthetic reasons. Of course, in the indefinite case some such cut-off is necessary. In particular such a choice guarantees that we do not need to worry about long cusps if we insist only on the condition $|f(x, y^2)| \le X$.

Both Theorems 1.1 and 1.2 apply to *indefinite* as well as definite forms. Although, for economy, we state the two cases together, there are some differences in the proof and, so far as we are aware, these give the first examples of asymptotic formulae for the number of prime values of indefinite nonhomogeneous polynomials of degree exceeding two.

Our proof, which further develops ideas in [10], yields the following general version of Theorem 1 of [10] or (1-2):

Theorem 1.2. *Let $f(x, y) = f_2x^2 + f_1xy + f_0y^2 \in \mathbb{Z}[x, y]$ be an irreducible and primitive binary quadratic form, with the property that $f(x, 1) \not\equiv x(x+1) \pmod 2$. Then for f positive definite we have*

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ 0< f(m,\ell^2)\le X}} \lambda(\ell)\lambda(f(m, \ell^2)) = \frac{\nu_f \mathfrak{S}_f X^{3/4}}{(\log X)^2}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right) \tag{1-7}$$

and for f indefinite we have

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ 0<f(m,\ell^2)\le X\\ 0<\ell\le X^{1/4}}} \lambda(\ell)\lambda(f(m,\ell^2)) = \frac{\nu_f \mathfrak{S}_f X^{3/4}}{(\log X)^2}\left(1+O\left(\frac{\log\log X}{\log X}\right)\right). \tag{1-8}$$

Theorem 1.2 implies that there are infinitely many integers x and primes p for which $f(x, p^2)$ is prime. We further note that the error term in Theorem 1.2 is slightly better than in (1-2), due to choosing a slightly different sieving parameter.

In [11], the key new insight is that the arithmetic of *ideal numbers* allows one to connect the multiplicative structure on the set of ideals of a ring of integers, which has unique factorization, to the arithmetic of the elements in a ring of integers which need not have unique factorization. This breaks a key barrier in [9] where the fact that $\mathbb{Z}[\sqrt[3]{2}]$ is a unique factorization domain is used in a crucial manner. The analytic estimates obtained by Heath-Brown in [9] can be applied with relatively few changes in the general setting [11].

In [12] we essentially pursued the same approach, although we did not state things in terms of ideal numbers but rather worked out an explicit composition law for binary quadratic forms, in the spirit of Gauss and Dirichlet. We have decided to adopt the approach of Heath-Brown and Moroz and use ideal numbers, as this is a more elegant and general approach.

In order to prove Theorems 1.1 and 1.2 we adopt an approach introduced by Heath-Brown in [9], which we call Heath-Brown's *comparison sieve*. This involves applying the same sieve procedure to two comparable sequences $\mathcal{A} = (a_n)$ and $\mathcal{B} = (b_n)$, producing cancellation at appropriate junctures. This was used again by Heath-Brown and Li in [10] for the proof of their result.

In order to prove Theorem 1.1 we choose our sequence $\mathcal{B}$ to simply be the set of prime ideals of the ring of integers $\mathcal{O}_K$, where $K = \mathbb{Q}(\sqrt{\Delta(f)})$ is the splitting field of our form f. The sequence $\mathcal{B}$ used by Heath-Brown and Li is exactly the sequence studied by Fouvry and Iwaniec in [3]. For positive definite forms f we may then apply the result in [12], and for indefinite forms we will need to prove an extension of our main result with Lam and Schindler in [12], which gives an asymptotic formula for the number of representation of primes by $f(x, p)$, with p prime.

For f positive definite put

$$\mathfrak{S}'_f = \text{Area}\{(x, y) \in \mathbb{R}^2 : f(x, y) \le 1\} \tag{1-9}$$

and for f indefinite put

$$\mathfrak{S}'_f = \lim_{X\to\infty} \frac{\text{Area}\{(x, y) \in \mathbb{R}^2 : 0 < f(x, y) < X,\ 0 < y < X^{1/2}\}}{X}.$$

Theorem 1.3. *Let $f(x, y) = f_2x^2 + f_1xy + f_0y^2 \in \mathbb{Z}[x, y]$ be an irreducible and primitive binary quadratic form, with the property that $f(x, 1) \not\equiv x(x+1) \pmod 2$. Then for f positive definite we have*

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ 0<f(m,\ell)\le X}} \Lambda(\ell)\Lambda(f(m,\ell)) = \nu_f \mathfrak{S}'_f X + O_A\left(\frac{X}{(\log X)^A}\right) \tag{1-10}$$

and for f indefinite we have

$$\sum_{\substack{m,\,\ell\in\mathbb{Z}\\ 0<f(m,\ell)\le X\\ 0<\ell\le X^{1/2}}} \Lambda(\ell)\Lambda(f(m,\ell)) = \nu_f \mathfrak{S}'_f X + O_A\left(\frac{X}{(\log X)^A}\right). \tag{1-11}$$

Here ν_f is as in Theorem 1.2 and $\mathfrak{S}'_f$ is as in (1-9).

Theorem 1.3 is stated with the von Mangoldt function rather than λ to emphasize that a substantially better error term, giving an arbitrary log-power saving, is possible.

Theorem 1.3 implies the following, which completely settles *Schinzel's hypothesis* for binary cubic forms:

Corollary 1.4. *Let $F(x, y)$ be a reducible binary cubic form of the shape $F(x, y) = L(x, y)Q(x, y)$, where Q is an irreducible binary quadratic form. Then there are infinitely many pairs of integers x, y such that $F(x, y)$ is divisible by exactly two primes.*

Corollary 1.4 is the final case of *Schinzel's hypothesis* in the setting of binary cubic forms. The hardest case, that of irreducible binary cubic forms, is settled by the work of Heath-Brown [9] and Heath-Brown and Moroz in [11]. The case with F reducible with a positive definite quadratic factor is settled by the author's joint work with Lam and Schindler in [12]. The special cases when the irreducible quadratic factor is x^2+y^2 was settled by Fouvry and Iwaniec in [3] and the special case when the quadratic factor is x^2+xy+y^2 was settled by M. Pandey [17]. The totally reducible case was settled by van der Corput [2], and was of course famously generalized by B. J. Green and T. Tao to cover arbitrarily long arithmetic progressions [8]; see also B. J. Green's work on Roth's theorem in the primes [7].

As in [11], our basic objective is to invoke composition laws involving ideal numbers of a fixed quadratic field in order to reduce the problem to one that is amenable to the analytic methods developed by Friedlander and Iwaniec in [5] and Heath-Brown and Li in [10]. In [11] the relevant analytic methods developed by Heath-Brown in [9] can be applied with only minor modifications once the relevant algebraic framework is established, since these estimates depend only on the $\mathbb{Z}$-module structure. However, the analytic estimates employed by Friedlander and Iwaniec in [5] are far more delicate and depend subtly on the fine arithmetic properties of the ring $\mathbb{Z}[i]$ rather than simply its structure as a rank-two $\mathbb{Z}$-module. Indeed, the obvious analogue of the so-called Jacobi–Kubota symbol introduced by Friedlander and Iwaniec in [5] does not seem to behave nicely and special care must be taken to define and work with the twisting factor $\xi_w(z)$ needed to recover multiplicativity. Again we emphasize that the definition and application of these generalized Jacobi–Kubota symbols and their twisting factors may be viewed as the most novel contribution of this paper. We give a rough explanation of this in the following section.

Organization of the paper. In Section 2 we give a brief overview of the ideas in this paper, emphasizing key new ingredients. In Section 3 we discuss our approach to implementing the asymptotic sieve for primes, in the manner introduced by Heath-Brown in [9] which we dub *Heath-Brown's comparison sieve*, also used by Heath-Brown and Moroz in [11] and Heath-Brown and Li in [10]. In Section 4 we introduce the necessary algebraic number theory involving the arithmetic ideal numbers, necessary to establish the framework needed to apply the analytic estimates in [5; 10]. In Section 5 we establish the needed level of distribution or type-I estimates. In Section 6 we will prove the necessary bilinear sum estimates to obtain the analogue of the main theorem of [12] in the indefinite case, which for us is needed to apply Heath-Brown's comparison sieve in the indefinite case. In Section 7 we establish the preliminary steps to proving our two key technical propositions, being Propositions 7.5 and 7.6, which are analogues of Heath-Brown and Li's Propositions 6 and 7 in [10]. In Section 8 we prove Proposition 7.5, the proof being identical to that of [10] except we avoid the language of Gaussian integers. In Sections 9 and 10 we modify Heath-Brown and Li's proof of their Proposition 7 in the setting of a general quadratic field K, thereby proving our Proposition 7.6, which then completes the proof of Theorem 1.2, conditioned on certain character sum estimates that they imported from [5]. Finally, in Section 11 we introduce the analogues of Friedlander and Iwaniec's notion of *Jacobi–Kubota symbols* in the setting of a general quadratic field, as well as the analogue of their symbol $[\,\cdot\,]$ which in some sense measures the "spin" of an ideal in $\mathbb{Z}[i]$, allowing us to prove versions of their Proposition 23.1 and Theorem ψ which are needed by Heath-Brown and Li. This may be of independent interest.

Notation. Throughout, we fix our binary quadratic form

$$f(x, y) = f_2x^2 + f_1xy + f_0y^2 \in \mathbb{Z}[x, y],$$

which satisfies the hypothesis that for all primes p there exist integers x_p, y_p such that $p \nmid f(x_p, y_p)$, and $f(x, 1) \not\equiv x(x+1) \pmod 2$. We will use both the Landau and Vinogradov notation $\ll$ and $O(\,\cdot\,)$.

2. Sketch of the main ideas

To sketch our ideas it is necessary to give a quick summary of the works of Friedlander–Iwaniec [5], Heath-Brown and Li [10], as well as the works of Heath-Brown [9] and Heath-Brown and Moroz [11]. In particular, we will see in this section that the paths to Theorems 1.1 and 1.2 are not as straightforward as going from [9] to [11].

We divide our arguments into ideas that are essentially algebraic, ideas which are essentially analytic in nature, and the final subsection is devoted to the new ingredient needed to tie these two bags of tools together to give the proof.

Analytic (sieve theoretic) ideas. In [5] the principal strategy is to verify the hypotheses of the asymptotic sieve for primes, introduced by Friedlander and Iwaniec in [4], hold for the sequence $\mathcal{A} = (a_n)$ defined by

$$a_n = \sum_{x^2+y^4=n} 1. \tag{2-1}$$

In order to do so they obtain an optimal level of distribution (or type-I estimates) for the sequence $\mathcal{A}$, which is extended and refined in their subsequent work [6]; see also [1]. The strength of their result relies on the remarkable property that roots of quadratic congruences are extraordinarily well-spaced modulo 1. However, the main obstacle they overcome in [5] is to obtain acceptable estimates for *bilinear sums* of the shape

$$\sum_{m} \alpha(m) \sum_{\substack{N \le n < 2N \\ mn \le X}} \beta(n) a_{mn}$$

for quite general complex sequences $(\alpha(m))$ and $(\beta(n))$.

To do so, Friedlander and Iwaniec converted the problem to one about estimating solutions to a family of quadratic congruences via Fourier analysis. They then succeeded in obtaining satisfactory estimates for the number of solutions after a herculean effort in [5]. This estimation constitutes the bulk of the work done in [5].

They partitioned their argument into estimating solutions with small, medium, or large moduli. As usual, the contribution from the small moduli is expected to be relatively straightforward since explicit asymptotic formulae are expected to exist. In [5] this was done from the ground up and in [10] this was obtained by applying the general Siegel–Walfisz type theorems of Mitsui [16].

One pillar of [5], which treats the lion's share of possible moduli in the middle, is their Proposition 14.1. There they cleverly used quadratic reciprocity to achieve *moduli flipping*, which allows one to swap a large modulus with a small one via complementary divisors as long as the modulus is not too large (so that its complementary divisor is not too small). This aspect of Friedlander–Iwaniec is imported without change in [10].

We state Proposition 14.1 in [5] here for convenience:

Proposition 2.1 [5, Proposition 14.1]. *Let D, R, $S \ge 1$. For any complex numbers α_{rs} with $\gcd(r, 2s) = 1$ supported in the box $R < r < 2R$, $S < s < 2S$ we have*

$$\sum_{D < d \le 2D} \sum_{a \,(\mathrm{mod}\, d)} \left| \sum\sum_{\bar{r}s \equiv a \,(\mathrm{mod}\, d)} \alpha_{rs} \left(\frac{r}{d}\right) \right|^2 \le \mathcal{N}(D, R, S) \sum_{r} \sum_{s} \tau(r) |\alpha_{rs}|^2,$$

where $\mathcal{N}(D, R, S)$ satisfies the bound

$$\mathcal{N}(D, R, S) \ll_\varepsilon D + D^{-1/2} RS + D^{1/3} (RS)^{2/3} (\log 2RS)^4 + (R+S)^{1/12} (RS)^{1/12+\varepsilon}$$

for any $\varepsilon > 0$. Here $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol if d is odd and is extended for d even via the Hilbert symbol.

It is important to emphasize that Proposition 14.1 in [5] only involves rational integers, and therefore only depends on the structure of $\mathbb{Z}^2$ as a $\mathbb{Z}$-module. The relevance to the present setting is that our particular quadratic field is of no concern when invoking this proposition.

The treatment of large moduli constitutes the bulk of the hard work in [5; 10]. Here the main obstacle to overcome is an acceptable estimate of a sum of the shape

$$\sum_{z_1, z_2} \beta'_{z_1} \beta'_{z_2} \tag{2-2}$$

where β' takes the form

$$\beta'_z = \beta_z i^{(x-1)/2}\left(\frac{y}{x}\right)$$

with $z = (x, y)$ and i is the imaginary unit. The function β_z is supported on a small region in $\mathbb{R}^2$. The key property needed here is that the sum (2-2) can be split as a product

$$\sum_{z_1, z_2} \beta'_{z_1} \beta'_{z_2} = \Bigl(\sum_{z_1} \beta'_{z_1}\Bigr)\Bigl(\sum_{z_2} \beta'_{z_2}\Bigr) \tag{2-3}$$

using properties of the Jacobi symbol and the arithmetic of $\mathbb{Z}[i]$. This will be discussed in detail later in Section 10.

We remark that two key results in [10], namely Corollaries 1 and 2 which are a refinement of the Barban–Davenport–Halberstam theorem and a Siegel–Walfisz type estimate, respectively, are not explicitly invoked here. This is because these two results are used in [10] to prove their Proposition 6 which, surprisingly, can be applied more or less without change in our case.

Algebraic ideas. The ideas in this subsection are introduced by Heath-Brown and Moroz [11], and are also related to the work of the author with Lam and Schindler [12].

The main role played by algebraic number theory in the present work is to obtain an analogue of equation (5.2) in [5], which we state here for convenience:

$$a_{mn} = \frac{1}{4} \sum_{|w|^2 = m} \sum_{|z|^2 = n} \mathfrak{Z}(\mathrm{Re}(\overline{w} z)),$$

where $\mathfrak{Z}$ is the indicator function for square integers. This crucial formula allows one to decompose the terms a_n given in (2-1) multiplicatively, which is the principal reason why such strong bilinear sum estimates can be obtained. In [5] they used the fact that the Gaussian integers $\mathbb{Z}[i]$ is a principal ideal domain, and more crucially has a *canonical basis*, in order to obtain their equation (5.2).

In general the quadratic order $\mathcal{O}_K$ we are working in is not a PID. Indeed if we are interested in general binary quadratic forms it is not enough to only work with $\mathcal{O}_K$ but rather *all* sets of ideal numbers simultaneously. If we denote by h the class number of $\mathcal{O}_K$ and $A_1, \dots, A_h$ the corresponding sets of ideal numbers, what we require is a choice of a basis of the A_j's as $\mathbb{Z}$-modules and a *composition law* connecting them, expressed in terms of the given bases.

These ideas are already expressed fully by Heath-Brown and Moroz in [11]. In [12] we adopted a more down-to-earth but ultimately more explicit approach. In the present paper we adapt the ideas in [11] instead.

The key algebraic result we will need is Proposition 4.1 which gives the analogue of equation (5.2) in [5]. This means that we again have sums which can be decomposed multiplicatively which enables us to obtain strong bilinear sum estimates.

The new input: a generalized Jacobi–Kubota symbol and its twist factor. So far we have discussed how Friedlander and Iwaniec relied on the fact that $\mathbb{Z}[i]$ is a PID in order to obtain their decomposition formula. Likewise, Heath-Brown had relied on the fact that $\mathbb{Z}[\sqrt[3]{2}]$ is a PID in order to obtain the decomposition formula he needed in [9]. The key new idea in [11] was to use the algebraic structure provided by Hecke's ideal numbers in order to reduce the problem of finding prime values of a binary cubic form F to estimating certain sums on $\mathbb{Z}$-modules. The latter constitutes the analytic portion of the argument. In essence, in [11] they successfully separated the algebraic and analytic arguments.

We are doing much the same, but there is one component of the arguments in [5] (and this is inherited by [10]) that is solidly wedged in between the algebraic and analytic worlds and requires separate treatment: the *Jacobi–Kubota symbol*.

The Jacobi–Kubota symbol as defined in [5] is essential in obtaining the decomposition property of (2-2) given as (2-3). Indeed, this is the lynchpin that holds together the arguments needed to obtain suitable estimates for the largest moduli in [5; 10].

The Jacobi–Kubota symbol is extremely specific to the Gaussian integers $\mathbb{Z}[i]$. In particular to obtain the nice properties derived in [5] one must use in an essential way the following properties of $\mathbb{Z}[i]$:

- the class number of $\mathbb{Z}[i]$ is 1;
- the norm of $\mathbb{Z}[i]$ is the same as the Euclidean norm on $\mathbb{C}$; and
- the odd rational primes that split in $\mathbb{Z}[i]$ are precisely those that are congruent to $1 \pmod 4$.

Clearly, no other ring of quadratic integers except for possibly suborders of $\mathbb{Z}[i]$ possess all three of these properties.

To see how these properties come into play, note that Friedlander and Iwaniec defined the Jacobi–Kubota symbol in [5, equation (20.1)] as

$$[z] = [r + is] = i^{(r-1)/2}\left(\frac{s}{|r|}\right)$$

where $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol. One sees right away that the choice of basis is relevant: the components s, r in the definition cannot make sense without a choice of basis. Right now it appears that the Jacobi–Kubota symbol depends only on the $\mathbb{Z}$-module structure of $\mathbb{Z}[i]$. This is indeed the case: once we have fixed a basis for our relevant $\mathbb{Z}$-module we can define the Jacobi–Kubota symbol analogously.

The trouble is that in order to have the desired multiplicative property, namely

$$[z][w] = \varepsilon[zw]\xi_w(z), \tag{2-4}$$

where $\varepsilon = \pm 1$ depending only on the quadrants containing z, w respectively, the "twist factor" $\xi_w(z)$ must satisfy nice properties that critically depend on the arithmetic structure of $\mathbb{Z}[i]$ as well as the niceness of the canonical basis. In particular, the twist factor $\xi_w(z)$ satisfies:

(1) It is multiplicative for each $w \in \mathbb{Z}[i]$: one has $\xi_w(z_1)\xi_w(z_2) = \xi_w(z_1 z_2)$.

(2) It is symmetric: $\xi_w(z) = \xi_z(w)$ for $w, z \in \mathbb{Z}[i]$.

(3) (Lemma 21.1 in [6]) For $q = |w_1 w_2|^2$ and $d = |\gcd(w_1, \overline{w_2})|^2$ one has

$$\sum_{\zeta \pmod q} \xi_{w_1}(\zeta)\xi_{w_2}(\zeta) = \begin{cases} q\varphi(d)\varphi(q/d) & \text{if } q, d \text{ are squares,} \\ 0 & \text{otherwise.} \end{cases}$$

(4) For $w = u + iv$ and $\omega \equiv -v\bar{u} \pmod q$ with $q = |w|^2$, one has

$$\xi_w(z) = \left(\frac{ur - vs}{q}\right) \quad \text{and} \quad \xi_w(z) = \left(\frac{r + \omega s}{q}\right),$$

where $z = r + is$.

In order for the Jacobi–Kubota symbol, defined analogously to [5], to have nice properties we must introduce an analogous twist factor to $\xi_w(z)$. One sees right away that this is a tall order. Simply writing down the definition will require much of the setup which will take place throughout the paper, so we defer this until Section 11.

3. Heath-Brown's comparison sieve

We describe the ideas given by Heath-Brown in [9] and expanded upon and refined in [11] and [10]. Heath-Brown's great insight is that quite often it is possible to establish the infinitude of primes in a sequence $\mathcal{A}$ by comparing it to a suitable sequence $\mathcal{B}$ known to contain infinitely many primes, suitably weighted. For example in [9] Heath-Brown compared the sequence of values of the binary cubic form $x^3 + 2y^3$ (weighted by multiplicity) and the sequence of values taken by the norm form of the cubic field $K = \mathbb{Q}(\sqrt[3]{2})$.

We shall consider two nonnegative sequences $\mathcal{A} = (a_n)$, $\mathcal{B} = (b_n)$ supported on positive integers $n \leq X$, and put

$$\pi(\mathcal{A}) = \sum_p a_p \quad \text{and} \quad \pi(\mathcal{B}) = \sum_p b_p, \tag{3-1}$$

where the summations run over primes. If one establishes an asymptotic relation of the form

$$\pi(\mathcal{A}) = \varkappa \pi(\mathcal{B})(1 + o(1))$$

say, then an asymptotic formula for $\pi(\mathcal{B})$ implies an asymptotic formula for $\pi(\mathcal{A})$. In particular, this allows us to avoid working through the difficult harmonic analysis in [5], and allows one to work with estimates that apply to general complex sequences rather than relying on properties of the Möbius function.

To simplify matters, we will restrict the variable of interest, namely ℓ, to a short interval of the shape $I(X) = (X^*, (1+\eta)X^*]$ where $\eta \asymp (\log X)^{-1}$ and $X^{1/2}(\log X)^{-4} \leq X^* \leq c_f X^{1/2}$ where

$$c_f = \begin{cases} \sup_{f(x,y) \leq 1} y & \text{if } f \text{ is definite,} \\ 1 & \text{if } f \text{ is indefinite.} \end{cases}$$

We then define

$$a_n = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} \mathfrak{Z}(\ell) \tag{3-2}$$

and

$$b_n = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} \Lambda(\ell). \tag{3-3}$$

Here

$$\mathfrak{Z}(\ell) = \begin{cases} 2p \log p & \text{if } \ell = p^2, \\ 0 & \text{otherwise,} \end{cases} \tag{3-4}$$

and Λ is the von Mangoldt function. In the definite case Lam, Schindler, and the author proved that $\pi(\mathcal{B})$ satisfies an asymptotic formula. We will extend this to the indefinite, irreducible case.

One notes that the sequences (a_n), (b_n) introduced in (3-2) and (3-3) are analogous to the sequences introduced in [10]. The analogous sequences $\mathcal{A}^{\spadesuit}$, $\mathcal{B}^{\spadesuit}$ for the purpose of Theorem 1.1 are

$$a_n^{\spadesuit} = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} \mathfrak{Z}^{\spadesuit}(\ell) \text{ and } b_n^{\spadesuit} = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} 1, \tag{3-5}$$

respectively, where

$$\mathfrak{Z}^{\spadesuit}(\ell) = \begin{cases} 2k & \text{if } \ell = k^2, \\ 0 & \text{otherwise.} \end{cases} \tag{3-6}$$

We emphasize that the integer k appearing in (3-6) is not required to be prime, unlike in (3-4).

Having established the asymptotic formula for $\pi(\mathcal{B})$, $\pi(\mathcal{B}^{\spadesuit})$, we will then prove an analogue of Proposition 1 in [10]. In [10] they introduced the quantity

$$\mu(I) = \int_I \sqrt{X - t^2}\, dt = \int_I \int_0^{\sqrt{X-t^2}} ds\, dt.$$

In other words, $\mu(I)$ is the area of the subset of the positive half-disk with y-coordinate restricted to I. We generalize this definition to

$$\mu_f(I) = \text{Area}\{(x,y) \in \mathbb{R}^2 : 0 < f(x,y) < X,\ y \in I(X)\} = \int_I \int_{0<f(s,t)<X} ds\, dt. \tag{3-7}$$

Observe that $\mu_f(I) \ll_f \sqrt{X} \cdot |I|$, where $|I|$ is the length of I. This brings us to the following statement:

Proposition 3.1. *Let $\mathcal{A} = (a_n)$, $\mathcal{B} = (b_n)$ be given as in (3-2) and (3-3). Then we have the asymptotic relation*

$$|\pi(\mathcal{A}) - \pi(\mathcal{B})| \ll_\varepsilon \frac{\mu_f(I) \log\log X}{(\log X)^2}.$$

Similarly, for $\mathcal{A}^{\spadesuit}$, $\mathcal{B}^{\spadesuit}$ given by (3-5) one has

$$|\pi(\mathcal{A}^{\spadesuit}) - \pi(\mathcal{B}^{\spadesuit})| \ll_\varepsilon \frac{\mu_f(I) \log\log X}{(\log X)^2}.$$

We will see that this is enough to prove Theorems 1.1 and 1.2 as in the proof of Theorem 1 from Proposition 1 in [10]. First we will prove that

$$\pi(\mathcal{B}) = \frac{\nu_f \mu_f(I)}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right), \tag{3-8}$$

this following from Theorem 1.3 via partial summation. In the case of $\mathcal{B}$ and f is definite we start with the asymptotic formula (1-10) and write it as

$$\sum_{q \leq X} \Lambda(q) \sum_{f(m,\ell)=q} \Lambda(\ell) = \nu_f \mathfrak{S}'_f X + O_A(X(\log X)^{-A}).$$

Writing $\Psi(q) = \sum_{f(m,\ell)=q} \Lambda(\ell)$ and replacing $\Lambda(q)$ with $\log q$ (supported on primes), we have by partial summation

$$\log X \sum_{q \leq X} \Psi(q) - \int_1^X \frac{1}{t}\left(\sum_{q \leq t} \Psi(q)\right) dt = \nu_f \mathfrak{S}'_f X + O_A(X(\log X)^{-A}).$$

An upper bound sieve gives that

$$\sum_{q \leq X} \Psi(q) = O\left(\frac{X}{\log X}\right),$$

hence

$$\log X \sum_{q \leq X} \Psi(q) = \nu_f \mathfrak{S}'_f X + O_A(X(\log X)^{-A}) + O\left(\int_1^X \frac{dt}{\log t}\right),$$

and thus

$$\sum_{q \leq X} \Psi(q) = \frac{\nu_f \mathfrak{S}'_f X}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right).$$

By replacing $\Psi(q)$ with

$$\Psi'(q) = \sum_{\substack{f(m,\ell)=q \\ \ell \in I(X)}} \Lambda(\ell),$$

we see from the same argument that

$$\sum_{q \leq X} \Psi'(q) = \frac{\nu_f \mu_f(I)}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right),$$

as desired. The same argument applies to the indefinite case, following (1-11).

Thus Proposition 3.1 gives

$$\pi(\mathcal{A}) = \frac{\nu_f \mu_f(I)}{\log X}\left(1 + O\left(\frac{\log \log X}{\log X}\right)\right). \tag{3-9}$$

We then proceed by partial summation as in [10]. We consider intervals $I_j = (X_j, X_j(1+\eta)]$ that form a partition of $(X^{1/2}(\log X)^{-4}, c_f X^{1/2}]$. Here $\eta \asymp (\log X)^{-1}$ is chosen so we have an exact partition. We let $\mathcal{A}_j$ be defined as in (3-2) with $I(X) = I_j$. The number of pairs (a, p) with $0 < f(a, p^2) \le X$ and $p|a$, $p^2 \in I$ is bounded by

$$\sum_{p^2 \in I} \frac{\sqrt{X}}{p} \ll_\varepsilon X^{1/2+\varepsilon}.$$

It follows that

$$\begin{aligned}
&\#\{(a, p) : 0 < f(a, p^2) \le X \text{ is prime }, \ p \text{ is prime}, \ p \le X^{1/4}\} \\
&\quad = \sum_j \frac{1}{\sqrt{X_j}\log X_j}\pi(\mathcal{A}_j)\Big(1 + O\Big(\frac{1}{\log X}\Big)\Big) + O\Big(\frac{X^{3/4}}{(\log X)^3}\Big) \\
&\quad = \frac{\nu_f + O((\log X)^{-1}\log\log X)}{(\log X)^2} \sum_j \frac{\mu_f(I_j)}{\sqrt{X_j}} + O\Big(\frac{X^{3/4}}{(\log X)^3}\Big) \\
&\quad = \frac{\nu_f + O((\log X)^{-1}\log\log X)}{(\log X)^2} \int_{\sqrt{X}/(\log X)^4}^{\sqrt{X}} \frac{1}{\sqrt{t}} \int_{0<f(s,t)<X} ds\, dt + O\Big(\frac{X^{3/4}}{(\log X)^3}\Big) \\
&\quad = \frac{\nu_f \mathfrak{S}_f X^{3/4}}{(\log X)^2}\Big(1 + O\Big(\frac{\log\log X}{\log X}\Big)\Big).
\end{aligned}$$

Thus Theorem 1.2 follows from Proposition 3.1. Next we do something similar to deduce Theorem 1.1. In this case it is trivial that

$$\pi(\mathcal{B}^\spadesuit) = \frac{\nu_f \mu_f(I)}{\log X}\Big(1 + O\Big(\frac{1}{\log X}\Big)\Big),$$

since this is a direct consequence of Landau's prime ideal theorem. Therefore Proposition 3.1 gives

$$\pi(\mathcal{A}^\spadesuit) = \frac{\nu_f \mu_f(I)}{\log X}\Big(1 + O\Big(\frac{1}{\log X}\Big)\Big).$$

We then proceed by partial summation as above, but noting that the weight is $2k$ rather than $2p\log p$. The same calculation then gives

$$\#\{(a, b) : 0 < f(a, b^2) \le X \text{ is prime}, \ b \le X^{1/4}\} = \frac{\nu_f \mathfrak{S}_f X^{3/4}}{\log X}\Big(1 + O\Big(\frac{1}{\log X}\Big)\Big),$$

which suffices to prove Theorem 1.1.

In order to establish Proposition 3.1 we apply the same sieve procedure to the pairs $(\mathcal{A}, \mathcal{B})$ and $(\mathcal{A}^\spadesuit, \mathcal{B}^\spadesuit)$, producing cancellation at key junctures and upper bounding the rest. For any complex sequence $\mathcal{C} = (c_n)$ supported on the positive integers put

$$S(\mathcal{C}, Z) = \sum_{\substack{n\in\mathbb{N} \\ p|n \Rightarrow p>Z}} c_n$$

and for each $d \in \mathbb{N}$ put

$$\mathcal{C}_d = \{c_{dn} : n \in \mathbb{N}\}.$$

We fix

$$\delta_1 = \delta_1(X) = (\log X)^{\varpi - 1} \quad \text{and} \quad \delta_2 = \delta_2(X) = \frac{A_1 \log\log X}{\log X} \tag{3-10}$$

for some large positive number A_1 and small number $0 < \varpi < 1$ which we specify later. In [10] they have a single parameter δ which is equal to our δ_1. The reason why we are having two separate parameters is to obtain the superior error term in Theorem 1.2 and the error term in Theorem 1.1.

We also fix $Y > X^{1/3}$, where the specific choice of Y will be made when it is relevant. Now put

$$S_1(\mathcal{C}) = S(\mathcal{C}, X^{\delta_1}), \quad S_2(\mathcal{C}) = \sum_{X^{\delta_1} \le p < Y} S(\mathcal{C}_p, p), \quad S_3(\mathcal{C}) = \sum_{Y \le p < X^{1/2 - \delta_2}} S(\mathcal{C}_p, p). \tag{3-11}$$

The astute reader will note that $S_1(\mathcal{C})$ is readily handled by the fundamental lemma of sieve theory, giving an asymptotic formula; see, for example, Corollary 6.10 in [6]. By Buchstab's identity, we have

$$\pi(\mathcal{C}) = S(\mathcal{C}, X^{1/2}) = S_1(\mathcal{C}) - S_2(\mathcal{C}) - S_3(\mathcal{C}) - \sum_{X^{1/2-\delta_2} \le p \le X^{1/2}} S(\mathcal{C}_p, p).$$

The last sum can be handled by Selberg's upper bound sieve, and we conclude:

Lemma 3.2. *For $Y = X^{17/48}$ and $\mathcal{C} = \mathcal{A}, \mathcal{B}, \mathcal{A}^{\spadesuit}, \mathcal{B}^{\spadesuit}$ we have*

$$\pi(\mathcal{C}) = S_1(\mathcal{C}) - S_2(\mathcal{C}) - S_3(\mathcal{C}) + O\left(\frac{\delta_2 \mu_f(I)}{\log X}\right).$$

We will see that $S_3(\mathcal{C})$ can be written in terms of appropriate bilinear forms, but $S_2(\mathcal{C})$ will require further treatment. Let us put

$$T^{(n)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1} \le p_n < \cdots < p_1 < Y \\ p_1 \cdots p_n < Y}} S(\mathcal{C}_{p_1 \cdots p_n}, X^{\delta_1})$$

and

$$U^{(n)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1} \le p_{n+1} < \cdots < p_1 < Y \\ p_1 \cdots p_n < Y \le p_1 \cdots p_{n+1}}} S(\mathcal{C}_{p_1 \cdots p_{n+1}}, p_{n+1}).$$

We then have:

Lemma 3.3. *For $n_0 = \left\lfloor \frac{\log Y}{\delta_1 \log X} \right\rfloor$ we have*

$$S_2(\mathcal{C}) = \sum_{1 \le n \le n_0} (-1)^{n-1} (T^{(n)}(\mathcal{C}) - U^{(n)}(\mathcal{C})).$$

The sums

$$|S_1(\mathcal{A}) - S_1(\mathcal{B})|, \quad |S_1(\mathcal{A}^{\spadesuit}) - S_1(\mathcal{B}^{\spadesuit})| \tag{3-12}$$

and

$$\sum_{1 \le n \le n_0} |T^{(n)}(\mathcal{A}) - T^{(n)}(\mathcal{B})|, \quad \sum_{1 \le n \le n_0} |T^{(n)}(\mathcal{A}^{\spadesuit}) - T^{(n)}(\mathcal{B}^{\spadesuit})| \tag{3-13}$$

can be handled by our type-I estimate Proposition 5.1 and the fundamental lemma; see Lemma 2 in [10]. To control these sums it suffices to prove:

Proposition 3.4. *Let $\mathfrak{Q}$ be a set of square-free numbers not exceeding $Y = X^{17/48}$. Then for any $A > 0$ we have*

$$\left|\sum_{q\in\mathfrak{Q}} S(\mathcal{A}_q, X^{\delta_1}) - \sum_{q\in\mathfrak{Q}} S(\mathcal{B}_q, X^{\delta_1})\right| \ll_A \frac{X}{(\log X)^A}$$

and

$$\left|\sum_{q\in\mathfrak{Q}} S(\mathcal{A}_q^{\spadesuit}, X^{\delta_1}) - \sum_{q\in\mathfrak{Q}} S(\mathcal{B}_q^{\spadesuit}, X^{\delta_1})\right| \ll_A \frac{X}{(\log X)^A}$$

By the definitions of $S_1(\mathcal{C})$ and $T^{(n)}(\mathcal{C})$, clearly Proposition 3.4 gives the bound of $O_A(X(\log X)^{-A})$ for both (3-12) and (3-13).

We now give a proof for Proposition 3.4, which depends on Proposition 5.1.

Proof of Proposition 3.4. The fundamental lemma allows us to give an asymptotic formula for the sum

$$\sum_{q\in\mathfrak{Q}} S(\mathcal{C}_q, X^{\delta_1})$$

for $\mathcal{C} = \mathcal{A}, \mathcal{B}, \mathcal{A}^{\spadesuit}, \mathcal{B}^{\spadesuit}$. Recall that $\delta_1 = (\log X)^{\varpi-1}$. Proposition 5.1 gives us a level of distribution of $X^{3/4}(\log X)^{-B}$ for some large B. We then apply an upper and lower bound sieve of level of distribution $X^{1/4}$, so that the sifting variable

$$s = \frac{\log D}{\log z} = \frac{\log X^{1/4}}{\log X^{\delta_1}} = \frac{1}{4\delta_1}.$$

We use the usual notation

$$V(z) = \prod_{p<z}(1-g(p)) = \prod_{p<z}\left(1 - \frac{\rho_f(p)}{p}\right),$$

where $\rho_f(p)$ counts the number of linear factors of f modulo p; see (5-2). We then have

$$R_d(\mathcal{C}) = |A_d(\mathcal{C}) - M_d(\mathcal{C})|$$

with $M_d(\mathcal{C})$ as in Proposition 5.1. By Corollary 6.10 in [6] and applying Proposition 5.1 we obtain

$$\begin{aligned}\sum_{q\in\mathfrak{Q}} S(\mathcal{C}_q, X^{\delta_1}) &= V(X^{\delta_1}) \sum_{q\in\mathfrak{Q}} \frac{\rho_f(q)}{q}\mu_f(I)\big(1 + O\big(\exp(-(4\delta)^{-1})\big)\big) + O\Big(\sum_{q\in\mathfrak{Q}}\sum_{d<X^{1/4}} R_{dq}(\mathcal{C})\Big)\\ &= V(X^{\delta_1}) \sum_{q\in\mathfrak{Q}} \frac{\rho_f(q)}{q}\mu_f(I)\left(1 + O\left(\frac{1}{(\log X)^A}\right)\right) + O\Big(\sum_{d<X^{3/4-1/8}} \tau(d) R_q(\mathcal{C})\Big)\\ &= V(X^{\delta_1}) \sum_{q\in\mathfrak{Q}} \frac{\rho_f(q)}{q}\mu_f(I)\left(1 + O\left(\frac{1}{(\log X)^A}\right)\right) + O_A(X(\log X)^{-A})\end{aligned}$$

for any $A > 0$. The last line is independent of whether $\mathcal{C} = \mathcal{A}, \mathcal{B}, \mathcal{A}^\spadesuit$ or $\mathcal{C} = \mathcal{B}^\spadesuit$. Since $V(X^{\delta_1}) \le 1$ it follows that

$$\sum_{q\in\mathfrak{Q}}(S(\mathcal{A}_q, X^{\delta_1}) - S(\mathcal{B}_q, X^{\delta_1})) \ll_A \frac{1}{(\log X)^A}\mu_f(I)\sum_{q\in\mathfrak{Q}}\frac{\rho_f(q)}{q} + X(\log X)^{-A}$$
$$\ll_A X(\log X)^{-A+2},$$

since $\rho_f(q) \ll \tau(q)$. Likewise,

$$\sum_{q\in\mathfrak{Q}}(S(\mathcal{A}_q^\spadesuit, X^{\delta_1}) - S(\mathcal{B}_q^\spadesuit, X^{\delta_1})) \ll_A X(\log X)^{-A+2}. \qquad \square$$

Thus it remains to show that

$$|S_3(\mathcal{A}) - S_3(\mathcal{B})| \ll_A \frac{X}{(\log X)^A} \quad \text{and} \quad |U^{(n)}(\mathcal{A}) - U^{(n)}(\mathcal{B})| \ll_A \frac{X}{(\log X)^A} \quad \text{for } n \ge 3 \tag{3-14}$$

and

$$|U^{(n)}(\mathcal{A}) - U^{(n)}(\mathcal{B})| \ll \frac{\delta_2\mu_f(I)}{\log X} \quad \text{for } n = 1, 2, \tag{3-15}$$

with analogous statements for $\mathcal{A}^\spadesuit, \mathcal{B}^\spadesuit$.

We proceed to reduce the verification of (3-14) and (3-15) to a bilinear sum estimate.

Reduction to a bilinear sum bound. Let us write $U^{(1)}$ and $U^{(2)}$ into a more convenient form, as in [10]. To do so let us put

$$U_1^{(1)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1}\le p_2<p_1<Y \\ Y\le p_1p_2<X^{1/2-\delta_2}}} S(\mathcal{C}_{p_1p_2}, p_2),$$

$$U_2^{(1)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1}\le p_2<p_1<Y \\ p_1p_2\ge X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1p_2}, p_2),$$

$$U_1^{(2)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1}\le p_3<\cdots<p_1<Y \\ p_1p_2<Y\le p_1p_2p_3<X^{1/2-\delta_2}}} S(\mathcal{C}_{p_1p_2p_3}, p_3),$$

$$U_2^{(2)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1}\le p_3<\cdots<p_1<Y \\ p_1p_2<Y\le p_1p_2p_3 \\ p_1p_2p_3\ge X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1p_2p_3}, p_3).$$

We now state Lemmas 6 and 7 from [10]. Their proofs apply equally well, but since for us δ_1, δ_2 are different we write out the proofs.

Lemma 3.5 [10, Lemma 6]. *For $\mathcal{C} = \mathcal{A}, \mathcal{B}$ we have $U^{(j)}(\mathcal{C})$ satisfies*

$$U^{(1)}(\mathcal{C}) = U_1^{(1)}(\mathcal{C}) + U_2^{(1)}(\mathcal{C}) + O\left(\frac{\delta_2\mu_f(I)}{\log X}\right) \tag{3-16}$$

and

$$U^{(2)}(\mathcal{C}) = U_1^{(2)}(\mathcal{C}) + U_2^{(2)}(\mathcal{C}) + O\left(\frac{\delta_2 \mu_f(I)}{\log X}\right). \tag{3-17}$$

Proof. To prove (3-16) it suffices to show

$$\sum_{\substack{X^{\delta_1} \le p_2 < p_1 < Y \\ X^{1/2-\delta_2} < p_1 p_2 \le X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1 p_2}, p_2) \ll \frac{\delta_2 \mu_f(I)}{\log X}.$$

In the sum above we have

$$p_2 \ge \frac{X^{1/2-\delta_2}}{p_1} > \frac{X^{1/2-\delta_2}}{Y} > X^{1/10}$$

so we may apply Selberg's upper bound sieve and our level of distribution to obtain

$$\begin{aligned}
\sum_{\substack{X^{\delta_1} \le p_2 < p_1 < Y \\ X^{1/2-\delta_2} < p_1 p_2 \le X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1 p_2}, p_2) &\ll \sum_{\substack{X^{\delta_1} \le p_2 < p_1 < Y \\ X^{1/2-\delta_2} < p_1 p_2 \le X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1 p_2}, X^{1/10}) \\
&\ll \frac{\mu_f(I)}{\log X} \sum_{\substack{X^{1/10} < p_2 < p_1 < Y \\ X^{1/2-\delta_2} < p_1 p_2 < X^{1/2+\delta_2}}} \frac{1}{p_1 p_2} \\
&\ll \frac{\delta_2 \mu_f(I)}{\log X}.
\end{aligned}$$

The proof for (3-17) follows similarly. □

Lemma 3.6 [10, Lemma 7]. *Let κ be a positive number satisfying $X^{-\delta_1} \le \kappa \le 1$. Let N_1, N_2 be positive numbers in the interval $[X^{\delta}, X^{1/3}]$. We then have, for any $A > 0$,*

$$\sum_{N_1 \le p_1 \le (1+\kappa)N_1} \sum_{N_2 \le p_2 \le (1+\kappa)N_2} \sum_{n \equiv 0 \,(\mathrm{mod}\, p_1 p_2)} c_n \tau(n) \ll_A \kappa^2 X (\log X)^{2^{17}} + \frac{X}{(\log X)^A}.$$

For $k \ge 3$, the condition of summation in $U^{(k)}(\mathcal{C})$ is

$$X^{17/48} = Y \le p_1 \cdots p_{k+1} < (p_1 \cdots p_k)^{(k+1)/k} \le Y^{4/3} < X^{1/2-\delta_2}.$$

Therefore, upon defining

$$U_*^{(k)}(\mathcal{C}) = \sum_{X^{\delta_1} \le p_{k+1} < \cdots < p_1 \cdots p_k < Y \le p_1 \cdots p_{k+1} < X^{1/2-\delta}} S(\mathcal{C}_{p_1 \cdots p_{k+1}}, p_{k+1})$$

we have

$$S_3(\mathcal{C}) = U_*^{(0)}(\mathcal{C}), \; U_1^{(1)}(\mathcal{C}) = U_*^{(1)}(\mathcal{C}), \; U_1^{(2)}(\mathcal{C}) = U_*^{(2)}(\mathcal{C})$$

and

$$U^{(k)}(\mathcal{C}) = U_*^{(k)}(\mathcal{C}) \quad \text{for } k \ge 3.$$

If $p \in \mathcal{J} = [V, (1+\kappa)V)$ and n is an integer is counted by $S(\mathcal{C}_{pq}, V)$ but not by $S(\mathcal{C}_{pq}, p)$, then n has at least two prime factors p_1, p_2 in $\mathcal{J}$. In our application we will have $V \le X^{1/2-\delta_2}$ and $n \ge X(\log X)^{-8}$. Note that n has a divisor exceeding one and coprime to $p_1 p_2$. It follows that

$$V^3 \le n \le X.$$

A given integer n may be counted multiple times by $U_*^{(k)}(\mathcal{C})$ but the multiplicity is bounded by the number of choices for $p_{k+1} < \cdots < p_1$ all dividing n, and therefore the multiplicity is at most $\tau(n)$. Applying Lemma 3.6 and setting

$$\mathcal{J}(r) = [V_r, V_{r+1}) = [X^{\delta_1}(1+\kappa)^r, X^{\delta_1}(1+\kappa)^{r+1}), \quad r \ge 0$$

and $R \ll \kappa^{-1} \log X$ satisfying $X^{\delta_1}(1+\kappa)^R > X$, we obtain

$$U_*^{(k)}(\mathcal{C}) = \sum_{0 \le r \le R} \sum_{p \in \mathcal{J}(r)} \sum_{p < p_k < \cdots < p_1 \cdots p_k < Y \le p_1 \cdots p_k p < X^{1/2-\delta_2}} S(\mathcal{C}_{p_1 \cdots p_k p}, V_r) + O_A\left(\kappa X (\log X)^{1+2^{17}} + \kappa^{-1} \frac{X}{(\log X)^{A-1}}\right). \tag{3-18}$$

We need to make sure that both

$$\kappa X (\log X)^{1+2^{17}}, \quad \kappa^{-1} \frac{X}{(\log X)^{A-1}}$$

are $O(X(\log X)^{-A'})$ for some $A' > 1$. This compels us to choose

$$\kappa = (\log X)^{-A/2}.$$

This gives

$$\kappa X (\log X)^{1+2^{17}} = X(\log X)^{1+2^{17}-A/2} \quad \text{and} \quad \kappa^{-1} \frac{X}{(\log X)^{A-1}} = \frac{X}{(\log X)^{A/2-1}}. \tag{3-19}$$

This procedure allows us to reduce our proof to estimations of certain bilinear sums since

$$\sum_{p \in \mathcal{J}(r)} \sum_{p < p_k < \cdots < p_1 \cdots p_k < Y \le p_1 \cdots p_k p < X^{1/2-\delta_2}} S(\mathcal{C}_{p_1 \cdots p_k p}, V_r) = \sum_{m,n} \alpha_m^{(r)} \beta_n^{(r)} c_{mn}, \tag{3-20}$$

where $\alpha_m^{(r)}$ is the characteristic function for the integers m all of whose prime factors are at least V_r and $\beta_n^{(r)}$ is the characteristic function for integers $n = p_1 \cdots p_k p$ satisfying

$$p \in \mathcal{J}(r), \quad p < p_k < \cdots < p_1 < Y \quad \text{and} \quad p_1 \cdots p_k < Y \le p_1 \cdots p_k p < X^{1/2-\delta_2}.$$

Observe that $\beta_n^{(r)}$ is supported on integers $n \in [Y, X^{1/2-\delta_2})$.

The procedure for $U_2^{(1)}(\mathcal{C})$ and $U_2^{(2)}(\mathcal{C})$ will be somewhat different. We may use Lemma 3.6 to replace $S(\mathcal{C}_{p_1p_2}, p_2)$ in $U_2^{(1)}(\mathcal{C})$ by $S(\mathcal{C}_{p_1p_2}, V_r)$ when $p_2 \in J(r)$. This yields

$$U_2^{(1)}(\mathcal{C}) = \sum_{0\le r\le R} \sum_{p_2\in J(r)} \sum_{\substack{p_1\ge X^{1/2-\delta_2}/p_2\\ p_2<p_1<Y}} S(\mathcal{C}_{p_1p_2}, V_r) + O(\kappa X(\log X)^{1+2^{17}}) + O\left(\kappa^{-1}\frac{X}{(\log X)^{A-1}}\right).$$

The sum on the right can be expressed as

$$\sum_{0\le r\le R} \sum_{m,n} \alpha_m^{(r)} \beta_n^{(r)} c(mn),$$

where we now take $\alpha_m^{(r)}$ to be the characteristic function for numbers $m = p_1p_2$ with $p_2 \in J(r)$, $p_2 < p_1 < Y$ and $p_1p_2 \ge X^{1/2+\delta_2}$, and $\beta_n^{(r)}$ to be the characteristic function for those numbers n all of whose prime factors are at least V_r. Since $c(n)$ is supported in

$$((X^*)^2, c_f X] \subseteq (X(\log X)^{-8}, c_f X]$$

we may assume that $\beta_n^{(r)}$ is supported in

$$(X(\log X)^{-8}Y^{-2}, X^{1/2-\delta_2}] \subseteq (X^{1/4+1/48}, X^{1/2-\delta_2}].$$

This is sufficient for our purposes. We may handle $U_2^{(2)}(\mathcal{C})$ in an analogous fashion.

On setting $\kappa = (\log X)^{-A/2}$ we find that each of

$$S_3(\mathcal{C}),\quad U_1^{(1)}(\mathcal{C}),\quad U_2^{(1)}(\mathcal{C}),\quad U_1^{(2)}(\mathcal{C}),\quad U_2^{(2)}(\mathcal{C}),\quad \text{and}\quad U^{(k)}(\mathcal{C})$$

for $k \ge 3$ can be expressed as a sum of $O(R)$ bilinear sums as in (3-20), together with an error term of $O_A(X(\log X)^{1+2^{17}-A/2})$. Thus it will be sufficient to prove:

Proposition 3.7 (main bilinear sum estimate). *Let $\xi > 0$ and suppose $X^{1/4+\xi} \le N < X^{1/2-\delta_2}$. Suppose (α_m), (β_n) are two complex sequences having sup-norm at most 1 supported on natural numbers with no prime factors less than X^{δ}. Then for any $A > 0$ we have*

$$\sum_{N<n\le 2N} \sum_{m<X/N} \alpha_m\beta_n(a_{mn} - b_{mn}) \ll_{A,\xi} \frac{X}{(\log X)^A} \tag{3-21}$$

and

$$\sum_{N<n\le 2N} \sum_{m<X/N} \alpha_m\beta_n(a^{\spadesuit}_{mn} - b^{\spadesuit}_{mn}) \ll_{A,\xi} \frac{X}{(\log X)^A} \tag{3-22}$$

It will be important that the sequences $\{\alpha_m\}$, $\{\beta_n\}$ are supported on those numbers whose prime factors all exceed X^{δ_1}, and in particular, they are supported on odd numbers.

The remainder of the paper is devoted to proving Proposition 3.7. In particular, Propositions 7.5 and 7.6 will imply Proposition 3.7. In order to get there, we need to decompose the terms c_{mn} for any positive integers m, n into components that resemble c_m, c_n. This turns out to be somewhat delicate and we will require the composition laws of the ideals of $\mathcal{O}_K$, expressed in terms of ideal numbers. This will be the primary focus of the next section.

4. Algebraic characterization of the multiplicative structure in terms of ideal numbers

The main purpose of this section is obtain an analogue of Proposition 2.3 in [12]. However, instead of using an explicit Dirichlet composition law as in [12] we will instead adopt the language of Hecke's *ideal numbers* as in [11].

Choose ideals $\mathfrak{a}_1, \dots, \mathfrak{a}_t$ whose classes generate the ideal class group of $\mathcal{O}_K$ so that every fractional ideal $\mathfrak{a} \subseteq \mathcal{O}_K$ has a unique decomposition

$$\mathfrak{a} = (\alpha)\mathfrak{a}_1^{\ell_1} \cdots \mathfrak{a}_t^{\ell_t}$$

where $\alpha \in K^*$ and $\ell_j \in \mathbb{Z}$ with $0 \leq \ell_j < h_j$, with h_j the smallest positive integer such that $\mathfrak{a}_j^{h_j} = (\alpha_j)$ is principal. Then the class number $h(K)$ of $\mathcal{O}_K$ is equal to

$$h(K) = \prod_{j=1}^{t} h_j. \tag{4-1}$$

Let us choose complex numbers $b_1, \dots, b_t$ so that

$$b_j^{h_j} = \alpha_j \quad \text{for } j = 1, \dots, t,$$

and $b_j^{(i)}$ are complex numbers such that

$$(b_j^{(i)})^{h_j} = \alpha_j^{(i)} \quad \text{for } i = 1, 2.$$

Now put $L = K(b_1, \dots, b_t)$ and $\mathfrak{J}(K)^*$ the subgroup of L^* generated by K^* and $\{b_j : 1 \leq j \leq t\}$. Then $\mathfrak{J}(K) = \{0\} \cup \mathfrak{J}(K)^*$ is the domain of *ideal numbers* of K. The quotient group $\mathfrak{J}(K)^*/\mathcal{O}_K^*$ is then isomorphic to the group of fractional ideals of K. Each $\gamma \in \mathfrak{J}(K)$ corresponds a unique fractional ideal $J(\gamma)$; the norm of the ideal $J(\gamma)$ is given by the product

$$N(\gamma) = N(J(\gamma)) = \gamma^{(1)}\gamma^{(2)}.$$

Note that $\gamma^{(2)}$ is the algebraic conjugate of $\gamma^{(1)} = \gamma$.

Further, we have $J(\gamma)$ is an integral ideal of $\mathcal{O}_K$ if and only if $\gamma \in \mathcal{O}_L$.

We thus have a correspondence between the ideal classes of $\mathcal{O}_K$ and a subset of integers in $\mathcal{O}_L$. Indeed, we can say that $\gamma, \gamma' \in \mathfrak{J}(K)$ belong to the same class if and only if the corresponding ideals $J(\gamma), J(\gamma') \subseteq \mathcal{O}_K$ are in the same ideal class. It follows that we may partition $\mathfrak{J}(K)$ into $h(K)$ classes, corresponding to the ideal classes of $\mathcal{O}_K$. Such a class of ideal numbers, say A, has an integral basis $\{w_1, w_2\}$ such that

$$A = \{a_1 w_1 + a_2 w_2 : (a_1, a_2) \in \mathbb{Q}^2\}$$

and

$$A \cap \mathcal{O}_L = \{a_1 w_1 + a_2 w_2 : (a_1, a_2) \in \mathbb{Z}^2\}.$$

Indeed, this follows by noting that

$$A = \{\gamma \beta_1^{\ell_1} \cdots \beta_t^{\ell_t} : \gamma \in K\},$$

where $(\ell_1, \dots, \ell_t)$ is a fixed tuple of nonnegative integers. If (v_1, v_2) is an integral basis of K, then we may take

$$w_1 = v_1 \beta_1^{\ell_1} \cdots \beta_t^{\ell_t}, \quad w_2 = v_2 \beta_1^{\ell_1} \cdots \beta_t^{\ell_t}.$$

Further, the discriminant of A, viewed as a $\mathbb{Z}$-lattice, is equal to $\Delta(K)$. Moreover for any basis $\{w_1, w_2\}$ of A and $\alpha \in A \setminus \{0\}$ we have that $\{\alpha^{-1} w_1, \alpha^{-1} w_2\}$ is a basis of $K/\mathbb{Q}$. This implies that there is a unique *dual basis* $\{\widetilde{w_1}, \widetilde{w_2}\}$ of A^{-1} defined by the condition

$$\operatorname{Tr}(w_i \widetilde{w_j}) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{4-2}$$

We use the notation $\operatorname{Cl} \mathfrak{a}, \operatorname{Cl} \alpha$ for the ideal class of the integral ideal $\mathfrak{a} \subset \mathcal{O}_K$ and the class of ideal numbers of the ideal number α.

Next we show that there is a correspondence between rank-two submodules of $\mathcal{O}_K$ and $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of irreducible integral binary quadratic forms having splitting field K. To establish this correspondence, first start with a rank-two submodule

$$\Lambda = \{a_1 \omega_1 + a_2 \omega_2 : a_1, a_2 \in \mathbb{Z}\}$$

with $\omega_1, \omega_2 \in \mathcal{O}_K$. Put

$$\mathfrak{S} = \gcd\{N_{K/\mathbb{Q}}(\omega_1 x + \omega_2) : x, y \in \mathbb{Z}\}.$$

Then the form

$$g(x, y) = N_{K/\mathbb{Q}}(\omega_1 x + \omega_2 y) \mathfrak{S}^{-1} \tag{4-3}$$

is an irreducible integral binary quadratic form with splitting field K.

Conversely, take an arbitrary irreducible integral binary quadratic form g which splits over K. Then there exists an integral nonsingular matrix M such that

$$g(x, y) = g^*((x, y)M),$$

where g^* is a primitive integral binary quadratic form with discriminant equal to $\Delta(K)$. Gauss's composition law then implies that g^* corresponds to an ideal class α, and in particular, can be expressed in the form

$$g^*(x, y) = N_{K/\mathbb{Q}}(\alpha_1 x + \alpha_2 y) N(\alpha^{-1})$$

with $\alpha = (\alpha_1, \alpha_2)$. Viewing α as a $\mathbb{Z}$-module and applying the transformation induced by M then gives the form g.

Now let $\mathfrak{f}$ be the $\mathbb{Z}$-module associated to f with basis $\{\nu_1, \nu_2\}$ so that

$$f(x, y) = N_{K/\mathbb{Q}}(\nu_1 x + \nu_2 y) N(\mathfrak{d}(f)^{-1}), \tag{4-4}$$

where $\mathfrak{d}(f) = (\nu_1, \nu_2)$ is the ideal generated by ν_1, ν_2. Let ψ_f be the ideal number of the ideal $\mathfrak{d}(f)$. Having identified $\mathfrak{f}$ we define the set of ideals

$$\mathfrak{A}(f) = \{(\nu_1 a_1 + \nu_2 a_2)\mathfrak{d}(f)^{-1} : a_1, a_2 \in \mathbb{Z},\ \gcd(a_1, a_2) = 1\}.$$

We now put $\mathcal{L}$ for the set of ideals in $\mathcal{O}_K$ which are not divisible by a rational prime. An integral ideal number $\gamma \in \mathfrak{I}(K)$ is said to be primitive if $J(\gamma) \in \mathcal{L}$. If K is a real quadratic field, put $\mathcal{L}_0$ the set of primitive ideal numbers γ satisfying the condition

$$\gamma = (N_{L/\mathbb{Q}}(\gamma))^{1/2} \varepsilon_0^z, \quad -\tfrac{1}{2} < z \le \tfrac{1}{2},\ \gamma > 0,$$

where $\varepsilon_0 > 1$ is a fundamental unit of $\mathcal{O}_K$. If K is an imaginary quadratic field we may simply take $\mathcal{L}_0$ to be the set of primitive ideal numbers.

We now want to use the above discussion to obtain a meaningful decomposition for

$$c_n = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} \Upsilon(\ell). \tag{4-5}$$

We follow the setup in [11] and introduce, for a given primitive vector $\boldsymbol{u} = (u_1, u_2)$ let $\mathfrak{F}(\boldsymbol{u})$ be the ideal in $\mathfrak{A}(f)$ given by $(\nu_1 u_1 + \nu_2 u_2) N(\mathfrak{d}(f)^{-1})$. We now put

$$\mathfrak{R}(X; n) = \{(u_1, u_2) \in \mathbb{Z}^2 : u_2 \in I(X),\ f(u_1, u_2) = n\}.$$

Note that $\mathfrak{R}(X; n)$ is finite for all $X > 0$ and $n \in \mathbb{Z}$. We then have

$$c_n = \sum_{\boldsymbol{u} \in \mathfrak{R}(X; n)} \Upsilon(u_2).$$

Via the correspondence

$$(u_1, u_2) \mapsto (\nu_1 u_1 + \nu_2 u_2) N(\mathfrak{d}(f)^{-1}) = \mathfrak{F}(u_1, u_2)$$

$\mathfrak{R}(X; n)$ corresponds to a set of ideals. For a given integer mn we then see that each element (u_1, u_2) of $\mathfrak{R}(X; mn)$ corresponds to a set of ideal factorizations of the form

$$\mathfrak{m}\mathfrak{n} = \mathfrak{F}(u_1, u_2) \tag{4-6}$$

with $N(\mathfrak{m}) = m$, $N(\mathfrak{n}) = n$. Now associate to $\mathfrak{m}$, $\mathfrak{n}$ ideal numbers $m^*, n^* \in \mathcal{L}_0$. Then (4-6) can be interpreted as multiplication in the set of ideal numbers. To make this concrete, first choose $\{w_1, w_2\}$ to be a basis for the ideal class $\mathrm{Cl}\,\mathfrak{d}(f)^{-1}$ such that $w_1 \psi_f^{-1} = z\nu_1$ and $w_2 \psi_f^{-1} = \nu_2$ for some integer z.

For each pair of ideal classes A, $B = A^{-1} \operatorname{Cl} f$ and any bases $\{a_1, a_2\}$, $\{b_1, b_2\}$ of A, B, respectively, we have a composition law

$$(a_1x_1 + a_2x_2)(b_1y_1 + b_2y_2) = \psi_f^{-1}(w_1 R_{A,B}(\boldsymbol{x};\, \boldsymbol{y}) + w_2 Q_{A,B}(\boldsymbol{x};\, \boldsymbol{y})).$$

By our choice of $\{w_1, w_2\}$ this is equivalent to

$$(a_1x_1 + a_2x_2)(b_1y_1 + b_2y_2) = z\nu_1 R_{A,B}(\boldsymbol{x};\, \boldsymbol{y}) + \nu_2 Q_{A,B}(\boldsymbol{x};\, \boldsymbol{y}).$$

This gives a bilinear mapping

$$\Phi_{A,B} : (\mathcal{L}_0 \cap A) \times (\mathcal{L}_0 \cap B) \to \{(x, y) \in \mathbb{R}^2 : y \in I(X)\},$$
$$\Phi_{A,B}(m_1, m_2; n_1, n_2) = (R_{A,B}(\boldsymbol{m}; \boldsymbol{n}),\, Q_{A,B}(\boldsymbol{m}; \boldsymbol{n}))$$

say. Let us write $A_0 = A \cap \mathcal{L}_0$ and $B_0 = B \cap \mathcal{L}_0$ for convenience. We then have

$$c_{mn} = \sum_{A\cdot B = \operatorname{Cl} f} \sum_{\substack{\boldsymbol{m}\in A,\, \boldsymbol{n}\in B \\ N(\boldsymbol{m})=m,\, N(\boldsymbol{n})=n \\ Q_{A,B}(\boldsymbol{m};\boldsymbol{n})\in I(X)}} \sum \Upsilon(Q_{A,B}(\boldsymbol{m}; \boldsymbol{n})). \tag{4-7}$$

This is the desired analogue to equation (5.2) in [5]. We summarize this below:

Proposition 4.1. *For $\mathcal{C} = \mathcal{A}, \mathcal{B}, \mathcal{A}^{\spadesuit}, \mathcal{B}^{\spadesuit}$, (4-7) holds.*

5. Type-I estimates

We will establish the necessary type-I estimate we need, following the work of Friedlander and Iwaniec in [6]. For this section, we shall put $\lambda(\ell)$ to be any function bounded by one supported on r-th powers of integers, and put

$$a_n = \sum_{\substack{f(\ell,m)=n \\ \ell\in I(X)}} \lambda(\ell). \tag{5-1}$$

We recall that

$$A_d(X) = \sum_{\substack{n\le X \\ n\equiv 0 \,(\operatorname{mod} d)}} a_n.$$

For a given positive integer ℓ put

$$\mathcal{I}(\ell;\, X) = \{x \in \mathbb{R}^2 : 0 < f(x, \ell) < X\}$$

and $\iota(\ell;\, X)$ to be the length of $\mathcal{I}(\ell;\, X)$. We then expect $A_d(X)$ to be well-approximated by

$$M_d(X) = \frac{\rho_f(d)}{d} \sum_{\substack{\ell\in I(X) \\ \gcd(\ell,d)=1}} \lambda(\ell) \frac{\varphi(\ell)}{\ell} \iota(\ell;\, X),$$

where φ is the Euler totient function and $\rho_f(d)$ is the number of solutions to the congruence

$$f(x, 1) \equiv 0 \pmod d. \tag{5-2}$$

Our goal is to establish:

Proposition 5.1. *Suppose that λ is supported on r-th powers. Then uniformly for $X^{1/2} \le D \le X^{(r+1)/(2r)}$ we have*

$$\sum_{d \le D} |A_d(X) - M_d(X)| \ll D^{1/4} X^{3(r+1)/(8r)} (\log X)^{24}.$$

As usual, our starting point is the following result from [1], which states that the roots of quadratic congruences are separated as much as possible:

Proposition 5.2 [1, Proposition 3]. *Let $F(x, y) = \alpha x^2 + \beta xy + \gamma y^2 \in \mathbb{Z}[x, y]$ be an arbitrary binary quadratic form whose discriminant is not a perfect square. For any sequence (α_n) of complex numbers and positive real numbers D, N we have*

$$\sum_{D \le d \le 2D} \sum_{F(1,\nu) \equiv 0 \,(\mathrm{mod}\, d)} \left| \sum_{n \le N} \alpha_n e\left(\frac{\nu n}{d}\right) \right|^2 \ll_F (D+N) \sum_n |\alpha_n|^2.$$

It is the fact that such a strong large sieve inequality exists for roots of quadratic congruences that enables such powerful results to be proved about thin variables as in [3; 5; 10]. We show how to derive the type-I estimates we need by following the same steps carried out in [6; 12]. We first replace $A_d(X)$, $M_d(X)$ with their smooth counterparts. Consider an auxiliary smooth function $\phi : \mathbb{R} \to \mathbb{R}$ satisfying

(1) $\phi(u) = 1$ if $0 < u \le X - Y$;

(2) $\phi^{(j)}(u) \ll Y^{-j}$ for $j \ge 0$; and

(3) $\phi(u) = 0$ if $u \ge X$.

Here $X^{7/8} \le Y \le X$ will be chosen later. We then introduce (by abuse of notation)

$$A_d(\phi) = \sum_{n \equiv 0 \,(\mathrm{mod}\, d)} a_n \phi(n) \tag{5-3}$$

and

$$M_d(\phi) = \frac{\rho_f(d)}{d} \sum_{\gcd(\ell, d)=1} \lambda(\ell) \frac{\varphi(d)}{d} \int_0^\infty \phi(f(\ell, t))\, dt. \tag{5-4}$$

We estimate the differences by elementary means as follows. Note that

$$\sum_{d \le D} |A_d(X) - A_d(\phi)| \le \sideset{}{'}\sum_{\substack{X - Y < f(m,\ell) \le X \\ \gcd(\ell, m)=1}} \lambda(\ell) \tau(f(m, \ell)) + O\big(\sqrt{X} \log X\big),$$

where $\sum'$ means that the terms with a value of ℓ closest to $\sqrt{X}$ are omitted. We then have the following consequence of Landreau's inequality [13], resulting in the bound

$$\sideset{}{'}\sum_{\ell \ll \sqrt{X}} \sum_{\substack{d \le X^{1/4} \\ \gcd(d, \ell)=1}} \tau(d)^8 \sum_{\substack{X - Y < f(m,\ell) \le X \\ f(m,\ell) \equiv 0 \,(\mathrm{mod}\, d)}} 1.$$

The conditions

$$X - Y < f(m,\ell) \le X \quad \text{and} \quad \ell \in I(X)$$

imply that m is restricted to an interval of length $O_f\big(Y/\sqrt{X+\ell^2}\big)$. Splitting into residue classes $m \equiv \alpha\ell \pmod d$ with α running over the roots of (5-2) we see that the above sum is bounded by

$$O\biggl(Y\biggl(\sum_{d\le X^{1/4}} \tau(d)^8 \frac{\rho_f(d)}{d}\biggr)\biggl(\sideset{}{'}\sum_{\ell\ll\sqrt{X}} |\lambda(\ell)|(X+\ell^2)^{-1/2}\biggr) + X^{1/4+1/(2r)}(\log X)^{256}\biggr).$$

We have the bounds

$$\sum_{d\le X^{1/4}} \tau(d)^8 \frac{\rho_f(d)}{d} \ll (\log X)^{256}$$

and

$$\begin{aligned}\sideset{}{'}\sum_{\ell\ll\sqrt{X}} |\lambda(\ell)|(X+\ell^2)^{-1/2} &\le \sideset{}{'}\sum_{k\ll X^{1/2r}} (X+k^{2r})^{-1/2}\\ &\ll X^{(1-2r)/(4r)} \sideset{}{'}\sum_{k\ll X^{1/2r}} (X^{1/(2r)}+k)^{-1/2}\\ &\ll X^{(1-2r+1)/(4r)} = X^{(1-r)/(2r)}.\end{aligned}$$

It follows that

$$\sum_{d\le D} |A_d(X) - A_d(\phi)| \ll Y X^{(1-r)/(2r)} (\log X)^{256}. \tag{5-5}$$

Similarly, we obtain

$$\sum_{d\le D} |M_d(X) - M_d(\phi)| \ll Y X^{(1-r)/(2r)} (\log X)^{256}. \tag{5-6}$$

We then proceed to decompose $A_d(\phi)$ as

$$\begin{aligned} A_d(\phi) &= \sum_{\substack{f(m,\ell)\equiv 0 \,(\mathrm{mod}\, d)\\ \gcd(\ell,m)=1}}\sum \lambda(\ell)\phi(f(m,\ell))\\ &= \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\, d)} \sum_{\ell} \lambda(\ell) \sum_{\substack{m\equiv\alpha\ell\,(\mathrm{mod}\, d)\\ \gcd(\ell,m)=1}} \phi(f(\ell,m))\\ &= \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\, d)} \sum_{a} \mu(a) \sum_{\ell} \lambda(a\ell) \sum_{m\equiv\alpha\ell\,(\mathrm{mod}\, d/\gcd(a,d))} \phi(a^2 f(m,\ell)), \end{aligned} \tag{5-7}$$

where we applied Möbius inversion to the inner sum to remove the awkward coprimality condition. We then apply Poisson's formula to the inner sum to obtain

$$\sum_{m\equiv\alpha\ell\,(\mathrm{mod}\, d/\gcd(a,d))} \phi(a^2 f(m,\ell)) = \frac{\gcd(a,d)}{d} \sum_{h\in\mathbb{Z}} e\biggl(\alpha h\ell \frac{\gcd(a,d)}{d}\biggr) \Phi_{a\ell}\biggl(\frac{h\gcd(a,d)}{d}\biggr),$$

where $\Phi_{a\ell}(v)$ is the Fourier integral

$$\Phi_{a\ell}(v) = \int_{-\infty}^{\infty} \phi(a^2 f(\ell, t)) e(-vt)\, dt. \tag{5-8}$$

The zero-frequency $h = 0$ gives exactly $M_d(\phi)$. Integration by parts yields

$$\Phi_{a\ell}(v) = (2\pi i v)^{-j} \int_{-\sqrt{X}/a}^{\sqrt{X}/a} e(-vt) \frac{\partial^j}{\partial t^j} \phi(a^2 f(\ell, t))\, dt.$$

Using our hypotheses on ϕ we estimate

$$\frac{\partial^j}{\partial t^j} \phi(a^2 f(\ell, t)) \ll \left(\frac{\sqrt{X}}{aY} \right)^j.$$

It follows that

$$\Phi_{a\ell}(v) \ll \frac{\sqrt{X}}{a} \left(\frac{\sqrt{X}}{aYv} \right)^j. \tag{5-9}$$

Now, for $R_d(\phi) = A_d(\phi) - M_d(\phi)$ we obtain from (5-7) that

$$\begin{aligned} R_d(\phi) &= \sum_{f(\alpha,1) \equiv 0 \,(\mathrm{mod}\, d)} \sum_a \mu(a) \sum_\ell \lambda(a\ell) \frac{\gcd(a,d)}{d} \sum_{h \neq 0} e\left(\alpha h \ell \frac{\gcd(a,d)}{d} \right) \Phi_{a\ell}\left(\frac{h \gcd(a,d)}{d} \right) \\ &= \frac{2}{d} \sum_a \mu(a) \sum_{\substack{bc = d \\ b \mid a}} b \sum_{\substack{\alpha \,(\mathrm{mod}\, bc) \\ f(\alpha,1) \equiv 0 \,(\mathrm{mod}\, bc)}} \sum_{h > 0} \sum_\ell \lambda(a\ell) e\left(\frac{\alpha h \ell}{c} \right) \Phi_{a\ell}\left(\frac{h}{c} \right) \\ &= \frac{2}{d} \sum_a \mu(a) \sum_{\substack{bc = d \\ b \mid a}} b \rho_f(b) \sum_{\substack{\alpha \,(\mathrm{mod}\, c) \\ f(\alpha,1) \equiv 0 \,(\mathrm{mod}\, c)}} \sum_{h > 0} \sum_\ell \lambda(a\ell) e\left(\frac{\alpha h \ell}{c} \right) \Phi_{a\ell}\left(\frac{h}{c} \right). \end{aligned} \tag{5-10}$$

Applying (5-9) for $h \geq a^{-1} Y^{-1} D X^{1/2 + \psi(r)} = H$ for some small $\psi(r) > 0$ and choosing $j = j(r)$ sufficiently large, we may assume that $\Phi_{a\ell}(h/c) \ll h^{-2} D^{-1}$. Bounding absolutely we then conclude that the tail is bounded by

$$O(\rho_f(d) d^{-1} \|\lambda\|_1).$$

Since $|a\ell| \ll \sqrt{X}$, we thus conclude that

$$\|\lambda\|_1 \ll X^{1/(2r)} \quad \text{and} \quad \sum_{D \leq d < 2D} \frac{\rho_f(d)}{d} \|\lambda\|_1 \ll X^{1/(2r)} \log D, \tag{5-11}$$

which is sufficiently small. To handle the remaining range, we apply a change of variables to obtain

$$\Phi_{a\ell}\left(\frac{h}{c} \right) = \frac{2\sqrt{X}}{ah} \int_0^\infty \phi\left(a^2 f\left(\ell, \frac{s\sqrt{X}}{ah} \right) \right) e\left(-\frac{s\sqrt{X}}{ac} \right) ds.$$

The integrand vanishes unless

$$\ell \ll \frac{\sqrt{X}}{a} \quad \text{and} \quad h \gg s.$$

It follows that

$$d\,|R_d(\phi)| \ll {\sum_a}^{\flat} \frac{\sqrt{X}}{a} \sum_{\substack{bc=d\\ b|a}} b\rho_f(b) \int_0^H \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\,c)} \biggl| \sum_{\substack{s<h<H\\ |\ell|\ll\sqrt{X}/a}} h^{-1}\lambda(a\ell)\phi\Bigl(a^2 f\Bigl(\ell, \frac{s\sqrt{X}}{ah}\Bigr)\Bigr) e\Bigl(\frac{\alpha h\ell}{c}\Bigr)\biggr| \, ds.$$

We reorganize the inner sum as

$$\sum_{\substack{s<h<H\\ |\ell|\ll\sqrt{X}/a}} h^{-1}\lambda(a\ell)\phi\Bigl(a^2 f\Bigl(\ell, \frac{s\sqrt{X}}{ah}\Bigr)\Bigr) e\Bigl(\frac{\alpha h\ell}{c}\Bigr)$$
$$= \sum_{n\ll H\sqrt{X}/a} \biggl(\sum_{\substack{h\ell=n\\ s<h<H}} \frac{1}{h}\lambda(a\ell)\phi\Bigl(a^2 f\Bigl(\ell, \frac{s\sqrt{X}}{ah}\Bigr)\Bigr)\biggr) e\Bigl(\frac{\alpha n}{c}\Bigr) = \sum_{n\ll H\sqrt{X}/a} \xi_n(s) e\Bigl(\frac{\alpha n}{c}\Bigr),$$

say. Next we write

$$\sum_{D\le d<2D} |R_d(\phi)| \ll O(X^{1/(4r)}\log D)$$
$$+ \frac{1}{D}\int_0^H \biggl({\sum_a}^{\flat} \frac{\sqrt{X}}{a} \sum_{b|a} \rho_f(b) b \sum_{D/b\le c<2D/b} \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\,c)} \biggl| \sum_{n\ll H\sqrt{X}/a} \xi_n(s) e\Bigl(\frac{\alpha n}{c}\Bigr)\biggr| \biggr)\, ds.$$

Applying Cauchy–Schwarz we obtain

$$\sum_{C\le c<2C} \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\,c)} \biggl| \sum_{n\ll H\sqrt{X}/a} \xi_n(s) e\Bigl(\frac{\alpha n}{c}\Bigr)\biggr|$$
$$\le C^{1/2} \biggl(\sum_{C\le c<2C} \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\,c)} \biggl| \sum_{n\ll H\sqrt{X}/a} \xi_n(s) e\Bigl(\frac{\alpha n}{c}\Bigr)\biggr|^2 \biggr)^{1/2}$$
$$\ll C^{1/2}\bigl(C + H\sqrt{X}/a\bigr)^{1/2} \|\xi\|_2 \tag{5-12}$$

by Proposition 5.2. Next we note that

$$\|\xi(s)\|_2^2 \le \frac{1}{s^2} \sum_{n\ll H\sqrt{X}/a} \biggl(\sum_{\substack{h\ell=n\\ s\le h<H}} \lambda(a\ell)\biggr)^2$$

It follows that

$$\sum_{D\le d<2D} |R_d(\phi)| \ll$$
$$\frac{\sqrt{X}}{D} {\sum_a}^{\flat} \frac{1}{a} \sum_{b|a} \rho_f(b) b \Bigl(\frac{D}{b}\Bigr)^{1/2} \Bigl(\frac{D}{b} + \frac{H\sqrt{X}}{a}\Bigr)^{1/2} \biggl| \sum_{n\ll H\sqrt{X}/a} \biggl(\sum_{\substack{h\ell=n\\ H\le h<2H}} \lambda(a\ell)\biggr)^2 \biggr|^{1/2} \int_0^H \frac{1}{s}\, ds. \tag{5-13}$$

Next we evaluate

$$\sum_{n \ll H\sqrt{X}/a} \Bigg(\sum_{\substack{h\ell = n \\ 0<h<H}} \lambda(a\ell) \Bigg)^2.$$

Since a is square-free and $a\ell$ is an r-th power, it follows that $\ell = a^{r-1}m^r$ with $m \le a^{-1}X^{1/(2r)} = M$, say. Therefore we see that the sum above is bounded by the number of solutions to

$$h_1 m_1^r = h_2 m_2^r$$

with $H \le h_1, h_2 < 2H$ and $m_1, m_2 \le M$. The solutions are parametrized by $m_1 = st_1, m_2 = st_2$ with $\gcd(t_1, t_2) = 1$, $st_1, st_2 \le M$. Observe that

$$s \le \frac{M}{\max(t_1, t_2)} \quad \text{and} \quad k \le \frac{2H}{\max(t_1^r, t_2^r)}.$$

This gives

$$\sum_n \Bigg(\sum_{\substack{h\ell = n \\ 0<h<H}} \lambda(a\ell) \Bigg)^2 \ll HM \sum_{t_1 \le t_2 \le M} \frac{1}{t_2^{r+1}} \ll HM \log M.$$

We thus obtain the upper bound of $O(Ha^{-1}X^{1/(2r)}(\log X))$. Inserting this into (5-13) and summing gives

$$\sum_{D \le d < 2D} |R_d(\phi)| \ll D^{-1/2}\big(D + H\sqrt{X}\big)^{1/2} H^{1/2} X^{1/2+1/(4r)} (\log X)^2. \tag{5-14}$$

Inserting

$$H \le DY^{-1}X^{1/2+\psi(r)}$$

into (5-14) gives the bound

$$\begin{aligned} \sum_{D \le d < 2D} |R_d(\phi)| &\ll X^{1/2} D^{-1} D^{1/2} H^{1/2} X^{1/4} H^{1/2} X^{1/(4r)} (\log X)^2 \\ &\ll_\varepsilon D^{-1/2}(Y^{-1}DX^{1/2+\psi(r)}) X^{3/4+1/(4r)} (\log X)^2 \\ &= D^{1/2} Y^{-1} X^{(5r+1)/(4r)+\psi(r)} (\log X)^2. \end{aligned} \tag{5-15}$$

This bound holds uniformly for $d \le D$. We may thus choose

$$Y = D^{1/4} X^{(7r-1)/(8r)-\psi(r)} (\log X)^{-26}.$$

This in turn gives the estimate

$$\sum_{d \le D} |A_d(\phi) - M_d(\phi)| \ll D^{1/4} X^{3(r+1)/(8r)} (\log X)^{24},$$

which is enough to prove Proposition 5.1.

6. Estimating $\pi(\mathcal{B})$: bilinear sum bounds

We will deal with the sum

$$P(X) = \sum_{n \le X} b_n \Lambda(n)$$

in the case of $\pi(\mathcal{B})$ via Vaughan's identity [18], which is an elegant combinatorial identity which decomposes the von Mangoldt function. The ideas recorded here are from [3]. Suppose $Y, Z \ge 1$ and suppose $n > Z$. Then

$$\Lambda(n) = \sum_{\substack{m|n \\ m \le Y}} \mu(m) \log \frac{n}{m} - \sum_{\substack{mc|n \\ m \le Y \\ c \le Z}} \mu(m)\Lambda(c) + \sum_{\substack{mc|n \\ b > Y \\ c > Z}} \mu(m)\Lambda(c) \tag{6-1}$$

and if $n \le Z$, the right hand side is zero. For $X > YZ$ then Vaughan's identity implies that

$$\begin{aligned}
P(X) &= P(Z) + \sum_{n \le X} b_n \Bigg(\sum_{\substack{m|n \\ m \le Y}} \mu(m) \log \frac{n}{m} - \sum_{\substack{mc|n \\ b \le Y \\ c \le Z}} \mu(m)\Lambda(c) + \sum_{\substack{mc|n \\ m > Y \\ c > Z}} \mu(m)\Lambda(c) \Bigg) \\
&= P(Z) + \sum_{m \le Y} \mu(m) \Bigg(\sum_{\substack{n \le X \\ m|n}} b_n \log n - \sum_{\substack{n \le X \\ m|n}} b_n \log m - \sum_{c \le Z} \Lambda(c) \sum_{\substack{n \le X \\ mc|n}} b_n \Bigg) + \sum_{m > Y} \mu(m) \sum_{c > Z} \Lambda(c) \sum_{\substack{n \le X \\ mc|n}} b_n \\
&= P(Z) + A(X; Y, Z) + \sum_{\substack{md \le X \\ m > Y}} \mu(m) \Bigg(\sum_{\substack{c|d \\ c > Z}} \Lambda(c) \Bigg) b_{md} \\
&= P(Z) + A(X; Y, Z) + B(X; Y, Z),
\end{aligned}$$

say. We can treat $P(Z)$ by applying trivial bounds provided that Z is sufficiently small with respect to X. The term $A(X; Y, Z)$ can be dealt with using the appropriate type-I estimates; see Proposition 5.1. The term $B(X; Y, Z)$, as expected, will require some type-II estimates. Given our treatment of the algebraic aspects of bilinear sums in Section 4, the treatment below is very similar to that given in [3; 12] so we will be fairly terse on the details.

Our target is the estimate

$$B(X; Y, Z) \ll \Delta X (\log X)^5,$$

with $\Delta = (\log X)^{-A}$ for any large, fixed $A > 5$. Recall that

$$B(X; Y, Z) = \sum_{Z < d < X/Y} \Bigg(\sum_{\substack{c|d \\ c > Z}} \Lambda(c) \Bigg) \sum_{Y < m \le X/d} \mu(m) b_{md}.$$

Using the trivial estimate

$$\sum_{\substack{c|d \\ c > Z}} \Lambda(c) \le \log X$$

we then find that

$$|B(X;Y,Z)| \le (\log X)\sum_{d>Z}\left|\sum_{Y<m\le X/d}\mu(m)b_{md}\right|.$$

We wish to break the sum into short sums of the shape

$$\mathcal{B}(M,N)=\sum_{M<m\le 2M}\left|\sum_{N<n\le N'}\mu(n)b_{mn}\right| \tag{6-2}$$

with $N'=e^{k\Delta}N$. Considering $M=2^jZ$ and $N=e^{\Delta k}y$ for various j,k, we then see that

$$|\mathcal{B}(X;Y,Z)| \le (\log X)\sum_{\substack{\Delta X<MN<X\\ M\ge Z\\ N\ge Y}}\sum \mathcal{B}(M,N)+O(\Delta X(\log X)^2), \tag{6-3}$$

where the error term $O(\Delta X(\log X)^2)$ represents a trivial bound for the contribution of $\mu(m)b_{md}$ with $md\le 2\Delta X$ or $e^{-2\Delta}X<md\le X$, where the terms are not covered exactly. There are at most $2\Delta^{-1}(\log X)^2$ short sums $\mathcal{B}(M,N)$ in (6-3) so it suffices to show that

$$\mathcal{B}(M,N)\ll \Delta^2X(\log X)^2 \tag{6-4}$$

for all M,N in the relevant range. We have a trivial bound

$$\mathcal{B}(M,N)\le \sum_{M<m\le 2M}\rho_f(m)\sum_{N<n\le N'}\rho_f(n)\ll \Delta MN,$$

and we can use this bound to obtain

$$\mathcal{B}(M,N)\le \sum_{d\le\Delta^{-1}}\mathcal{B}_d(M,N)+O(\Delta^2X),$$

where $\mathcal{B}_d(M,N)$ consists of the subsum of $\mathcal{B}(M,N)$ where $\gcd(m,n)=d$. The error term $O(\Delta^2X)$ comes from the trivial bound and the condition $d>\Delta^{-1}$. Next observe that

$$\mathcal{B}_d(M,N)\le \mathcal{B}_1(dM,N/d),$$

and so it suffices to show

$$\mathcal{B}_1(M,N)\ll \Delta^3X(\log X)^2 \tag{6-5}$$

for M,N satisfying $M\ge Z$, $N\ge\Delta Y$ and $\Delta X<MN<X$.

Applying (4-7) to (6-2) we then obtain

$$\mathcal{B}_1(M,N)\le \sum_{A\cdot B=\mathrm{Cl}\,f}\ \sum_{\substack{\boldsymbol{m}\in A\\ M<N(J(\boldsymbol{m}))\le 2M}}\left|\sum_{\substack{\boldsymbol{n}\in B\\ N<N(J(\boldsymbol{n}))\le N'\\ \gcd(N(J(\boldsymbol{n})),N(J(\boldsymbol{m})))=1}}\mu(N(J(\boldsymbol{n})))\Lambda(Q_{A,B}(\boldsymbol{m};\boldsymbol{n}))\right|.$$

Removing the coprimality condition via Möbius inversion as in [3; 12], as well as partitioning the sum $\mathcal{B}_1(M,N)$ based on the classes A, B, it suffices to show that the sums

$$\mathcal{C}_r(M,N)=\sum_{\substack{M<g_1(x_1,x_2)\le 2M\\(x_1,x_2)\in\mathcal{K}_1}}\Bigg|\sum_{\substack{N<g_2(y_1,y_2)\le N'\\(y_1,y_2)\in\mathcal{K}_2}}\mu(rg_2(y_1,y_2))\Lambda(Q(x_1,x_2;y_1,y_2))\Bigg| \tag{6-6}$$

are bounded by $O(\Delta^5X(\log X)^2)$ for every r, M, N satisfying

$$r<\Delta^{-2},\quad M\ge Z,\quad N\ge\Delta^3Y\quad\text{and}\quad \Delta X<MN<X$$

and $\mathcal{K}_1,\mathcal{K}_2$ domains which are contained in $[-CX,CX]^2$ for some absolute constant C depending only on our choices of fundamental domains.

If we write

$$Q(x_1,x_2;y_1,y_2)=x_1\ell_1(y_1,y_2)+x_2\ell_2(y_1,y_2)$$

for linear forms $\ell_1,\ell_2\in\mathbb{Z}[x,y]$ then the condition that $Q(\boldsymbol{x};\boldsymbol{y})=0$ implies that $(\ell_1(y_1,y_2),\ell_2(y_1,y_2))$ is proportional to $(-x_2,x_1)$. We then make a change of variables in the inner sum, obtaining

$$\mathcal{C}_r(M,N)=\sum_{\substack{M<g_1(x_1,x_2)\le 2M\\(x_1,x_2)\in\mathcal{K}_1}}\Bigg|\sum_{\substack{N<g_2^*(z_1,z_2)\le N'\\(y_1,y_2)\in\mathcal{K}_2}}\mu(rg_2^*(z_1,z_2))\Lambda(x_1z_1+x_2z_2)\Bigg|,$$

where $z_i=\ell_i(y_1,y_2)$ and g_2^* is such that $g_2^*(z_1,z_2)=g_2(y_1,y_2)$. We are then left with the bilinear sum

$$C(\alpha,\beta;\lambda)=\sum_{\boldsymbol{z}}{\sum_{\boldsymbol{w}}}^{*}\alpha(\boldsymbol{z})\beta(\boldsymbol{w})\lambda(\mathcal{Q}(\boldsymbol{z};\boldsymbol{w})), \tag{6-7}$$

where α is supported in a disk of radius R_1 and β supported on an annulus $\mathbb{A}(R_2,2R_2)$ having inner radius R_2 and outer radius $2R_2$, say. Further, we assume that λ is supported on $|\ell|\le CAB$ for some absolute constant C depending only on f, so in particular the ℓ^2-norm of λ is finite. Applying the Cauchy–Schwarz inequality we obtain

$$|\mathcal{C}(\alpha,\beta;\lambda)|\le\sum_{\ell}|\lambda(\ell)|{\sum_{\boldsymbol{y}}}^{*}|\beta(\boldsymbol{y})|\Bigg|\sum_{\mathcal{Q}(\boldsymbol{x};\boldsymbol{y})=\ell}\alpha(\boldsymbol{x})\Bigg|\le\|\lambda\|_2\cdot\|\beta\|_2\mathcal{D}(\alpha)^{1/2}, \tag{6-8}$$

where $\|\cdot\|_2$ denotes the ℓ^2-norm and

$$\mathcal{D}(\alpha)={\sum_{\boldsymbol{y}}}^{*}\mathcal{G}(\boldsymbol{y})\sum_{\ell}\Bigg|\sum_{\mathcal{Q}(\boldsymbol{x};\boldsymbol{y})=\ell}\alpha(\boldsymbol{x})\Bigg|^2,$$

where $\mathcal{G}$ is any nonnegative function with $\mathcal{G}(\boldsymbol{y})\ge 1$ on the annulus $\mathbb{A}(R_2,2R_2)$. As in [3; 12] it will be convenient to suppose that $\mathcal{G}$ is a radial, compactly supported, and smooth function. Squaring out we obtain

$$\mathcal{D}(\alpha)={\sum_{\boldsymbol{y}}}^{*}\mathcal{G}(\boldsymbol{y})\sum_{\mathcal{Q}(\boldsymbol{x};\boldsymbol{y})=0}(\alpha*\alpha)(\boldsymbol{x}), \tag{6-9}$$

with

$$(\alpha * \alpha)(\boldsymbol{x}) = \sum_{\boldsymbol{u}-\boldsymbol{v}=\boldsymbol{x}} \alpha(\boldsymbol{u})\bar{\alpha}(\boldsymbol{v}).$$

Note that

$$(\alpha * \alpha)(0) = \|\alpha\|_2^2.$$

The orthogonality relation $\boldsymbol{x} \cdot \boldsymbol{y} = 0$ for a primitive $\boldsymbol{x}$ in (6-9) is equivalent to the statement that $\boldsymbol{y}$ is a rational integer multiple of $\boldsymbol{x}' = (-x_2, x_1)$. It follows that

$$\mathcal{D}(\alpha) = \sum_{c \in \mathbb{Z}} \sum_{\boldsymbol{y}}{}^{*} \mathcal{G}(\boldsymbol{y})(\alpha * \alpha)(c\boldsymbol{y}) = \mathcal{D}_0(\alpha) + 2\mathcal{D}^*(\alpha), \tag{6-10}$$

where $\mathcal{D}_0(\alpha)$ denotes the contribution with $c = 0$ and $\mathcal{D}^*(\alpha)$ that of all $c > 0$. Thus

$$\mathcal{D}_0(\alpha) = \|\alpha\|_2^2 \sum_{\boldsymbol{y}}{}^{*} \mathcal{G}(\boldsymbol{y}) \ll \|\alpha\|_2^2 B^2$$

and

$$\mathcal{D}^*(\alpha) = \sum_{\boldsymbol{x} \neq \boldsymbol{0}} \mathcal{G}(\boldsymbol{x}^*)(\alpha * \alpha)(z),$$

where $\boldsymbol{x}^*$ is a primitive vector proportional to $\boldsymbol{x}$. Again, we may apply Möbius inversion to remove the primitivity conditions, and obtain

$$\mathcal{D}^*(\alpha) = \sum_{b,c>0} \sum \mu(b)\mathcal{D}(\alpha; bc)$$

where

$$\mathcal{D}(\alpha; bc) = \sum_{\boldsymbol{x} \equiv 0 \,(\mathrm{mod}\, bc)} \mathcal{G}(c^{-1}\boldsymbol{x})(\alpha * \alpha)(\boldsymbol{x}).$$

From here, the treatment is identical to the one given in [3; 12] as no structure of the Gaussian integers or even an imaginary quadratic field is necessary. This completes our treatment for $\pi(\mathcal{B})$.

7. Type-II estimates for $\pi(\mathcal{A}) - \pi(\mathcal{B})$: preliminary steps

We discuss the proof of Proposition 3.7. We note that Proposition 3.7 is exactly analogous to Proposition 5 in [10], though our sequences $\mathcal{A}, \mathcal{B}$ are different. We have largely divorced the arithmetic of our field K with the analysis of bilinear sums in Section 4, and so we are in good shape to import results from [10] directly. We will make clear which components of [10] can be used without change, and where we need to make suitable modifications.

We substitute (4-7) into (3-21) to obtain

$$\sum_{N<n\le 2N} \sum_{m<X/N} \alpha_m \beta_n (a_{mn} - b_{mn}) = \sum_{A \cdot B = \mathrm{Cl}\, f} \sum_{\substack{w \in A_0 \\ N < N(J(w)) \le 2N}} \beta_w \sum_{\substack{v \in B_0 \\ N(J(v)) < X/N}} \alpha_v (\mathfrak{Z}(Q_{A,B}(v,w)) - \Lambda(Q_{A,B}(v,w)), \tag{7-1}$$

where $\alpha_v = \alpha_{N(J(v))}$, $\beta_w = \beta_{N(J(w))}$. Writing each bilinear form Q above as $w_1\ell_1(v_1, v_2) + w_2\ell_2(v_1, v_2)$ say and applying a linear change of variables to the inner sum, we transform the inner sum

$$\sum_{\substack{v \in B_0 \\ N(J(v)) < X/N}} \alpha_v(\mathfrak{Z}(Q(v, w)) - \Lambda(Q(v, w)) = \sum_z \alpha_z(\mathfrak{Z}(w_1z_1 + w_2z_2) - \Lambda(w_1z_1 + w_2z_2))$$

say, with the support of z being the image of the support of the sum on the left under the linear transformation. The linear transformation depends only on Q and not X.

After applying these linear transformations, we have now changed all of our bilinear forms Q to

$$Q_0(x_1, x_2; y_1, y_2) = x_1y_1 + x_2y_2.$$

We write $\mathcal{S}_1(X) \times \mathcal{S}_2(X)$ for the union of the images of the supports of w, v in (7-1), so that (7-1) becomes

$$h(K) \sum_{\boldsymbol{w} \in \mathcal{S}_1(X)} \sum_{\boldsymbol{v} \in \mathcal{S}_2(X)} \alpha_{\boldsymbol{w}}\beta_{\boldsymbol{v}}(\mathfrak{Z}(w_1v_1 + w_2v_2) - \Lambda(w_1v_1 + w_2v_2)). \tag{7-2}$$

Remark 7.1. Since the linear transformations depend only on the class $1 \le j \le h(K)$ and the corresponding choice of fundamental domain, the image of the set $\mathcal{F}_j(X)$ with $N < N(J(w)) \le 2N$ is contained in the annulus $\mathbb{A}(c_1N, c_2N)$ for some positive numbers c_1, c_2 independent of N. Similarly, the image of $\mathcal{F}'_j(X)$ with $N(J(v)) \le X/N$ is contained in the disk $\mathbb{D}(c_3X/N)$ for some $c_3 > 0$ depending at most on f. This observation is crucial because we will use the Euclidean norm and the corresponding geometry to treat our sums when we wish to import estimates from [5; 10], and switch to using the norm on $\mathcal{O}_K$ and the corresponding induced norm on ideal numbers when the arithmetic of K is relevant.

Since we are looking to save an arbitrary power of log, it suffices to further subdivide the support of (7-2), and consider sums of the shape

$$\sum_{\substack{\boldsymbol{w} \\ N < \|\boldsymbol{w}\|_2 \le 2N}} \alpha_{\boldsymbol{w}} \sum_{\substack{z \\ \|z\|_2 \le X/N}} \beta_z(\mathfrak{Z}(w_1z_1 + w_2z_2) - \Lambda(w_1z_1 + w_2z_2)).$$

Remark 7.2. We abuse notation and refer to the terms β_n for some positive integer n as well as β_z for some vector $z \in \mathbb{Z}^2$. In the former case we interpret the support of β_n to be a set of ideal numbers of $\mathcal{O}_K$ in a fixed class having norm equal to n, and in the latter we simply interpret the set of ideal numbers as a $\mathbb{Z}$-module.

Put

$$S_1(z, \boldsymbol{w}) = \sum_{\substack{p^2 \in I \\ w_1z_1 + w_2z_2 = p^2}} 2p \log p \quad \text{and} \quad S_2(z, \boldsymbol{w}) = \sum_{\substack{p \in I \\ w_1z_1 + w_2z_2 = p}} \log p \tag{7-3}$$

and

$$S_1^{\spadesuit}(z, \boldsymbol{w}) = \sum_{\substack{k^2 \in I(X) \\ w_1z_1 + w_2z_2 = k^2}} 2k \quad \text{and} \quad S_2^{\spadesuit}(z, \boldsymbol{w}) = \sum_{\substack{k \in I \\ w_1z_1 + w_2z_2 = k}} 1.$$

Our aim is to obtain the estimates

$$\sum_{\substack{\boldsymbol{w} \\ N<\|\boldsymbol{w}\|_2\le 2N}} \sum_{\substack{z \\ \|z\|_2\le X/N}} \alpha_{\boldsymbol{w}}\beta_z(S_1(z,\boldsymbol{w}) - S_2(z,\boldsymbol{w})) \ll_A \frac{X}{(\log X)^A} \tag{7-4}$$

and

$$\sum_{\substack{\boldsymbol{w} \\ N<\|\boldsymbol{w}\|_2\le 2N}} \sum_{\substack{z \\ \|z\|_2\le X/N}} \alpha_{\boldsymbol{w}}\beta_z(S_1^{\spadesuit}(z,\boldsymbol{w}) - S_2^{\spadesuit}(z,\boldsymbol{w})) \ll_A \frac{X}{(\log X)^A}.$$

We are almost ready to import the remaining argument from [10]. Let us put

$$\mathcal{R}(N;X) = \left\{ z \in \mathbb{Z}^2 : N \le \|z\|_2 < 2N, \left|\arg(z) - \frac{k\pi}{2}\right| \le (\log X)^{-A} \text{ for some } k \in \mathbb{Z} \right\}.$$

We note that, as we will use repeatedly later (and we will remind the reader of this again when this becomes relevant), that once we subdivide the regions into small dyadic ranges that the conditions $\|z\|_2 \sim N$ and $N(z) \sim N$ are nearly identical. Here $z = \hat{z}$ is the vector associated to z, viewed as an ideal number of K.

The following results from [10] can now be imported without change:

Lemma 7.3 [10, Lemma 9]. *Suppose that both z and q are fixed. Then the number of possible $\boldsymbol{w}$ with $q = w_1 z_1 + w_2 z_2$ is $O((M/N)^{1/2})$, where $M = X/N$.*

Lemma 7.4 [10, Lemma 10]. *We have*

$$\sum_{z \in \mathcal{R}(N;X)} \sum_{\boldsymbol{w}} \beta_z \alpha_{\boldsymbol{w}} S_j(z,\boldsymbol{w}) \ll_A X(\log X)^{-A}$$

for $j = 1, 2$.

We remark that Lemma 7.4 apply equally well with $S_j(z,\boldsymbol{w})$ replaced with $S_j^{\spadesuit}(z,\boldsymbol{w})$.

As is standard at this juncture (see [3; 5; 10]), we apply Cauchy–Schwarz to obtain

$$\left(\sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} \sum_{z} \beta_z(S_1(z,\boldsymbol{w}) - S_2(z,\boldsymbol{w}))\right)^2 \le \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}}^2 \sum_{\boldsymbol{w}} \left(\sum_{z} \beta_z(S_1(z,\boldsymbol{w}) - S_2(z,\boldsymbol{w}))\right)^2.$$

It is then sufficient to show that

$$\sum_{\boldsymbol{y},z} \beta_{\boldsymbol{y}}\beta_z \sum_{\boldsymbol{w}} (S_1^{\spadesuit}(\boldsymbol{y},\boldsymbol{w}) - S_2^{\spadesuit}(\boldsymbol{y},\boldsymbol{w}))(S_1^{\spadesuit}(z,\boldsymbol{w}) - S_2^{\spadesuit}(z,\boldsymbol{w})) \ll_A \frac{XN}{(\log X)^A} \tag{7-5}$$

and

$$\sum_{\boldsymbol{y},z} \beta_{\boldsymbol{y}}\beta_z \sum_{\boldsymbol{w}} (S_1(\boldsymbol{y},\boldsymbol{w}) - S_2(\boldsymbol{y},\boldsymbol{w}))(S_1(z,\boldsymbol{w}) - S_2(z,\boldsymbol{w})) \ll_A \frac{XN}{(\log X)^A} \tag{7-6}$$

for any $A > 0$.

We emphasize, as this will be relevant later, that the vectors $z, \boldsymbol{y}$ represent elements in the same ideal class.

Next we consider the diagonal contribution coming from $\boldsymbol{y} = z$. This gives the sums

$$\sum_z \beta_z \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} (S_1^{\spadesuit}(z, \boldsymbol{w}) - S_2^{\spadesuit}(z, \boldsymbol{w}))^2 = \sum_z \beta_z \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} (S_1^{\spadesuit}(z, \boldsymbol{w})^2 - 2S_1^{\spadesuit}(z, \boldsymbol{w}) S_2^{\spadesuit}(z, \boldsymbol{w}) + S_2^{\spadesuit}(z, \boldsymbol{w})^2)$$

and

$$\sum_z \beta_z \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} (S_1(z, \boldsymbol{w}) - S_2(z, \boldsymbol{w}))^2 = \sum_z \beta_z \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} (S_1(z, \boldsymbol{w})^2 - 2S_1(z, \boldsymbol{w}) S_2(z, \boldsymbol{w}) + S_2(z, \boldsymbol{w})^2).$$

Clearly,

$$S_1(z, \boldsymbol{w}) S_2(z, \boldsymbol{w}) = S_1^{\spadesuit}(z, \boldsymbol{w}) S_2^{\spadesuit}(z, \boldsymbol{w}) = 0$$

since their supports are incompatible. Next we have the trivial estimate

$$\begin{aligned}
\sum_z \sum_{\boldsymbol{w}} S_1(z, \boldsymbol{w}) &\ll \sum_{N < \|z\|_2 \le 2N} \sum_{p^2 \in I} p \log p \sum_{\substack{\boldsymbol{w} \\ w_1 z_1 + w_2 z_2 = p^2}} 1 \\
&\ll \sqrt{\frac{M}{N}} \sum_{p^2 \in I} p \log p \sum_{N \le \|z\|_2 < 2N} 1 \\
&\ll \sqrt{MN} \sum_{p^2 \in I} p \log p \\
&\ll_\varepsilon \sqrt{MN} X^{1/2+\varepsilon} \ll_\varepsilon X^{1+\varepsilon}.
\end{aligned}$$

Similarly, we conclude

$$\begin{aligned}
\sum_z \sum_{\boldsymbol{w}} S_2(z, \boldsymbol{w}) &\ll_\varepsilon X^{1+\varepsilon}, \\
\sum_z \sum_{\boldsymbol{w}} S_1^{\spadesuit}(z, \boldsymbol{w}) &\ll_\varepsilon X^{1+\varepsilon}, \\
\sum_z \sum_{\boldsymbol{w}} S_2^{\spadesuit}(z, \boldsymbol{w}) &\ll_\varepsilon X^{1+\varepsilon}.
\end{aligned}$$

From here we obtain

$$\begin{aligned}
\sum_z \sum_{\boldsymbol{w}} S_1(z, \boldsymbol{w})^2 + S_2(z, \boldsymbol{w})^2 &\ll X^{1/4} \log X \sum_z \sum_{\boldsymbol{w}} S_1(z, \boldsymbol{w}) + \log X \sum_z \sum_{\boldsymbol{w}} S_2(z, \boldsymbol{w}) \\
&\ll_\varepsilon X^{5/4+\varepsilon}.
\end{aligned}$$

and

$$\sum_z \sum_{\boldsymbol{w}} S_1^{\spadesuit}(z, \boldsymbol{w})^2 + S_2^{\spadesuit}(z, \boldsymbol{w})^2 \ll_\varepsilon X^{5/4+\varepsilon}.$$

At this stage, we expunge the references to the Gaussian domain $\mathbb{Z}[i]$ in [10] to make it clear that much of their treatment of bilinear sums apply equally well in our situation, despite the fact that our number field is different from $\mathbb{Q}(i)$. For $\boldsymbol{y}, z \in \mathbb{Z}^2$ put $\Delta(\boldsymbol{y}, z) = y_1 z_2 - y_2 z_1$. Given $\boldsymbol{w}, \boldsymbol{y}, z \in \mathbb{Z}^2$ such that

$$w_1 y_1 + w_2 y_2 = q_1 \quad \text{and} \quad w_1 z_1 + w_2 z_2 = q_2,$$

we have

$$\begin{bmatrix} y_1 & y_2 \\ z_1 & z_2 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}.$$

Inverting the matrix on the left we see that

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{\Delta(\boldsymbol{z}, \boldsymbol{y})}\begin{bmatrix} z_2 & -y_2 \\ -z_1 & y_1 \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}.$$

Since $\boldsymbol{w} = (w_1, w_2) \in \mathbb{Z}^2$, it follows that

$$q_1 z_2 - q_2 y_2 \equiv q_1 z_1 - q_2 y_1 \equiv 0 \pmod{\Delta(\boldsymbol{z}, \boldsymbol{y})}. \tag{7-7}$$

Let $C(q_1, q_2, \boldsymbol{z}, \boldsymbol{y})$ be the statement that $q_1, q_2, \boldsymbol{z}, \boldsymbol{y}$ satisfy (7-7). Next we have

$$\begin{aligned} \|q_1(z_1, z_2) - q_2(y_1, y_2)\|_2 &= \sqrt{(q_1 z_1 - q_2 y_1)^2 + (q_1 z_2 - q_2 y_1)^2} \\ &= \sqrt{(w_1 \Delta(\boldsymbol{z}, \boldsymbol{y}))^2 + (w_2 \Delta(\boldsymbol{z}, \boldsymbol{y}))^2} \\ &= \Delta(\boldsymbol{z}, \boldsymbol{y})\sqrt{w_1^2 + w_2^2} \le \Delta(\boldsymbol{z}, \boldsymbol{y}) M. \end{aligned} \tag{7-8}$$

We also wish to impose the condition that $\Delta(\boldsymbol{z}, \boldsymbol{y})$ is small. In particular, we wish to only consider those $\boldsymbol{z}, \boldsymbol{y}$ with

$$\Delta(\boldsymbol{z}, \boldsymbol{y}) > \mathfrak{D}_0 = N(\log X)^{-A-6}. \tag{7-9}$$

For brevity, let us write

$$h^{\dagger}(q) = \begin{cases} 2p \log p & \text{if } q = p^2 \in I(X), \\ 0 & \text{otherwise}, \end{cases} \qquad h^{\ddagger}(q) = \begin{cases} \log p & \text{if } q = p \in I(X), \\ 0 & \text{otherwise}, \end{cases}$$

and

$$h(q) = h^{\dagger}(q) - h^{\ddagger}(q).$$

Similarly, let us write

$$h^{\spadesuit,\dagger}(q) = \begin{cases} 2p \log p & \text{if } q = p^2 \in I(X), \\ 0 & \text{otherwise}, \end{cases} \qquad h^{\spadesuit,\ddagger}(q) = \begin{cases} \log p & \text{if } q = p \in I(X), \\ 0 & \text{otherwise}, \end{cases}$$

and

$$h^{\spadesuit}(q) = h^{\spadesuit,\dagger}(q) - h^{\spadesuit,\ddagger}(q).$$

For any subinterval $J \subset I(X)$ we have

$$\sum_{q \in J} h(q) = O_C\left(\frac{X^{1/4}}{(\log X)^C}\right)$$

for any $C > 0$. This is a consequence of our choice of weights.

As in [10], we want to carve up the support of $\boldsymbol{z}, \boldsymbol{y}$ into regions of the form

$$\mathcal{U} = \mathcal{U}(c, \theta_0) = \left\{\boldsymbol{z} : c\sqrt{N} < \|\boldsymbol{z}\|_2 \le c(1 + \omega_1)\sqrt{N},\ \theta_0 < \arg(\boldsymbol{z}) \le \theta_0 + \omega_2\right\}, \tag{7-10}$$

for fixed $1 \le c \le \sqrt{2}$ and θ_0. We may choose ω_1 and ω_2 so that the regions $\mathcal{U}$ form a partition of the region

$$\{z : N \le \|z\|_2 < 2N,\ z_1 > 0\} \setminus \mathcal{R}.$$

The number of regions needed for the sum over z, y is $O((\log X)^{4L})$. Here, as in [10], we allow the parameters ω_1 and ω_2, both of order $(\log X)^{-A}$, to be different in order to perfectly cover our region.

As in [10] let us write $\mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ as the condition that all $(z, y, q_1, q_2) \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ satisfy (7-8) and (7-9). We remark that such tuples are the most intricate to estimate; in fact it is only in the treatment of these tuples where we must diverge from the argument given in [10].

Similarly, let $\mathfrak{C}_2(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ denote the condition that there exists some tuple $(z, y, q_1, q_2) \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ which satisfies (7-8) and there exists some tuple $(z', y', q_1', q_2') \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ which does not satisfy (7-8). Finally, let $\mathfrak{C}_3(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ be the condition that all tuples $(z, y, q_1, q_2) \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ satisfy (7-8) but there exists some tuple $(z, y, q_1, q_2) \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ which does not satisfy (7-9).

Recall that $C(q_1, q_2, z, y)$ is the condition that z, y, q_1, q_2 satisfy (7-7). We also introduce the condition $\sum^\flat$ to indicate a summation over primitive $z \notin \mathcal{R}(N; X)$. Observe that we do not insist that $z \equiv (1, 0) \pmod 2$ as in [10]. For $\mathcal{U}_1, \mathcal{U}_2, J_1, J_2$ satisfying $\mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ put

$$T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) = \sideset{}{^\flat}\sum_{\substack{z \in \mathcal{U}_1 \\ y \in \mathcal{U}_2}} \beta_z \beta_y \sum_{\substack{q_1 \in J_1 \\ q_2 \in J_2 \\ C(q_1, q_2, z, y)}} h(q_1) h(q_2), \tag{7-11}$$

and otherwise set $T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) = 0$. Further, let

$$T'(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) = \sideset{}{^\flat}\sum_{\substack{z \in \mathcal{U}_1 \\ y \in \mathcal{U}_2}} \sum_{\substack{q_1 \in J_1 \\ q_2 \in J_2 \\ C(q_1, q_2, z, y)}} |h(q_1) h(q_2)|. \tag{7-12}$$

Similarly, define

$$T_\spadesuit(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \text{ and } T'_\spadesuit(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$$

analogously with h replaced with $h^\spadesuit$. Then to obtain (7-5) and (7-6) it suffices to show that

$$\sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T_\spadesuit(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) + \sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_2(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \text{ or } \mathfrak{C}_3(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T'_\spadesuit(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \ll_A \frac{XN}{(\log X)^A} \tag{7-13}$$

and

$$\sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) + \sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_2(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \text{ or } \mathfrak{C}_3(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T'(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \ll_A \frac{XN}{(\log X)^A}. \tag{7-14}$$

As in [10], we will show that the contribution from $\mathfrak{C}_i(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ is negligible for $i = 2, 3$. Indeed, we shall obtain:

Proposition 7.5. *We have*

$$\sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_2(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \text{ or } \mathfrak{C}_3(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T'_{\spadesuit}(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \ll_A \frac{XN}{(\log X)^A}$$

and

$$\sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_2(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \text{ or } \mathfrak{C}_3(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T'(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \ll_A \frac{XN}{(\log X)^A}.$$

In fact, Proposition 7.5 is exactly analogous to Proposition 6 in [10]. More strikingly, the proof does not need to be modified and we can simply apply Proposition 6 of [10]. However, given that our setups are not identical we will explain why our situations are indeed interchangeable.

We will also need the following analogue of Proposition 7 in [10]:

Proposition 7.6. *For fixed J_1, J_2 and $L = 6A + 52$ we have*

$$\sum_{\substack{\mathcal{U}_1, \mathcal{U}_2 \\ \mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T_{\spadesuit}(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \ll_A \frac{XN}{(\log X)^{A+2L}}.$$

and

$$\sum_{\substack{\mathcal{U}_1, \mathcal{U}_2 \\ \mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \ll_A \frac{XN}{(\log X)^{A+2L}}.$$

Unlike Proposition 7.5 we cannot simply import Proposition 7 from [10]. This is because Proposition 7.5, by the definition of $T'(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$, is insensitive to the nature of the coefficients β_z and so the treatment in [10] is directly applicable to our situation. However in order to prove Proposition 7 in [10] the specific shape of β_z was needed. That said, the modifications needed to adapt their proof to our case are minor, and we will still be able to follow their argument for the most part.

In the next few sections we will give proofs for Propositions 7.5 and 7.6. We will largely follow the structure of the argument given in [10].

8. Proof of Proposition 7.5

First we have the following lemma, which is Lemma 12 from [10]:

Lemma 8.1. *We have the bounds*

$$\sideset{}{^\flat}\sum_{z,y} \sum_{\substack{q_1 \in J_1, q_2 \in J_2 \\ C(q_1, q_2, z, y) \\ \gcd(q_1 q_2, \Delta(z,y)) > 1}} |h^{\spadesuit}(q_1) h^{\spadesuit}(q_2)| \ll N^2 \sqrt{X} (\log X)^3$$

and

$$\sideset{}{^\flat}\sum_{z,y} \sum_{\substack{q_1 \in J_1, q_2 \in J_2 \\ C(q_1, q_2, z, y) \\ \gcd(q_1 q_2, \Delta(z,y)) > 1}} |h(q_1) h(q_2)| \ll N^2 \sqrt{X} (\log X)^3.$$

Proof. See Section 7 in [10]. □

Lemma 8.1 allows us, as in [10], to write

$$T(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)=\sum_{D\le 2N}\ \sum_{a\,(\mathrm{mod}\,D)}^{*}\mathcal{Y}(a,D;h,h)\mathcal{Z}(a,D)+O\big(N^2\sqrt{X}(\log X)^3\big) \tag{8-1}$$

where

$$\mathcal{Z}(a,D)=\sum^{\flat}_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ \Delta(z,y)=D\\ a\,y\equiv z\,(\mathrm{mod}\,D)}}\beta_z\beta_y$$

and

$$\mathcal{Y}(a,D;h_1,h_2)=\sum_{\substack{q_1\in J_1,\,q_2\in J_2\\ q_1\equiv aq_2\,(\mathrm{mod}\,D)\\ \gcd(q_1q_2,D)=1}}h_1(q_1)h_2(q_2).$$

Similarly, we have

$$T_{\spadesuit}(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)=\sum_{D\le 2N}\ \sum_{a\,(\mathrm{mod}\,D)}^{*}\mathcal{Y}^{\spadesuit}(a,D;h^{\spadesuit},h^{\spadesuit})\mathcal{Z}(a,D)+O\big(N^2\sqrt{X}(\log X)^3\big) \tag{8-2}$$

where

$$\mathcal{Y}^{\spadesuit}(a,D;h_1^{\spadesuit},h_2^{\spadesuit})=\sum_{\substack{q_1\in J_1,\,q_2\in J_2\\ q_1\equiv aq_2\,(\mathrm{mod}\,D)\\ \gcd(q_1q_2,D)=1}}h_1^{\spadesuit}(q_1)h_2^{\spadesuit}(q_2)$$

This crucial decomposition allows us to separate $T(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)$ and $T_{\spadesuit}(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)$ into components $\mathcal{Z}(a,D)$ containing the coefficients β_z, β_y and a congruence sum which no longer has anything to do with the coefficients β. To treat (7-14) requires a treatment of $\mathcal{Y}(a,D)$ involving primes. For this purpose they needed a refinement of the Barban–Davenport–Heilbronn theorem, which we will not go into more detail here as we can use their Proposition 6 directly.

The following lemma is critical to the proof of Proposition 7.5:

Lemma 8.2. *Let*

$$\widetilde{\mathcal{Z}}(a,D)=\sum^{\flat}_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ \Delta(z,y)=D\\ a\,y\equiv z\,(\mathrm{mod}\,D)}}1.$$

We then have the bounds

$$\sum_{D}\tau(D)\sum_{a\,(\mathrm{mod}\,D)}^{*}\widetilde{\mathcal{Z}}(a,D)\ll\omega^4N^2(\log X)^{16}, \tag{8-3}$$

$$\sum_{\mathcal{U}_1,\mathcal{U}_2}\ \sum_{a\,(\mathrm{mod}\,D)}^{*}\widetilde{\mathcal{Z}}(a,D)\ll N, \tag{8-4}$$

and

$$\sum_{a \,(\mathrm{mod}\, D)}^{*} \widetilde{\mathcal{Z}}(a, D) \ll (\log X)^3 \frac{N^2}{D} \tau(D)^6. \tag{8-5}$$

Proof. See Lemma 13 in [10]. □

For an interval J and a function $\mathfrak{h}$, put

$$\mathcal{Y}(J, \mathfrak{h}; D) = \sum_{\substack{q \in J \\ \gcd(q, D)=1}} \mathfrak{h}(q)$$

and

$$\mathcal{Y}_{\mathfrak{h}_1,\mathfrak{h}_2}(D) = \mathcal{Y}(D) = \frac{1}{\varphi(D)} \mathcal{Y}(J_1, \mathfrak{h}_1; D)\mathcal{Y}(J_2, \mathfrak{h}_2; D).$$

Recall that q_1, q_2 appearing in $\mathcal{Y}(a, D; \mathfrak{h}_1, \mathfrak{h}_2)$ satisfy $\gcd(q_1 q_2, D) = 1$. If h_1 or h_2 is equal to $h^{\ddagger}$, then $\mathcal{Y}(D)$ is the expected value of $\mathcal{Y}(a, D; h_1, h_2)$. If $h_1 = h_2 = h^{\dagger}$, note that $p_1^2 \equiv a p_2^2 \pmod{D}$ implies that $p_1 \equiv b p_2 \pmod{D}$ for some b such that $a \equiv b^2 \pmod{D}$. Here, $\mathcal{Y}(a, D; h_1, h_2) = 0$ if a is not a square modulo D. Therefore

$$\sum_{a \,(\mathrm{mod}\, D)}^{*} \mathcal{Y}(a, D; h_1, h_2)\mathcal{Z}(a, D) = \sum_{b \,(\mathrm{mod}\, D)}^{*} \mathcal{Y}_{h^{\dagger}}(b, D)\mathcal{Z}(b^2, D)$$

where

$$\mathcal{Y}_{h^{\dagger}}(b, D) = \sum_{\substack{p_1^2 \in J_1,\, p_2^2 \in J_2 \\ p_1 \equiv b p_2 \,(\mathrm{mod}\, D) \\ \gcd(p_1 p_2, D)=1}} h^{\dagger}(p_1^2) h^{\dagger}(p_2^2).$$

When $h_1 = h_2 = h^{\dagger}$, then $\mathcal{Y}(D)$ is the expected value of $Y_{h^{\dagger}}(b, D)$. Now put

$$\mathcal{E}(N) = \sum_{D \le 2N} \sum_{a \,(\mathrm{mod}\, D)}^{*} |\mathcal{Y}(a, D; h_1, h_2) - \mathcal{Y}_{h_1,h_2}(D)| \, \widetilde{\mathcal{Z}}(a, D)$$

if either $h_1 = h^{\dagger}$ or $h_2 = h^{\dagger}$, and

$$\mathcal{E}_{h^{\dagger}}(N) = \sum_{D \le 2N} \sum_{b \,(\mathrm{mod}\, D)}^{*} |\mathcal{Y}_{h^{\dagger}}(b, D) - \mathcal{Y}_{h^{\dagger},h^{\dagger}}(D)| \, \widetilde{\mathcal{Z}}(b^2, D)$$

if $h_1 = h_2 = h^{\dagger}$. We then have the following proposition, which is Proposition 8 from [10]:

Proposition 8.3. *For any $C > 0$ we have*

$$\mathcal{E}(N) \ll_C \frac{XN}{(\log X)^C} \quad \textit{and} \quad \mathcal{E}_{h^{\dagger}}(N) \ll_C \frac{XN}{(\log X)^C}.$$

With this proposition in hand, we may proceed to prove Proposition 7.5 in the exact same way as Proposition 6 in [10]. We will not repeat the details.

We now move to the proof of Proposition 7.6. Most of the arguments can be adapted from the proof of Proposition 7 in [10], but since we rely on some properties of the coefficients β_z in this argument we cannot follow all of the arguments in [10] verbatim. We will especially emphasize those points where modifications are required.

9. Proof of Proposition 7.6: some maneuvers

Supposing that one of the functions h_1, h_2 is $h^{\ddagger}$, we have according to Proposition 8.3 that

$$\sum_{D\le 2N}\ \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Y}(a, D; h_1, h_2)\mathcal{Z}(a, D) = \sum_{D\le 2N}\ \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Y}_{h_1,h_2}(D)\mathcal{Z}(a, D) + O_C\biggl(\frac{XN}{(\log X)^C}\biggr)$$

for any $C > 0$. In the remaining case with $h_1 = h_2 = h^{\dagger}$, we have

$$\sum_{D\le 2N}\ \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Y}(a, D; h^{\dagger}, h^{\dagger})\mathcal{Z}(a, D) = \sum_{D\le 2N}\ \sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D)} \mathcal{Y}_{h^{\dagger},h^{\dagger}}(D)\mathcal{Z}(b^2, D) + O_C\biggl(\frac{XN}{(\log X)^C}\biggr).$$

As in [10] we may replace $\mathcal{Y}_{h^{\dagger},h^{\dagger}}(D)$ by $|J_1|\,|J_2|/\varphi(D)$ in each case, with a total error of

$$O\bigl(X\exp\bigl(\sqrt{-\log X}\,\bigr)N(\log X)^2\bigr).$$

Our remaining task is the inequality

$$|J_1|\,|J_2| \sum_{\substack{\mathcal{U}_1,\mathcal{U}_2\\ \mathfrak{C}_1(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)}}\ \sum_{D\le 2N} \frac{1}{\varphi(D)}\biggl(\ \sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D))} \mathcal{Z}(b^2, D) - \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Z}(a, D)\biggr) \ll \frac{XN}{(\log X)^{A+2L}},$$

or

$$\mathcal{E}' = \sum_{\substack{\mathcal{U}_1,\mathcal{U}_2\\ \mathfrak{C}_1(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)}}\ \sum_{D} \frac{1}{\varphi(D)}\biggl(\ \sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D)} \mathcal{Z}(b^2, D) - \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Z}(a, D)\biggr) \ll \frac{N}{(\log X)^{A}}.$$

Here we dropped the condition $D \le 2N$, which follows automatically since β_z is supported on $\|z\|_2 \le 2N$.

We may follow Heath-Brown and Li's arguments in [10] to conclude that it suffices to obtain the estimate

$$\mathcal{E}_1(\mathcal{U}_1, \mathcal{U}_2) = \sum_{D} \frac{D}{\varphi(D)}\biggl(\ \sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D)} \mathcal{Z}(b^2, D) - \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Z}(a, D)\biggr) \ll \frac{N^2}{(\log X)^{C_1}} \tag{9-1}$$

for any $C_1 > 0$ and for fixed $\mathcal{U}_1, \mathcal{U}_2$. By Möbius inversion we deduce that

$$\mathcal{E}_1(\mathcal{U}_1, \mathcal{U}_2) = \sum_{D=1}^{\infty}\sum_{k=1}^{\infty} \frac{D\mu(k)}{\varphi(D)}\biggl(\ \sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D)} W(b^2, k, D) - \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} W(a, k, D)\biggr),$$

where

$$W(a, k, D) = \sideset{}{^\flat}\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ kD\mid\Delta(z,y)\\ ay\equiv z\,(\mathrm{mod}\,D)}} \beta_z\beta_y.$$

Here the condition $\Delta(z, y) = D$ which appears in the definitions of $\mathcal{Z}(a, D)$, $\widetilde{\mathcal{Z}}(a, D)$ is replaced with a divisibility condition via Möbius inversion.

If kD divides $\Delta(\mathbf{z}, \mathbf{y})$, there is a unique integer $c = c(\mathbf{z}, \mathbf{y}; kD)$ modulo kD such that $c\mathbf{y} \equiv \mathbf{z} \pmod{kD}$, and conversely this congruence implies that kD divides $\Delta(\mathbf{z}, \mathbf{y})$. We have $\gcd(c, kD) = 1$ and

$$\#\{b \pmod D : b^2\mathbf{y} \equiv \mathbf{z} \pmod D\} = \#\{b \pmod D : b^2 \equiv c \pmod D\} = \sum_{\substack{\chi \,(\mathrm{mod}\, D) \\ \chi^2 = \chi_0}} \chi(c).$$

It now follows that

$$\sum_{b\,(\mathrm{mod}\, D)}^{*} W(b^2, k, D) - \sum_{a\,(\mathrm{mod}\, D)}^{*} W(a, k, D) = \sum_{\substack{\chi \,(\mathrm{mod}\, D) \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{c\,(\mathrm{mod}\, kD)}^{*} \sum_{\substack{(\mathbf{z},\mathbf{y}) \in \mathcal{U}_1 \times \mathcal{U}_2 \\ c\mathbf{y} \equiv \mathbf{z}\,(\mathrm{mod}\, kD)}}^{\flat} \beta_{\mathbf{z}} \beta_{\mathbf{y}} \chi(c),$$

and hence

$$\mathcal{E}_1(\mathcal{U}_1, \mathcal{U}_2) = \sum_{D=1}^{\infty} \sum_{k=1}^{\infty} \frac{D\mu(k)}{\varphi(D)} \sum_{\substack{\chi \,(\mathrm{mod}\, D) \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{c\,(\mathrm{mod}\, kD)}^{*} \sum_{\substack{(\mathbf{z},\mathbf{y}) \in \mathcal{U}_1 \times \mathcal{U}_2 \\ c\mathbf{y} \equiv \mathbf{z}\,(\mathrm{mod}\, kD)}}^{\flat} \beta_{\mathbf{z}} \beta_{\mathbf{y}} \chi(c).$$

Let $d = d(\chi)$ be the conductor of χ and write $D = de$ and $ek = \mathfrak{k}$, giving

$$\mathcal{E}(\mathcal{U}_1, \mathcal{U}_2) = \sum_{d>1} \sum_{\mathfrak{k}} C(d, \mathfrak{k}) \sum_{\substack{\chi \,(\mathrm{mod}\, d) \\ \chi^2 = \chi_0}}^{*} \sum_{c\,(\mathrm{mod}\, d\mathfrak{k})}^{*} \sum_{\substack{(\mathbf{z},\mathbf{y}) \in \mathcal{U}_1 \times \mathcal{U}_2 \\ c\mathbf{y} \equiv \mathbf{z}\,(\mathrm{mod}\, d\mathfrak{k})}}^{\flat} \beta_{\mathbf{z}} \beta_{\mathbf{y}} \chi(c),$$

where

$$C(d, \mathfrak{k}) = \sum_{d_1 k = d_2} \frac{de\mu(k)}{\varphi(de)} = \frac{d}{\varphi(d)} \sum_{ek = \mathfrak{k}} \frac{\varphi(d)e\mu(k)}{\phi(de)}.$$

The sum for $\chi \pmod d$ is empty unless $d = d_1, 4d_1, 8d_1$ with d_1 odd and square-free, in which cases there are at most two possible characters χ. For fixed d the function

$$\varphi_d(e) = \frac{\varphi(d)e}{\varphi(de)}$$

is multiplicative in e. Further, for $v \geq 1$ we have

$$(\varphi_e * \mu)(p^v) = \begin{cases} (p-1)^{-1} & \text{if } v = 1 \text{ and } p \nmid d, \\ 0 & \text{otherwise.} \end{cases}$$

We then see that

$$C(d, \mathfrak{k}) = \frac{d\mu^2(\mathfrak{k})}{\varphi(d\mathfrak{k})}$$

if $\gcd(d, \mathfrak{k}) = 1$ and $C(d, \mathfrak{k}) = 0$ otherwise. This gives the expression

$$\mathcal{E}_1(\mathcal{U}_1, \mathcal{U}_2) = \sum_{\substack{\mathfrak{k}, d \\ \gcd(d, \mathfrak{k}) = 1}} \frac{d\mu^2(\mathfrak{k})}{\varphi(d\mathfrak{k})} \sum_{\substack{\chi \,(\mathrm{mod}\, d) \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}}^{*} \Bigg(\sum_{c\,(\mathrm{mod}\, d\mathfrak{k})}^{*} \sum_{\substack{(\mathbf{z},\mathbf{y}) \in \mathcal{U}_1 \times \mathcal{U}_2 \\ c\mathbf{y} \equiv \mathbf{z}\,(\mathrm{mod}\, d\mathfrak{k})}}^{\flat} \beta_{\mathbf{z}} \beta_{\mathbf{y}} \chi(c) \Bigg). \tag{9-2}$$

We proceed to show that large values of $\mathfrak{k}$ make a negligible contribution. Since $d\mathfrak{k} \mid \Delta(z, y)$ we have $d\mathfrak{k} \leq 2N$. Since $0 \leq \beta_z \leq 1$ we find that

$$\begin{aligned}\sum_{\mathfrak{k}>\mathfrak{K}} \sum_{\substack{\mathfrak{k} \\ \gcd(d,\mathfrak{k})=1}} \frac{d\mu^2(\mathfrak{k})}{\varphi(d\mathfrak{k})} \sideset{}{^*}\sum_{\substack{\chi \,(\mathrm{mod}\, d) \\ \chi^2=\chi_0}} \Bigg| \sideset{}{^*}\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})} \sideset{}{^\flat}\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2 \\ cy\equiv z \,(\mathrm{mod}\, d\mathfrak{k})}} \beta_z\beta_y\chi(c)\Bigg| &\ll (\log X) \sum_{\mathfrak{k}>\mathfrak{K}} \mathfrak{k}^{-1} \sum_{d\leq 2N/\mathfrak{k}} \sum_{\substack{d\mathfrak{k}\mid D \\ D\leq 2N}} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, D)} \widetilde{\mathcal{Z}}(a, D) \\ &\ll (\log X) \sum_{\mathfrak{k}>\mathfrak{K}} \mathfrak{k}^{-1} \sum_{d\leq 2N/\mathfrak{k}} \sum_{\substack{d\mathfrak{k}\mid D \\ D\leq 2N}} N \\ &\ll \frac{N^2(\log X)^2}{\mathfrak{K}}.\end{aligned}$$

Choosing

$$\mathfrak{K} = (\log X)^{C_1+2}$$

and applying Lemma 8.2 then gives a satisfactory bound.

Observe that the argument above only depends on the property that $0 \leq \beta_z \leq 1$, and so no modification is necessary from the argument given by Heath-Brown and Li in [10]. As in [10] we divide into three ranges for d, namely

$$d \leq D_1, \quad D_1 < d \leq D_2, \quad \text{and} \quad d > D_2,$$

where

$$D_1 = \mathfrak{K}^{10}(\log X)^{2C_1+14} \quad \text{and} \quad D_2 = \frac{N}{\mathfrak{K}^{15}(\log X)^{3C_1+21}}.$$

Next we handle the middle range of d. The treatment given here is identical to that in [10], since again the specific shape of β_z is of no consequence in this part. Set

$$\mathcal{E}_1(D) = \sum_{\mathfrak{k}\leq\mathfrak{K}} \mathfrak{k}^{-1}\mu^2(\mathfrak{k}) \sum_{\substack{D<d\leq 2D \\ \gcd(d,\mathfrak{k})=1}} \sideset{}{^*}\sum_{\substack{\chi \,(\mathrm{mod}\, d) \\ \chi^2=\chi_0}} \Bigg| \sideset{}{^*}\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})} \sideset{}{^\flat}\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2 \\ cy\equiv z \,(\mathrm{mod}\, d\mathfrak{k})}} \beta_z\beta_y\chi(c)\Bigg|.$$

We remark on the significance that the sum is over primitive characters in the definition of $\mathcal{E}_1(D)$. Indeed, as seen in [10] this property is necessary to decompose the characters into Jacobi symbols.

Heath-Brown and Li obtains the following bound, which we summarize in the following lemma:

Lemma 9.1. *For any $\varepsilon > 0$ we have*

$$\mathcal{E}_1(D) \ll_\varepsilon \mathfrak{K}^5(\log X)^6(D + D^{-1/2}N + D^{1/3}N^{2/3} + N^{23/24+\varepsilon})N.$$

Summing over dyadic ranges of D, we see that the values of d in the range $D_1 \leq d \leq D_2$ make a satisfactory contribution given our choices of D_1, D_2.

It then remains to give estimates for the small and large ranges of d, where we must depart somewhat from Heath-Brown and Li's treatment due to the dependence on the specific shapes of the coefficients β_z.

10. Proof of Proposition 7.6: remaining ranges

Large d. We will obtain the bound

$$\sum_{\substack{d>D_2\\ \gcd(d,\mathfrak{k})=1}} \sideset{}{^*}\sum_{\substack{\chi \,(\mathrm{mod}\, d)\\ \chi^2=\chi_0}} \Bigg(\sideset{}{^*}\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})} \sideset{}{^\flat}\sum_{\substack{(z,\, y)\in\mathcal{U}_1\times\mathcal{U}_2\\ c y\equiv z \,(\mathrm{mod}\, d\mathfrak{k})}} \beta_z \beta_y \chi(c) \Bigg) \ll_C \frac{N^2}{(\log X)^C}$$

for any $C > 0$ and $\mathfrak{k} \le \mathfrak{K}$. There is still more mileage we can get from the argument given in [10]. In particular, we follow their argument in Section 11 of [10] and decompose d as $d_1 d_2$, as well as $\chi = \chi_1\chi_2$. We have $d\mathfrak{k} \mid \Delta(z, y)$ and thus we may set $\Delta(z, y) = d_1 e t$ where e is odd and t is a power of 2. Our conditions on $\mathcal{U}_1, \mathcal{U}_2$ guarantee that $0 < \Delta(z, y) \le 2N$, hence $1 \le et \le 16N/D_2 \ll (\log X)^{18C_1+51}$. We split the sums over z, y into congruence classes $z \equiv u \pmod{8et}$, $y \equiv v \pmod{8et}$ and fix the parameters

$$\mathfrak{k}, \quad d_2, \quad \chi_2, \quad e, \quad u, \quad v, \quad \text{and} \quad t. \tag{10-1}$$

Each admissible pair u, v corresponds to a unique integer $k \pmod{\Delta(z, y)}$ with the property that $k y \equiv z \pmod{\Delta(z, y)}$, and then

$$\chi(c) = \chi(k) = \chi_2(k)\left(\frac{k}{d_1}\right),$$

where $\chi_2(k)$ is determined by the parameters (10-1). The number of choices for the parameters (10-1) is bounded by a fixed power of $\log X$ and so it suffices to show that

$$\sum_{\substack{d_1>D_2/d_2\\ \gcd(d_2,2\mathfrak{k})=1}} \frac{d_1\mu^2(d_1)}{\varphi(d_1)} \Bigg(\sideset{}{^*}\sum_{k \,(\mathrm{mod}\, d_1 e t)} \sideset{}{^\flat}\sum_{z,\, y} \beta_z\beta_y \left(\frac{k}{d_1}\right)\Bigg) \ll_C \frac{N^2}{(\log X)^C}$$

for every $C > 0$, where the sum over z, y satisfies the conditions

$$(z, y)\in\mathcal{U}_1\times\mathcal{U}_2, \quad k y\equiv z \pmod{\Delta(z, y)}, \quad z\equiv u \pmod{8et}, \quad y\equiv v \pmod{8et}, \quad \text{and} \quad \Delta(z, y) = d_1 e t.$$

Following the same analysis in Section 11.1 of [10], we conclude that it is sufficient to obtain the bound

$$\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ z\equiv u,\ y\equiv v \,(\mathrm{mod}\, 8etn)\\ \Delta(z,y)>etD_2/d_2}} \beta'_z\beta'_y \ll_C \frac{N^2}{(\log X)^C}$$

where

$$\beta'_z = \beta_z(-1)^{(z_1-1)/2}\left(\frac{z_2}{z_1}\right).$$

for every fixed $C > 0$, for each choice of parameters $e, t, n \le (\log X)^C$, and for each u, v. Further subdividing into congruence classes it suffices to handle

$$\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ z\equiv u,\ y\equiv v \,(\mathrm{mod}\, 8etn)}} \beta'_z\beta'_y = \Bigg(\sum_{\substack{z\in\mathcal{U}_1\\ z\equiv u \,(\mathrm{mod}\, 8etn)}} \beta'_z\Bigg)\Bigg(\sum_{\substack{z\in\mathcal{U}_2\\ z\equiv v \,(\mathrm{mod}\, 8etn)}} \beta'_z\Bigg). \tag{10-2}$$

At this stage, we must diverge from Heath-Brown and Li's treatment. We briefly discuss why this is necessary. In order to proceed, Heath-Brown and Li rely on the crucial property that their β_z are supported on Gaussian integers z such that $N(z)$ has no small prime factors. The analogous condition for us is that the ideal number $\gamma(\boldsymbol{z})$ has norm (equal to the norm of the ideal $J(\gamma(\boldsymbol{z}))$ in $\mathcal{O}_K$) without small prime factors. Thus, now going to the perspective that $\boldsymbol{z}$ represents an ideal number γ, we see that $N(\gamma) = N(J(\gamma))$ is automatically coprime to $8etn$ and therefore we may assume that υ, ν (the ideal numbers corresponding to $\boldsymbol{u}, \boldsymbol{v}$, respectively) are coprime to $8etn$. This allows us to pick out the congruence condition $\gamma \equiv \upsilon, \nu \pmod{8etn}$ using multiplicative characters. In order to make this precise, we borrow from the algebraic treatment given in [11] and put

$$\mathfrak{J}(q) = \{\alpha \in \mathfrak{J} : \gcd(\alpha, q) = 1\}$$

and $\mathfrak{J}_1(q) = \mathfrak{J}(q) \cap K$. Further, put

$$\mathfrak{J}_0(q) = \{\alpha \in K : \alpha \equiv 1 \pmod{q}\}.$$

Then our congruence conditions can be picked out using characters of the quotient group $\mathfrak{J}_1(q)/\mathfrak{J}_0(q)$, and we conclude that

$$\sum_{\substack{\boldsymbol{a}\in\mathcal{U}_j \\ \alpha\equiv\nu \pmod{8etn}}} = \frac{1}{\varphi_K(8etn)} \sum_{\chi \pmod{8etn}} \bar{\chi}(\nu)\mathcal{S}(\chi, \mathcal{U}_j),$$

where φ_K is the Euler-φ function for $\mathcal{O}_K$ and

$$\mathcal{S}(\chi, \mathcal{U}) = \sum_{\boldsymbol{a}\in\mathcal{U}} \beta'_\alpha \chi(\alpha).$$

In order to obtain acceptable estimates for $\mathcal{S}(\chi, \mathcal{U})$, we will need to generalize certain results from [5] to apply to general quadratic fields. This work may be of independent interest and is recorded in the next section; see Propositions 11.7 and 11.9 in particular. We emphasize that these results rely on the setup in (10-2): in particular, we need $\boldsymbol{z}, \boldsymbol{w}$ to come from the same ideal class and that they satisfy a congruence condition modulo $8etn$.

We now proceed to pick out the condition that we are constrained in a narrow sector using a twice-differentiable periodic function $\upsilon(\theta)$, where

$$\upsilon(\theta) = \begin{cases} 1 & \text{if } \theta \in (\theta_0, \theta_0 + \varpi_2) \pmod{2\pi}, \\ 0 & \text{if } \theta \notin [\theta_0 - (\log X)^{-C}, \theta_0 + \varpi_2 + (\log X)^{-C}] \pmod{2\pi}, \end{cases}$$

and where $|\upsilon''(\theta)| \ll (\log X)^{-2C}$. Then

$$\mathcal{S}(\chi, \mathcal{U}) = \sum_{N' < N(z) \le N'(1+\varpi)} \beta'_z \chi(z) \upsilon(\arg z) + O\left(\frac{N}{(\log X)^C}\right).$$

The Fourier coefficients of υ satisfy $c_k \ll k^{-2}(\log X)^{2C}$ for $k \neq 0$, and so

$$\upsilon(\arg z) = \sum_k c_k \left(\frac{z}{|z|}\right)^k = \sum_{|k| \leq (\log X)^{3C}} c_k \left(\frac{z}{|z|}\right)^k + O((\log X)^{-C}).$$

It then suffices to show that

$$\mathcal{S}(\chi, N', k) = \sum_{N' < N(z) \leq N'(1+\varpi)} \beta'_z \chi(z) \left(\frac{z}{|z|}\right)^k \ll_C N(\log X)^{-4C}$$

for any $C > 0$, and for $|k| \leq (\log X)^{3C}$. As in [10] we can obtain in fact a small power-saving in N. We recall that $\beta_z = \beta_{N(z)}$ is the indicator function of a set of one of the shapes

$$Q_j = \{p_1 \cdots p_{j+1} \in (N', N'(1+\varpi)] : p_{j+1} \in J,\ p_{j+1} < \cdots < p_1,\ p_1 \cdots p_j < Y \leq p_1 \cdots p_{j+1} < X^{1/20\delta}\}$$

or

$$R = \{n \in (N', N'(1+\varpi)] : \gcd(n, P(V)) = 1\}.$$

Here we will have $0 \leq j \leq n_0 = \lfloor \log Y/(\delta \log X) \rfloor$, and $J = [V, V(1+\kappa)) \subseteq [X^\delta, X^{1/2-\delta})$. In particular we interpret Q_0 to be $\{p : p \in J \cap (N', N'(1+\varpi)]\}$.

We now write

$$\lambda(n) = \sum_{N(z)=n}^{\wedge} \chi(z) \left(\frac{z}{|z|}\right)^k u^{(x-1)/2}\left(\frac{z_2}{z_1}\right),$$

where $\sum^\wedge$ denotes a sum over primitive ideal numbers z in a fixed class of ideal numbers, with $\hat{z} = (z_1, z_2)$. We then have

$$\mathcal{S}(\chi, N', k) = \sum_n \lambda(n),$$

where n runs over R or Q_j for some j. As in [10], the treatment for R and Q_j are similar. To begin, we first handle the contribution from those n whose largest prime factor, say $\mathcal{P}(n)$, exceeds $N^{99/100}$. The contribution from such integers is

$$\sum_{m \leq 2N^{1/100}} \sum_{\substack{p > \max\{\mathcal{P}(m), N^{99/100}\} \\ mp \in Q_j}} \lambda(mp).$$

Since p is the largest prime factor of mp one sees from the definition of the set Q_j that one may rewrite the conditions $p > \mathcal{P}(m)$ and $mp \in Q_j$ to say that p runs over an interval $I_j(m) \subseteq [N/m, 2N/m)$. We may then apply Proposition 11.9 to conclude that

$$\begin{aligned}\sum_{m \leq 2N^{1/100}} \sum_{\substack{p > \max\{\mathcal{P}(m), N^{99/100}\} \\ mp \in Q_j}} \lambda(mp) &\ll q_0(|k|+1) \sum_{m \leq 2N^{1/100}} m(N/m)^{76/77} \\ &\ll q_0(|k|+1) N^{76/77 + (78/77)/100}.\end{aligned}$$

Since $76/77 + (78/77)/100 < 1$, this gives the required power-saving bound.

Next we deal with the terms where every prime factor is at most $N^{99/100}$. To do so we rewrite our sum in terms of bilinear sums. Suppose $n = p_1 \cdots p_{j+1}$ as in the description of the set Q_j, and divide the range of each prime p_i into intervals of the shape $(B_i, 2B_i]$. This will give us at most $(2\log N)^{1+n_0}$ sets of dyadic ranges, and since $n_0 \ll \delta^{-1} = (\log X)^{1-\varpi}$ there will be at most $O_\varepsilon(N^\varepsilon)$ such ranges. Moreover we may suppose

$$\prod_{i=1}^{j+1} B_i \ll N \ll 2^{j+1} \prod_{i=1}^{j+1} B_i.$$

Since we may now assume that $B_1 \le N^{99/100}$ there will be an index u such that

$$N^{1/100} \le \prod_{i=1}^{u} B_i \le N^{99/100}.$$

Fixing such an index u we split $n = n_1 n_2$ with

$$n_1 = \prod_{i=1}^{u} p_i \quad \text{and} \quad n_2 = \prod_{i=u+1}^{j+1} p_i,$$

so that $n_1 \le N_1$ and $n_2 \le N_2$ with

$$N_1 = 2^{1+n_0} \prod_{i=1}^{u} B_i \quad \text{and} \quad N_2 = 2^{1+n_0} \prod_{i=u+1}^{j+1} B_i.$$

It follows that

$$N_1 N_2 \ll_\varepsilon N^{1+\varepsilon} \quad \text{and} \quad N_1, N_2 \ll_\varepsilon N^{99/100+\varepsilon}.$$

This implies that

$$N_1 N^{-\varepsilon} \ll n_1 \le N_1 \quad \text{and} \quad N_2 N^{-\varepsilon} \ll n_2 \le N_2.$$

We may thus reinterpret our description of Q_j by requiring that $n_1 \in Q_{j,u}$ and $n_2 \in Q'_{j,u}$ for appropriate sets $Q_{j,u}$, $Q'_{j,u}$, together with the conditions that

$$n_1 n_2 \in I = (N', N'(1+\varpi)] \cap [Y, X^{1/2-\delta}),\ p_{j+1}^{-1} n_1 n_2 < Y,\ \text{and}\ p_{u+1} < p_u). \tag{10-3}$$

In other words, we put

$$Q_{j,u} = \{n_1 = p_1 \cdots p_u : p_i \in (B_i, 2B_i],\ p_u < \cdots < p_1\}$$

and

$$Q'_{j,u} = \{n_2 = p_{u+1} \cdots p_{j+1} : p_i \in (B_i, 2B_i],\ p_{j+1} \in J,\ p_{j+1} < \cdots < p_{u+1} < Y\}.$$

In order to separate the variables n_1, n_2 completely we subdivide the available ranges for n_1, n_2, p_{j+1}, p_u, and p_{u+1} into intervals of the shape $(A, A + A/L)$, $(A', A' + A'/L]$, $(B'_{j+1}, B'_{j+1} + B'_{j+1}/L]$, $(B'_u, B'_u + B'_u/L]$ and $(B'_{u+1}, B'_{u+1} + B'_{u+1}/L]$. Here the parameter L will be chosen to be a small power of N. These intervals may have length less than one. Indeed such an interval may contain no integers at all.

There will be $O(L^5(\log X)^2)$ such intervals and there will be some for which $n_1n_2 \in I$, $p_{j+1}^{-1}n_1n_2 < Y$ and $p_{u+1} < p_u$ for every choice of $p_1, \dots, p_{j+1}$ satisfying

$$n_1 \in (A, A + A/L], \quad n_2 \in (A', A' + A'/L],$$
$$p_{j+1} \in (B'_{j+1}, B'_{j+1} + B'_{j+1}/L], \quad p_u \in (B'_u, B'_u + B'_u/L], \quad p_{u+1} \in (B'_{u+1}, B'_{u+1} + B'_{u+1}/L],$$

and

$$p_i \in I_i \quad \text{with } i \neq 1, u, u+1.$$

This case gives the subsum

$$\sum_{\substack{n_1 \in Q_{j,u} \cap (A, A+N_1/L] \\ p_u \in (B'_u, B'_u + B'_u/K]}} \sum_{\substack{n_2 \in Q'_{j,u} \cap (A', A'+A'/L] \\ p_{j+1} \in (B'_{j+1}, B'_{j+1} + B'_{j+1}/L] \\ p_{u+1} \in (B'_{u+1}, B'_{u+1} + B'_{u+1}/L]}} \lambda(n_1n_2),$$

so that we have separated the variables n_1, n_2. For such sums we can apply Proposition 11.7 which gives the bound

$$O_\varepsilon((N_1 + N_2)^{1/12}(N_1N_2)^{1/12+\varepsilon}) = O_\varepsilon(N^{(9/100)\cdot(1/12)} \cdot N^{1/12+\varepsilon}) = O_\varepsilon(N^{1-1/1200+\varepsilon}).$$

Since there are $O_\varepsilon(L^5N^\varepsilon)$ such subsums the overall contribution will be $O(L^5N^{1-1/200+\varepsilon})$.

It remains to consider the contribution from the remaining "bad" sets of ranges which are not exclusively contained in the region given by (10-3). First suppose that the interval I is given by $[e_1, e_2]$ say, and that there are integers $n_1, n'_1 \in (A, A + A/L]$ and $n_2, n'_2 \in (A', A' + A'/L]$ such that $n_1n_2 \in I$ but $n'_1n'_2 \notin I$. Then we must have $n_1n_2 = (1 + O(L^{-1})e_1$ or $n_1n_2 = (1 + O(L^{-1}))e_2$. We now consider the total contribution from integers $n \in Q_j$ for all such "bad" choices of intervals $(A, A + A/L)$, $(A', A' + A'/L]$, $(B'_{j+1}, B'_{j+1} + B'_{j+1}/L]$, $(B'_u, B'_u + B'_u/L]$ and $(B'_{u+1}, B'_{u+1} + B'_{u+1}/L]$. Since each integer n occurs at most once, and $\lambda(n) = O(\tau(n))$, the contribution will be

$$O_\varepsilon\biggl(\sum_{n=(1+O(L^{-1}))e_1} \tau(n)\biggr) = O_\varepsilon(N^{1+\varepsilon}L^{-1}).$$

Similarly, if we have $p_{j+1}^{-1}n_1n_2 < Y$ but $(p'_{j+1})^{-1}n'_1n'_2 \geq Y$, then $p_{j+1}^{-1}n_1n_2 = (1 + O(L^{-1}))Y$. This gives

$$B_{j+1}Y \asymp AA' \leq N_1N_2 \ll_\varepsilon N^{1+\varepsilon},$$

so any n which is counted in this case will have a prime factor $p \ll N^{1+\varepsilon}/Y$ and such that $p^{-1}n = (1 + O(L^{-1}))Y$. Thus, on writing $n = pm$, we see that the total contribution in this case is

$$O\biggl(\sum_{p \ll N^{1+\varepsilon}/Y} \sum_{m=(1+O(L^{-1}))Y} \tau(pm)\biggr) = O_\varepsilon(N^{1+\varepsilon}Y^{-1}(1 + L^{-1}Y)) = O_\varepsilon(N^{1+\varepsilon}L^{-1}),$$

for $L \leq Y$.

Finally, if $B_u = B_{u+1}$, then it may happen that the condition $p_{u+1} < p_u$ is satisfied by some, but not all, pairs of primes (p_u, p_{u+1}) from the intervals $(B'_u, B'_u + B'_u/L]$ and $(B'_{u+1}, B'_{u+1} + B'_{u+1}/L]$. Clearly this problem cannot arise when $L \geq 2P_u$ since then the intervals $(B'_u, B'_u + B'_u/L]$ and $(B'_{u+1}, B'_{u+1} + B'_{u+1}/L]$ contain at most one prime each. It follows that any such n to be counted in this case must have two prime factors $p' > p \geq P_u \geq L/2$ with $p' = (1 + O(L^{-1}))p$. Hence the corresponding contribution is

$$O\Bigg(\sum_{\substack{p'>p\geq L/2 \\ p'=(1+O(L^{-1}))p}} \sum_{\substack{n\ll N \\ p'p|n}} \tau(n) \Bigg) = O_\varepsilon\Bigg(\sum_{\substack{p'>p\geq L/2 \\ p'=(1+O(L^{-1}))p}} \frac{N^{1+\varepsilon}}{p'p} \Bigg) = O_\varepsilon(N^{1+\varepsilon}L^{-1}).$$

We therefore find that our sum is bounded by

$$O_\varepsilon(L^5 N^{1-1/1200+\varepsilon} + N^{1+\varepsilon}L^{-1}),$$

whenever $L \leq Y$. We may then choose $L = N^{10^{-5}}$ say, to achieve the claimed power saving in the case of large d.

Small d. To handle small d it suffices to show that for any $\mathfrak{k} \leq \mathcal{C}$, $d \leq D_1$, and any nonprincipal $\chi \pmod{d}$,

$$\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})}^{*} \sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2 \\ cz\equiv y \,(\mathrm{mod}\, d\mathfrak{k})}}^{\flat} \beta_z\beta_y\chi(c) \ll_C \frac{N^2}{(\log X)^C}$$

for every $C > 0$. As is usually the case in prime number theory, the case of small moduli can be handled using some type of Siegel–Walfsiz theorem; we shall use the results of Mitsui [16] following the argument in [10].

Since

$$\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})}^{*} \chi(c) = 0,$$

it suffices to prove that if $\mathcal{U} = \mathcal{U}_1$ or $\mathcal{U}_2$ then there is a number $\mathfrak{M} = \mathfrak{M}(\mathcal{U}, d\mathfrak{k})$ such that

$$\sum_{\substack{z\in\mathcal{U} \\ z\equiv\alpha \,(\mathrm{mod}\, 2d\mathfrak{k})}} \beta_z = \mathfrak{M} + O_C\left(\frac{N}{(\log X)^C}\right)$$

for any $\gcd(\alpha, 2d\mathfrak{k}) = 1$ and $C > 0$, since β_z is supported on those z free of small prime factors, and $2d\mathfrak{k}$ is small. As before we may drop the summation condition $\flat$. For notational convenience, we set $q = 2d\mathfrak{k}$ and note that $q \leq (\log X)^{C_0}$ for some $C_0 > 0$.

As in the previous subsection we may assume that $\beta_z = \beta_{N(z)}$, where β_n is the indicator function of either Q_j or R. We describe the procedure for Q_j, the method for R being similar. We decompose z as s_1s_2 with $N(s_1)$ being the largest prime factor of $N(s_1s_2)$. The requirement that $n \in Q_j$ is then equivalent to a condition of the form $N(s_2) \in Q'_j$ together with a restriction of the type $N(s_1) \in I(s_2)$ for some real interval $I(s_2)$. Specifically, we have

$$Q'_{j+1} = \{p_2 \cdots p_{j+1} : p_{j=1} \in J,\ p_{j+1} < \cdots < p_2\}$$

and

$$I(s_2) = (p_2, \infty) \cap \left(\frac{N'}{N(s_2)}, \frac{N'(1+\varpi)}{N(s_2)} \right] \cap \left[\frac{Y}{N(s_2)}, \frac{X^{1/2-\delta}}{N(s_2)} \right),$$

where p_2 is the largest prime factor of $N(s_2)$. When $\mathcal{U}$ is given by (7-10) the condition on the size of $N(s_1 s_2)$ is exactly the condition

$$N(s_1) \in \left(\frac{N'}{N(s_2)}, \frac{N'(1+\omega)}{N(s_2)} \right],$$

and we have $\theta_0 < \arg z \leq \theta_0 + \omega_2$ exactly when

$$\theta_1(s_2) < \arg s_1 \leq \theta_1(s_2) + \omega_2,$$

with $\theta_1(s_2) = \theta_1 - \arg s_2$. It follows that

$$\sum_{\substack{z \in \mathcal{U} \\ z \equiv \alpha \,(\mathrm{mod}\, q)}} \beta_z = \sum_{\substack{N(s_2) \in Q'_j \\ \gcd(s_2, q) = 1}} \mathcal{N}(s_2, \alpha), \tag{10-4}$$

where $\mathcal{N}(s_2, \alpha)$ is the number of ideal numbers s_1 satisfying

$$s_1 s_2 \equiv \alpha \pmod q, \quad N(s_1) \in I(s_2), \quad \text{and} \quad \theta_1(s_2) < \arg s_1 \leq \theta_1(s_2) + \omega_2$$

and for which $N(s_1)$ is prime. We can estimate $\mathcal{N}(s_2, \alpha)$ using a form of the prime number theorem for arithmetic progressions over number fields, due to Mitsui [16]. As we remarked earlier, we can easily redivide our sectors in accordance with the condition $N(z) \sim N$ as opposed to $\|z\|_2 \sim N$, so we may apply Mitsui's theorem without worry in each of our sectors. If we put $\pi(X; q, \alpha, \theta)$ for the number of prime ideal numbers $\mathfrak{p}$ in a fixed ideal class satisfying $\mathfrak{p} \equiv \alpha \pmod q$ and having norm at most X with $0 \leq \arg \mathfrak{p} \leq \theta$, then Mitsui's theorem gives the estimate

$$\pi(X; q, \alpha, \theta) = \frac{w \theta R_K}{2^{r_1} h_K \varphi_K(\mathfrak{a})} \operatorname{Li}(X) + O_K\big(X \exp\big(-c\sqrt{\log X}\,\big)\big), \tag{10-5}$$

where r_1 is the number of real embeddings of K, w the number of roots of unity in K, R_K the regulator of K, and h_K the class number of K. Here c is an absolute constant. Since we do not care about dependence on K, we may take the implied constant in (10-5) as an absolute constant. We emphasize that (10-5) holds uniformly for $\theta \in [0, 2\pi]$ and for all $q \leq (\log X)^A$.

Applying (10-5) with $q = 2d\mathfrak{k}$ to estimate $\mathcal{N}(s_2, \alpha)$, we have $I(s_2) \subseteq (0, 2N/N(s_2)]$ and so we will need to know that $q = 2d\mathfrak{k} \leq (\log 2N/N(s_2))^A$ for some constant A. This holds whenever p divides an element of Q_j then one has $p \geq X^{\delta_1}$ with $\delta = (A \log\log X)/\log X$. Thus we will have $2N/N(s_2) \geq X^{\delta_1}$ and so

$$\delta_1 \log X \leq \log\left(\frac{N}{N(s_2)}\right),$$

which implies that

$$\log X \leq \left(\log\left(\frac{N}{N(s_2)}\right)\right)^{1/\varpi}.$$

Therefore whenever $2d\mathfrak{k} \le (\log X)^{C_0}$ we have

$$2d\mathfrak{k} \le (\log X)^{C_0} \le \left(\log\left(\frac{2N}{N(s_2)}\right)\right)^{C_0/\varpi}.$$

The required condition therefore holds when $\mathfrak{k} \le \mathfrak{K}$ and $d \le D_1$.

We may then conclude, as in [10], that

$$\mathcal{N}(s_2, \alpha) = \mathfrak{M}(s_2, d\mathfrak{k}, j, \mathcal{U}) + O\left(\frac{N}{N(s_2)} \exp(-c(\log X)^{\varpi/2})\right),$$

where the main term crucially is independent of α. Feeding this into (10-4) then completes our treatment of small d, and hence the proof of Proposition 3.7.

11. A generalization of the Jacobi–Kubota symbol and consequences

We will introduce and prove analogues of Proposition 23.1 and Theorem ψ in [5]. We introduce, for an ideal number α in a fixed class A, the vector

$$\widehat{\alpha} = (a_1, a_2) \in \mathbb{Z}^2$$

corresponding to the class A with basis produced as in Section 4. We then introduce the *Jacobi–Kubota symbol*

$$[\alpha] = i^{(a_1-1)/2}\left(\frac{a_2}{|a_1|}\right), \tag{11-1}$$

where $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol. Our generalized Jacobi–Kubota symbol $[\,\cdot\,]$ depends on the class A *and the choice of basis*, which we have suppressed.

Our goal is to obtain an analogue of Lemma 20.1 in [5], which shows that while $[\,\cdot\,]$ is not multiplicative, a suitable result exists to separate $[zw]$ into $[z][w]\kappa(zw)$, where $|\kappa(zw)| = 1$ and κ can be described explicitly. To do so we need to introduce an analogue of the "twist factor" $\xi_w(z)$ in [5]. Defining the analogue of $\xi_w(z)$ in the present setting is tricky, due to the fact that in general $\mathcal{O}_K$ need not be a unique factorization domain. In fact the situation is even more delicate than that; in order for our $\xi_w(z)$ to have nice properties, we must restrict the ideal classes of w, z as well as requiring w, z to satisfy a congruence condition like in (10-2).

To prepare for our definition, we first gather several of the key properties satisfied by Friedlander and Iwaniec's $\xi_w(z)$ in [5]. In particular, it satisfies the following:

(1) It satisfies an equation of the form

$$[z][w] = \varepsilon[zw]\xi_w(z),$$

where $\varepsilon = \pm 1$ depending only on the quadrants containing z, w, respectively.

(2) It is multiplicative for each $w \in \mathbb{Z}[i]$: one has $\xi_w(z_1)\xi_w(z_2) = \xi_w(z_1 z_2)$.

(3) It is symmetric: $\xi_w(z) = \xi_z(w)$ for $w, z \in \mathbb{Z}[i]$.

(4) (Lemma 21.1 in [5]) For $q = |w_1 w_2|^2$ and $d = |\gcd(w_1, \overline{w_2})|^2$ one has

$$\sum_{\zeta \pmod q} \xi_{w_1}(\zeta)\xi_{w_2}(\zeta) = \begin{cases} q\varphi(d)\varphi(q/d) & \text{if } q, d \text{ are squares,} \\ 0 & \text{otherwise.} \end{cases}$$

(5) For $w = u + iv$ and $\omega \equiv -v\bar{u} \pmod q$ with $q = |w|^2$, one has

$$\xi_w(z) = \left(\frac{ur - vs}{q}\right) \quad \text{and} \quad \xi_w(z) = \left(\frac{r + \omega s}{q}\right),$$

where $z = r + is$.

We would like to define our function $\xi_\alpha(z)$ to have the same properties. Unfortunately, it seems that at least some of these properties require special structures of the Gaussian integers $\mathbb{Z}[i]$. Thus, some more preparatory work is needed before we can define our stand-in for the symbol $\xi_w(z)$. We then check that our analogous symbol has the necessary properties to carry out the proofs of analogous statements in [5].

First we note that our symbol $\xi_\alpha(z)$ depends on α, and in particular, depends on the class A of α. This of course is a trivial point when $K = \mathbb{Q}(i)$, since $\mathbb{Z}[i]$ has unique factorization. Next we will also need to restrict the class of the *inputs* z, in order for our symbol to be well-behaved. This is far from ideal and is likely too restrictive, but it suffices for our purposes in this paper. Indeed, later we will see that it is necessary to define a separate symbol ξ for each class of ideal numbers *along with a basis* of said ideal numbers.

The most important property turns out to be (1), so we define our symbol with this in mind. To simplify matters we will assume that in our composition law the bilinear form $Q_{A,B}(w, z)$ is given by $w_1 z_1 + w_2 z_2$. In particular, we fix bases $\{\alpha_1, \alpha_2\} \subset A$, $\{\beta_1, \beta_2\} \subset B$, $\{\gamma_1, \gamma_2\} \subset C = A \cdot B$ so that

$$(\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 y_1 + \beta_2 y_2) = (x_1 \ell_1(y_1, y_2) + x_2 \ell_2(y_1, y_2))\gamma_1 + (x_1 y_1 + x_2 y_2)\gamma_2.$$

We begin with the Jacobi symbol

$$\left(\frac{w_1 z_1 + w_2 z_2}{|w_1 \ell_1(z_1, z_2) + w_2 \ell_2(z_1, z_2)|}\right),$$

where $R_{A,B}(w, z) = w_1 \ell_1 + w_2 \ell_2$. We can extend the definition of the Jacobi symbol by setting

$$\left(\frac{a}{b}\right) = \left(\frac{a}{|b|}\right)(a, b)_\infty,$$

where

$$(a, b)_\infty = \begin{cases} -1 & \text{if } a, b < 0, \\ 1 & \text{otherwise,} \end{cases}$$

is the Hilbert symbol. Next we note quadratic reciprocity, which states for a, b odd and coprime that

$$\left(\frac{a}{|b|}\right)\left(\frac{b}{|a|}\right) = (-1)^{((a-1)/2)\cdot((b-1)/2)}(a, b)_\infty \tag{11-2}$$

and for $d > 0$ odd we have

$$\left(\frac{2}{d}\right) = (-1)^{(d^2-1)/4}.$$

Clearly, not both $Q_{A,B}$, $R_{A,B}$ can be even otherwise the corresponding ideal number is not primitive. Without loss of generality, let us suppose that $w_1z_1 + w_2z_2$ is odd. Let 2^k be the highest power of 2 dividing $w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)$. Then

$$\left(\frac{w_1z_1 + w_2z_2}{|w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)|}\right) = \left(\frac{w_1z_1 + w_2z_2}{2^{-k}|w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)|}\right).$$

We put

$$u = w_1z_1 + w_2z_2, \quad v = w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)$$

for simplicity. Applying quadratic reciprocity (11-2) then gives

$$\begin{aligned}\left(\frac{w_1z_1 + w_2z_2}{2^{-k}(w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2))}\right)(u, v)_\infty &= \left(\frac{2^{-k}v}{u}\right)(-1)^{((u-1)/2)\cdot((2^{-k}v-1)/2)} \\ &= \left(\frac{2^k}{u}\right)(-1)^{((u-1)/2)\cdot((2^{-k}v-1)/2)}\left(\frac{v}{u}\right).\end{aligned}$$

Let us write $u_1 = \gcd(u, z_2)$ and $u_2 = u/u_1$. From the definition we see that $u_1 = \gcd(w_1, z_2)$. Put $w_1 = u_1w_1^*$, $z_2 = u_1z_2^*$ with $\gcd(w_1^*, z_2^*) = 1$. We now make use of the fact

$$w_1z_1 + w_2z_2 \equiv 0 \pmod{u}. \tag{11-3}$$

We treat the congruence modulo u_2 first. Plainly, $\gcd(z_2, u_2) = 1$. (11-3) then implies

$$w_2 \equiv -z_2^{-1}w_1z_1 \pmod{u_2}.$$

Substituting this into $R_{A,B}(w, z)$ gives

$$\begin{aligned}w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2) &\equiv w_1\ell_1(z_1, z_2) - z_2^{-1}w_1z_1\ell_2(z_1, z_2) \pmod{u_2} \\ &\equiv z_2^{-1}w_1(z_2\ell_1(z_1, z_2) - z_1\ell_2(z_1, z_2)) \pmod{u_2}.\end{aligned}$$

Modulo u_1 we must have

$$\begin{aligned}w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2) &\equiv w_2\ell_2(z_1, z_2) \pmod{u_1} \\ &\equiv -z_1^{-1}w_2(z_2\ell_1(z_1, z_2) - z_1\ell_2(z_1, z_2)) \pmod{u_1}.\end{aligned}$$

In both cases, we see that $R_{A,B}(w, z)$ is congruent to a multiple of the quadratic form

$$g(z_1, z_2) = z_2\ell_1(z_1, z_2) - z_1\ell_2(z_1, z_2),$$

which we now interpret. By definition, our composition law gives the relation

$$\begin{aligned}(z_2\alpha_1 - z_1\alpha_2)(z_1\beta_1 + z_2\beta_2) &= R_{A,B}(z_2, -z_1; z_1, z_2)\gamma_1 + Q_{A,B}(z_2, -z_1; z_1, z_2)\gamma_2 \\ &= (z_2\ell_1(z_1, z_2)) - z_1\ell_2(z_1, z_2)\gamma_1 + (z_2z_1 - z_1z_2)\gamma_2 \\ &= g(z_1, z_2)\gamma_1.\end{aligned} \tag{11-4}$$

Dividing both sides by γ_1 we then see that $g(z_1, z_2)$ must be equivalent to the norm form of $\mathcal{O}_K$.

We must now relate $g(z_1, z_2)$ to $N(z) = N(J(z_1\beta_1 + z_2\beta_2))$. Note that

$$g(z_1, z_2) = \gamma_1^{-1}(\alpha_1 z_2 - \alpha_2 z_1)(\beta_1 z_1 + \beta_2 z_2)$$

is divisible by $z = \beta_1 z_1 + \beta_2 z_2$, which implies that $g(z_1, z_2)$ is a rational integer divisible by $N(z)$. By primitivity we then see that $g(z_1, z_2)$ must be a constant multiple of $N(z)$, the constant depending only on the classes A, B. We summarize this as a lemma:

Lemma 11.1. *Let $g(x, y)$ be the integral binary quadratic form which arises from the composition law (11-4). Then $g(z_1, z_2)$ is a constant multiple of $N(J(\beta_1 z_1 + \beta_2 z_2))$, with the constant depending only on the classes A, B and choices of bases of A, B, $A \cdot B$.*

Similarly, since $v = w_1\ell_1 + w_2\ell_2$ is divisible by 2^k, we may assume without loss of generality that ℓ_1 is odd to obtain

$$w_1 \equiv -\ell_1^{-1} w_2 \ell_2 \pmod{2^k}$$

and this implies that

$$\begin{aligned} w_1 z_1 + w_2 z_2 &\equiv -\ell_1^{-1} w_2 \ell_2 z_1 + w_2 z_2 \pmod{2^k} \\ &\equiv -\ell_1^{-1} w_2 (z_2 \ell_1 - z_1 \ell_2) \pmod{2^k} \\ &\equiv -\ell_1^{-1} w_2 g(z_1, z_2) \pmod{2^k}. \end{aligned}$$

Since $u = w_1 z_1 + w_2 z_2$ is odd by assumption, it follows that $g(z_1, z_2)$ must be odd as well.

Continuing on, with $u = u_1 u_2$ as before, we have

$$\begin{aligned} z_2^{-1} w_1 (z_2 \ell_1(z_1, z_2) - z_1 \ell_2(z_1, z_2)) &\equiv z_2^{-1} w_1 g(z_1, z_2) \pmod{u_2}, \\ -z_1^{-1} w_2 (z_2 \ell_1(z_1, z_2) - z_1 \ell_2(z_1, z_2)) &\equiv -z_1^{-1} w_2 g(z_1, z_2) \pmod{u_1} \end{aligned}$$

which implies that

$$\begin{aligned} \left(\frac{v}{u}\right) &= \left(\frac{-z_1^{-1} w_2 g(z_1, z_2)}{u_1}\right)\left(\frac{z_2^{-1} w_1 g(z_1, z_2)}{u_2}\right) \\ &= \left(\frac{-z_1 w_2}{u_1}\right)\left(\frac{z_2^* w_1^*}{u_2}\right)\left(\frac{g(z_1, z_2)}{u}\right). \end{aligned}$$

Observe that, by definition, we have

$$u_2 = w_1^* z_1 + w_2 z_2^*.$$

Since u is odd, it follows that exactly one of the pairs $\{w_1, z_1\}$, $\{w_2, z_2\}$ consists of two odd numbers. Without loss of generality, we assume that w_1, z_1 are both odd. We write $z_2^* = 2^{k_2} \nu_2$ with ν_2 odd. Applying

(11-2) we find that

$$\begin{aligned}\left(\frac{w_1^* z_2^*}{u_2}\right) &= \left(\frac{2^{k_2}}{u_2}\right)\left(\frac{w_1^* v_2}{w_1^* z_1 + w_2 z_2^*}\right)\\ &= \left(\frac{2^{k_2}}{u_2}\right)(-1)^{(u_2-1)/2}(-1)^{(w_1^* v_2-1)/2}\left(\frac{w_1^* z_1 + w_2 z_2^*}{|w_1^* v_2|}\right)\\ &= \left(\frac{2^{k_2}}{u_2}\right)(-1)^{(u_2-1)/2}(-1)^{(w_1^* v_2-1)/2}\left(\frac{w_2 z_2^*}{|w_1^*|}\right)\left(\frac{w_1^* z_1}{|v_2|}\right)\\ &= \left(\frac{2^{k_2}}{u_2}\right)(-1)^{(u_2-1)/2}(-1)^{(w_1^* v_2-1)/2}\left(\frac{w_2}{|w_1^*|}\right)\left(\frac{z_2^*}{|w_1^*|}\right)\left(\frac{w_1^*}{|v_2|}\right)\left(\frac{z_1}{|v_2|}\right).\end{aligned}$$

Applying (11-2) repeatedly we obtain

$$\left(\frac{w_1^* z_2^*}{u_2}\right) = \varepsilon_1\left(\frac{w_2}{|w_1^*|}\right)\left(\frac{v_2}{|z_1|}\right) \tag{11-5}$$

where

$$\varepsilon_1 = \left(\frac{2^{k_2}}{u_2}\right)\left(\frac{2^{k_2}}{|w_1^*|}\right)(-1)^{(u_2-1)/2}(-1)^{(w_1^* v_2-1)/2}(-1)^{(w_1^*-1)/2}(-1)^{(z_1-1)/2}(w_1^*, v_2)_\infty(v_2, z_1)_\infty. \tag{11-6}$$

Next we note that

$$\left(\frac{-z_1 w_2}{u_1}\right) = \left(\frac{-1}{u_1}\right)\left(\frac{z_1}{u_1}\right)\left(\frac{w_2}{u_1}\right).$$

It follows that

$$\begin{aligned}\left(\frac{-z_1 w_2}{u_1}\right)\left(\frac{z_2^* w_1^*}{u_2}\right) &= \varepsilon_1\left(\frac{-1}{u_1}\right)\left(\frac{u_1}{|z_1|}\right)(-1)^{(w_2-1)/2}(-1)^{(z_1-1)/2}(w_1, u_1)_\infty\left(\frac{w_2}{|u_1|}\right)\left(\frac{w_2}{|w_1^*|}\right)\left(\frac{v_2}{|z_1|}\right)\\ &= \varepsilon_2\left(\frac{z_2}{|z_1|}\right)\left(\frac{w_2}{|w_1|}\right).\end{aligned}$$

Here we have

$$\varepsilon_2 = \varepsilon_1(-1)^{(w_2-1)/2}(-1)^{(z_1-1)/2}(w_1, u_1)_\infty\left(\frac{-1}{u_1}\right)\left(\frac{2^{k_2}}{|z_1|}\right). \tag{11-7}$$

Finally, by (11-2) we have

$$\left(\frac{u}{g(z_1, z_2)}\right) = (-1)^{(g(z_1,z_2)-1)/2}(-1)^{(u-1)/2}\left(\frac{g(z_1, z_2)}{u}\right).$$

Collecting these calculations we conclude that

$$\left(\frac{w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)}{|w_1 z_1 + w_2 z_2|}\right) = \left(\frac{w_2}{|w_1|}\right)\left(\frac{z_2}{|z_1|}\right)\left(\frac{w_1 z_1 + w_2 z_2}{g(z_1, z_2)}\right)\varepsilon(w, z), \tag{11-8}$$

where

$$\varepsilon(w, z) = \varepsilon_2(-1)^{(g(z_1,z_2)-1)/2}(-1)^{(u-1)/2}. \tag{11-9}$$

From (11-9) we make the following conclusion:

Lemma 11.2. *Let $\varepsilon(w, z)$ be given as in (11-9). Then $\varepsilon(w, z) \in \{\pm 1\}$ and its value is determined by the quadrants of (z_1, z_2), (w_1, w_2), the congruence classes of w_1, w_2, z_1, z_2 modulo 8, and whether 2 divides z_2 an odd or an even number of times.*

Since we have insisted that w, z belong to fixed congruence classes modulo $8etn$ as in (10-2) it follows that $\varepsilon(w, z)$ can be determined as a function of the congruence class alone, except for the condition on whether z_2 is divisible by an even or odd power of 2. We can treat the two cases separately, and in each case assume that $\varepsilon(w, z)$ is constant.

These calculations compel us to define our analogue of the twist factor in the multiplication law for the Jacobi–Kubota symbol as

$$\xi_w(z) = \left(\frac{w_1 z_1 + w_2 z_2}{g(w_1, w_2)}\right). \tag{11-10}$$

Note that $\xi_w(z)$ depends on the ideal classes of w, z and a choice of basis for the ideal classes.

Next we observe for w, z satisfying (10-2), w, z are in the same class and therefore $R_{A,B}(w, z) = R_{A,A}(w, z)$ must be symmetric in w, z. From here it follows that

$$\begin{aligned} z_2^{-1} w_1 g(z_1, z_2) &\equiv R_{A,A}(w, z) \pmod{u} \\ &\equiv R_{A,A}(z, w) \pmod{u} \\ &\equiv w_1^{-1} z_2 g(w_1, w_2) \pmod{u}. \end{aligned}$$

This implies that

$$\left(\frac{g(z_1, z_2)}{u}\right)\left(\frac{g(w_1, w_2)}{u}\right) = 1.$$

Thus, up to a factor ε depending at most on congruence classes and signs of w, z, we have

$$\xi_w(z) = \varepsilon \xi_z(w). \tag{11-11}$$

Summarizing, we obtain the following analogue of Lemma 20.1 in [6]:

Lemma 11.3. *Let w, z satisfy the hypothesis given in (10-2). Then there exist numbers $\theta(w, z) \in \{-1, 1\}$ depending only on the signs and congruence classes of w, z modulo $8etn$ such that*

$$\left(\frac{Q_{A,B}(w, z)}{|R_{A,B}(w, z)|}\right) = \theta(w, z)\left(\frac{w_2}{|w_1|}\right)\left(\frac{z_2}{|z_1|}\right)\xi_z(w). \tag{11-12}$$

Next we show that the analogue of Lemma 21.1 in [5] holds:

Lemma 11.4. *For fixed elements w, v in the class A and*

$$q = g(w_1, w_2) g(v_1, v_2) \quad \text{and} \quad d = \gcd(g(w_1, w_2), g(v_1, v_2)),$$

we have

$$\sum_{z \,(\mathrm{mod}\, q)} \xi_{w_1}(z) \xi_{w_2}(z) = \begin{cases} q\varphi(d)\varphi(q/d) & \text{if } q, d \text{ are squares,} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We have

$$\sum_{z \pmod q} \xi_w(z)\xi_v(z) = \sum_{z \pmod q} \left(\frac{w_1z_1+w_2z_2}{g(w_1,w_2)}\right)\left(\frac{v_1z_1+v_2z_2}{g(v_1,v_2)}\right)$$
$$= \sum_{z \pmod q} \left(\frac{(w_1z_1+w_2z_2)(v_1z_1+v_2z_2)}{d}\right)\left(\frac{w_1z_1+w_2z_2}{g(w_1,w_2)/d}\right)\left(\frac{v_1z_1+v_2z_2}{g(v_1,v_2)/d}\right).$$

From here we see that the final sum is zero unless each of the summands is equal to 1 or 0 identically. This is only the case when d, $g(w_1,w_2)/d$, $g(w_1,w_2)/d$ are all squares. Since $d \mid \gcd(g(w_1,w_2), g(v_1,v_2))$ and $d \nmid \Delta(f)$ it follows that w_1x+w_2y, v_1x+v_2y are not proportional modulo d. From here we see that, modulo d, the number of solutions to $\gcd(w_1x+w_2y,d) = \gcd(v_1x+v_2y,d) = 1$ is equal to $\varphi(d)^2$. Similarly, modulo $g(w_1,w_2)/d$ and $g(v_1,v_2)/d$ there are $g(w_1,w_2)\varphi(g(w_1,w_2)/d)/d$ solutions to $\gcd(w_1x+w_2y, g(w_1,w_2)/d) = 1$ and $\gcd(v_1x+v_2y, g(v_1,v_2)/d) = 1$, respectively. Lifting to the modulus q yields

$$\frac{q^2}{d^2}\cdot\varphi(d)^2\cdot\frac{g(w_1,w_2)g(v_1,v_2)}{d^2}\varphi(g(w_1,w_2)/d)\varphi(g(v_1,v_2)/d) = q\varphi(d)\varphi(q/d),$$

since $\gcd(q/d^2,d)=1$. This completes the proof. □

Lemma 11.4 is analogous to Lemma 21.1 in [5].

We now prove the following analogue of Lemma 21.2 in [5]:

Proposition 11.5. *Let A, B be classes of ideal numbers. Put*

$$\mathcal{Q}(M,N) = \sum_w{}^{*} \sum_z \alpha_w\beta_z\xi_w(z), \tag{11-13}$$

where α_w, β_z are bounded real coefficients supported in appropriate fundamental domains for A, B having norm bounded by M, N, respectively. Then for all $\varepsilon > 0$ we have

$$\mathcal{Q}(M,N) \ll_\varepsilon (M+N)^{1/12}(MN)^{1/12+\varepsilon}. \tag{11-14}$$

Proof. Applying Cauchy's inequality we obtain

$$|\mathcal{Q}(M,N)|^2 \le \|\beta\|_2^2 \sum_z \left|\sum_w{}^{*} \alpha_w\xi_w(z)\right|^2$$
$$= \|\beta\|_2^2 \sum_{w_1}{}^{*}\sum_{w_2}{}^{*} \alpha_{w_1}\alpha_{w_2}\sum_z \xi_{w_1}(z)\xi_{w_2}(z).$$

We then find that splitting z into congruence classes modulo $q = g(w_1)g(w_2)$ that

$$\sum_z \xi_{w_1}(z)\xi_{w_2}(z) = \sum_{\zeta \pmod q} \xi_{w_1}(\zeta)\xi_{w_2}(\zeta)\cdot\left(\frac{c_fN}{q^2} + O_f\left(\frac{\sqrt{N}}{q}+1\right)\right)$$

where

$$c_f = \lim_{s\to 1}(s-1)\zeta_K(s).$$

We obtain, by Lemma 11.1 and using (11-11) if necessary,

$$\mathcal{Q}(M,N)^2 \ll N^2 \sum_{\substack{m_1,m_2\le M\\ m_1m_2=\square}}\sum \tau(m_1m_2) + NM^4\bigl(\sqrt{N}+M^2\bigr), \tag{11-15}$$

which gives the bound

$$\mathcal{Q}(M,N) \ll_\varepsilon (M^3N^{1/2}+M^2N^{3/4}+M^{1/2}N)(MN)^\varepsilon.$$

In the next step we shall apply Hölder's inequality to obtain

$$\mathcal{Q}(M,N)^k \ll M^{k-1}\sum_w{}^{*}\Bigl|\sum_z \beta_z\xi_w(z)\Bigr|^k = M^{k-1}\widetilde{\mathcal{Q}}(M,N^k),$$

say. In [5] the next step is to argue that $\widetilde{\mathcal{Q}}(M,N^k)$ can be written as a bilinear form of the shape (11-13), using the fact that in the case $K=\mathbb{Q}(i)$ that $\xi_w(z)$ is multiplicative in z. In general this is not the case. However, we are free to choose a basis for the class B^k for each positive integer k, which allows one to write

$$\xi_w(z_1)\cdots\xi_w(z_k) = \xi_w^{(k)}(z_1\cdots z_k) \tag{11-16}$$

in a consistent way. Recalling (11-10), we note that

$$\xi_w(z_1)\xi_w(z_2) = \left(\frac{Q_{B,B}(z_1)Q_{B,B}(z_2)}{g(w_1,w_2)}\right).$$

The numerator is a bilinear form in z_1, z_2. Using composition laws to write

$$z_1z_2 = R_{B^2}(z_1,z_2)\gamma_1^{(2)} + Q_{B^2}(z_1,z_2)\gamma_2^{(2)}$$

as ideal numbers, we see that we can apply a change of variables, depending only on w, the class B, and the choice of bases, so that the numerator $Q_{B,B}(z_1)Q_{B,B}(z_2)$ is a linear form in $R_{B^2}(z_1,z_2)$, $Q_{B^2}(z_1,z_2)$. Inductively, we then find that

$$\xi_w(z_1)\cdots\xi_w(z_k) = \left(\frac{L_w(z_1\cdots z_k)}{g(w_1,w_2)}\right),$$

where L_w is a linear form in two variables with coefficients depending at most on w and evaluates $z_1\cdots z_k$ in terms of its representation as an element in the lattice of the corresponding ideal numbers. Defining the right-hand side as $\xi_w^{(k)}(z_1\cdots z_k)$ we obtain (11-16). Replacing $\xi_w(\cdot)$ with $\xi_w^{(k)}(\cdot)$ in (11-13) shows that (11-15) holds, and therefore we may proceed as in [5] after applying Hölder's inequality to conclude

$$\mathcal{Q}(M,N)^k \ll_\varepsilon M^{k-1}(M^3N^{k/2}+M^2N^{3k/4}+M^{1/2}N^k)(MN)^\varepsilon,$$

which upon taking k-th roots gives us the bound

$$\mathcal{Q}(M,N) \ll_\varepsilon (M^{1+2/k}N^{1/2}+M^{1+1/k}N^{3/4}+M^{1-1/2k}N)(MN)^\varepsilon$$

for all positive $k \in \mathbb{N}$. Switching the roles of M, N and applying Lemma 11.4, we obtain as in [5] that

$$\mathcal{Q}(M, N) \ll_\varepsilon (M+N)^{1/12}(MN)^{1/12+\varepsilon}$$

upon setting $k = 6$. □

Next we move on to proving the analogue of Proposition 22.1 in [5]. We define, for any ideal number z, a rational integer k, and a character χ modulo $4d$ the Hecke character

$$\psi(z) = \chi(z)\left(\frac{z}{|z|}\right)^k. \tag{11-17}$$

Consider the sum

$$\mathcal{K}(N) = \sum_{z\in\mathfrak{B}}^{\wedge} \psi(z)[wz]$$

and

$$\mathcal{K}^*(N) = \sum_{\substack{z\in\mathfrak{B}\\ \gcd(z,w)=1}}^{\wedge} \psi(z)[wz],$$

where $\mathfrak{B}$ is a narrow sector contained in the intersection of a fundamental domain for the ideal class numbers containing z having norm bounded N. We treat w as a fixed primitive ideal number. Our analogue of Proposition 22.1 in [5] is thus:

Proposition 11.6. *Given ψ and w as above we have*

$$\mathcal{K}(N) \ll d(|k|+1)|w|N^{3/4}\log(|w|N) \tag{11-18}$$

and

$$\mathcal{K}^*(N) \ll d(|k|+1)|w|\tau(N(w))N^{3/4}\log(|w|N). \tag{11-19}$$

Proof. Just like the proof of Proposition 22.1 in [5], the key result needed to obtain the necessary cancellation is the Polya–Vinogradov theorem, which asserts that

$$\sum_{n\le N}\chi(n) \ll \sqrt{q}\log q$$

for every nontrivial Dirichlet character $\chi \pmod q$ with an absolute implied constant. To estimate $\mathcal{K}(N)$ we apply Lemma 11.3 to obtain

$$\mathcal{K}(N) = [w]\sum_{z\in\mathfrak{B}}^{\wedge} \varepsilon(w,z)\psi(z)[z]\xi_w(z),$$

and by breaking the sum up to finitely many congruence classes if necessary, we may factor the ε-factor out (because it will be constant) to obtain

$$\mathcal{K}(N) = [w]\varepsilon\sum_{z\in\mathfrak{B}}^{\wedge} \psi(z)[z]\xi_w(z).$$

Breaking the sum up into a double sum over rational integers forming vectors running over $\mathfrak{B}$ as in [5] and applying Polya–Vinogradov we obtain (11-18) and (11-19) as required. □

Put

$$\lambda(n) = \sum_{N(z)=n}^{\wedge} \psi(z)[z],$$

the sum restricted to a fundamental domain of ideal numbers so each ideal is represented at most once. Consider the sum

$$\mathcal{L}(M, N) = \sum_{m} \sum_{n} \alpha(m)\beta(n)\lambda(cmn), \tag{11-20}$$

where α, β are complex coefficients having norm at most 1 and supported on $1 \le m \le M$ and $n \le N$. Likewise, let $\mathcal{L}^*(M, N)$ be the subsum of (11-20) restricted to $\gcd(m, n) = 1$. Combining Proposition 11.5 and Lemma 11.3 then gives the following analogue of Proposition 23.1 in [5]:

Proposition 11.7. *For any complex coefficients $\alpha(m)$, $\beta(n)$ as above and for any positive integer c,*

$$\mathcal{L}(M, N) \ll \tau(c)(M + N)^{1/12}(MN)^{1/12+\varepsilon}. \tag{11-21}$$

We also introduce the analogues of $\mathcal{K}(N)$, $\mathcal{K}^*(N)$:

$$\mathcal{L}(N) = \sum_{n \le N} \lambda(mn), \quad \mathcal{L}^*(N) = \sum_{\substack{n \le N \\ \gcd(m,n)=1}} \lambda(mn) \tag{11-22}$$

and obtain the following analogue of Proposition 23.2 in [5] by applying Proposition 11.6:

Proposition 11.8. *For ψ as defined by (11-17) and positive integer m we have the bounds*

$$\mathcal{L}(N) \ll d(|k| + 1)\tau(m)^4 \sqrt{m} N^{3/4} \log(mN) \tag{11-23}$$

and

$$\mathcal{L}^*(N) \ll d(|k| + 1)\tau(m)^2 \sqrt{m} N^{3/4} \log(mN). \tag{11-24}$$

These estimates then imply the following analogue of Theorem ψ in [5]:

Proposition 11.9. *For any $c \ge 1$ we have*

$$\sum_{n \le X} \Lambda(n)\lambda(cn) \ll cd(|k| + 1)X^{6/77} \tag{11-25}$$

with the absolute constant dependent only on f.

Proof. This is the same as the proof of Theorem ψ in [5] with Propositions 23.1 and 23.2 replaced by Propositions 11.7 and 11.8, respectively. □

Acknowledgements

The author owes an incalculable debt of gratitude to John Friedlander, whose encouragement and guidance made this paper possible. The author also thanks D. R. Heath-Brown whose work on prime number theory is an inspiration for the present work, D. Schindler and J. Maynard for helpful discussions, to C. L. Stewart for a careful reading of an earlier version of this paper and for providing instrumental advice, and to S. Yamagishi whose collaboration and friendship was instrumental in the author's pursuit of prime number theory. The author would like to express his deepest appreciation for the hard work done by the referee and for the helpful suggestions made.

References

[1] A. Balog, V. Blomer, C. Dartyge, and G. Tenenbaum, "Friable values of binary forms", *Comment. Math. Helv.* **87**:3 (2012), 639–667. MR Zbl

[2] J. G. van der Corput, "Über Summen von Primzahlen und Primzahlquadraten", *Math. Ann.* **116**:1 (1939), 1–50. MR Zbl

[3] E. Fouvry and H. Iwaniec, "Gaussian primes", *Acta Arith.* **79**:3 (1997), 249–287. MR Zbl

[4] J. Friedlander and H. Iwaniec, "Asymptotic sieve for primes", *Ann. of Math.* (2) **148**:3 (1998), 1041–1065. MR Zbl

[5] J. Friedlander and H. Iwaniec, "The polynomial $X^2 + Y^4$ captures its primes", *Ann. of Math.* (2) **148**:3 (1998), 945–1040. MR Zbl

[6] J. B. Friedlander and H. Iwaniec, "Gaussian sequences in arithmetic progressions", *Funct. Approx. Comment. Math.* **37** (2007), 149–157. MR Zbl

[7] B. Green, "Roth's theorem in the primes", *Ann. of Math.* (2) **161**:3 (2005), 1609–1636. MR Zbl

[8] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions", *Ann. of Math.* (2) **167**:2 (2008), 481–547. MR Zbl

[9] D. R. Heath-Brown, "Primes represented by $x^3 + 2y^3$", *Acta Math.* **186**:1 (2001), 1–84. MR Zbl

[10] D. R. Heath-Brown and X. Li, "Prime values of $a^2 + p^4$", *Invent. Math.* **208**:2 (2017), 441–499. MR Zbl

[11] D. R. Heath-Brown and B. Z. Moroz, "Primes represented by binary cubic forms", *Proc. Lond. Math. Soc.* (3) **84**:2 (2002), 257–288. MR Zbl

[12] P. C.-H. Lam, D. Schindler, and S. Y. Xiao, "On prime values of binary quadratic forms with a thin variable", *J. Lond. Math. Soc.* (2) **102**:2 (2020), 749–772. MR Zbl

[13] B. Landreau, "Majorations de fonctions arithmétiques en moyenne sur des ensembles de faible densité", exposé 13 in *Séminaire de Théorie des Nombres* (Talence, France, 1987-1988), Univ. Bordeaux I, Talence, France, 1988. MR Zbl

[14] X. Li, "Prime values of a sparse polynomial sequence", *Duke Math. J.* **171**:1 (2022), 101–208. MR Zbl

[15] J. Maynard, "Primes represented by incomplete norm forms", *Forum Math. Pi* **8** (2020), art. id. e3. MR Zbl

[16] T. Mitsui, "Generalized prime number theorem", *Jpn. J. Math.* **26** (1956), 1–42. MR Zbl

[17] M. Pandey, "On Eisenstein primes", *Integers* **18** (2018), art. id. A59. MR Zbl

[18] R. C. Vaughan, "Mean value theorems in prime number theory", *J. Lond. Math. Soc.* (2) **10** (1975), 153–162. MR Zbl

Communicated by Andrew Granville
Received 2022-07-17 Revised 2023-08-30 Accepted 2023-10-12

stanleyyao.xiao@unbc.ca *Department of Mathematics and Statistics, University of Northern British Columbia, Prince George, Canada*

ALGEBRA AND NUMBER THEORY 18:9 (2024)
https://doi.org/10.2140/ant.2024.18.1681

Affine Deligne–Lusztig varieties with finite Coxeter parts

Xuhua He, Sian Nie and Qingchao Yu

We study affine Deligne–Lusztig varieties $X_{w(b)}$ when the finite part of the element w in the Iwahori–Weyl group is a partial σ-Coxeter element. We show that such w is a cordial element and $X_{w(b)} \neq \varnothing$ if and only if b satisfies a certain Hodge–Newton indecomposability condition. Our main result is that for such w and b, $X_{w(b)}$ has a simple geometric structure: the σ-centralizer of b acts transitively on the set of irreducible components of $X_{w(b)}$; and each irreducible component is an iterated fibration over a classical Deligne–Lusztig variety of Coxeter type, and the iterated fibers are either $\mathbb{A}^1$ or $\mathbb{G}_m$.

1. Introduction

1A. ***Classical/affine Deligne–Lusztig varieties.*** The classical Deligne–Lusztig varieties were introduced by Deligne and Lusztig in [3]. They play a crucial role in the representation theory of finite reductive groups. They are defined for a connected reductive group $\boldsymbol{G}$ over a finite field $\mathbb{F}_q$. For any w in the (finite) Weyl group of $\boldsymbol{G}(\bar{\mathbb{F}}_q)$, the corresponding Deligne–Lusztig variety X_w is a certain locally closed subvariety of the flag variety of $\boldsymbol{G}(\bar{\mathbb{F}}_q)$, and it admits a natural action of the finite reductive group $\boldsymbol{G}(\mathbb{F}_q)$. It is known that

(a) the classical Deligne–Lusztig variety X_w is smooth and of dimension equal to the length of w, and the finite reductive group $\boldsymbol{G}(\mathbb{F}_q)$ acts transitively on the set of irreducible components of X_w.

Affine Deligne–Lusztig varieties were introduced by Rapoport in [27] as the affine analog of classical Deligne–Lusztig varieties. They serve as group-theoretic models for Shimura varieties and shtukas. They are defined for a connected reductive group $\boldsymbol{G}$ over a nonarchimedean local field F. Let $\breve{F}$ be the

MSC2020: 11G25, 20G25.
Keywords: affine Deligne–Lusztig varieties, Coxeter elements.

completion of the maximal unramified extension of F. For any element w in the Iwahori–Weyl group $\tilde{W}$ of $\boldsymbol{G}(\breve{F})$ and any element $b \in \boldsymbol{G}(\breve{F})$, the corresponding affine Deligne–Lusztig variety $X_w(b)$ is a certain locally closed subscheme of finite type in the affine flag variety of $\boldsymbol{G}(\breve{F})$, and it admits a natural action of the σ-centralizer $\boldsymbol{J}_b(F)$ of b. Unlike classical Deligne–Lusztig varieties, which have the nice geometric structure described in (a), the geometric structures of affine Deligne–Lusztig varieties are very complicated:

- For many pairs (w, b), $X_w(b)$ are empty.
- Even if $X_w(b)$ is nonempty, it is not equidimensional in general, and it is very difficult to determine its dimension.
- In general, the group $\boldsymbol{J}_b(F)$ does not act transitively on the set of irreducible components of $X_w(b)$, and very little is known about this $\boldsymbol{J}_b(F)$-action.
- The irreducible components of $X_w(b)$, in general, have a very complicated geometric structure.

We refer to the survey article [15] and [16] for recent developments regarding the nonemptiness pattern and the dimension formula for $X_w(b)$.

1B. ***Main result.*** Milićević and Viehmann [25] introduced the notion of cordial elements. The geometry of the affine Deligne–Lusztig varieties associated with cordial elements is "well-behaved" in the following sense. If w is a cordial element, then the elements b with $X_w(b) \neq \varnothing$ form a saturated set in the sense of [25, Theorem 1.1], and for any such b, there is a simple dimension formula for $X_w(b)$. Moreover, $X_w(b)$ is equidimensional. Schremmer gave a classification of the cordial elements in [28].

However, even for a cordial element w, very little is known about the $\boldsymbol{J}_b(F)$-orbits on the set of irreducible components of $X_w(b)$ or about the geometric structure of the irreducible components of $X_w(b)$.

On the other hand, by [12], there is a family of elements in the Iwahori–Weyl group whose associated affine Deligne–Lusztig varieties have very simple geometric structures. We denote by σ the Frobenius morphism on $\boldsymbol{G}(\breve{F})$ and the induced group automorphism on the Iwahori–Weyl group $\tilde{W}$. Suppose that w is a minimal length element in its σ-conjugacy class of $\tilde{W}$; then there exists a unique σ-conjugacy class $[b]$ with $X_w(b) \neq \varnothing$. In this case, there exist a parahoric subgroup P of $\boldsymbol{J}_b(F)$ and a classical Deligne–Lusztig variety X (associated with the reductive quotient of P) such that $X_w(b) \cong \boldsymbol{J}_b(F) \times^P X$. Such a simple geometric structure has been used in the study of certain Shimura varieties with simple geometric structure (see [4; 7; 8]). However, these minimal length elements w form only a tiny fraction of the whole Iwahori–Weyl group, and such a simple geometric structure only occurs in a few cases.

We will focus on another family of elements in $\tilde{W}$. For any $w \in \tilde{W}$, we define its finite part to be the image of w under the map $\eta_\sigma : \tilde{W} \to W$ (see Section 2E). Our main result is that if the finite part of w is a σ-Coxeter element of W, then the associated affine Deligne–Lusztig variety $X_w(b)$ for any b has a simple geometric structure.

Theorem 1.1 (see Theorem 2.6). *Let $w \in \tilde{W}$ such that $\eta_\sigma(w)$ is a σ-Coxeter element of W. Then*

(1) *w is a cordial element*;

(2) *$X_w(b) \neq \varnothing$ if and only if b satisfies a certain Hodge–Newton indecomposability condition*;

(3) *for any b with $X_w(b) \neq \varnothing$, there exists a parahoric subgroup P of $\boldsymbol{J}_b(F)$ and a classical Deligne–Lusztig variety X of Coxeter type, and an iterated fibration $Y \to X$ whose iterated fibers are either $\mathbb{A}^1$ or $\mathbb{G}_m$ such that $X_w(b) \cong \boldsymbol{J}_b(F) \times^P Y$.*

We refer to Section 2 for the precise statement and definitions of the notions used here. The special case of part (3) where $\boldsymbol{G} = \mathrm{GL}_n$, b is basic and w is a certain element with finite Coxeter part was studied by Shimada [29].

1C. *Strategy.* One major tool used in the study of affine Deligne–Lusztig varieties is the Deligne–Lusztig reduction method [12]. Based on the Deligne–Lusztig reduction, a close relationship between affine Deligne–Lusztig varieties and the class polynomials of affine Hecke algebras was established in [12]. One remarkable property of these class polynomials is that they are polynomials in $(\boldsymbol{q} - 1)$ with nonnegative integral coefficients. With each element $w \in \tilde{W}$, we may associate a reduction tree, which encodes the information on the reduction steps and determines the class polynomials associated with w. However, obtaining an explicit description of the reduction trees is quite challenging.

Another key ingredient is the Chen–Zhu conjecture. This conjecture predicts the $\boldsymbol{J}_b(F)$-action on the top-dimensional irreducible components of affine Deligne–Lusztig varieties in the affine Grassmannian. This conjecture was verified recently in [20; 26; 33]. Part of the Chen–Zhu conjecture predicts the isotropy group for the $\boldsymbol{J}_b(F)$-action, which gives some information about the end points of the reduction trees.

Combining the above two ingredients, in Section 5, we show that the end points of each reduction tree for w with finite Coxeter part must be certain σ-Coxeter elements, and each path, which corresponds to a σ-conjugacy class $[b]$ of b, in a reduction tree provides a $\boldsymbol{J}_b(F)$-orbit of irreducible components of $X_w(b)$. It remains to show that for any b, there is at most one path in the reduction tree that corresponds to $[b]$ (i.e., the "multiplicity one" result). For the (unique) maximal σ-conjugacy class $[b]$ with $X_w(b) \neq \varnothing$, this "multiplicity one" result is obvious. For the basic σ-conjugacy class $[b]$, one may deduce the "multiplicity one" result by showing that any path corresponding to $[b]$ is of a unique type. See Section 5F.

It is more challenging to determine the numbers of reduction paths for other σ-conjugacy classes in a reduction tree. We use the following indirect approach to establish the "multiplicity one" result. We first interpret the class polynomials as the number of rational points for certain admissible subsets. We then use the positivity property of the class polynomials to show that the "multiplicity one" result for all b is equivalent to the single combinatorial identity

$$\sum_{[b] \in B(\boldsymbol{G}, \mu)_{\mathrm{indec}}} (\boldsymbol{q} - 1)^{?} \boldsymbol{q}^{-??} = 1. \tag{$*$}$$

Here $B(\boldsymbol{G}, \mu)_{\mathrm{indec}}$ is the set of all Hodge–Newton indecomposable σ-conjugacy classes (see Section 2C), and the powers "?" and "??" are certain nonnegative integers determined by w and b (see Section 6A).

Verifying the combinatorial identity $(*)$ is the most technical part of this paper and is done in Section 6. We first establish natural bijections between the sets $B(\boldsymbol{G}, \mu)_{\mathrm{indec}}$ for various pairs $(\boldsymbol{G}, \mu)$, which is of independent interest. In combination with other techniques, we reduce the verification of $(*)$ to the case for simply laced, $\breve{F}$-simple and split groups and for fundamental coweights.

For classical groups, we may further reduce to the case where μ is minuscule. In this case, the "multiplicity one" result follows from the Chen–Zhu conjecture. For exceptional groups, we use a computer to verify $(*)$. The most complicated case for the exceptional group is $(E_8, \omega_4^\vee)$. In this case, the left-hand side of $(*)$ involves a summation of 729 terms. It is also worth mentioning that in the case $(A_{n-1}, \omega_i^\vee)$, we may write $(*)$ explicitly as

$$\sum_{\substack{k\geqslant 1\\ 1>a_1/b_1>\cdots>a_k/b_k>0\\ a_i+\cdots+a_k=i\\ b_1+\cdots+b_k=n}} (\boldsymbol{q}-1)^{k-1}\boldsymbol{q}^{1-k+\left(\sum_{1\leqslant l_1<l_2\leqslant k}(a_{l_1}b_{l_2}-a_{l_2}b_{l_1})+\sum_{1\leqslant l\leqslant k}\gcd(a_l,b_l)\right)/2}=\boldsymbol{q}^{(i(n-i)-n)/2+1}.$$

We do not know if there is a purely combinatorial proof of this identity.

2. Preliminaries

2A. ***Reductive groups.*** Let F be a nonarchimedean local field with residue field $\mathbb{F}_q$ and let $\breve{F}$ be the completion of the maximal unramified extension of F. We write Γ for $\mathrm{Gal}(\bar{F}/F)$, and Γ_0 for the inertia subgroup of Γ.

Let $\boldsymbol{G}$ be a quasisplit connected reductive group over F. We set $\breve{G}=\boldsymbol{G}(\breve{F})$. Let σ be the Frobenius morphism of $\breve{F}$ over F. We use the same symbol σ for the induced Frobenius morphism on $\breve{G}$. Let S be a maximal $\breve{F}$-split torus of $\boldsymbol{G}$ defined over F, which contains a maximal F-split torus. Let T be the centralizer of S in $\boldsymbol{G}$. Then T is a maximal torus. We denote by N the normalizer of T in $\boldsymbol{G}$. Let $W=N(\breve{F})/T(\breve{F})$ be the *relative Weyl group*. We fix a σ-stable Iwahori subgroup $\breve{\mathcal{I}}$ of $\breve{G}$. Let $\tilde{W}=N(\breve{F})/T(\breve{F})\cap\breve{\mathcal{I}}$ be the *Iwahori–Weyl group*. The action σ on $\breve{G}$ induces a natural action on $\tilde{W}$ and W, which we still denote by σ. For any $w\in\tilde{W}$, we choose a representative $\dot{w}$ in $N(\breve{F})$. We have the splitting

$$\tilde{W}=X_*(T)_{\Gamma_0}\rtimes W=\{t^\lambda w\mid \lambda\in X_*(T)_{\Gamma_0},\ w\in W\}.$$

Here $X_*(T)_{\Gamma_0}$ denotes the Γ_0-coinvariants of $X_*(T)$.

Since $\boldsymbol{G}$ is quasisplit over F, σ acts naturally on $X_*(T)_{\Gamma_0}$ and on W. We denote by ℓ the length function on $\tilde{W}$ and on W, and by $\leqslant$ the Bruhat order on $\tilde{W}$ and on W. Let $\tilde{\mathbb{S}}$ be the index set of simple reflections in $\tilde{W}$ and let $\mathbb{S}\subset\tilde{\mathbb{S}}$ be the index set of simple reflections in W. In other words, the set of simple reflections in $\tilde{W}$ is $\{s_i\mid i\in\tilde{\mathbb{S}}\}$.

For any $w\in W$, we denote by $\mathrm{supp}(w)$ the set of i such that s_i occurs in some (or, equivalently, any) reduced expressions of w, and we set $\mathrm{supp}_\sigma(w)=\bigcup_{l\in\mathbb{N}}\sigma^l(\mathrm{supp}(w))$.

An element $c\in W$ is called a *(full) σ-Coxeter element* if it is a product of simple reflections, one from each σ-orbit of $\mathbb{S}$. An element $c\in W$ is called a *partial σ-Coxeter element* if it is a product of simple reflections, at most one from each σ-orbit of $\mathbb{S}$.

Let Φ be the reduced root system underlying the relative root system of $\boldsymbol{G}$ over $\breve{F}$ (the *échelonnage root system*). For any $i \in \mathbb{S}$, we denote by α_i and $\alpha_i^\vee$ the corresponding (relative) simple root and simple coroot, respectively.

2B. ***The σ-conjugacy classes of $\breve{G}$.*** The σ-conjugation action on $\breve{G}$ is defined by $g \cdot_\sigma g' = gg'\sigma(g)^{-1}$. For $b \in \breve{G}$, we denote by $[b]$ the σ-conjugacy class of b. Let $B(\boldsymbol{G})$ be the set of σ-conjugacy classes on $\breve{G}$. The classification of the σ-conjugacy classes is due to Kottwitz in [21; 22]. Any σ-conjugacy class $[b]$ is determined by two invariants:

- the element $\kappa([b]) \in \pi_1(\boldsymbol{G})_\sigma$;
- the Newton point $\nu_b \in ((X_*(T)_{\Gamma_0,\mathbb{Q}})^+)^{\langle\sigma\rangle}$.

Here $\pi_1(\boldsymbol{G})_\sigma$ denotes the σ-coinvariants of $\pi_1(\boldsymbol{G})$, and $(X_*(T)_{\Gamma_0,\mathbb{Q}})^+$ denotes the intersection of $X_*(T)_{\Gamma_0} \otimes \mathbb{Q} = X_*(T)^{\Gamma_0} \otimes \mathbb{Q}$ with the set $X_*(T)_{\mathbb{Q}}^+$ of dominant elements in $X_*(T)_{\mathbb{Q}}$. Define

$$V = X_*(T)_{\Gamma_0} \otimes \mathbb{R}.$$

For any $v \in V$, define

$$I(v) = \{i \in \mathbb{S} \mid \langle v, \alpha_i \rangle = 0\}.$$

Here $\langle\cdot,\cdot\rangle : V \times \mathbb{R}\Phi \to \mathbb{R}$ is the natural pairing. Let V^+ be the set of dominant vectors $v \in V$, that is, $\langle v, \alpha_i \rangle \geqslant 0$ for $i \in \mathbb{S}$.

The set $B(\boldsymbol{G})$ is equipped with a natural partial order: $[b] \leqslant [b']$ if and only if $\kappa([b]) = \kappa([b'])$ and $\nu_b \leqslant \nu_{b'}$. Here $\leqslant$ is the dominance order on the set of dominant elements in $X_*(T)_{\mathbb{Q}}$, that is, $\nu \leqslant \nu'$ if $\nu' - \nu$ is a nonnegative rational linear combination of positive relative coroots. It is proved in [2, Theorem 7.4] that the poset $B(\boldsymbol{G})$ is ranked. For any $[b] \leqslant [b']$ in $B(\boldsymbol{G})$, we denote by $\operatorname{length}([b], [b'])$ the length of any maximal chain between $[b]$ and $[b']$.

Let μ be a dominant coweight. Let $\mu^\diamond$ be the average of the σ-orbit of μ. The set of *neutrally acceptable* σ-conjugacy classes is defined by

$$B(\boldsymbol{G}, \mu) = \{[b] \in B(\boldsymbol{G}) \mid \kappa([b]) = \kappa(\mu),\ \nu_b \leqslant \mu^\diamond\}.$$

For any $i \in \mathbb{S}$, let $\omega_i \in \mathbb{R}\Phi$ be the corresponding fundamental weight. For any σ-orbit $\mathcal{O}$ of $\mathbb{S}$, let $\omega_{\mathcal{O}} = \sum_{i \in \mathcal{O}} \omega_i$. The following length formula is due to Chai (see [2, Theorem 7.4; 31, Theorem 3.4]):

(a) For $[b] \in B(\boldsymbol{G}, \mu)$, $\operatorname{length}([b], [t^\mu]) = \sum_{\mathcal{O} \in \mathbb{S}/\langle\sigma\rangle} \lceil \langle \mu - \nu_b, \omega_{\mathcal{O}} \rangle \rceil$.

2C. ***Hodge–Newton indecomposable/irreducible set.*** For any σ-stable subset J of $\mathbb{S}$, we denote by $\boldsymbol{M}_J$ the standard Levi subgroup of $\boldsymbol{G}_{\breve{F}}$ associated with J. Let $W_J \subseteq W$ be the parabolic subgroup generated by the simple reflections in J. Then W_J is the Weyl group of $\boldsymbol{M}_J$. Let $b \in \breve{G}$. We say that (μ, b) is *Hodge–Newton decomposable* with respect to $\boldsymbol{M}_J$ if $I(\nu_b) \subseteq J$ and $\mu^\diamond - \nu_b \in \sum_{j \in J} \mathbb{R}_{\geqslant 0} \alpha_j^\vee$. If (μ, b) is not Hodge–Newton decomposable with respect to any proper σ-stable standard Levi subgroup of $\boldsymbol{G}_{\breve{F}}$, then we say that $[b]$ is *Hodge–Newton indecomposable*. Set

$$B(\boldsymbol{G}, \mu)_{\text{indec}} = \{[b] \in B(\boldsymbol{G}, \mu) \mid [b] \text{ is Hodge–Newton indecomposable}\}.$$

We say that (μ, b) is *Hodge–Newton J-irreducible* if $\mu^\diamond - \nu_b \in \sum_{j\in J} \mathbb{R}_{>0}\alpha_j^\vee$. Set

$$B(\boldsymbol{G}, \mu)_{J\text{-irr}} = \{[b] \in B(\boldsymbol{G}, \mu) \mid [b] \text{ is Hodge–Newton } J\text{-irreducible}\}.$$

We simply write $B(\boldsymbol{G}, \mu)_{\text{irr}} = B(\boldsymbol{G}, \mu)_{\mathbb{S}\text{-irr}}$.

We say that μ is *essentially noncentral* with respect to $\boldsymbol{M}_J$ if it is noncentral on every σ-orbit of connected components of J. It is easy to see that $B(\boldsymbol{G}, \mu)_{J\text{-irr}} \neq \varnothing$ if and only if μ is essentially noncentral with respect to $\boldsymbol{M}_J$. We may simply say that μ is essentially noncentral if it is essentially noncentral with respect to $\boldsymbol{G}$. If μ is essentially noncentral, then $B(\boldsymbol{G}, \mu)_{\text{irr}} = B(\boldsymbol{G}, \mu)_{\text{indec}}$.

Let $\mathfrak{M}_\mu$ be the set of σ-stable subsets $J \subseteq \mathbb{S}$ such that μ is essentially noncentral in J. Note that if $B(\boldsymbol{G}, \mu)_{J\text{-irr}} \neq \varnothing$, then $J \in \mathfrak{M}_\mu$. By definition, any $[b] \in B(\boldsymbol{G}, \mu)$ lies in some $B(\boldsymbol{G}, \mu)_{J-\text{irr}}$. Then we have

$$B(\boldsymbol{G}, \mu) = \bigsqcup_{J\in\mathfrak{M}_\mu} B(\boldsymbol{G}, \mu)_{J\text{-irr}}.$$

Let $J = \sigma(J) \subseteq \mathbb{S}$. For $b \in \boldsymbol{M}_J(\breve{F})$, we denote by $[b]_{\boldsymbol{M}_J}$ the σ-conjugacy class of b in $\boldsymbol{M}_J(\breve{F})$, and denote by $\nu_b^{\boldsymbol{M}_J}$ its $\boldsymbol{M}_J$-dominant Newton point.

Lemma 2.1. *Let μ be a dominant coweight and J be a σ-stable subset of $\mathbb{S}$. Then*

(1) *the map $\phi_J : B(\boldsymbol{M}_J, \mu) \to B(\boldsymbol{G}, \mu)$, $[b]_{\boldsymbol{M}_J} \mapsto [b]$, is injective*;

(2) *the image of ϕ_J consists of $[b] \in B(\boldsymbol{G}, \mu)$ with $\mu^\diamond - \nu_b \in \sum_{i\in J} \mathbb{R}_{\geqslant 0}\alpha_i^\vee$*;

(3) *for $[b]_{\boldsymbol{M}_J} \in B(\boldsymbol{M}_J, \mu)$, $\text{length}_{\boldsymbol{G}}([b], [t^\mu]) = \text{length}_{\boldsymbol{M}_J}([b]_{\boldsymbol{M}_J}, [t^\mu]_{\boldsymbol{M}_J})$.*

Proof. Let $[b]_{\boldsymbol{M}_J} \in B(\boldsymbol{M}_J, \mu)$. Then $\mu^\diamond - \nu_b^{\boldsymbol{M}_J} \in \sum_{i\in J} \mathbb{R}_{\geqslant 0}\alpha_i^\vee$, which implies that $\nu_b^{\boldsymbol{M}_J}$ is dominant with respect to $\boldsymbol{G}$, and hence $\nu_b^{\boldsymbol{M}_J} = \nu_b$. Now the Newton point and the Kottwitz point of $[b]_{\boldsymbol{M}_J}$ are determined by $[b]$ and μ, respectively. Hence ϕ_J is injective.

Part (2) follows from [2, §7.1; 19, Lemma 3.5]. Part (3) follows from part (2) and Chai's length formula Section 2B(a). □

As a consequence, we have the following.

Corollary 2.2. *Let $J \in \mathfrak{M}_\mu$. Then the map ϕ_J in Lemma 2.1 induces a bijection $B(\boldsymbol{M}_J, \mu)_{\text{irr}} \cong B(\boldsymbol{G}, \mu)_{J\text{-irr}}$.*

2D. *Affine Deligne–Lusztig varieties.* Let $\text{Fl} = \breve{G}/\breve{\mathcal{I}}$ be an *affine flag variety*. For any $b \in \breve{G}$ and $w \in \tilde{W}$, we define the corresponding *affine Deligne–Lusztig variety* in the affine flag variety

$$X_w(b) = \{g\breve{\mathcal{I}} \in \breve{G}/\breve{\mathcal{I}} \mid g^{-1}b\sigma(g) \in \breve{\mathcal{I}}\dot{w}\breve{\mathcal{I}}\} \subset \text{Fl}\,.$$

Let $\boldsymbol{k}$ be the residue field of $\breve{F}$. In the equal characteristic setting, the affine Deligne–Lusztig variety $X_w(b)$ is the set of $\boldsymbol{k}$-valued points of a locally closed subscheme of the affine flag variety, equipped with the reduced scheme structure. In the mixed characteristic setting, we consider $X_w(b)$ as the $\boldsymbol{k}$-valued points of a perfect scheme in the sense of Zhu [34] and Bhatt and Scholze [1], a locally closed perfect subscheme of the p-adic partial flag variety.

We denote by $\Sigma^{\text{top}}(X_w(b))$ the set of top-dimensional irreducible components of $X_w(b)$. Let $\boldsymbol{J}_b$ be the σ-centralizer of b and let $\boldsymbol{J}_b(F) = \{g \in \breve{G} \mid gb\sigma(g)^{-1} = b\}$ be the group of F-points of $\boldsymbol{J}_b$. The left action of $\boldsymbol{J}_b(F)$ on $X_w(b)$ induces an action of $\boldsymbol{J}_b(F)$ on $\Sigma^{\text{top}}(X_w(b))$.

We denote by $\boldsymbol{J}_b(F)\backslash\Sigma^{\text{top}}(X_w(b))$ the set of $\boldsymbol{J}_b(F)$-orbits on $\Sigma^{\text{top}}(X_w(b))$.

If b and b' are σ-conjugate in $\breve{G}$, then $X_w(b)$ and $X_w(b')$ are isomorphic. Thus the affine Deligne–Lusztig variety $X_w(b)$ (up to isomorphism) depends only on the element w in the Iwahori–Weyl group $\tilde{W}$ and the σ-conjugacy class $[b]$ in $B(\boldsymbol{G})$. We set

$$B(\boldsymbol{G})_w = \{[b] \in B(\boldsymbol{G}) \mid X_w(b) \neq \varnothing\}.$$

There is a unique maximal σ-conjugacy class in $B(\boldsymbol{G})_w$, which we denote by $[b_w]$. By [25, Lemma 3.2], $\dim X_w(b_w) = \ell(w) - \langle \nu_{b_w}, 2\rho\rangle$. Here ρ is the half sum of the positive roots in Φ.

2E. *Cordial elements.* For $b \in \breve{G}$, the *defect* of b is defined to be

$$\operatorname{def}(b) = \operatorname{rank}_F \boldsymbol{G} - \operatorname{rank}_F \boldsymbol{J}_b,$$

where rank_F denotes the F-rank of a reductive group over F.

By [31, Theorem 3.4], we have the following length formula:

(a) For $[b] \in B(\boldsymbol{G}, \mu)$, $\operatorname{length}([b], [t^\mu]) = \langle \mu - \nu_b, \rho\rangle + \frac{1}{2}\operatorname{def}(b)$.

Let ${}^{\mathbb{S}}\tilde{W}$ be the set of minimal length representatives for the cosets in $W\backslash\tilde{W}$. Any element $w \in \tilde{W}$ can be written in a unique way as $w = xt^\mu y$ with μ dominant, $x, y \in W$ such that $t^\mu y \in {}^{\mathbb{S}}\tilde{W}$. We have $\ell(w) = \ell(x) + \langle \mu, 2\rho\rangle - \ell(y)$. In this case, we set $\eta_\sigma(w) = \sigma^{-1}(y)x$. The *virtual dimension* is defined to be

$$d_w(b) = \tfrac{1}{2}(\ell(w) + \ell(\eta_\sigma(w)) - \operatorname{def}(b) - \langle \nu_b, 2\rho\rangle).$$

For $w = t^\mu y \in {}^{\mathbb{S}}\tilde{W}$, it is easy to see that $d_w(b) = \langle \mu - \nu_b, \rho\rangle - \frac{1}{2}\operatorname{def}(b)$.

By [12, Theorem 10.3; 13, Theorem 2.30], we have

(b) for $w \in \tilde{W}$ and $b \in \breve{G}$, $\dim X_w(b) \leq d_w(b)$.

In the special case $[b] = [b_w]$, (b) implies that

$$\ell(w) - \ell(\eta_\sigma(w)) \leq \langle \nu_{b_w}, 2\rho\rangle - \operatorname{def}(b_w).$$

Cordial elements were introduced by Milićević and Viehmann in [25]. By definition, an element $w \in \tilde{W}$ is *cordial* if $\dim X_w(b_w) = d_w(b_w)$. This condition is equivalent to the condition that $\ell(w) - \ell(\eta_\sigma(w)) = \langle \nu_{b_w}, 2\rho\rangle - \operatorname{def}(b_w)$. The following nice properties of the cordial elements are established in [25].

Theorem 2.3. *Let $w \in \tilde{W}$ be a cordial element. Then*

(1) *$B(\boldsymbol{G})_w$ is saturated, that is, if $[b_1] \leqslant [b_2] \leqslant [b_3]$ in $B(\boldsymbol{G})$ and $[b_1], [b_3] \in B(\boldsymbol{G})_w$, then $[b_2] \in B(\boldsymbol{G})_w$;*

(2) *for each $[b] \in B(\boldsymbol{G})_w$, $X_w(b)$ is equidimensional of dimension equal to $d_w(b)$.*

2F. ***Minimal length elements.*** We consider the σ-conjugation action on $\tilde{W}$ defined by $w \cdot_\sigma w' = ww'\sigma(w)^{-1}$. Let $B(\tilde{W}, \sigma)$ be the set of σ-conjugacy classes of $\tilde{W}$. For any σ-conjugacy class $\mathcal{O}$ of $\tilde{W}$, we let $\mathcal{O}_{\min}$ be the set of minimal length elements in $\mathcal{O}$, and we write $\ell(\mathcal{O}) = \ell(w)$ for any $w \in \mathcal{O}_{\min}$.

For $w, w' \in \tilde{W}$ and $i \in \tilde{\mathbb{S}}$, we write $w \xrightarrow{s_i}_\sigma w'$ if $w' = s_i w \sigma(s_i)$ and $\ell(w') \leqslant \ell(w)$. We write $w \to_\sigma w'$ if there is a sequence $w = w_0, w_1, \dots, w_n = w'$ of elements in $\tilde{W}$ such that for any k, $w_{k-1} \xrightarrow{s_i}_\sigma w_k$ for some $s_i \in \tilde{\mathbb{S}}$. We write $w \approx_\sigma w'$ if $w \to_\sigma w'$ and $w' \to_\sigma w$.

We call $w, w' \in \tilde{W}$ *elementarily strongly σ-conjugate* if $\ell(w) = \ell(w')$ and there exists $x \in \tilde{W}$ such that $w' = xw\sigma(x)^{-1}$ and $\ell(xw) = \ell(x) + \ell(w)$ or $\ell(w\sigma(x)^{-1}) = \ell(x) + \ell(w)$. We call w, w' *strongly σ-conjugate* if there is a sequence $w = w_0, w_1, \dots, w_n = w'$ such that for each i, w_{i-1} is elementarily strongly σ-conjugate to w_i. We write $w \sim_\sigma w'$ if w and w' are strongly σ-conjugate.

The following result is proved in [17, Theorem 2.10].

Theorem 2.4. *Let $\mathcal{O}$ be a σ-conjugacy class in $\tilde{W}$. Then the following hold*:

(1) *For each element $w \in \mathcal{O}$, there exists $w' \in \mathcal{O}_{\min}$ such that $w \to_\sigma w'$.*

(2) *Let $w, w' \in \mathcal{O}_{\min}$. Then $w \sim_\sigma w'$.*

2G. ***Decompositions of affine Deligne–Lusztig varieties.*** We recall the Deligne–Lusztig reduction method on affine Deligne–Lusztig varieties.

Proposition 2.5 [20, Proposition 3.3.1]. *Let $w \in \tilde{W}$, $i \in \tilde{\mathbb{S}}$, and $b \in \breve{G}$. If* $\mathrm{char}(F) > 0$, *then the following two statements hold*:

(1) *If $\ell(s_i w \sigma(s_i)) = \ell(w)$, then there exists a $\boldsymbol{J}_b(F)$-equivariant morphism $X_w(b) \to X_{s_i w \sigma(s_i)}(b)$ which is a universal homeomorphism.*

(2) *If $\ell(s_i w \sigma(s_i)) = \ell(w) - 2$, then $X_w(b) = X_1 \sqcup X_2$, where X_1 is a $\boldsymbol{J}_b(F)$-stable open subscheme X of $X_w(b)$ and X_2 is its closed complement satisfying the following conditions*:

- *X_1 is $\boldsymbol{J}_b(F)$-equivariant universally homeomorphic to a Zariski-locally trivial $\mathbb{G}_m$-bundle over $X_{s_i w}(b)$.*
- *X_2 is $\boldsymbol{J}_b(F)$-equivariant universally homeomorphic to a Zariski-locally trivial $\mathbb{A}^1$-bundle over $X_{s_i w \sigma(s_i)}(b)$.*

If $\mathrm{char}(F) = 0$, *then the above two statements still hold, but with $\mathbb{A}^1$ and $\mathbb{G}_m$ replaced by the perfections $\mathbb{A}^{1,\mathrm{perf}}$ and $\mathbb{G}_m^{\mathrm{perf}}$, respectively.*

For any $a_1, a_2 \in \mathbb{N}$, we say that a scheme X is an *iterated fibration* of type (a_1, a_2) over a scheme Y if there exist morphisms

$$X = Y_0 \to Y_1 \to \cdots \to Y_{a_1+a_2} = Y$$

such that for any i with $0 \leqslant i < a_1 + a_2$, Y_i is a Zariski-locally trivial $\mathbb{A}^{1,(\mathrm{perf})}$-bundle or $\mathbb{G}_m^{(\mathrm{perf})}$-bundle over Y_{i+1}, and there are exactly a_1 locally trivial $\mathbb{G}_m^{(\mathrm{perf})}$-bundles in the sequence. In this case, there are exactly a_2 locally trivial $\mathbb{A}^{1,(\mathrm{perf})}$-bundles in the sequence.

2H. ***Classical Deligne–Lusztig varieties.*** Let $\mathbb{F}_q$ be a finite field and let $\bar{\mathbb{F}}_q$ an algebraic closure of $\mathbb{F}_q$. Let $\boldsymbol{H}$ be a connected reductive group over $\mathbb{F}_q$ and $H = \boldsymbol{H}(\bar{\mathbb{F}}_q)$. Let $\sigma_{\boldsymbol{H}}$ be the Frobenius morphism on H. Let B be a $\sigma_{\boldsymbol{H}}$-stable Borel subgroup of H. Let W_H be the Weyl group of H and let $\mathbb{S}_H$ be the set of simple reflections. Then $\sigma_{\boldsymbol{H}}$ induces a group automorphism on W_H preserving $\mathbb{S}_H$. (Classical) Deligne–Lusztig varieties were introduced in [3]. They are defined as follows. For $x \in W_H$, we set

$$X_x^{\boldsymbol{H}} = \{hB \in H/B \mid h^{-1}\sigma(h) \in B\dot{x}B\}.$$

If x is a $\sigma_{\boldsymbol{H}}$-Coxeter element of W_H, then we say that $X_x^{\boldsymbol{H}}$ is a *classical Deligne–Lusztig variety of Coxeter type* for $\boldsymbol{H}$. It is well known that classical Deligne–Lusztig varieties of Coxeter type are irreducible.

It is proved in [12, Theorem 4.8] that if $w \in \tilde{W}$ is a minimal length element in its σ-conjugacy class, then $X_w(b) \neq \varnothing$ if and only if b and $\dot{w}$ are in the same σ-conjugacy class of $\breve{G}$. In this case,

$$X_w(b) \cong \boldsymbol{J}_b(F) \times^P X. \tag{2-1}$$

Here $\cong$ means a $\boldsymbol{J}_b(F)$-equivariant universal homeomorphism, P is a parahoric subgroup of $\boldsymbol{J}_b(F)$, and X is (the perfection of) a classical Deligne–Lusztig variety for some connected reductive group $\boldsymbol{H}$ over $\mathbb{F}_q$ with $\boldsymbol{H}(\mathbb{F}_q)$ isomorphic to the reductive quotient of P. The group P acts on $\boldsymbol{J}_b(F) \times X$ by $p \cdot (g, z) = (gp^{-1}, p \cdot z)$, and $\boldsymbol{J}_b(F) \times^P X$ is the quotient space.

2I. ***Very special parahoric subgroups.*** We follow [20, §2] for the definition of very special parahoric subgroups.

Let $\boldsymbol{G}_1$ be a (not necessarily quasisplit) connected reductive group over F and let σ_1 be its Frobenius morphism. Let $\tilde{W}_1$ be the Iwahori–Weyl group $\boldsymbol{G}_1$. We still denote by σ_1 the action on $\tilde{W}_1$ induced from the Frobenius morphism on $\boldsymbol{G}_1$. Let $\tilde{\mathbb{S}}_1$ be the set of simple reflections in $\tilde{W}_1$.

A parahoric subgroup P of $\boldsymbol{G}_1(F)$ is called *very special* if it is of maximal volume among all the parahoric subgroups of $\boldsymbol{G}_1(F)$. A σ_1-stable subset $\breve{K} \subseteq \tilde{\mathbb{S}}_1$ is called *very special* with respect to σ_1 if the parabolic subgroup $W_{\breve{K}}$ generated by $\breve{K}$ is finite and the longest element of $W_{\breve{K}}$ is of maximal length among all such σ_1-stable subsets of $\tilde{\mathbb{S}}_1$. By [20, Proposition 2.2.5],

(a) a σ_1-stable subset $\breve{K}$ of $\tilde{\mathbb{S}}_1$ is very special if and only if $\breve{P}_{\breve{K}} \cap \boldsymbol{G}_1(F)$ is a very special parahoric subgroup of $\boldsymbol{G}_1(F)$, where $\breve{P}_{\breve{K}}$ is the parahoric subgroup of $\boldsymbol{G}_1(\breve{F})$ corresponding to $\breve{K}$.

2J. ***Statement of the main result.*** Let $w \in \tilde{W}$. We say that w *has finite σ-Coxeter part* if $\eta_\sigma(w)$ is a partial σ-Coxeter element. In this case, we set $J(w) = \mathrm{supp}_\sigma(\eta_\sigma(w))$. For any σ-stable subset J of $J(w)$, denote by J'_μ (resp. J''_μ) the union of all σ-orbits of connected components of J in which μ is noncentral (resp. central). Then μ is essentially noncentral in J'_μ. By Section 2C, $B(\boldsymbol{G}, \mu)_{J'_\mu\text{-irr}} \neq \varnothing$, and we have a natural bijection $B(\boldsymbol{M}_{J'_\mu}, \mu)_{\mathrm{irr}} \cong B(\boldsymbol{G}, \mu)_{J'_\mu\text{-irr}}$.

Let $J_0(w)$ be the subset of $\mathbb{S}$ with $\mu - \nu_{b_w} \in \sum_{i \in J_0(w)} \mathbb{R}_{>0}\alpha_i^\vee$. Then $J_0(w) = J_0(w)'_\mu$. By definition, $[\dot{w}] \leqslant [b_w]$. Thus $\nu_{\dot{w}} \leqslant \nu_{b_w}$ and $J_0(w) \subseteq J(w)$.

For any σ-stable $J_1, J_2 \subseteq \mathbb{S}$, denote by $[J_1, J_2]$ the set of σ-stable subsets $J \subseteq \mathbb{S}$ such that $J_1 \subseteq J \subseteq J_2$ and denote by $[J_1, J_2]_\mu$ the set of $J \in [J_1, J_2]$ such that μ is essentially noncentral in J (or equivalently, $J = J'_\mu$).

Now we state our main result.

Theorem 2.6. *Let $w \in W t^\mu W$ such that $\eta_\sigma(w)$ is a partial σ-Coxeter element. Then*

(1) *w is a cordial element*;

(2) $B(\boldsymbol{G})_w = \bigsqcup_{J \in [J_0(w), J(w)]_\mu} B(\boldsymbol{G}, \mu)_{J\text{-irr}}$;

(3) *for any $[b] \in B(\boldsymbol{G})_w$, we have*

$$X_w(b) \cong \boldsymbol{J}_b(F) \times^P Y,$$

where P is a parahoric subgroup of $\boldsymbol{J}_b(F)$, and Y is an iterated fibration, whose iterated fibers are either $\mathbb{A}^1$ or $\mathbb{G}_m$, over (the perfection of) a classical Deligne–Lusztig variety of Coxeter type for some connected reductive group $\boldsymbol{H}$ over $\mathbb{F}_q$ with $\boldsymbol{H}(\mathbb{F}_q)$ isomorphic to the reductive quotient of P.

Remark 2.7. (1) In particular, $\boldsymbol{J}_b(F)$ acts transitively on the set of irreducible components of $X_w(b)$.

(2) Combining Remark 3.10 with Theorem 7.1, we have a detailed description of the classical Deligne–Lusztig variety in Theorem 2.6(3). If $w \in {}^{\mathbb{S}}\tilde{W}$ and $\eta_\sigma(w)$ is a (full) σ-Coxeter element, then the parahoric subgroup P in Theorem 2.6(3) is very special; see Section 5E.

(3) In Theorem 7.1, we explicitly compute the numbers of $\mathbb{G}_m$-bundles and $\mathbb{A}^1$-bundles appearing in the iterated fibration.

Parts (1) and (2) will be proved in Section 4. Part (3) is the most difficult part of this paper and will be proved in Sections 6 and 7. The proof is based on a deep analysis of the reduction tree of w, which will be introduced in Section 3.

3. Class polynomials and reduction trees

We recall the class polynomials of Hecke algebras and the connection with affine Deligne–Lusztig varieties discovered in [12]. We then introduce the reduction tree, which encodes more information than the class polynomials.

3A. ***Hecke algebras and their cocenters.*** Let $\boldsymbol{q}$ be an indeterminate. Let H be the Hecke algebra associated with $\tilde{W}$, that is, it is the $\mathbb{Z}[\boldsymbol{q}^{\pm 1}]$-algebra generated by T_w for $w \in \tilde{W}$ subject to the relations

- $T_w T_{w'} = T_{ww'}$ if $\ell(ww') = \ell(w) + \ell(w')$;
- $(T_{s_i} + 1)(T_{s_i} - \boldsymbol{q}) = 0$ for $i \in \tilde{\mathbb{S}}$.

The action of σ on $\tilde{W}$ induces an action on H, which we still denote by σ. The *σ-commutator* $[H, H]_\sigma$ is the $\mathbb{Z}[\boldsymbol{q}^{\pm 1}]$-submodule generated by $hh' - h'\sigma(h)$ for $h, h' \in H$. The *σ-cocenter* of H is defined to be

$$\bar{H}_\sigma = H/[H, H]_\sigma.$$

By Theorem 2.4(2), for any σ-conjugacy class $\mathcal{O}$ of $\tilde{W}$ and $w, w' \in \mathcal{O}_{\min}$, we have $T_w + [H, H]_\sigma = T_{w'} + [H, H]_\sigma$. We write $T_{\mathcal{O}} = T_w + [H, H]_\sigma \in \bar{H}_\sigma$ for any $w \in \mathcal{O}_{\min}$.

Theorem 3.1 [17, Theorem 6.7]. *$\bar{H}_\sigma$ is a free $\mathbb{Z}[\boldsymbol{q}^{\pm 1}]$-module with basis $\{T_{\mathcal{O}}\}_{\mathcal{O} \in B(\tilde{W}, \sigma)}$.*

By [12, §2.3; 13, §2.8.2], for any $w \in \tilde{W}$ and $\mathcal{O} \in B(\tilde{W}, \sigma)$, there exists a unique polynomial $F_{w,\mathcal{O}} \in \mathbb{N}[q-1]$ such that

$$T_w + [H, H]_\sigma = \sum_{\mathcal{O} \in B(\tilde{W}, \sigma)} F_{w,\mathcal{O}} T_{\mathcal{O}} \in \bar{H}_\sigma.$$

The polynomials $F_{w,\mathcal{O}}$ are called *class polynomials*.[1]

3B. *Class polynomials and affine Deligne–Lusztig varieties.* The class polynomials encode a lot of information about affine Deligne–Lusztig varieties.

Let $\tilde{W} \to B(\boldsymbol{G})$ be the map sending $w \in \tilde{W}$ to the σ-conjugacy class $[\dot{w}]$ of $\breve{G}$. It is known that this map is independent of the choice of the representative $\dot{w}$ of w, and it induces a map

$$\Psi : B(\tilde{W}, \sigma) \to B(\boldsymbol{G}).$$

By [12, Theorem 3.7], the map Ψ is surjective.

Let $w \in \tilde{W}$ and $[b] \in B(\boldsymbol{G})$. We set

$$F_{w,[b]} = \sum_{\substack{\mathcal{O} \in B(\tilde{W}, \sigma) \\ \Psi(\mathcal{O}) = [b]}} \boldsymbol{q}^{\ell(\mathcal{O})} F_{w,\mathcal{O}} \in \mathbb{N}[\boldsymbol{q}-1].$$

Here $\ell(\mathcal{O}) = \ell(x)$ for any $x \in \mathcal{O}_{\min}$.

The following "dimension = degree" theorem is established in [12, Theorem 6.1].

Theorem 3.2. *Let $w \in \tilde{W}$ and $b \in \breve{G}$. Then* $\dim X_w(b) = \deg F_{w,[b]} - \langle \nu_b, 2\rho \rangle$.

Remark 3.3. Here, by convention, $\dim \varnothing = \deg 0 = -\infty$.

We have the following "leading coefficients = irreducible components" theorem. This is established in [13, Theorem 2.19]. See also [20, Theorem 3.3.9 and Corollary 3.3.11].

Theorem 3.4. *For $w \in \tilde{W}$ and $b \in \breve{G}$, the cardinality of $\boldsymbol{J}_b(F) \backslash \Sigma^{\mathrm{top}}(X_w(b))$ equals the leading coefficient of $F_{w,[b]}$.*

Although not needed in this paper, it is also worth mentioning that in the superbasic case, the class polynomial gives the number of rational points of an affine Deligne–Lusztig variety. This is established in [12, Proposition 8.3].

[1] The polynomials we use here coincide with those in [13] and differ from the polynomials used in [12] by a certain monomial. See [13, footnote on p. 106].

Proposition 3.5. *Suppose that the residue field of F is $\mathbb{F}_q$. Assume that $\boldsymbol{G} = \mathrm{PGL}_n$ split over F and $b \in \boldsymbol{G}(F)$ is a superbasic element in $\breve{G}$. Then*

$$\sharp X_w(b)^\sigma = n F_{w,[b]}|_{\boldsymbol{q}=q}.$$

3C. ***An identity on the class polynomials.*** We have the following identity on the class polynomials. We first give a proof using representations of Hecke algebras. Then we provide a geometric interpretation of this identity.

Proposition 3.6. *Let $w \in \tilde{W}$. Then*

$$\boldsymbol{q}^{\ell(w)} = \sum_{\mathcal{O} \in B(\tilde{W},\sigma)} \boldsymbol{q}^{\ell(\mathcal{O})} F_{w,\mathcal{O}} = \sum_{[b] \in B(\boldsymbol{G})} F_{w,[b]}.$$

Proof. We prove the first equality. The second follows from the definition.

Let $\pi : H \to \mathbb{Z}[\boldsymbol{q}^{\pm 1}]$ be the homomorphism of $\mathbb{Z}[\boldsymbol{q}^{\pm 1}]$-algebras sending T_{s_i} to $\boldsymbol{q}$ for any $i \in \tilde{\mathbb{S}}$. As $\pi \circ \sigma = \pi$, $\pi([H, H]_\sigma) = 0$ and thus π induces a homomorphism of the $\mathbb{Z}[\boldsymbol{q}^{\pm 1}]$-modules $\bar{H}_\sigma \to \mathbb{Z}[\boldsymbol{q}^{\pm 1}]$, which we still denote by π.

We have $T_w + [H, H]_\sigma = \sum_{\mathcal{O} \in B(\tilde{W},\sigma)} F_{w,\mathcal{O}} T_{\mathcal{O}}$. Applying π to both sides, we obtain $\boldsymbol{q}^{\ell(w)} = \sum_{\mathcal{O} \in B(\tilde{W},\sigma)} \boldsymbol{q}^{\ell(\mathcal{O})} F_{w,\mathcal{O}}$. □

In the rest of this subsection, we assume that $F = \mathbb{F}_q((\epsilon))$ and that $\boldsymbol{G}$ is split over F. We give a geometric interpretation of the above identity.

For any $n \in \mathbb{N}$, let $\breve{I}_n$ be the n-th congruence subgroup of $\breve{I}$. Following [5, §2.10], we call a subset X of $\breve{G}$ *admissible* if for any $w \in \tilde{W}$, there exists $n \in \mathbb{N}$ such that $X \cap \breve{I}\dot{w}\breve{I}$ is stable under right multiplication of $\breve{I}_n$. In this case, the action of $\breve{I}_n$ on $X \cap \breve{I}\dot{w}\breve{I}$ is free, and $\frac{\sharp((X \cap \breve{I}\dot{w}\breve{I})/\breve{I}_n)^\sigma}{\sharp(\breve{I}/\breve{I}_n)^\sigma}$ is independent of the choice of such n. We set

$$\sharp_{\breve{I}}(X \cap \breve{I}\dot{w}\breve{I}) = \frac{\sharp((X \cap \breve{I}\dot{w}\breve{I})/\breve{I}_n)^\sigma}{\sharp(\breve{I}/\breve{I}_n)^\sigma}.$$

An admissible subset X is called *bounded* if $X \cap \breve{I}\dot{w}\breve{I} = \varnothing$ for all but finitely many $w \in \tilde{W}$. For any bounded admissible subset X, we set

$$\sharp_{\breve{I}} X = \sum_{w \in \tilde{W}} \sharp_{\breve{I}}(X \cap \breve{I}\dot{w}\breve{I})$$

and call it the *normalized cardinality of the rational points* of X. By [14, Theorem A.1], each σ-conjugacy class of $\breve{G}$ is admissible. We have the following geometric interpretation of the class polynomials.

Proposition 3.7. *Let $w \in \tilde{W}$ and $[b] \in B(\boldsymbol{G})$. Then*

$$\sharp_{\breve{I}}([b] \cap \breve{I}\dot{w}\breve{I}) = F_{w,[b]}|_{\boldsymbol{q}=q}.$$

Proof. We argue by induction on $\ell(w)$.

If w is a minimal length element in its σ-conjugacy class in $\tilde{W}$, then

$$[b] \cap \breve{I}\dot{w}\breve{I} = \begin{cases} \breve{I}\dot{w}\breve{I} & \text{if } [b] = [\dot{w}], \\ \varnothing & \text{otherwise.} \end{cases}$$

On the other hand, by [13, §2.8.2],

$$F_{w,[b]} = \begin{cases} \boldsymbol{q}^{\ell(w)} & \text{if } [b] = [\dot{w}], \\ 0 & \text{otherwise.} \end{cases}$$

Thus the proposition holds in this case.

Now we assume that w is not a minimal length element in its σ-conjugacy class. By Theorem 2.4(1), there exist $w' \in \tilde{W}$ and $i \in \tilde{\mathbb{S}}$ such that $w \approx_\sigma w'$ and $s_i w' \sigma(s_i) < w'$. Set $w_1 = s_i w'$ and $w_2 = s_i w' \sigma(s_i)$. Then $\ell(w_1), \ell(w_2) < \ell(w)$. By the proof of [13, Theorem 2.16],

$$\sharp_{\breve{I}}([b] \cap \breve{I}\dot{w}\breve{I}) = \sharp_{\breve{I}}([b] \cap \breve{I}\dot{w}'\breve{I}) = (q-1)\sharp_{\breve{I}}([b] \cap \breve{I}\dot{w}_1\breve{I}) + q\sharp_{\breve{I}}([b] \cap \breve{I}\dot{w}_2\breve{I}).$$

On the other hand, by [13, §2.8.2], $F_{w,[b]} = (\boldsymbol{q}-1)F_{w_1,[b]} + \boldsymbol{q}F_{w_2,[b]}$. Now the statement for w follows from the inductive hypothesis on w_1 and w_2. □

We have the decomposition $\breve{I}\dot{w}\breve{I} = \bigsqcup_{[b] \in B(\boldsymbol{G})} [b] \cap \breve{I}\dot{w}\breve{I}$. Thus

$$\boldsymbol{q}^{\ell(w)} = \sharp_{\breve{I}}(\breve{I}\dot{w}\breve{I}) = \sum_{[b] \in B(\boldsymbol{G})} \sharp_{\breve{I}}([b] \cap \breve{I}\dot{w}\breve{I}) = \sum_{[b] \in B(\boldsymbol{G})} F_{w,[b]}.$$

This gives an alternative proof of Proposition 3.6 and a geometric interpretation of Proposition 3.6 in the case where the σ-action on $\tilde{W}$ is trivial via counting (the normalized cardinality of) the rational points of $\breve{I}\dot{w}\breve{I}$.

3D. *Reduction trees.* Let $w \in \tilde{W}$. We construct the reduction tree for w, which encodes the Deligne–Lusztig reduction for the affine Deligne–Lusztig varieties associated with w (and with all $b \in \breve{G}$).

The vertices of the graphs are the elements of $\tilde{W}$, and the (oriented) edges are of the form $x \rightharpoonup y$, where $x, y \in \tilde{W}$, and there exist $x' \in \tilde{W}$ and $i \in \tilde{\mathbb{S}}$ with $x \approx_\sigma x'$, $s_i x' \sigma(s_i) < x'$ and $y \in \{s_i x', s_i x' \sigma(s_i)\}$. Some elements of $\tilde{W}$ may occur more than once in a reduction tree.

The reduction trees are constructed inductively.

Suppose that w is of minimal length in its σ-conjugacy class of $\tilde{W}$. Then the reduction tree of w consists of a single vertex w and no edges.

Suppose that w is not of minimal length in its σ-conjugacy class of $\tilde{W}$ and that a reduction tree is given for any $z \in \tilde{W}$ with $\ell(z) < \ell(w)$. By Theorem 2.4(1), there exist $w' \in \tilde{W}$ and $i \in \tilde{\mathbb{S}}$ with $w \approx_\sigma w'$ and $s_i w' \sigma(s_i) < w'$. The reduction tree of w is the graph containing the given reduction tree for $s_i w'$ and the reduction tree for $s_i w' \sigma(s_i)$, and the edges $w \rightharpoonup s_i w'$ and $w \rightharpoonup s_i w' \sigma(s_i)$.

The reduction trees of w are not unique. They depend on the choices of w' and s_i in the construction. We will see in the rest of this section that the reduction trees encode more information than the class polynomials.

3E. ***Reduction path.*** Let $\mathcal{T}$ be a reduction tree of w. An *end point* of the tree $\mathcal{T}$ is a vertex x of $\mathcal{T}$ such that there is no edge of the form $x \rightharpoonup x'$ in $\mathcal{T}$. By Theorem 2.4, each end point is of minimal length in its σ-conjugacy class. A *reduction path* in $\mathcal{T}$ is a path $\underline{p}: w \rightharpoonup w_1 \rightharpoonup \cdots \rightharpoonup w_n$, where w_n is an end point of $\mathcal{T}$. The *length* $\ell(\underline{p})$ of the reduction path $\underline{p}$ is the number of edges in $\underline{p}$. We also write $\operatorname{end}(\underline{p}) = w_n$ and $[b]_{\underline{p}} = \Psi(\operatorname{end}(\underline{p})) \in B(\boldsymbol{G})$.

If $x \rightharpoonup y$, then $\ell(x) - \ell(y) \in \{1, 2\}$. We say that the edge $x \rightharpoonup y$ is of type I if $\ell(x) - \ell(y) = 1$ and of type II if $\ell(x) - \ell(y) = 2$. For any reduction path $\underline{p}$, we denote by $\ell_{\mathrm{I}}(\underline{p})$ the number of type-I edges in $\underline{p}$ and by $\ell_{\mathrm{II}}(\underline{p})$ the number of type-II edges in $\underline{p}$. Then $\ell(\underline{p}) = \ell_{\mathrm{I}}(\underline{p}) + \ell_{\mathrm{II}}(\underline{p})$.

The following relation between class polynomials and reduction trees follows easily from the inductive construction, and we omit the details of its proof.

Lemma 3.8. *Let $w \in \tilde{W}$ and let $\mathcal{T}$ be a reduction tree of w. Then, for any σ-conjugacy class $\mathcal{O}$ of $\tilde{W}$,*

$$F_{w,\mathcal{O}} = \sum_{\underline{p}} (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(\underline{p})} \boldsymbol{q}^{\ell_{\mathrm{II}}(\underline{p})},$$

where $\underline{p}$ runs over all the reduction paths in $\mathcal{T}$ with $\Psi(\operatorname{end}(\underline{p})) = \mathcal{O}$.

Combining Proposition 2.5 with the construction of the reduction trees, we obtain the following result.

Proposition 3.9. *Let $w \in \tilde{W}$ and $\mathcal{T}$ be a reduction tree of w. Then, for any $b \in \breve{G}$, there exists a decomposition*

$$X_w(b) = \bigsqcup_{\substack{\underline{p} \text{ is a reduction path of } \mathcal{T} \\ [b]_{\underline{p}} = [b]}} X_{\underline{p}},$$

where $X_{\underline{p}}$ is a locally closed subscheme of $X_w(b)$ and is $\boldsymbol{J}_b(F)$-equivariant universally homeomorphic to an iterated fibration of type $(\ell_{\mathrm{I}}(\underline{p}), \ell_{\mathrm{II}}(\underline{p}))$ over $X_{\operatorname{end}(\underline{p})}(b)$.

Remark 3.10. Since $\operatorname{end}(\underline{p})$ is a minimal length element in its σ-conjugacy class, by (2-1) we have $X_{\operatorname{end}(\underline{p})}(b) \cong \boldsymbol{J}_b(F) \times^P X$, where P is a parahoric subgroup of $\boldsymbol{J}_b(F)$ and X is (the perfection of) an irreducible component of a classical Deligne–Lusztig variety. Thus each irreducible component Y of $X_{\underline{p}}$ is universally homeomorphic to an iterated fibration of type $(\ell_{\mathrm{I}}(\underline{p}), \ell_{\mathrm{II}}(\underline{p}))$ over X. We have a natural action of $\boldsymbol{J}_b(F)$ on $X_w(b)$, and $X_{\underline{p}}$ is stable under this action. In this case, $X_{\underline{p}} \cong \boldsymbol{J}_b(F) \times^P Y$.

4. Cordiality and the set $B(\boldsymbol{G})_w$

4A. ***Maximal Hodge–Newton irreducible elements.*** Recall that $[b_{t^\mu}]$ is the unique maximal element of $B(\boldsymbol{G}, \mu)$. By [32, Corollary 7.6], there is a unique maximal element $[b_{\mu,\boldsymbol{G}\text{-indec}}]$ in $B(\boldsymbol{G}, \mu)_{\text{indec}}$, whose Newton point is denoted by $\nu_{b_{\mu,\boldsymbol{G}\text{-indec}}}$. We give an explicit description of $\nu_{b_{\mu,\boldsymbol{G}\text{-indec}}}$ below.

Following [2], for any subset $E \subseteq (V^+)^\sigma$, we set

$$C_{\geqslant E} = \{v \in (V^+)^\sigma \mid v \geqslant v', \forall v' \in E\}.$$

By [2, Theorem 6.5], $C_{\geqslant E}$ has a unique minimal element, which we denote by $\min C_{\geqslant E}$.

Proposition 4.1. *We have* $\nu_{b_{\mu,\boldsymbol{G}\text{-indec}}} = \min C_{\geqslant E_0}$, *where* $E_0 = \{e_i \mid i \in \mathbb{S}\}$ *with* $e_i \in \mathbb{R}\omega_i^\vee$ *such that* $\langle e_i, \omega_i \rangle = \frac{1}{\sharp\mathcal{O}_i} \max\{0, \langle \mu, \omega_{\mathcal{O}_i} \rangle - 1\}$.

Proof. By [2, Theorem 6.5], there exists a unique σ-conjugacy class $[b] \in B(\boldsymbol{G}, \mu)$ such that $\nu_b = \min C_{\geqslant E_0}$ and $\langle \mu - \nu_b, \omega_{\mathcal{O}_k} \rangle = 1$ for any $k \in \mathbb{S} - I(\nu)$.

Suppose that $(\mu, [b])$ is Hodge–Newton decomposable with respect to some standard Levi subgroup $\boldsymbol{M}_J$ with $J = \sigma(J) \subsetneq \mathbb{S}$. Let $j \in \mathbb{S} - J$. By definition, $\langle \mu - \nu_b, \omega_{\mathcal{O}_j} \rangle = 0$. On the other hand, as $I(\nu_b) \subseteq J$ we have $\langle \nu_b, \alpha_j \rangle \neq 0$. Thus $\langle \mu - \nu_b, \omega_{\mathcal{O}_j} \rangle = 1$, which is a contradiction. Therefore $[b] \in B(\boldsymbol{G}, \mu)_{\text{indec}}$.

On the other hand, let $[b'] \in B(\boldsymbol{G}, \mu)_{\text{indec}}$. For any $i \in \mathbb{S} - I(\nu_{b'})$, we denote by $\text{pr}_{(i)} : V = \mathbb{R}\omega_i^\vee \oplus \sum_{j \neq i} \mathbb{R}\alpha_j^\vee \to \mathbb{R}\omega_i^\vee$ the natural projection. Set $e_i' = \text{pr}_{(i)}(\nu_{b'}) \in \mathbb{R}\omega_i^\vee$. Let $E' = \{e_i' \mid i \in \mathbb{S} - I(\nu_{b'})\}$. Again by [2, Theorem 6.5], we have $\nu_{b'} = \min C_{\geqslant E'}$. By Section 2B(a), we have $\langle \mu - \nu_{b'}, \omega_{\mathcal{O}_i} \rangle \in \mathbb{Z}_{\geqslant 1}$ for $i \in \mathbb{S} - I(v')$. This means that $e_i' \leqslant e_i$ for $i \in \mathbb{S} - I(v')$. So $\nu_b \in C_{\geqslant E'}$ and $\nu_b \geqslant \min C_{\geqslant E'} = \nu_{b'}$. Hence $[b]$ is the unique maximal element of $B(\boldsymbol{G}, \mu)_{\text{indec}}$. □

Corollary 4.2. *Suppose that* μ *is essentially noncentral. Then*

$$\text{length}([b_{\mu,\boldsymbol{G}\text{-indec}}], [t^\mu]) = \sharp(\mathbb{S}/\langle\sigma\rangle).$$

Proof. As μ is essentially noncentral, we have $B(\boldsymbol{G}, \mu)_{\text{indec}} = B(\boldsymbol{G}, \mu)_{\text{irr}}$ and hence $0 < \langle \mu - \nu_{b_{\mu,\boldsymbol{G}\text{-indec}}}, \omega_{\mathcal{O}_i} \rangle$ for any $i \in \mathbb{S}$. On the other hand, we have $\langle \mu - \nu_{b_{\mu,\boldsymbol{G}\text{-indec}}}, \omega_{\mathcal{O}_i} \rangle \leqslant 1$ by Proposition 4.1. Then the statement follows directly from Section 2B(a). □

4B. ***Proof of Theorem 2.6(1) and (2) for $w = t^\mu c$.*** Now we prove Theorem 2.6(1) and (2) in the case $w = t^\mu c \in {}^{\mathbb{S}}\tilde{W}$, where c is a partial σ-Coxeter element. If μ is central over some connected components of $\text{supp}_\sigma(c)$, then the element $t^\mu c$ would not be in ${}^{\mathbb{S}}\tilde{W}$. Thus μ is essentially noncentral in $\text{supp}_\sigma(c)$. Set $J = \text{supp}_\sigma(c)$. Reviewing the definition of $J(w)$ in Section 2J, we have $J(w) = J(w)'_\mu = J_0(w) = J$. We need to show that w is cordial and $B(\boldsymbol{G})_w = B(\boldsymbol{G}, \mu)_{J\text{-irr}}$.

By Section 2C, we have a natural bijection $B(\boldsymbol{M}_J, \mu)_{\text{irr}} \cong B(\boldsymbol{G}, \mu)_{J\text{-irr}}$. By [6, Theorem 3.3.1], we have $B(\boldsymbol{G})_w = B(\boldsymbol{M}_J)_w$. Note that $\mu^\diamond - \nu_{b_w} \in \sum_{j \in J} \mathbb{R}\alpha_j^\vee$. Hence $\langle \mu - \nu_{b_w}, \rho \rangle = \langle \mu - \nu_{b_w}, \rho_J \rangle$, where ρ_J is the half sum of positive roots of $\boldsymbol{M}_J$. By definition, w is cordial in $\boldsymbol{G}$ if and only if it is cordial in $\boldsymbol{M}_J$. Hence we may assume that $J = \mathbb{S}$ and that c is a σ-Coxeter element.

We first show that

(a) $B(\boldsymbol{G})_w \subseteq B(\boldsymbol{G}, \mu)_{\text{indec}}$.

Suppose that $X_w(b) \neq \varnothing$ and (μ, b) is Hodge–Newton decomposable with respect to some proper standard Levi subgroup $\boldsymbol{M}$. By [7, Theorem 1.11], there is some $u \in W$ such that $u^{-1} w \sigma(u)$ lies in $\tilde{W}_{\boldsymbol{M}}$, which contradicts the fact that c is σ-Coxeter. Thus (a) is proved.

Next we show that

(b) $[b_w] = [b_{\mu,\boldsymbol{G}\text{-indec}}]$ and w is a cordial element.

By Section 2E(b), we have

$$\langle \mu, 2\rho\rangle - 2\ell(c) = \ell(w) - \ell(\eta_\sigma(w)) \leqslant \langle \nu_{b_w}, 2\rho\rangle - \mathrm{def}(b_w).$$

Combined with Section 2E(a), we get $\mathrm{length}([b_w], [t^\mu]) \leqslant \ell(c) = \sharp(\mathbb{S}/\langle\sigma\rangle)$. However, by Corollary 4.2, we have $\mathrm{length}([b_w], [b_{t^\mu}]) \geqslant \mathrm{length}([b_{\mu,\boldsymbol{G}\text{-indec}}], [b_{t^\mu}]) = \sharp(\mathbb{S}/\langle\sigma\rangle)$. Thus we must have $[b_w] = [b_{\mu,\boldsymbol{G}\text{-indec}}]$ and $\dim X_w(b_w) = d_w(b_w)$. Thus (b) is proved.

Finally, we show that

(c) $B(\boldsymbol{G})_w = B(\boldsymbol{G}, \mu)_{\mathrm{indec}} = B(\boldsymbol{G}, \mu)_{\mathrm{irr}}$.

Let $[b_{\min}]$ be the unique basic σ-conjugacy class in $B(\boldsymbol{G})$ with $\kappa(b_{\min}) = \kappa(\mu)$. Then $[b_{\min}]$ is the unique minimal element in $B(\boldsymbol{G}, \mu)_{\mathrm{indec}}$. Note that c is a σ-Coxeter element of W. Thus $\nu_{\dot{w}}$ is central and $[\dot{w}] = [b_{\min}]$. In particular, $[b_{\min}] \in B(\boldsymbol{G})_w$. By Theorem 2.3, $B(\boldsymbol{G})_w$ is saturated and hence must be equal to $B(\boldsymbol{G}, \mu)_{\mathrm{indec}}$.

This completes the proof of the $t^\mu c$ case.

4C. ***Partial conjugation.*** To handle the general case, we use the partial conjugation method introduced in [11].

By partial conjugation, we mean conjugating by elements in the finite Weyl group W. For any $x \in {}^{\mathbb{S}}\tilde{W}$, set

$$I(x) = \max\{J \subseteq \mathbb{S} \mid \mathrm{Ad}(x)\sigma(J) = J\}.$$

This is well-defined. Indeed, if $\mathrm{Ad}(x)\sigma(J_i) = J_i$ for $i = 1, 2$, then $\mathrm{Ad}(x)\sigma(J_1 \cup J_2) = J_1 \cup J_2$. Let $W_{I(x)}$ be the subgroup of W generated by the simple reflections in $I(x)$. Then $\mathrm{Ad}(x) \circ \sigma$ gives a length-preserving group automorphism on $W_{I(x)}$. By [11, Proposition 2.4], we have

$$\tilde{W} = \bigsqcup_{x \in {}^{\mathbb{S}}\tilde{W}} W \cdot_\sigma (W_{I(x)} x) = \bigsqcup_{x \in {}^{\mathbb{S}}\tilde{W}} W \cdot_\sigma (x W_{\sigma(I(x))}).$$

Moreover, by [11, Proposition 3.4], we have the following:

(a) For any $w \in \tilde{W}$, there exist $x \in {}^{\mathbb{S}}\tilde{W}$ and $u \in W_{I(x)}$ such that $w \to_\sigma ux$ and all the simple reflections involved in the conjugations are in $\mathbb{S}$.

By [12, Proposition 4.9], we have the following:

(b) Let $x \in {}^{\mathbb{S}}\tilde{W}$ and $u \in W_{I(x)}$. Then $B(G)_{ux} = B(G)_x$ and $\dim X_{ux}(b) = \dim X_x(b) + \ell(u)$ for any $[b] \in B(\boldsymbol{G})_x$.

Similar to Section 3D, we may consider partial reductions. By partial reduction, we mean reduction $w \rightharpoonup s_i w$ or $w \rightharpoonup s_i w \sigma(s_i)$ with $i \in \mathbb{S}$. We show that partial reduction preserves elements with finite Coxeter parts.

Lemma 4.3. *Let $w \in \tilde{W}$ with $\eta_\sigma(w)$ a partial σ-Coxeter element of W. Let $i \in \mathbb{S}$ with $s_i w < w$. Then*:

(1) *$\eta_\sigma(s_i w \sigma(s_i))$ is a partial σ-Coxeter element of W and* $\mathrm{supp}_\sigma(\eta_\sigma(s_i w \sigma(s_i))) = \mathrm{supp}_\sigma(\eta_\sigma(w))$.

(2) *If, moreover,* $s_i w\sigma(s_i) < w$, *then* $\eta_\sigma(s_i w)$ *is a partial* σ*-Coxeter element of* W *and* $\mathrm{supp}_\sigma(\eta_\sigma(s_i w)) = \mathrm{supp}_\sigma(\eta_\sigma(w)) - \{\sigma^l(i') \mid l \in \mathbb{Z}\}$ *for some* $i' \in \mathrm{supp}_\sigma(\eta_\sigma(w))$.

Proof. We prove part (1). The proof of part (2) is similar, and we skip the details.

Write $w = xt^\mu y$ with $t^\mu y \in {}^{\mathbb{S}}\tilde{W}$. Set $c = \eta_\sigma(w) = \sigma^{-1}(y)x$. If $y\sigma(s_i) \in {}^{\mathbb{S}}\tilde{W}$, then $\eta_\sigma(s_i w\sigma(s_i)) = \eta_\sigma(w)$, and the statement is obvious. Now assume that $y\sigma(s_i) = s_{i'}y$ for some $i' \in I(\mu)$. Then we have $x^{-1}(\alpha_i) < 0$ and $\sigma^{-1}(y)(\alpha_i) > 0$. Thus $(\sigma^{-1}(y)x)(-x^{-1}(\alpha_i)) < 0$. It follows that $\sigma^{-1}(y)s_i x = (\sigma^{-1}(y)x)(x^{-1}s_i x) < \sigma^{-1}(y)x$. By the cancellation property of Coxeter groups, we conclude that $\sigma^{-1}(y)s_i x$ is a partial σ-Coxeter element and $\ell(\sigma^{-1}(y)s_i x) = \ell(\sigma^{-1}(y)x) - 1$. Write $c' = \sigma^{-1}(y)s_i x$. Notice that $\sigma^{-1}(s_{i'})c = \sigma^{-1}(s_{i'}y)x = \sigma^{-1}(y)s_i x = c'$. It follows that $\sigma^{-1}(i')$ is not in the σ-support of c'. Hence $\eta_\sigma(s_i w\sigma(s_i)) = \sigma^{-1}(y)s_i x s_{i'} = c's_{i'}$ is a partial σ-Coxeter element, and

$$\mathrm{supp}_\sigma(\eta_\sigma(s_i w\sigma(s_i))) = \mathrm{supp}_\sigma(\eta_\sigma(w)). \qquad \square$$

4D. ***Proof of Theorem 2.6(1) and (2): general case.*** Let $w = xt^\mu y \in \tilde{W}$ with $t^\mu y \in {}^{\mathbb{S}}\tilde{W}$. We assume that $\eta_\sigma(w)$ a partial σ-Coxeter element. We prove Theorem 2.6(1) and (2) by induction on $\ell(x)$.

The case $\ell(x) = 0$ has already been proved in Section 4B. Assume that $\ell(x) > 0$. Let $i \in \mathbb{S}$ such that $s_i x < x$. There are three different cases.

Case (1): $\ell(s_i w\sigma(s_i)) < \ell(w)$. Write $w_1 = s_i w$ and $w_2 = s_i w\sigma(s_i)$. By Lemma 4.3, $\eta_\sigma(w_1)$ and $\eta_\sigma(w_2)$ are both partial σ-Coxeter elements. The inductive hypothesis applies for w_1 and w_2. Since $w_1 > w_2$, we have $[b_w] = \max([b_{w_1}], [b_{w_2}]) = [b_{w_1}]$. Then $d_w(b_w) = d_{w_1}(b_w) + 1$ and $\dim X_{w_1}(b_w) + 1 = \dim X_w(b_w)$. Thus w is cordial.

Observe that

$$B(\boldsymbol{G})_w = B(\boldsymbol{G})_{w_1} \cup B(\boldsymbol{G})_{w_2} = \bigsqcup_{J \in [J_0(w_1), J(w_1)]_\mu \cup [J_0(w_2), J(w_2)]_\mu} B(\boldsymbol{G}, \mu)_{J\text{-irr}}.$$

Note that $J(w_2) = J(w)$, $J_0(w_1) = J_0(w)$ and $J(w_1) \subset J(w)$. By Section 2E(b), $B(\boldsymbol{G})_w$ is saturated. Hence we must have

$$[J_0(w_1), J(w_1)]_\mu \cup [J_0(w_2), J(w_2)]_\mu = [J_0(w), J(w)]_\mu.$$

This proves part (2) of Theorem 2.6.

Case (2): $y\sigma(s_i) < y$. Write $w' = s_i w\sigma(s_i) = s_i x t^\mu y\sigma(s_i)$. The inductive hypothesis applies for w'. Note that $w \approx_\sigma w'$, in particular, $B(\boldsymbol{G})_w = B(\boldsymbol{G})_{w'}$. Also $J(w) = J(w')$. Hence the statements hold for w.

Case (3): $y\sigma(s_i) = s_{i'}y$ for some $i' \in I(\mu)$ and $\ell(s_i x s_{i'}) = \ell(x)$. Write $w' = s_i w\sigma(s_i) = s_i x s_{i'} t^\mu y$. Then $w \approx_\sigma w'$. By Lemma 4.3, $\eta_\sigma(w') = \sigma^{-1}(y)s_i x s_{i'}$ is a partial σ-Coxeter element with length equal to $\ell(\sigma^{-1}(y)x)$. Hence the statements hold for w if and only if they hold for w'. We continue the procedure until case (1) or (2) happens. If case (3) happens all the time and the procedure does not end, then $x \in W_{I(t^\mu y)}$, and both x and y are partial σ-Coxeter elements. Then the statements follow from Section 4C(b).

5. Analyzing the reduction paths

5A. *σ-Conjugacy classes of $\tilde{W}$.* We first recall the definition of elliptic conjugacy classes. Let W_1 be a Coxeter group and let S_1 be the index set of simple reflections in W_1. Let δ be a length-preserving group automorphism on W_1. A δ-conjugacy class C of W_1 is called *elliptic* if it contains no elements in any proper δ-stable standard parabolic subgroup of W_1.

Let $x \in \tilde{W}$. We regard $x\sigma$ as an element in $\tilde{W} \rtimes \langle\sigma\rangle$. There exists a positive integer such that $(x\sigma)^n = t^\lambda$ for some $\lambda \in X_*(T)_{\Gamma_0}$. Then we set $\nu_x = \lambda/n \in V$. It is easy to see that ν_x is independent of the choice of n. Moreover, the unique dominant element in the W-orbit of ν_x equals the (dominant) Newton point $\nu_{\dot{x}}$ for $\dot{x} \in \breve{G}$.

We follow [13, §1.8.3]. Let $J \subseteq \mathbb{S}$. Let $\tilde{W}_J = X_*(T)_{\Gamma_0} \rtimes W_J$ be the Iwahori–Weyl group of the standard Levi subgroup $\boldsymbol{M}_J$ of $\boldsymbol{G}$. Let ${}^J\tilde{W}$ be the set of minimal length representatives for cosets in $W_J\backslash\tilde{W}$. Let $\tilde{J} \supseteq J$ be the set of simple reflections for the Iwahori–Weyl group $\tilde{W}_J$.

We say that $(J, x, \breve{K}, C)$ is a *standard quadruple* if

(1) $\sigma(J) = J$;

(2) $x \in {}^J\tilde{W}$ such that ν_x is dominant, $J = I(\nu_x)$, and $\mathrm{Ad}(x)\circ\sigma$ preserves $\tilde{J}$;

(3) $\breve{K} \subseteq \tilde{J}$ with $W_{\breve{K}}$ finite and $\mathrm{Ad}(x)(\sigma(\breve{K})) = \breve{K}$;

(4) C is an elliptic $(\mathrm{Ad}(x)\circ\sigma)$-conjugacy class of $W_{\breve{K}}$.

We say that the standard quadruples $(J, x, \breve{K}, C)$ and $(J', x', \breve{K}', C')$ are *equivalent* in $\tilde{W}$ if $J = J'$, there exists a length-zero element τ of $\tilde{W}_J$ with $x' = \tau x\sigma(\tau)^{-1}$, and there exists $w \in \tilde{W}_J$ with $x'\sigma(w)(x')^{-1} = w$ and $C' = w\tau C(w\tau)^{-1}$.

By [13, Theorem 1.19], we have the following:

(a) The map $(J, x, \breve{K}, C) \mapsto \tilde{W}\cdot_\sigma Cx$ induces a bijection between the equivalence classes of standard quadruples and the set of σ-conjugacy classes of $\tilde{W}$.

Let $\mathcal{O} \in B(\tilde{W}, \sigma)$ and let $(J, x, \breve{K}, C)$ be a standard quadruple associated with $\mathcal{O}$. We say that $\mathcal{O}$ is *Coxeter* (resp. *elliptic*) associated with $[b] \in B(\boldsymbol{G})$ if $\Psi(\mathcal{O}) = [b]$, $\breve{K} \subset \tilde{J}$ is very special with respect to $\mathrm{Ad}(\dot{x})\circ\sigma$, and C is an $(\mathrm{Ad}(x)\circ\sigma)$-Coxeter (resp. elliptic) conjugacy class of $W_{\breve{K}}$. Namely, C contains an $(\mathrm{Ad}(x)\circ\sigma)$-Coxeter (resp. elliptic) element of the finite Coxeter group $W_{\breve{K}}$. We say that $w \in \tilde{W}$ is a σ-Coxeter element associated with $[b]$ if it is a minimal length element in a Coxeter σ-conjugacy class associated with $[b]$.

5B. *Description of reduction trees.* For any $[b] \in B(\boldsymbol{G}, \mu)$, set

$$\ell_{\mathrm{I}}(\mu, [b]) = \sharp(\mathbb{S}/\langle\sigma\rangle) - \sharp(I(\nu_b)/\langle\sigma\rangle),$$
$$\ell_{\mathrm{II}}(\mu, [b]) = \mathrm{length}([b], [t^\mu]) - \sharp(\mathbb{S}/\langle\sigma\rangle).$$

We have the following description of reduction trees.

Theorem 5.1. *Let c be a σ-Coxeter element of W such that $t^\mu c \in {}^{\mathbb{S}}\tilde{W}$. Let $\mathcal{T}$ be a reduction tree of $t^\mu c$. Then, for any reduction path $\underline{p}$ in $\mathcal{T}$, we have*

(1) $\ell_{\mathrm{I}}(\underline{p}) = \ell_{\mathrm{I}}(\mu, [b]_{\underline{p}})$ *and* $\ell_{\mathrm{II}}(\underline{p}) = \ell_{\mathrm{II}}(\mu, [b]_{\underline{p}})$;

(2) $\mathrm{end}(\underline{p})$ *is a σ-Coxeter element associated with* $[b]_{\underline{p}}$.

Moreover, for any $[b] \in B(\boldsymbol{G}, \mu)_{\mathrm{indec}}$, there exists a unique reduction path $\underline{p}$ in $\mathcal{T}$ with $[b]_{\underline{p}} = [b]$.

Combining Theorem 5.1 with Proposition 3.9 and Remark 3.10, we obtain part (3) of Theorem 2.6 for $w = t^\mu c$. We will describe the reduction trees of the elements with finite partial σ-Coxeter part in Section 7 and deduce Theorem 2.6(3) for such elements.

In the rest of this section, we shall prove parts (1) and (2) of Theorem 5.1. The "moreover" part (i.e., the multiplicity-one result) is the most difficult part and will be proved in Section 6.

5C. *Estimate ℓ_{I}.* Let $\mathrm{Aff}(V)$ be the group of affine transformations on V. For any $g \in \mathrm{Aff}(V)$, define $V^g = \{v \in V \mid g(v) = v\}$. We have a natural projection map $p : \tilde{W} \rtimes \langle\sigma\rangle \to \mathrm{Aff}(V) \to \mathrm{GL}(V)$. For any $w \in \tilde{W}$, define $V_w = \{v \in V \mid w\sigma(v) = v + \nu_w\}$. We have $\dim V_w = \dim V^{p(w\sigma)}$.

It is easy to see that for any $g \in \mathrm{GL}(V)$ and any reflection $r \in \mathrm{GL}(V)$, we have

$$|\dim V^{rg} - \dim V^g| \leqslant 1.$$

In particular, for any $w \in \tilde{W}$ and $i \in \tilde{\mathbb{S}}$, we have

$$|\dim V_{s_i w} - \dim V_w| \leqslant 1.$$

Now let $w = t^\mu c$ be as in Theorem 5.1. Let $\mathcal{T}$ be a reduction tree of w and let $\underline{p}$ be a reduction path in $\mathcal{T}$. Set $e = \mathrm{end}(\underline{p})$ and $[b] = \Psi(e) \in B(\boldsymbol{G})$. Let $(J, x, \breve{K}, C)$ be a standard quadruple associated with the σ-conjugacy class of e.

Consider the variation of $\dim V_?$ along the reduction path $\underline{p}$, where ? stands for any element in $\tilde{W}$. Type-II edges do not change $\dim V_?$, and type-I edges change $\dim V_?$ by at most 1. Therefore

$$\ell_{\mathrm{I}}(\underline{p}) \geqslant |\dim V_e - \dim V_w|.$$

Since $p(w) = c$ is a σ-Coxeter element, $V^{p(w\sigma)} = V^{p(\tilde{W} \rtimes \langle\sigma\rangle)}$. Since C is $\mathrm{Ad}(x) \circ \sigma$-elliptic in $\breve{K}$, we have

$$\dim V_e = \dim V_x - \sharp(\breve{K}/\langle \mathrm{Ad}(x) \circ \sigma\rangle).$$

Hence

$$\ell_{\mathrm{I}}(\underline{p}) \geqslant \dim V_x - \sharp(\breve{K}/\langle \mathrm{Ad}(x) \circ \sigma\rangle). \tag{5-1}$$

By [23, §1.9], $\mathrm{def}(b) = \sharp(\mathbb{S}/\langle\sigma\rangle) - \dim V_x$. Note that $\ell_{\mathrm{I}}(\underline{p}) + 2\ell_{\mathrm{II}}(\underline{p}) = \ell(t^\mu c) - \ell(e)$. Moreover,

$$\ell(e) \geqslant \langle \nu_b, 2\rho\rangle + \sharp(\breve{K}/\langle \mathrm{Ad}(x) \circ \sigma\rangle), \tag{5-2}$$

with equality holding if and only if C is an $(\mathrm{Ad}(x)\circ\sigma)$-Coxeter conjugacy class in $\breve{K}$. We have

$$\begin{aligned}\dim X_{\underline{p}} &= \ell_{\mathrm{I}}(\underline{p}) + \ell_{\mathrm{II}}(\underline{p}) + \ell(e) - \langle \nu_b, 2\rho\rangle \\ &= \tfrac{1}{2}(\ell_{\mathrm{I}}(\underline{p}) + \langle \mu, 2\rho\rangle - \sharp(\mathbb{S}/\langle\sigma\rangle) + \ell(e)) - \langle \nu_b, 2\rho\rangle \\ &\geqslant \langle \mu - \nu_b, \rho\rangle + \tfrac{1}{2}(\dim V_x - \sharp(\mathbb{S}/\langle\sigma\rangle)) \\ &= \langle \mu - \nu_b, \rho\rangle - \tfrac{1}{2}\mathrm{def}(b) = d_w(b).\end{aligned}$$

By Section 2E(b), we have $\dim X_{\underline{p}} \leqslant \dim X_w(b) \leqslant \dim d_w(b)$. Thus the inequalities in (a) and (b) are equalities, and $\dim X_{\underline{p}} = \dim X_w(b)$. In particular, C is an $(\mathrm{Ad}(x)\circ\sigma)$-Coxeter conjugacy class in $\breve{K}$.

5D. ***Affine Deligne–Lusztig varieties in the affine Grassmannian.*** It remains to show that $\breve{K}$ occurring in Section 5C is very special. To do this, we need some information on affine Deligne–Lusztig varieties in the affine Grassmannian.

Let $\breve{P} \subseteq \breve{G}$ be a special parahoric subgroup containing $\breve{I}$. The affine Deligne–Lusztig variety in the affine Grassmannian $\breve{G}/\breve{P}$ is defined by

$$X_\mu(b) = \{g \in \breve{G}/\breve{P} \mid g^{-1} b\sigma(g) \in \breve{P} t^\mu \breve{P}\}.$$

The following dimension formula is proved for split groups [5; 30], for unramified groups [9; 34], and in general [13, Theorem 2.29].

Theorem 5.2. *Suppose that* $[b] \in B(\boldsymbol{G}, \mu)$. *Then* $\dim X_\mu(b) = \langle \mu - \nu_b, \rho\rangle - \frac{1}{2}\mathrm{def}(b)$.

Let $\Sigma^{\mathrm{top}}(X_\mu(b))$ be the set of top-dimensional irreducible components of $X_\mu(b)$.

Let $\widehat{\boldsymbol{G}}$ be the Langlands dual of $\boldsymbol{G}$ over the complex number field $\mathbb{C}$. Let $\widehat{T}$ be the maximal torus dual to T. Then σ acts on $\widehat{T}$ in a natural way, and we denote by $\widehat{T}^\sigma$ the σ-fixed points of $\widehat{T}$. Let $\lambda_b \in X^*(\widehat{T}^\sigma)$ be the "best integral approximation" of the Newton point of b in the sense of [10, Definition 2.1]. Let V_μ be the irreducible representation of $\widehat{\boldsymbol{G}}$ with highest weight μ. Write $V_\mu(\lambda_b)$ for the corresponding λ_b-weight subspace of $\widehat{T}^\sigma$. The following result was conjectured by M. Chen and X. Zhu, and is proved in [20; 26; 33].

Theorem 5.3. *The number of* $\boldsymbol{J}_b(F)$*-orbits on* $\Sigma^{\mathrm{top}}(X_\mu(b))$ *equals* $\dim V_\mu(\lambda_b)$. *The stabilizer of each element in* $\Sigma^{\mathrm{top}}(X_\mu(b))$ *is a very special parahoric subgroup of* $\boldsymbol{J}_b(F)$.

We also need the following result that connects affine Deligne–Lusztig varieties in the affine flag and in the affine Grassmannian.

Lemma 5.4. *The* $\boldsymbol{J}_b(F)$*-equivariant projection map* $X_{t^\mu c}(b) \to X_\mu(b)$ *is injective.*

Proof. Let $g\breve{I}, g'\breve{I} \in \breve{G}/\breve{I}$ be in the same fiber of the natural projection map $X_{t^\mu c}(b) \to X_\mu(b)$. Then $g'^{-1}g \in \breve{P}$. We have $g'^{-1}g \in \breve{I}\dot{x}\breve{I}$ for some $x \in W$. Since $(g'^{-1}g)(g^{-1}b\sigma(g)) = (g'^{-1}b\sigma(g'))\sigma(g'^{-1}g)$, $(\breve{I}\dot{x}\breve{I})(\breve{I}t^\mu\dot{c}\breve{I}) \cap (\breve{I}t^\mu\dot{c}\breve{I})\sigma(\breve{I}\dot{x}\breve{I}) \neq \varnothing$. Since $t^\mu c \in {}^{\mathbb{S}}\tilde{W}$, $(\breve{I}\dot{x}\breve{I})(\breve{I}t^\mu\dot{c}\breve{I}) = \breve{I}\dot{x}t^\mu\dot{c}\breve{I}$. Thus we have $xt^\mu c = t^\mu c\sigma(x)$ and $\mathrm{supp}_\sigma(x) \subset I(t^\mu c)$. As c is σ-elliptic, we conclude that $\mathrm{supp}_\sigma(x) = \varnothing$ and hence $g'g^{-1} \in \breve{I}$ as desired. □

5E. ***Proof of Theorem 5.1(1) and (2).*** We continue our analysis of reduction paths. All the notation is the same as in Section 5C.

Equalities (5-1) and (5-2) hold and $\operatorname{def}(b) = \sharp(I(\nu_b)/\langle\sigma\rangle) - \sharp(\breve{K}/\langle \operatorname{Ad}(x)\circ\sigma\rangle)$. It follows that

$$\ell_{\mathrm{I}}(\underline{p}) = \sharp(\mathbb{S}/\langle\sigma\rangle) - \sharp(I(\nu_b)/\langle\sigma\rangle) = \ell_{\mathrm{I}}(\mu, [b]).$$

Using Section 2E(a) and the simple fact that $\ell_{\mathrm{I}}(\underline{p}) + 2\ell_{\mathrm{II}}(\underline{p}) = \ell(w) - \ell(e)$, one can prove that $\ell_{\mathrm{II}}(\underline{p}) = \ell_{\mathrm{II}}(\mu, [b])$. This proves part (1) of Theorem 5.1.

By (2-1), the stabilizer in $\boldsymbol{J}_b(F)$ of any irreducible component of $X_{\underline{p}}$ is isomorphic to the parahoric subgroup $\breve{P}_{\breve{K}} \cap \boldsymbol{J}_b(F) \subseteq \boldsymbol{J}_b(F)$. By Section 5C, we have $\dim X_{\underline{p}} = d_w(b) = \dim X_\mu(b)$. By Lemma 5.4, the image of each irreducible component Z of $X_{\underline{p}}$ in $X_\mu(b)$ is an open dense subset of some top-dimensional irreducible component Y of $X_\mu(b)$. Thus the stabilizers of Z and Y coincide. By Theorem 5.3, $\breve{P}_{\breve{K}} \cap \boldsymbol{J}_b(F)$ is a very special parahoric subgroup of $\boldsymbol{J}_b(F)$. Hence, by Section 2I(a), $\breve{K} \subset \tilde{J}$ is very special with respect to $\operatorname{Ad}(\dot{x})\circ\sigma$. By Section 5C, C is an $(\operatorname{Ad}(x)\circ\sigma)$-Coxeter conjugacy class of $W_{\breve{K}}$. This proves part (2) of Theorem 5.1.

5F. ***The extreme cases.*** Let $\mathcal{T}$ be a reduction tree of $t^\mu c$. Let $[b] \in B(\boldsymbol{G})_{t^\mu c}$ and let $\underline{p}$ be a path in $\mathcal{T}$ such that $[b]_{\underline{p}} = [b]$.

If $[b] = [b_{\mu, \boldsymbol{G}\text{-indec}}]$, then

$$\ell_{\mathrm{II}}(\underline{p}) = \ell_{\mathrm{II}}(\mu, [b]) = 0.$$

Therefore $\underline{p}$ consists only of type-I edges and is unique.

If $[b]$ is basic, then

$$I(\nu_b) = \mathbb{S} \quad \text{and} \quad \ell_{\mathrm{I}}(\underline{p}) = \ell_I(\mu, [b]) = 0.$$

Therefore $\underline{p}$ consists only of type-II edges and is unique.

This proves the "moreover" part of Theorem 2.6 for these two extreme cases.

6. Some combinatorial identities

6A. ***Reduction to combinatorial identities.*** In this section, we assume that μ is essentially noncentral. Let c be a σ-Coxeter element of W such that $t^\mu c \in {}^{\mathbb{S}}\tilde{W}$. Let $\mathcal{T}$ be a reduction tree of $t^\mu c$. For any $[b] \in B(\boldsymbol{G}, \mu)_{\mathrm{indec}}$, let $n_{[b]}$ be the number of reduction paths $\underline{p}$ in $\mathcal{T}$ with $[b]_{\underline{p}} = [b]$. By Theorem 2.6(2), $n_{[b]} \geqslant 1$ for all $[b] \in B(\boldsymbol{G}, \mu)_{\mathrm{indec}}$. By Section 5F, $n_{[b]} = 1$ if $[b]$ is either the minimal or the maximal element in $B(\boldsymbol{G}, \mu)_{\mathrm{indec}}$.

Combining Theorem 5.1(1) and (2) with Proposition 3.6 and Lemma 3.8, we have

$$\boldsymbol{q}^{\langle\mu, 2\rho\rangle - \sharp(\mathbb{S}/\langle\sigma\rangle)} = \sum_{[b]\in B(\boldsymbol{G},\mu)_{\mathrm{indec}}} n_{[b]} (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(\mu,[b])} \boldsymbol{q}^{\ell_{\mathrm{II}}(\mu,[b]) + \ell_{[b]}},$$

where $\ell_{[b]} = \langle\nu_b, 2\rho\rangle + \sharp(I(\nu_b)/\langle\sigma\rangle) - \operatorname{def}(b)$ and equals $\ell(\mathcal{O})$ for any σ-Coxeter class $\mathcal{O}$ associated with $[b]$.

Note that $(\boldsymbol{q}-1)^a\boldsymbol{q}^{a'} \in \mathbb{N}[\boldsymbol{q}-1]$ for all $a, a' \in \mathbb{N}$. Thus, to show that $n_{[b]} = 1$ for all $[b]$, it suffices to show that

$$\sum_{[b]\in B(\boldsymbol{G},\mu)_{\mathrm{indec}}} (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(\mu,[b])}\boldsymbol{q}^{\ell_{\mathrm{II}}(\mu,[b])+\ell_{[b]}} = \boldsymbol{q}^{\langle\mu,2\rho\rangle-\sharp(\mathbb{S}/\langle\sigma\rangle)}. \tag{$\spadesuit$}$$

(In fact, it is enough to prove the inequality $\geqslant$.)

Using Section 2E(a), one computes that

$$\ell_{\mathrm{II}}(\mu,[b]) + \ell_{[b]} - (\langle\mu,2\rho\rangle - \sharp(\mathbb{S}/\langle\sigma\rangle)) = \sharp(I(\nu_b)/\langle\sigma\rangle) - \mathrm{length}([b],[t^\mu]).$$

Thus, ($\spadesuit$) is equivalent to

$$\sum_{[b]\in B(\boldsymbol{G},\mu)_{\mathrm{indec}}} (\boldsymbol{q}-1)^{\sharp(\mathbb{S}/\langle\sigma\rangle)-\sharp(I(\nu_b)/\langle\sigma\rangle)}\boldsymbol{q}^{\sharp(I(\nu_b)/\langle\sigma\rangle)-\mathrm{length}([b],[t^\mu])} = 1. \tag{$\spadesuit'$}$$

6B. ***Reduction to unramified adjoint groups.*** Let $\boldsymbol{G}_{\mathrm{ad}}$ be the adjoint group of $\boldsymbol{G}$ and let T_{ad} be the image of T in $\boldsymbol{G}_{\mathrm{ad}}$. We denote by μ_{ad} the image of μ in $X_*(T_{\mathrm{ad}})_{\Gamma_0}$. For any $b \in \breve{G}$, we denote by b_{ad} its image in $\breve{G}_{\mathrm{ad}}$. By [22, Proposition 4.10], the map $\boldsymbol{G} \to \boldsymbol{G}_{\mathrm{ad}}$ identifies the reduced root system of $\boldsymbol{G}_{\mathrm{ad}}$ with that of $\boldsymbol{G}$ and induces an isomorphism of posets

$$B(\boldsymbol{G},\mu)_{\mathrm{indec}} \cong B(\boldsymbol{G}_{\mathrm{ad}},\mu_{\mathrm{ad}})_{\mathrm{indec}}, \quad [b] \mapsto [b_{\mathrm{ad}}].$$

Therefore, ($\spadesuit'$) for $\boldsymbol{G}$ is equivalent to that for $\boldsymbol{G}_{\mathrm{ad}}$. We can therefore assume that $\boldsymbol{G}$ is adjoint. In this case, it is convenient to work with the reduced root system Φ of $\boldsymbol{G}$. Define

$$B(\Phi,\sigma,\mu) = \{\nu \in (V^+)^\sigma \mid \langle\mu^\diamond - \nu, \omega_{\mathcal{O}_i}\rangle \in \mathbb{Z}_{\geqslant 0} \text{ for any } i \in \mathbb{S} - I(\nu)\},$$

where $\mathcal{O}_i$ denotes the σ-orbit of i. By [19, Lemma 3.5], the map $[b] \mapsto \nu_b$ identifies $B(\boldsymbol{G},\mu)$ with $B(\Phi,\sigma,\mu)$ as posets. For any $\nu \in B(\Phi,\sigma,\mu)$, we set $\mathrm{length}(\nu,\mu^\diamond) = \sum_{\mathcal{O}\in\mathbb{S}/\langle\sigma\rangle}\lceil\langle\mu^\diamond-\nu,\omega_{\mathcal{O}}\rangle\rceil$. Then by Section 2B(a),

$$\mathrm{length}([b],[t^\mu]) = \mathrm{length}(\nu_b,\mu^\diamond) \quad \text{for any } [b] \in B(\boldsymbol{G},\mu).$$

We set

$$f_{\Phi,\sigma,\mu}(\nu) = (\boldsymbol{q}-1)^{\sharp(\mathbb{S}/\langle\sigma\rangle)-\sharp(I(\nu)/\langle\sigma\rangle)}\boldsymbol{q}^{\sharp(I(\nu)/\langle\sigma\rangle)-\mathrm{length}(\nu,\mu^\diamond)}.$$

Now ($\spadesuit'$) can be reformulated in a purely combinatorial way as

$$\sum_{\nu\in B(\Phi,\sigma,\mu)_{\mathrm{indec}}} f_{\Phi,\sigma,\mu}(\nu) = 1. \tag{$\spadesuit''$}$$

As any triple (Φ,σ,μ) arises from an unramified group, it suffices to prove ($\spadesuit''$) for the triples (Φ,σ,μ) arising from unramified adjoint groups. In the rest of this section, we assume that $\boldsymbol{G}$ is an unramified adjoint group.

6C. ***Reduction to F-simple groups.*** Write $\boldsymbol{G} = \boldsymbol{G}_1 \times \cdots \times \boldsymbol{G}_l$, where $\boldsymbol{G}_i$ are F-simple adjoint groups. Write $\mu = (\mu_1, \dots, \mu_l)$, where μ_i is a dominant coweight of $\boldsymbol{G}_i$. Also we have $\Phi = \Phi_1 \sqcup \cdots \sqcup \Phi_l$, where Φ_i is the root system of $\boldsymbol{G}_i$. It is easy to see that

$$B(\Phi, \sigma, \mu)_{\text{indec}} = B(\Phi_1, \sigma_1, \mu_1)_{\text{indec}} \times \cdots \times B(\Phi_l, \sigma_l, \mu_l)_{\text{indec}},$$

where σ_i is the restriction of σ on Φ_i. It is clear that $f_{\Phi,\sigma,\mu} = \prod_{i=1}^{l} f_{\Phi_i,\sigma_i,\mu_i}$. In the rest of this section, we assume that $\boldsymbol{G}$ is an F-simple unramified adjoint group.

6D. ***Reduction to $\breve{F}$-simple groups.*** We have $\Phi = \Phi_1 \sqcup \cdots \sqcup \Phi_l$, where $\Phi_1 \cong \cdots \cong \Phi_l$ are irreducible root systems and σ induces an isomorphism from Φ_i to Φ_{i+1}. Here, by convention, we set $\Phi_{l+1} = \Phi_1$. Then the map

$$\nu = (\nu_1, \dots, \nu_l) \mapsto |\nu| = \nu_l + \sigma(\nu_{l-1}) + \cdots + \sigma^{l-1}(\nu_1)$$

induces an isomorphism of posets

$$B(\Phi, \sigma, \mu)_{\text{indec}} \xrightarrow{\sim} B(\Phi_l, \sigma^l, |\mu|)_{\text{indec}}.$$

There is a natural bijection $\mathbb{S}/\langle\sigma\rangle \cong \mathbb{S}_l/\langle\sigma^l\rangle$, where $\mathbb{S}_l$ is the set of simple reflections for Φ_l. Thus $f_{\Phi,\sigma,\mu} = f_{\Phi_l,\sigma^l,|\mu|}$. In the rest of this section, we assume that $\boldsymbol{G}$ is an $\breve{F}$-simple unramified adjoint group.

6E. ***Reduction to split groups.*** Let $\mathcal{O}$ be a σ-orbit of $\mathbb{S}$. If all the simple roots in $\mathcal{O}$ commute with each other, we define $\alpha'_{\mathcal{O}} = \sum_{i\in\mathcal{O}} \alpha_i$. If $\mathcal{O} = \{i_0, j_0\}$ with $\langle\alpha^\vee_{j_0}, \alpha_{i_0}\rangle = \langle\alpha^\vee_{i_0}, \alpha_{j_0}\rangle = -1$, we define $\alpha'_{\mathcal{O}} = 2(\alpha_{i_0} + \alpha_{j_0})$. Let $\mathbb{S}' = \mathbb{S}/\langle\sigma\rangle$ and let Φ' be the root system generated by $\alpha'_{\mathcal{O}}$ for all $\mathcal{O} \in \mathbb{S}'$. The coroot corresponding to $\mathcal{O}$ is given by $\alpha'^\vee_{\mathcal{O}} = \frac{1}{\sharp\mathcal{O}} \sum_{i\in\mathcal{O}} \alpha_i^\vee$, and the fundamental weight corresponding to $\mathcal{O}$ is given by $\omega'_{\mathcal{O}} = \sum_{i\in\mathcal{O}} \omega_i$. For any $\nu \in (V^+)^\sigma$, $\nu \in B(\Phi, \sigma, \mu)$ if and only if $\langle\mu^\diamond - \nu, \omega'_{\mathcal{O}}\rangle \in \mathbb{Z}_{\geqslant 0}$ for any $\mathcal{O} \in \mathbb{S}/\langle\sigma\rangle$ such that $\langle\nu, \alpha'_{\mathcal{O}}\rangle \neq 0$, which is also equivalent to $\nu \in B(\Phi', \text{id}, \mu^\diamond)$. Hence we have the following:

(a) The natural identification $(\mathbb{R}\Phi^\vee)^\sigma = \mathbb{R}\Phi'^\vee$ induces an bijection of posets

$$B(\Phi, \sigma, \mu)_{\text{indec}} \xrightarrow{\sim} B(\Phi', \text{id}, \mu^\diamond)_{\text{indec}}.$$

It follows from (a) that $f_{\Phi,\sigma,\mu} = f_{\Phi',\text{id},\mu^\diamond}$. Therefore, it suffices to prove (♠″) for the triples (Φ, σ, μ) arising from split groups. In the rest of this section, we assume that $\boldsymbol{G}$ is a split $\breve{F}$-simple adjoint group. We identify $B(\boldsymbol{G}, \mu)$ with $B(\Phi, \sigma, \mu)$ and write $f_{\boldsymbol{G},\mu}(\nu) = f_{\Phi,\sigma,\mu}(\nu)$.

6F. ***Reduction to simply laced groups.*** By Section 6, (♠) is equivalent to the condition that in some (or, equivalently, any) reduction tree of $t^\mu c$, there exists only one reduction path whose end point is associated with a given $[b] \in B(\boldsymbol{G}, \mu)_{\text{indec}}$.

There exists an irreducible, simply laced, extended affine Weyl group $(\tilde{W}', \tilde{\mathbb{S}}')$ of adjoint type and a length-preserving automorphism δ on $\tilde{W}'$ such that $\tilde{W} = (\tilde{W}')^\delta$. We have a natural bijection between $\tilde{\mathbb{S}}$ and $\tilde{\mathbb{S}}'/\langle\delta\rangle$. We may assume that the simple reflections in each δ-orbit in $\tilde{\mathbb{S}}'$ commute. More explicitly,

- if $\tilde{W}$ is of type $\tilde{B}_n$, then we take $\tilde{W}'$ to be of type $\tilde{D}_{n+1}$ and δ is of order 2;

- if $\tilde{W}$ is of type $\tilde{C}_n$, then we take $\tilde{W}'$ to be of type $\tilde{A}_{2n-1}$ and δ is of order 2;
- if $\tilde{W}$ is of type $\tilde{F}_4$, then we take $\tilde{W}'$ to be of type $\tilde{E}_6$ and δ is of order 2;
- if $\tilde{W}$ is of type $\tilde{G}_2$, then we take $\tilde{W}'$ to be of type $\tilde{D}_4$ and δ is of order 3.

Let $\iota : \tilde{W} \to \tilde{W}'$ be the natural embedding. For each $i \in \tilde{\mathbb{S}}$, we have $\iota(s_i) = s_{i'_1} \cdots s_{i'_k}$, where $i'_1, \ldots, i'_k$ are the δ-orbits of i in $\tilde{\mathbb{S}}'$. Let $w \in \tilde{W}$ and $w \rightharpoonup s_i w$ be a type-I reduction edge (see Section 3D). Then one can construct a k-step reduction path

$$\iota(w) \rightharpoonup s_{i'_k}\iota(w) \rightharpoonup s_{i'_{k-1}} s_{i'_k}\iota(w) \rightharpoonup \cdots \rightharpoonup s_{i'_1} \cdots s_{i'_k}\iota(w) = \iota(s_i w)$$

in $\tilde{W}'$. Similarly, a type-II reduction edge $w \rightharpoonup s_i w s_i$ corresponds to a k-step reduction path from $\iota(w)$ to $\iota(sws)$, whose edges are all of type II. Now considering a reduction tree Γ of w, we can construct a reduction tree $\mathcal{T}'$ of $\iota(w)$ such that $\mathcal{T}$ can be viewed as a subtree of $\mathcal{T}'$ in the above way. Hence the multiplicity-one result of $\iota(t^\mu c) \in \tilde{W}'$ implies the multiplicity-one result of $t^\mu c \in \tilde{W}$.

In the rest of this section, we assume that $\boldsymbol{G}$ is a split $\breve{F}$-simple simply laced adjoint group. We will then reduce to the case where μ is a fundamental coweight. We first need a combinatorial identity on finite graphs.

6G. ***A combinatorial identity on graphs.*** Let X be a finite graph and $Y \subseteq X$. Denote by $\mathcal{A}(Y, X)$ the set of subsets $J \subseteq X$ such that none of the connected components of $X - J$ lies in Y. Define

$$f_{Y,X} = \sum_{J \in \mathcal{A}(Y,X)} (\boldsymbol{q} - 1)^{\sharp(J - Y \cap J^\circ)} \boldsymbol{q}^{\sharp(Y \cap J^\circ)} \in \mathbb{Z}[\boldsymbol{q}],$$

where $J^\circ = \{i \in J \mid i \text{ has no neighbors in } X - J\}$ is the interior of J.

Lemma 6.1. *We have $f_{Y,X} = \boldsymbol{q}^{\sharp X}$ for any $Y \subseteq X$.*

Proof. Define

$$\alpha : \{(J, K) \mid J \in \mathcal{A}(Y, X),\ K \subseteq Y \cap J^\circ\} \to \{\text{subsets of } X\}, \quad (J, K) \mapsto J - K.$$

We construct the inverse map β of α as follows. Let $H \subseteq X$. Let C be the union of connected components of $X - H$ that are contained in Y. Then we define $\beta(H) = (H \sqcup C, C)$. By definition, $\alpha \circ \beta = \mathrm{id}$. On the other hand, for any $J \in \mathcal{A}(Y, X)$ and $K \subseteq Y \cap J^\circ$, $\alpha((J, K)) = J - K$. Moreover, $X - (J - K) = (X - J) \sqcup K$. As $K \subseteq Y \cap J^\circ$, K and $X - J$ are not connected with each other. Hence K is the union of connected components of $X - (J - K)$ contained in Y. Therefore $\beta \circ \alpha((J, K)) = \beta(J - K) = ((J - K) \sqcup K, K) = (J, K)$ and hence $\beta \circ \alpha = \mathrm{id}$. Therefore α is a bijection. Using the binomial expansion, we get

$$\begin{aligned} f_{Y,X} &= \sum_{J \in \mathcal{A}(Y,X)} (\boldsymbol{q} - 1)^{\sharp(J - Y \cap J^\circ)} \boldsymbol{q}^{\sharp(Y \cap J^\circ)} = \sum_{J \in \mathcal{A}(Y,X)} (\boldsymbol{q} - 1)^{\sharp(J - Y \cap J^\circ)} \sum_{K \subseteq Y \cap J^\circ} (\boldsymbol{q} - 1)^{\sharp(Y \cap J^\circ - K)} \\ &= \sum_{J \in \mathcal{A}(X,Y)} \sum_{K \subseteq Y \cap J^\circ} (\boldsymbol{q} - 1)^{\sharp(J - K)} = \sum_{H \subseteq X} (\boldsymbol{q} - 1)^{\sharp H} = \boldsymbol{q}^{\sharp X}. \end{aligned}$$

□

6H. *Reduction to fundamental coweights.* Assume that μ is not a fundamental coweight. Then there exist $i, j \in \mathbb{S}$ (here i and j are not necessarily distinct) such that $\mu - \omega_i^\vee - \omega_j^\vee$ is also dominant. Let X be the (unique) shortest path in the Dynkin diagram of $\mathbb{S}$ with end points i, j.

Let $\lambda = \mu - \sum_{k \in X} \alpha_k^\vee$. Set

$$Y = \{i \in X \cap I(\lambda) \mid i \text{ has no neighbors in } \mathbb{S} - X\}.$$

Let $\mathcal{A}$ be the set of subsets $J \subseteq X$ such that λ is noncentral on each connected component of $\mathbb{S} - J$.

Lemma 6.2. *We have*

(1) λ *is dominant*;

(2) $B(\boldsymbol{G}, \mu)_{\text{indec}} = \bigsqcup_{J \in \mathcal{A}} B(\boldsymbol{G}, \lambda)_{(\mathbb{S}-J)\text{-irr}}$;

(3) $\mathcal{A} = \mathcal{A}(Y, X)$, *where* $\mathcal{A}(Y, X)$ *is defined as in Section 6G.*

Proof. Let $l \in \mathbb{S}$. If $l \in \mathbb{S} - X$; then $\langle \lambda, \alpha_l \rangle \geqslant -\langle \sum_{k \in X} \alpha_k^\vee, \alpha_l \rangle \geqslant 0$. If $l \in X - \{i, j\}$, then $\langle \lambda, \alpha_l \rangle = -\langle \sum_{k \in X} \alpha_k^\vee, \alpha_l \rangle = -\langle \alpha_k^\vee + \alpha_l^\vee + \alpha_{k'}^\vee, \alpha_l \rangle \geqslant -(-1+2-1) = 0$, where k and k' are the neighbors of l in X. If $l \in \{i, j\}$ and $i \neq j$, then $\langle \lambda, \alpha_l \rangle = \langle \mu, \alpha_l \rangle - \langle \sum_{k \in X} \alpha_k^\vee, \alpha_l \rangle \geqslant 1 - 1 = 0$. If $l = i = j$, then $\langle \lambda, \alpha_l \rangle = \langle \mu, \alpha_l \rangle - \langle \sum_{k \in X} \alpha_k^\vee, \alpha_l \rangle \geqslant 2 - 2 = 0$. This proves part (1).

By definition, $B(\boldsymbol{G}, \lambda)_{(\mathbb{S}-J)\text{-irr}} \subseteq B(\boldsymbol{G}, \mu)_{\text{indec}}$ for any $J \subseteq X$. Now we prove the other direction. Let $E_0 = \{e_i; i \in \mathbb{S}\}$ be as in Proposition 4.1. Then, for each $i \in \mathbb{S}$, by the definition of λ we have

$$\langle \lambda, \omega_i \rangle \geqslant \max\{0, \langle \mu, \omega_i \rangle - 1\} = \langle e_i, \omega_i \rangle.$$

Hence $\lambda \in C_{\geqslant E_0}$, and it follows from Proposition 4.1 that $\lambda \geqslant \min C_{\geqslant E_0} = [b]_{\mu, \boldsymbol{G}\text{-indec}}$.

For any $\nu \in B(\boldsymbol{G}, \mu)_{\text{indec}}$, there exists a unique subset $J \subseteq X$ such that $\lambda - \nu \in \sum_{l \in \mathbb{S}-J} \mathbb{R}_{>0} \alpha_l^\vee$. By Lemma 2.1, $\nu \in B(\boldsymbol{G}, \lambda)_{(\mathbb{S}-J)\text{-irr}}$. Part (2) is proved.

To prove (3) we first claim that

(a) λ is noncentral on each connected component of $\mathbb{S} - X$.

Let H be a connected component of $\mathbb{S} - X$. Choose $a \in H$, $b \in X$ and let $a = i_0, i_1, \dots, i_n = b$ be the shortest path in the Dynkin diagram of $\mathbb{S}$. Then there exists $1 \leqslant m \leqslant n$ such that $i_m \in X$ and $i_0, \dots, i_{m-1} \in \mathbb{S} - X$. As H is a connected component of $\mathbb{S} - X$ containing i_0, it follows that $i_0, i_1, \dots, i_{m-1} \in H$. Since $\langle -\alpha_m^\vee, \alpha_{m-1} \rangle > 0$, we deduce that $\lambda = \mu - \sum_{k \in X} \alpha_k^\vee$ is strictly dominant on H. The claim (a) is proved.

Let $J \subseteq X$ and A be a connected component of $\mathbb{S} - J$. Note that $\mathbb{S} - J = (\mathbb{S} - X) \sqcup (X - J)$. Hence $A = A_1 \sqcup A_2$, where A_1 (resp. A_2) is a union of connected components of $\mathbb{S} - X$ (resp. $X - J$). In view of (a), λ is central on A if and only if $A_1 = \varnothing$ and λ is central on A_2, i.e., $A = A_2$ is a connected component of $X - J$ contained in Y. On the other hand, any connected component of $X - J$ contained in Y is also a connected component of $\mathbb{S} - J$. Therefore, $\mathcal{A} = \mathcal{A}(Y, X)$ and part (3) is proved. □

Proposition 6.3. *Suppose that* (♠″) *holds for all fundamental coweights. Then it holds for all dominant coweights.*

Proof. We argue by induction on the semisimple rank of $\boldsymbol{G}$ and the number $\langle \rho, \mu \rangle$.

Suppose μ is not fundamental. Let the notation be as in Section 6H. By Corollary 2.2, we identify $B(\boldsymbol{G}, \lambda)_{(\mathbb{S}-J)\text{-irr}}$ with $B(\boldsymbol{M}_{\mathbb{S}-J}, \mu)_{\text{irr}}$ for $J \in \mathcal{A}$. We show that

(a) for any $J \in \mathcal{A}$ and $\nu \in B(\boldsymbol{M}_{\mathbb{S}-J}, \mu)_{\text{irr}}$, we have

$$f_{\boldsymbol{G},\mu}(\nu) = \boldsymbol{q}^{-\sharp X} (\boldsymbol{q}-1)^{\sharp (J - Y \cap J^\circ)} \boldsymbol{q}^{\sharp (Y \cap J^\circ)} f_{\boldsymbol{M}_{\mathbb{S}-J},\lambda}(\nu).$$

By definition, $f_{\boldsymbol{M}_K,\mu}(\nu) = (\boldsymbol{q}-1)^{\sharp(K - K\cap I(\nu))} \boldsymbol{q}^{\sharp K \cap I(\nu) - \text{length}_{\boldsymbol{M}_K}(\mu,\nu)}$ for any $K \subseteq \mathbb{S}$.

Note that $\text{length}_{\boldsymbol{G}}(\lambda, \mu) = \langle \mu - \lambda, \rho \rangle = \sharp X$, and hence

$$\text{length}_{\boldsymbol{G}}(\nu, \mu) = \text{length}_{\boldsymbol{G}}(\nu, \lambda) + \text{length}_{\boldsymbol{G}}(\lambda, \mu) = \text{length}_{\boldsymbol{G}}(\nu, \lambda) + \sharp X.$$

By Lemma 2.1, we have $\text{length}_{\boldsymbol{G}}(\nu, \mu) = \text{length}_{\boldsymbol{M}}(\nu, \mu)$. To show (a), it remains to show that

$$I(\nu) - (\mathbb{S} - J) \cap I(\nu) = J \cap I(\nu) = Y \cap J^\circ.$$

Let $l \in J$. Write $\nu = \lambda - \delta$ for some $\delta \in \sum_{k \in \mathbb{S}-J} \mathbb{R}_{>0} \alpha_k^\vee$. Then $l \in J \cap I(\nu)$ if and only if $\langle \nu, \alpha \rangle = \langle \lambda, \alpha_l \rangle - \langle \delta, \alpha_l \rangle = 0$, which is equivalent to $\langle \lambda, \alpha_l \rangle = \langle \delta, \alpha_l \rangle = 0$, that is, $l \in Y \cap J^\circ$ as desired. Hence (a) is proved.

Now, by Lemma 6.1 and the inductive hypothesis on $\boldsymbol{M}_{\mathbb{S}-J}$ and λ, we have

$$\begin{aligned}
\sum_{\nu \in B(\boldsymbol{G},\mu)_{\text{indec}}} f_{\boldsymbol{G},\mu}(\nu) &= \sum_{J \in \mathcal{A}} \sum_{\nu \in B(\boldsymbol{M}_{\mathbb{S}-J},\lambda)_{\text{irr}}} f_{\boldsymbol{G},\mu}(\nu) \\
&= \boldsymbol{q}^{-\sharp X} \sum_{J \in \mathcal{A}} \sum_{\nu \in B(\boldsymbol{M}_{\mathbb{S}-J},\lambda)_{\text{irr}}} (\boldsymbol{q}-1)^{\sharp(J - Y \cap J^\circ)} \boldsymbol{q}^{\sharp (Y \cap J^\circ)} f_{\boldsymbol{M}_{\mathbb{S}-J},\lambda}(\nu) \\
&= \boldsymbol{q}^{-\sharp X} \sum_{J \in \mathcal{A}} (\boldsymbol{q}-1)^{\sharp(J - Y \cap J^\circ)} \boldsymbol{q}^{\sharp (Y \cap J^\circ)} = 1. \qquad \square
\end{aligned}$$

Now it is sufficient to deal with the fundamental coweights of $\boldsymbol{G}$.

6I. ***Proof for the minuscule coweights.*** Let μ a be (nonzero) minuscule coweight. Then $\dim V_\mu(\lambda_b) = 1$ for any $[b] \in B(\boldsymbol{G}, \mu)$. By Theorem 5.3, we conclude that $\sharp(\boldsymbol{J}_b(F) \backslash \Sigma^{\text{top}}(X_\mu(b))) = 1$ for any $[b] \in B(\boldsymbol{G}, \mu)$. As in Section 6A, we have

$$\boldsymbol{q}^{\langle \mu, 2\rho \rangle - \sharp(\mathbb{S}/\langle \sigma \rangle)} = \sum_{[b] \in B(\boldsymbol{G},\mu)_{\text{indec}}} n_{[b]} (\boldsymbol{q}-1)^{\ell_{\text{I}}(\mu,[b])} \boldsymbol{q}^{\ell_{\text{II}}(\mu,[b]) + \ell_{[b]}},$$

where $n_{[b]}$ is the number of reduction paths $\underline{p}$ in a given reduction tree $\mathcal{T}$ of $t^\mu c$ with $[b]_{\underline{p}} = [b]$. At the end of Section 5C, we showed that $\dim X_{\underline{p}} = \dim X_{t^\mu c}([b]_{\underline{p}}) = \dim X_\mu([b]_{\underline{p}})$ for any reduction path $\underline{p}$. Using Lemma 5.4, we conclude that all $n_{[b]} = 1$. This implies the combinatorial identity (♠), and then (♠″) follows.

In particular, the combinatorial identity (♠″) holds for type A, since all the fundamental coweights are minuscule. For $(A_{n-1}, \omega_i^\vee)$, we may write (♠) explicitly as

$$\sum_{\substack{k\geqslant 1, 1>a_1/b_1>\cdots>a_k/b_k>0;\\ a_i+\cdots+a_k=i, b_1+\cdots+b_k=n}} (\boldsymbol{q}-1)^{k-1}\boldsymbol{q}^{1-k+\left(\sum_{1\leqslant l_1<l_2\leqslant k}(a_{l_1}b_{l_2}-a_{l_2}b_{l_1})+\sum_{1\leqslant l\leqslant k}\gcd(a_l,b_l)\right)/2} = \boldsymbol{q}^{(i(n-i)-n)/2+1}.$$

We do not know if there is a purely combinatorial proof of this identity.

6J. ***Type-D_n case.*** In this subsection, we assume that $\boldsymbol{G}$ is of type D_n $(n\geqslant 4)$. Note that the fundamental coweights $\omega_1^\vee, \omega_{n-1}^\vee, \omega_n^\vee$ are minuscule and have already been dealt with in Section 6I. Here we deal only with $\omega_i^\vee$ for $2\leq i\leq n-2$.

For any integer k, denote by $[1,k]$ the set $\{m\in\mathbb{Z};\ 1\leqslant m\leqslant k\}$. Set $\omega_0^\vee=0$. Since $\omega_{i-2}^\vee<\omega_i^\vee$, we have a natural embedding $B(\boldsymbol{G},\omega_{i-2}^\vee)\to B(\boldsymbol{G},\omega_i^\vee)$. For $i\geqslant 3$, Set

$$B_{\mathrm{I}} = \bigsqcup_{i-2\leqslant k\leqslant n-3} B(\boldsymbol{G},\omega_{i-2}^\vee)_{[1,k]\text{-irr}} \cong \bigsqcup_{i-2\leqslant k\leqslant n-3} B(\boldsymbol{M}_{[1,k]},\omega_{i-2}^\vee)_{\mathrm{irr}},$$
$$B_{\mathrm{II}} = \bigsqcup_{J\subseteq\{n-1,n\}} B(\boldsymbol{G},\omega_{i-2}^\vee)_{(\mathbb{S}-J)\text{-irr}} \cong \bigsqcup_{J\subseteq\{n-1,n\}} B(\boldsymbol{M}_{\mathbb{S}-J},\omega_{i-2}^\vee)_{\mathrm{irr}}.$$

For $i=2$, set $B_{\mathrm{I}}=\{0\}$ and $B_{\mathrm{II}}=\varnothing$.

For $i\leqslant k\leqslant n-2$, the adjoint group of $\boldsymbol{M}_{\mathbb{S}-\{k\}}$ is of type $A_{k-1}\times D_{n-k}$. Here, by convention, type D_3 is the same as type A_3, and type D_2 is the same as type $A_1\times A_1$. Set $\mu_k=(1^{i-1},0^{k-i+1},1,0^{n-k-1})$. Then μ and $\omega_i^\vee$ are in the same W-orbit. The restriction of μ_k to $\boldsymbol{M}_{\mathbb{S}-\{k\}}$ is (dominant) minuscule, and its projection to the adjoint group of $\boldsymbol{M}_{\mathbb{S}-\{k\}}$ is the coweight $(\omega_{i-1}^\vee,\omega_1^\vee)$ if $k<n-2$ and $(\omega_{i-1}^\vee,\omega_1^\vee,\omega_1^\vee)$ if $k=n-2$. As in Section 2C, we identify $B(\boldsymbol{M}_{\mathbb{S}-\{k\}},\mu_k)$ with its natural image in $B(\boldsymbol{G},\omega_i)$. Set

$$B_{\mathrm{III}} = \bigsqcup_{i\leqslant k\leqslant n-2} B(\boldsymbol{G},\mu_k)_{(\mathbb{S}-\{k\})\text{-irr}} \cong \bigsqcup_{i\leqslant k\leqslant n-2} B(\boldsymbol{M}_{\mathbb{S}-\{k\}},\mu_k)_{\mathrm{irr}}.$$

By direct computation, B_{I}, B_{II}, and B_{III} are disjoint in $B(\boldsymbol{G},\omega_i^\vee)$.

We give an example to illustrate the subsets $B_{\mathrm{I}}, B_{\mathrm{II}}, B_{\mathrm{III}}$. Let $\boldsymbol{G}$ be of type D_6 and $\mu=\omega_3^\vee$. We have

$$B_{\mathrm{I}} = \left\{\left(\tfrac{1}{2},\tfrac{1}{2},0,0,0,0\right),\left(\tfrac{1}{3},\tfrac{1}{3},\tfrac{1}{3},0,0,0\right),\left(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{4},0,0\right)\right\};$$
$$B_{\mathrm{II}} = \left\{\left(\tfrac{1}{5},\tfrac{1}{5},\tfrac{1}{5},\tfrac{1}{5},\tfrac{1}{5},0\right),\left(\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\right),\left(\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},-\tfrac{1}{6},\right),\left(0,0,0,0,0,0\right)\right\};$$
$$B_{\mathrm{III}} = \left\{\left(\tfrac{2}{3},\tfrac{2}{3},\tfrac{2}{3},0,0,0\right),\left(\tfrac{2}{4},\tfrac{2}{4},\tfrac{2}{4},\tfrac{2}{4},0,0\right)\right)\right\}.$$

In this case, one may check directly that $B(\boldsymbol{G},\omega_3^\vee)_{\mathrm{indec}}=B_{\mathrm{I}}\sqcup B_{\mathrm{II}}\sqcup B_{\mathrm{III}}$.

Now we come back to the general situation. Intuitively, when $3\leqslant i\leqslant n-2$, B_{III} consists of Newton vectors whose coordinates sum up to $i-1$, B_{I} consists of Newton vectors with the last two coordinates 0 and all the coordinates sum up to $i-2$, B_{II} consists of the rest of elements. This partition makes computation of the sum $\sum f_{\boldsymbol{G},\omega_i^\vee}(\nu)$ easier.

By Section 6A, we have

$$\sum_{\nu\in B_{\mathrm{I}}\sqcup B_{\mathrm{II}}\sqcup B_{\mathrm{III}}} f_{\boldsymbol{G},\omega_i^\vee}(\nu) \leqslant \sum_{\nu\in B(\boldsymbol{G},\omega_i^\vee)_{\mathrm{indec}}} f_{\boldsymbol{G},\omega_i^\vee}(\nu) \leqslant 1.$$

In the rest of this section, we will show that

$$\sum_{\nu\in B_{\mathrm{I}}\sqcup B_{\mathrm{II}}\sqcup B_{\mathrm{III}}} f_{\boldsymbol{G},\omega_i^\vee}(\nu) = 1. \tag{$**$}$$

The equality $(**)$ will, in particular, imply that $B_{\mathrm{I}} \sqcup B_I \sqcup B_{\mathrm{III}} = B(\boldsymbol{G}, \omega_i^\vee)_{\mathrm{indec}}$.

It can be checked directly that $(**)$ holds for D_4. Using induction, we may assume that $(**)$ holds for groups of type D with semisimple rank less than n. Note also that $(\spadesuit'')$ for type A has already been proved in Section 6I. Therefore we have

(a) $\sum_{\nu\in B_{\mathrm{II}}} f_{\boldsymbol{G},\omega_i^\vee}(\nu) = \boldsymbol{q}^{-\langle\alpha^\vee,\rho\rangle}\cdot(1+2(\boldsymbol{q}-1)^2+(\boldsymbol{q}-1)^2) = \boldsymbol{q}^{-\langle\alpha^\vee,\rho\rangle}\cdot \boldsymbol{q}^2$.

Next we handle B_{I} and B_{III}. Let

$$\alpha^\vee = \omega_i^\vee - \omega_{i-2}^\vee = \alpha_{i-1}^\vee + 2\alpha_i^\vee + \cdots + 2\alpha_{n-2}^\vee + \alpha_{n-1}^\vee + \alpha_n^\vee.$$

Note that $\langle\alpha^\vee, \rho\rangle = 2n - 2i + 1$. We claim that:

(b) For $k \in [i-2, n-3]$ and $\nu \in B(\boldsymbol{M}_{[1,k]}, \omega_{i-2}^\vee)_{\mathrm{irr}}$, we have

$$f_{\boldsymbol{G},\omega_i^\vee}(\nu) = (\boldsymbol{q}-1)\cdot \boldsymbol{q}^{-\langle\alpha^\vee,\rho\rangle+n-k-1}\cdot f_{\boldsymbol{G},\omega_{i-2}^\vee}(\nu).$$

(c) For $k \in [i, n-2]$ and $\nu \in B(\boldsymbol{M}_{\mathbb{S}-\{k\}}, \mu_k)_{\mathrm{irr}}$, we have

$$f_{\boldsymbol{G},\omega_i^\vee}(\nu) = (\boldsymbol{q}-1)\cdot \boldsymbol{q}^{-\langle\alpha^\vee,\rho\rangle+2n-k-i}\cdot f_{\boldsymbol{M}_{\mathbb{S}-\{k\}},\mu_k}(\nu).$$

We prove (c) here. The proof of (b) is similar.

By definition, we have

$$f_{\boldsymbol{G},\omega_i^\vee}(\nu) = (\boldsymbol{q}-1)\cdot \boldsymbol{q}^{n-k-n_0}\cdot f_{\boldsymbol{M}_{\mathbb{S}-\{k\}},\mu_k}(\nu),$$

where $n_0 = \mathrm{length}_{\boldsymbol{G}}(\nu, \omega_i^\vee) - \mathrm{length}_{\boldsymbol{M}_{\boldsymbol{G}-\{k\}}}(\nu, \mu_k)$. It can be checked directly that

$$\langle \mu_k - \nu, \rho_{\boldsymbol{M}_{\mathbb{S}-\{k\}}}\rangle = \langle \omega_{i-1}^\vee - \nu, \rho\rangle + \langle (0^k, 1, 0^{n-k-1}), \rho_{\boldsymbol{M}_{[k+1,n]}}\rangle = \langle \omega_{i-1}^\vee - \nu, \rho\rangle + 1.$$

Note also that $\mathrm{def}_{\boldsymbol{G}}(\nu) = \mathrm{def}_{\boldsymbol{M}_{\mathbb{S}-\{k\}}}(\nu)$. Therefore, by the length formula, we have

$$\begin{aligned} n_0 &= \langle \omega_i^\vee - \nu, \rho\rangle - \langle \mu_k - \nu, \rho_{\boldsymbol{M}_{\mathbb{S}-\{k\}}}\rangle + \tfrac{1}{2}\left(\mathrm{def}_{D_n}(\nu) - \mathrm{def}_{\boldsymbol{M}_{\mathbb{S}-\{k\}}}(\nu)\right) \\ &= \langle \omega_i^\vee - \omega_{i-1}^\vee, \rho\rangle + 1 = n - i + 1 = \langle \alpha^\vee, \rho\rangle - n + i. \end{aligned}$$

The statement (c) follows.

	μ	$\omega_1^\vee$	$\omega_2^\vee$	$\omega_3^\vee$	$\omega_4^\vee$	$\omega_5^\vee$	$\omega_6^\vee$	$\omega_7^\vee$	$\omega_8^\vee$
E_6		7	15	30	15				
E_7	$\sharp B(\boldsymbol{G},\mu)_{\mathrm{indec}}$	13	26	50	125	69	32		
E_8		56	126	254	729	424	220	94	27

Table 1. The numbers of elements in $B(\boldsymbol{G},\mu)_{\mathrm{indec}}$ for all the fundamental, nonminuscule coweights in type E.

Combining (a), (b), and (c) with the combinatorial identity (♠) for type A and for type D_l with $l<n$, we have

$$\sum_{\nu\in B_{\mathrm{I}}\sqcup B_{\mathrm{II}}\sqcup B_{\mathrm{III}}} f_{\boldsymbol{G},\omega_i^\vee}(\nu)=\boldsymbol{q}^{-\langle\alpha^\vee,\rho\rangle}\cdot\Bigl(\boldsymbol{q}^2+(\boldsymbol{q}-1)\Bigl(\sum_{k=i-2}^{n-3}\boldsymbol{q}^{n-k-1}+\sum_{k=i}^{n-2}\boldsymbol{q}^{2n-k-i}\Bigr)\Bigr)$$
$$=\boldsymbol{q}^{-\langle\alpha^\vee,\rho\rangle}\cdot\boldsymbol{q}^{2n-2i+1}=1.$$

This completes the proof of (∗∗).

6K. *Type-E case.* In this subsection, we assume that $\boldsymbol{G}$ is of type E_n. We verify the combinatorial identity (♠′′) by computer. Recall that a vector $v\in(V^+)^\sigma$ lies in $B(\boldsymbol{G},\mu)$ if and only if $\langle\mu-v,\omega_i\rangle\in\mathbb{Z}_{\geqslant 0}$ for any $i\in\mathbb{S}-I(v)$. As a consequence, we have the following characterization of $B(\boldsymbol{G},\mu)_{\mathrm{indec}}$:

(a) The set $B(\boldsymbol{G},\mu)_{\mathrm{indec}}$ equals the set of dominant vectors of the form

$$\nu=\mathrm{pr}_{\mathrm{I}}\bigl(\mu-\sum_{i\in I}c_i\alpha_i^\vee\bigr),$$

where I is a σ-stable subset of $\mathbb{S}$, $c_i\in\mathbb{Z}$, $1\le c_i\le\langle\mu,\omega_i\rangle$, and $\mathrm{pr}_{\mathrm{I}}:V\to\bigoplus_{i\in I}\mathbb{R}\omega_i^\vee$ is the natural orthogonal projection.

On the basis of (a), we can use a computer program to list all the elements in $B(\boldsymbol{G},\mu)_{\mathrm{indec}}$ and then verify the combinatorial identity (♠′′) directly.

In Table 1, we provide the numbers of elements in $B(\boldsymbol{G},\mu)_{\mathrm{indec}}$ for all the fundamental, nonminuscule coweights in type E. The most complicated case is E_8, and $\mu=\omega_4^\vee$, in which $B(\boldsymbol{G},\mu)_{\mathrm{indec}}$ contains 729 elements.

7. The general case

7A. *Description of the reduction trees.* Let $w\in Wt^\mu W$ with finite partial σ-Coxeter part, that is, $\eta_\sigma(w)$ is a partial σ-Coxeter element. For any $J\in[J_0(w),J(w)]_\mu$ and $[b]\in B(\boldsymbol{G},\mu)_{J\text{-irr}}$, we set

$$J^{\flat,w}=\{i\in I(\mu^\diamond)\cap(J(w)-J)\mid i\text{ commutes with }J\},$$
$$\ell_{\mathrm{I}}(w,[b],J)=\sharp(J(w)/\langle\sigma\rangle)-\sharp(J^{\flat,w}/\langle\sigma\rangle)-\sharp(I(\nu_b)\cap J/\langle\sigma\rangle),$$
$$\ell_{\mathrm{II}}(w,[b],J)=\mathrm{length}([b],[t^\mu])-\sharp(J_0(w)/\langle\sigma\rangle).$$

By Lemma 2.1, there exists a unique σ-conjugacy class $[b]_{M_J} \in B(M_J, \mu)$ such that $[b]_{M_J} \subseteq [b]$. We similarly define $[b]_{M_{J^\flat,w\cup J}} \in B(M_{J^\flat,w\cup J}, \mu)$. In this case, a σ-Coxeter element associated with $[b]_{M_{J^\flat,w\cup J}}$ is equal to the product of a σ-Coxeter element of $W_{J^\flat,w}$ and a σ-Coxeter element associated with $[b]_{M_J}$.

The main result of this section is the following description of the reduction tree of w.

Theorem 7.1. *Let $w \in Wt^\mu W$ with $\eta_\sigma(w)$ a partial σ-Coxeter element. Let $\mathcal{T}$ be a reduction tree of w. Then, for any $J \in [J_0(w), J(w)]_\mu$ and $[b] \in B(G, \mu)_{J\text{-irr}}$, there exists a unique reduction path $\underline{p}$ in $\mathcal{T}$ with $[b]_{\underline{p}} = [b]$. Moreover,*

(1) *$\ell_{\mathrm{I}}(\underline{p}) = \ell_{\mathrm{I}}(w, [b], J)$ and $\ell_{\mathrm{II}}(\underline{p}) = \ell_{\mathrm{II}}(w, [b], J)$;*

(2) *$\mathrm{end}(\underline{p})$ is a σ-Coxeter element associated with $[b]_{M_{J^\flat,w\cup J}}$.*

Combining Theorem 7.1 with Proposition 3.9 and Remark 3.10, we obtain Theorem 2.6(3) for w.

7B. *Strategy.* The strategy for proving Theorem 7.1 is very different from that adopted for the proof of Theorem 5.1. In the latter case, we used the Chen–Zhu conjecture and the dimension formula to determine the end points of the reduction trees of $t^\mu c$. However, such a method, when applied to general w, cannot determine the end points.

The approach we use here is as follows. We first apply the partial reduction method and the class polynomials for $t^\mu c$ to calculate the class polynomials for w. As we mentioned earlier, class polynomials, in general, contain less information than reduction trees. Fortunately, for the elements w we consider here, by combining the information on the class polynomials and the estimates on the type-I edges, we obtain the required information for *any* reduction tree.

The information about the class polynomials we need is contained in the following equality on the σ-cocenter of the Iwahori–Hecke algebra H:

$$T_w + [H, H]_\sigma = \sum_{\substack{J\in[J_0(w),J(w)]_\mu \\ [b]\in B(G,\mu)_{J\text{-irr}}}} (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(w,[b],J)} \boldsymbol{q}^{\ell_{\mathrm{II}}(w,[b],J)} T_{\mathcal{O}_{w,[b]}} + [H, H]_\sigma, \qquad (\diamondsuit)$$

where $\mathcal{O}_{w,[b]}$ is the σ-conjugacy class containing a σ-Coxeter element associated with $[b]_{M_{J^\flat,w\cup J}}$.

7C. *Proof of* ($\diamondsuit$).

7C1. We consider the case where $w = t^\mu c \in {}^{\mathbb{S}}\tilde{W}$ for some partial σ-Coxeter element c of W. Let $J = \mathrm{supp}_\sigma(c)$. In this case, $J_0(w) = J(w) = J$.

Suppose $J = \mathbb{S}$, that is, c is a (full) σ-Coxeter element. Let $\mathcal{T}$ be a reduction tree of w. By the description of the reduction tree in Theorem 5.1, each end point of $\mathcal{T}$ is a σ-Coxeter element associated with some $[b] \in B(G, \mu)_{\mathrm{irr}}$. Then by Lemma 3.8, we get $F_{w,\mathcal{O}} = (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(w,[b],J)} \boldsymbol{q}^{\ell_{\mathrm{II}}(w,[b],J)}$ if $\mathcal{O}$ contains an end point e of $\mathcal{T}$ with $\Phi(e) = [b]$ and $F_{w,\mathcal{O}} = 0$ otherwise. Then ($\diamondsuit$) follows.

Assume $J \subsetneq \mathbb{S}$. It follows from [18, Theorem 7.3] that $F_{w,\mathcal{O}} = \sum_{\mathcal{O}^J \subseteq \mathcal{O}} F^J_{w,\mathcal{O}^J}$, where $\mathcal{O}^J$ denotes a σ-conjugacy class of $\tilde{W}_J$ and $F^J_{w,\mathcal{O}^J}$ denotes the corresponding class polynomial for M_J. Using the description of each $F^J_{w,\mathcal{O}^J}$ we get in the $J = \mathbb{S}$ case, we conclude that $F_{w,\mathcal{O}} = (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(w,[b],J)} \boldsymbol{q}^{\ell_{\mathrm{II}}(w,[b],J)}$ if $\mathcal{O}$ contains a σ-Coxeter element associated with $[b]_{M_J}$, and $F_{w,\mathcal{O}} = 0$ otherwise. This proves ($\diamondsuit$).

7C2. We consider the case where $w = c_1 t^\mu c_2$ for some partial σ-Coxeter elements c_1, c_2 of W such that c_1 commutes with c_2, $t^\mu c_2 \in {}^{\mathbb{S}}\tilde{W}$ and $c_1 \in W_{I(t^\mu c_2)}$.

Set $J_1 = \operatorname{supp}_\sigma(c_1)$, $J_2 = \operatorname{supp}_\sigma(c_2)$ and $J = J_1 \sqcup J_2 = J(w)$. One can construct a reduction tree $\mathcal{T}$ of $t^\mu c_2$ in $\tilde{W}_{J_2}$ such that c_1 commutes with all the vertices in $\mathcal{T}$. In particular, if $w_1 \rightharpoonup w_2$ is one edge in $\mathcal{T}$, then we have $c_1 w_1 \rightharpoonup c_1 w_2$ in $\tilde{W}_J$. Note that c_1 is a minimal length element in its σ-conjugacy class in $\tilde{W}_{J_1}$, and the simple reflections in $\tilde{W}_{J_1}$ commute with the simple reflections in $\tilde{W}_{J_2}$. It is easy to see that if w_1 is an end point in $\mathcal{T}$ (and hence is a minimal length element in its σ-conjugacy class in $\tilde{W}_{J_2}$ and commutes with c_1), then $c_1 w_1$ is a minimal length element in its σ-conjugacy class in $\tilde{W}_J$.

Let $\mathcal{T}'$ be the tree with vertices $c_1 w_1$ for all vertices $w_1 \in \mathcal{T}$ and the edges $c_1 w_1 \rightharpoonup c_1 w_2$ for all edges $w_1 \rightharpoonup w_2$ in $\mathcal{T}$. Then, from the above discussion, $\mathcal{T}'$ is a reduction tree of $c_1 t^\mu c_2$ in $\tilde{W}_J$. Note that

$$\ell_{\mathrm{I}}(w, [b], J_2) = \ell_{\mathrm{I}}(t^\mu c_2, [b], J_2) \quad \text{and} \quad \ell_{\mathrm{II}}(w, [b], J_2) = \ell_{\mathrm{II}}(t^\mu c_2, [b], J_2) \qquad \text{for any } [b] \in B(\boldsymbol{G}, \mu)_{J_2-\mathrm{irr}}.$$

On the other hand, we have

$$F_{w,\mathcal{O}} = \sum_{\mathcal{O}^J \subseteq \mathcal{O}} F^J_{w,\mathcal{O}^J}.$$

Hence $(\diamond)$ for w follows from $(\diamond)$ for $t^\mu c_2$ established in Section 7C1.

7C3. We consider the case where $w = c_1 t^\mu c_2$ for partial σ-Coxeter elements c_1 and c_2 in W such that $t^\mu c_2 \in {}^{\mathbb{S}}\tilde{W}$ and $\operatorname{supp}_\sigma(c_1) \cap \operatorname{supp}_\sigma(c_2) = \varnothing$.

We prove $(\diamond)$ by induction on $\ell(c_1)$. The case $\ell(c_1) = 0$ is proved in Section 7C1.

Assume that $\ell(c_1) > 0$. Let $i \in \mathbb{S}$ such that $s_i c_1 < c_1$. There are two cases as follows.

Case (1): $c_2 \sigma(s_i) \in {}^{I(\mu)}W$. Write $w_1 = s_i w$ and $w_2 = s_i w \sigma(s_i) = s_i c_1 t^\mu c_2 \sigma(s_i)$. Then $T_w + [H, H]_\sigma = (\boldsymbol{q}-1) T_{w_1} + \boldsymbol{q} T_{w_2} + [H, H]_\sigma$. Note that $J(w_2) = J(w)$, $J_0(w_1) = J_0(w)$, $J(w_1) = J(w) - \{\sigma^\ell(s_i) \mid \ell \in \mathbb{Z}\}$, and $J_0(w_2) = J_0(w) \sqcup \{\sigma^\ell(s_i) \mid \ell \in \mathbb{Z}\}$. Then

$$[J_0(w), J(w)]_\mu = [J_0(w_1), J(w_1)]_\mu \sqcup [J_0(w_2), J(w_2)]_\mu.$$

By the induction hypothesis, it suffices to prove that

(a) if $J \in [J_0(w_1), J(w_1)]_\mu$, then $J^{\flat, w} = J^{\flat, w_1}$; and

(b) if $J \in [J_0(w_2), J(w_2)]_\mu$, then $J^{\flat, w} = J^{\flat, w_2}$.

Statement (b) is obvious since $J(w_2) = J(w)$. Let us prove (a). Suppose $J^{\flat, w} \neq J^{\flat, w_1}$; then $i \in I(\mu^\diamond)$ and s_i commutes with J. Since $J \supseteq J_0(w)$, s_i also commutes with $J_0(w)$. Then μ is not essentially noncentral over $J_0(w_2) = J_0(w) \sqcup \{\sigma^l(i) \mid l \in \mathbb{Z}\}$, which is a contradiction. This completes the proof.

Case (2): $\sigma(i) \in I(\mu)$, and $\sigma(s_i)$ commutes with c_2. Then $(\diamond)$ holds for w if and only if it holds for $s_i w \sigma(s_i)$. We continue with the procedure until case (1) happens. If case (2) happens all the time and the procedure does not stop, then $c_1 \in W_{I(t^\mu c_2)}$, and $(\diamond)$ follows from Section 7C2.

7D. ***Proof of Theorem 7.1.***

7D1. We consider the case $w = ct^{\mu}$, where c is a partial σ-Coxeter element of W. Let $\mathcal{T}$ be a reduction tree of w. By Lemma 3.8 and ($\diamondsuit$) for w, we have that, for any σ-conjugacy class $\mathcal{O}$ of $\tilde{W}$,

$$\sum_{\underline{p};\mathrm{end}(\underline{p})\in\mathcal{O}} (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(\underline{p})}\boldsymbol{q}^{\ell_{\mathrm{II}}(\underline{p})} = \begin{cases} (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(w,[b],J)}\boldsymbol{q}^{\ell_{\mathrm{II}}(w,[b],J)} & \text{if } \mathcal{O}=\mathcal{O}_{w,[b]} \text{ for some } [b], \\ 0 & \text{otherwise.} \end{cases}$$

Let $\underline{p}$ be a path in $\mathcal{T}$. Set $e = \mathrm{end}(\underline{p})$ and $[b] = [b]_{\underline{p}}$. Assume that $[b] \in B(\boldsymbol{G}, \mu)_{J\text{-irr}}$ for some J. As $(\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(\underline{p})}\boldsymbol{q}^{\ell_{\mathrm{II}}(\underline{p})} \in \mathbb{N}[\boldsymbol{q}-1]$, there is no cancellation involved in the left-hand side of the above equality. Therefore, e must be contained in $\mathcal{O}_{w,[b]}$.

As in Section 5C, we have

$$\begin{aligned}\ell_{\mathrm{I}}(\underline{p}) &\geqslant \dim V_e - \dim V_w = \sharp(J(w)/\langle\sigma\rangle) - \sharp(J^{\flat,w}/\langle\sigma\rangle) - \sharp(I(\nu_b)\cap J/\langle\sigma\rangle) \\ &= \ell_{\mathrm{I}}(w,[b],J).\end{aligned}$$

Note that $\ell_{\mathrm{I}}(\underline{p}) + 2\ell_{\mathrm{II}}(\underline{p}) = \ell_{\mathrm{I}}(w,[b],J) + 2\ell_{\mathrm{II}}(w,[b],J)$. Thus

$$\deg_{\boldsymbol{q}}(\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(\underline{p})}\boldsymbol{q}^{\ell_{\mathrm{II}}(\underline{p})} \geqslant \deg_{\boldsymbol{q}}(\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(w,[b],J)}\boldsymbol{q}^{\ell_{\mathrm{II}}(w,[b],J)},$$

with equality holding if and only if $\ell_{\mathrm{I}}(\underline{p}) = \ell_{\mathrm{I}}(w,[b],J)$. Again, since there is no cancellation involved in $\sum_{\underline{p};\mathrm{end}(\underline{p})\in\mathcal{O}}(\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(\underline{p})}\boldsymbol{q}^{\ell_{\mathrm{II}}(\underline{p})}$, we must have $\ell_{\mathrm{I}}(\underline{p}) = \ell_{\mathrm{I}}(w,[b],J)$. In this case, $(\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(\underline{p})}\boldsymbol{q}^{\ell_{\mathrm{II}}(\underline{p})} = (\boldsymbol{q}-1)^{\ell_{\mathrm{I}}(w,[b],J)}\boldsymbol{q}^{\ell_{\mathrm{II}}(w,[b],J)}$. This also shows that for each $\mathcal{O} = \mathcal{O}_{w,[b]}$, there is a unique reduction path $\underline{p}$ with $\mathrm{end}(\underline{p}) \in \mathcal{O}$.

This completes the proof of Theorem 7.1 for $w = ct^{\mu}$.

7D2. Now we consider the general case. Let $w = xt^{\mu}y$ with $t^{\mu}y \in {}^{\mathbb{S}}\tilde{W}$. Set $c = \sigma^{-1}(y)x$ and $w' = ct^{\mu}$. We relate w and w' as in the proof of [12, Theorem 10.3]. Let $y = s_1 s_2 \cdots s_r$ be a reduced expression. Let $w^{(0)} = w'$, $w^{(1)} = \sigma^{-1}(s_1)w^{(0)}s_1$, $w^{(2)} = \sigma^{-1}(s_2)w^{(1)}s_2, \dots, w^{(r)} = w$. We have $w^{(0)} \to_\sigma w^{(1)} \to_\sigma \cdots \to_\sigma w^{(r)}$. This give a path $w' \rightharpoonup w$, consisting of $\frac{1}{2}(\ell(w') - \ell(w))$ type-II edges. We denote this path by $\underline{p}_0$.

Let $\mathcal{T}$ be a reduction tree of w. One may construct a reduction tree $\mathcal{T}'$ of w' containing the concatenation $\underline{p}_0 \circ \mathcal{T}$ as a subgraph. In particular, for any reduction path $\underline{p}$ in $\mathcal{T}$, the concatenation $\underline{p}' := \underline{p}_0 \circ \underline{p}$ is a reduction path in $\mathcal{T}'$. By definition, $\ell_{\mathrm{I}}(\underline{p}) = \ell_{\mathrm{I}}(\underline{p}')$ and $\ell_{\mathrm{II}}(\underline{p}) + \frac{1}{2}(\ell(w') - \ell(w)) = \ell_{\mathrm{II}}(\underline{p}')$. It is obvious that $J(w) = J(w')$ and $J_0(w') = \varnothing$. Hence $\ell_{\mathrm{I}}(w,[b],J) = \ell_{\mathrm{I}}(w',[b],J)$. On the other hand, we have $\sharp(J_0(w)/\langle\sigma\rangle) = \mathrm{length}([b_w],[t^{\mu}]) = \frac{1}{2}(\ell(\eta_\sigma(w) + \ell(y) - \ell(x))) = \frac{1}{2}(\ell(w') - \ell(w))$, where the second equality follows from the definition of cordial elements. Hence $\ell_{\mathrm{II}}(w,[b],J) + \frac{1}{2}(\ell(w') - \ell(w)) = \ell_{\mathrm{II}}(w',[b],J)$. The statements for $\mathcal{T}$ can now be deduced from the statements for $\mathcal{T}'$.

Acknowledgments

He is partially supported by Hong Kong RGC Grant 14300220, by the Xplorer prize and by New Cornerstone Science Foundation. Nie is partially supported by National Key R&D Program of China, No. 2020YFA0712600, CAS Project for Young Scientists in Basic Research, Grant No. YSBR-003, NSFC grants Nos. 11922119, 12288201 and 12231001. We thank the referees for a careful reading of the paper and many helpful suggestions.

On 03/08/2023, Dong Gyu Lim [24] informed us that he found a purely combinatorial proof of $(*)$ for all quasisplit reductive groups $\boldsymbol{G}$ in the introduction. Lim's approach is based on a probability-theoretic interpretation of a variant of $(*)$ and is independent of our approach.

References

[1] B. Bhatt and P. Scholze, "Projectivity of the Witt vector affine Grassmannian", *Invent. Math.* **209**:2 (2017), 329–423. MR Zbl

[2] C.-L. Chai, "Newton polygons as lattice points", *Amer. J. Math.* **122**:5 (2000), 967–990. MR Zbl

[3] P. Deligne and G. Lusztig, "Representations of reductive groups over finite fields", *Ann. of Math.* (2) **103**:1 (1976), 103–161. MR Zbl

[4] U. Görtz and X. He, "Basic loci of Coxeter type in Shimura varieties", *Camb. J. Math.* **3**:3 (2015), 323–353. MR Zbl

[5] U. Görtz, T. J. Haines, R. E. Kottwitz, and D. C. Reuman, "Dimensions of some affine Deligne–Lusztig varieties", *Ann. Sci. École Norm. Sup.* (4) **39**:3 (2006), 467–511. MR Zbl

[6] U. Görtz, X. He, and S. Nie, "$\mathbb{P}$-alcoves and nonemptiness of affine Deligne–Lusztig varieties", *Ann. Sci. École Norm. Sup.* (4) **48**:3 (2015), 647–665. MR Zbl

[7] U. Görtz, X. He, and S. Nie, "Fully Hodge–Newton decomposable Shimura varieties", *Peking Math. J.* **2**:2 (2019), 99–154. MR Zbl

[8] U. Görtz, X. He, and S. Nie, "Basic loci of Coxeter type with arbitrary parahoric level", *Canad. J. Math.* **76**:1 (2024), 126–172. MR Zbl

[9] P. Hamacher, "The geometry of Newton strata in the reduction modulo p of Shimura varieties of PEL type", *Duke Math. J.* **164**:15 (2015), 2809–2895. MR Zbl

[10] P. Hamacher and E. Viehmann, "Irreducible components of minuscule affine Deligne–Lusztig varieties", *Algebra Number Theory* **12**:7 (2018), 1611–1634. MR Zbl

[11] X. He, "Minimal length elements in some double cosets of Coxeter groups", *Adv. Math.* **215**:2 (2007), 469–503. MR Zbl

[12] X. He, "Geometric and homological properties of affine Deligne–Lusztig varieties", *Ann. of Math.* (2) **179**:1 (2014), 367–404. MR Zbl

[13] X. He, "Hecke algebras and p-adic groups", pp. 73–135 in *Current developments in mathematics* (Cambridge, 2015), edited by D. Jerison et al., Int. Press, Somerville, MA, 2016. MR Zbl

[14] X. He, "Kottwitz–Rapoport conjecture on unions of affine Deligne–Lusztig varieties", *Ann. Sci. École Norm. Sup.* (4) **49**:5 (2016), 1125–1141. MR Zbl

[15] X. He, "Some results on affine Deligne–Lusztig varieties", pp. 1345–1365 in *Proceedings of the International Congress of Mathematicians*, *II*, edited by B. Sirakov et al., World Sci., Hackensack, NJ, 2018. MR Zbl

[16] X. He, "Cordial elements and dimensions of affine Deligne–Lusztig varieties", *Forum Math. Pi* **9** (2021), art. id. e9. MR Zbl

[17] X. He and S. Nie, "Minimal length elements of extended affine Weyl groups", *Compos. Math.* **150**:11 (2014), 1903–1927. MR Zbl

[18] X. He and S. Nie, "P-alcoves, parabolic subalgebras and cocenters of affine Hecke algebras", *Selecta Math.* (*N.S.*) **21**:3 (2015), 995–1019. MR Zbl

[19] X. He and S. Nie, "On the acceptable elements", *Int. Math. Res. Not.* **2018**:3 (2018), 907–931. MR Zbl

[20] X. He, R. Zhou, and Y. Zhu, "Stabilizers of irreducible components of affine Deligne–Lusztig varieties", preprint, 2021. arXiv 2109.02594

[21] R. E. Kottwitz, "Isocrystals with additional structure", *Compos. Math.* **56**:2 (1985), 201–220. MR Zbl

[22] R. E. Kottwitz, "Isocrystals with additional structure, II", *Compos. Math.* **109**:3 (1997), 255–339. MR Zbl

[23] R. E. Kottwitz, "Dimensions of Newton strata in the adjoint quotient of reductive groups", *Pure Appl. Math. Q.* **2**:3 (2006), 817–836. MR Zbl

[24] D. G. Lim, "A combinatorial proof of the general identity of He–Nie–Yu", preprint, 2023. arXiv 2302.13260

[25] E. Milićević and E. Viehmann, "Generic Newton points and the Newton poset in Iwahori-double cosets", *Forum Math. Sigma* **8** (2020), art. id. e50. MR Zbl

[26] S. Nie, "Irreducible components of affine Deligne–Lusztig varieties", *Camb. J. Math.* **10**:2 (2022), 433–510. MR Zbl

[27] M. Rapoport, "A guide to the reduction modulo p of Shimura varieties", pp. 271–318 in *Automorphic forms*, *I* (Paris, 2000), edited by J. Tilouine et al., Astérisque **298**, Soc. Math. France, Paris, 2005. MR Zbl

[28] F. Schremmer, "Generic Newton points and cordial elements", preprint, 2022. arXiv 2205.02039

[29] R. Shimada, "On some simple geometric structure of affine Deligne–Lusztig varieties for GL_n", *Manuscripta Math.* **173**:3-4 (2024), 977–1001. MR Zbl

[30] E. Viehmann, "The dimension of some affine Deligne–Lusztig varieties", *Ann. Sci. École Norm. Sup.* (4) **39**:3 (2006), 513–526. MR Zbl

[31] E. Viehmann, "On the geometry of the Newton stratification", pp. 192–208 in *Shimura varieties*, edited by T. Haines and M. Harris, Lond. Math. Soc. Lect. Note Ser. **457**, Cambridge Univ. Press, 2020. MR Zbl

[32] E. Viehmann, "On Newton strata in the B_{dR}^{+}-Grassmannian", *Duke Math. J.* **173**:1 (2024), 177–225. MR Zbl

[33] R. Zhou and Y. Zhu, "Twisted orbital integrals and irreducible components of affine Deligne–Lusztig varieties", *Camb. J. Math.* **8**:1 (2020), 149–241. MR Zbl

[34] X. Zhu, "Affine Grassmannians and the geometric Satake in mixed characteristic", *Ann. of Math.* (2) **185**:2 (2017), 403–492. MR Zbl

Communicated by Shou-Wu Zhang
Received 2022-10-23 Revised 2023-06-28 Accepted 2023-10-12

xuhuahe@hku.hk *Department of Mathematics, University of Hong Kong, Pokfulam, Hong Kong*

niesian@amss.ac.cn *Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China*

School of Mathematical Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

yuqingchao@bicmr.pku.edu.cn *Beijing International Center for Mathematical Research, Beijing University, Beijing, China*

ALGEBRA AND NUMBER THEORY 18:9 (2024)
https://doi.org/10.2140/ant.2024.18.1715

Semistable models for some unitary Shimura varieties over ramified primes

Ioannis Zachos

We consider Shimura varieties associated to a unitary group of signature $(n-2, 2)$. We give regular p-adic integral models for these varieties over odd primes p which ramify in the imaginary quadratic field with level subgroup at p given by the stabilizer of a selfdual lattice in the hermitian space. Our construction is given by an explicit resolution of a corresponding local model.

1. Introduction

1A. This paper is a contribution to the problem of constructing regular integral models for Shimura varieties over places of bad reduction. There are several implicit examples of constructions of such regular integral models in special cases; see, for example, work of de Jong [1993], Genestier [2000], Pappas [2000b], Faltings [1997] and the very recent work of Pappas and Zachos [2022]. Here, we consider Shimura varieties associated to unitary groups of signature (r, s) over an imaginary quadratic field F_0. These Shimura varieties are of PEL type, so they can be written as a moduli space of abelian varieties with polarization, endomorphisms and level structure. Shimura varieties have canonical models over the "reflex" number field E. In the cases we consider here the reflex field is the field of rational numbers $\mathbb{Q}$ if $r = s$ and $E = F_0$ otherwise.

Constructing such well-behaved integral models is an interesting and hard problem whose solution has many applications to number theory. The behavior of these depends very much on the "level subgroup". Here, the level subgroup is the stabilizer of a selfdual lattice in the hermitian space. This stabilizer, by what follows below, is not connected when n is even, so not parahoric. However, by using work of

MSC2020: 11G18, 11G25, 14F40, 14G35, 14K10.

Keywords: Shimura varieties, regular integral models, local models.

Rapoport and Zink [1996] and Pappas [2000b] we construct p-adic integral models, which have simple and explicit moduli descriptions, and are étale locally around each point isomorphic to certain simpler schemes the *naive local models*. Inspired by the work of Pappas and Rapoport [2005] and Krämer [2003], we consider a variation of the above moduli problem where we add in the moduli problem an additional subspace in the deRham filtration $\mathrm{Fil}^0(A) \subset H^1_{\mathrm{dR}}(A)$ of the universal abelian variety A, which satisfies certain conditions. This is essentially an instance of the notion of a "linear modification" introduced in [Pappas 2000b]. We then show that the blow-up of this model along a smooth (non-Cartier) divisor produces a semistable integral model of the corresponding Shimura variety, i.e., it is regular and the irreducible components of the special fiber are smooth divisors crossing normally. We expect that our construction will find applications to the study of arithmetic intersections of special cycles and Kudla's program; see [Zhang 2021; Bruinier et al. 2020; He et al. 2023] for important applications of integral models of unitary Shimura varieties to number theory.

1B. To explain our results, we need to introduce some notation. We consider the group G of unitary similitudes for a hermitian vector space (W, ϕ) of dimension $n > 3$ over an imaginary quadratic field $F_0 \subset \mathbb{C}$, and fix a conjugacy class of homomorphisms $h : \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\, \mathbb{G}_m \to G_{\mathbb{R}}$ corresponding to a Shimura datum (G, X_h) of signature $(r, s) = (n-2, 2)$ (see Section 6). Let us mention here that the case $(r, s) = (1, 2)$, when $n = 3$, was studied in [Pappas 2000b, 4.5, 4.15]; see also [Pappas and Rapoport 2009, Section 6].

We assume that $F_0/\mathbb{Q}$ is ramified over p, where p is an odd prime number. Let $F_1 = F_0 \otimes \mathbb{Q}_p$ and $V = W \otimes_{\mathbb{Q}} \mathbb{Q}_p$. We fix a square root π of p and we set $k = \bar{\mathbb{F}}_p$. We assume that the hermitian form ϕ on V is split, i.e., that there is a basis $e_1, \dots, e_n$ such that $\phi(e_i, e_{n+1-j}) = \delta_{ij}$ for $i, j \in \{1, \dots n\}$.

In addition, we denote by Λ the standard lattice $O^n_{F_1}$ in V and we let $\mathcal{L}$ be the self-dual multichain consisting of $\{\pi^k \Lambda\}_{k\in\mathbb{Z}}$. Denote by K the stabilizer of Λ in $G(\mathbb{Q}_p)$ and let $\mathcal{G}$ be the (smooth) group scheme of automorphisms of the polarized chain $\mathcal{L}$ over $\mathbb{Z}_p$; see [Pappas and Rapoport 2009, Section 1.5]. Then $\mathcal{G}(\mathbb{Z}_p) = K$ and the group scheme $\mathcal{G}$ has $G \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ as its generic fiber. It turns out that when n is odd the stabilizer K is a parahoric subgroup. When n is even, K is not a parahoric subgroup since it contains a parahoric subgroup with index 2 and the corresponding parahoric group scheme is its connected component K°; see [Pappas and Rapoport 2009, Section 1.2] for more details.

Choose also a sufficiently small compact open subgroup K^p of the prime-to-p finite adelic points $G(\mathbb{A}^p_f)$ of G and set $\boldsymbol{K} = K^p K$. The Shimura variety $\mathrm{Sh}_{\boldsymbol{K}}(G, X)$ with complex points

$$\mathrm{Sh}_{\boldsymbol{K}}(G, X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/\boldsymbol{K}$$

is of PEL type. We set $\mathcal{O} = O_{E_v}$ where v the unique prime ideal of E above (p).

Next, we follow [Rapoport and Zink 1996, Definition 6.9] to define the moduli scheme $\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}}$ over $\mathcal{O}$ whose generic fiber agrees with $\mathrm{Sh}_{\boldsymbol{K}}(G, X)$ (see also Section 6). A point of $\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}}$ with values in the $\mathcal{O}$-scheme S is the isomorphism class of the following set of data $(A, \bar{\lambda}, \bar{\eta})$:

(1) An $\mathcal{L}$-set of abelian varieties $A = \{A_\Lambda\}$.

(2) A $\mathbb{Q}$-homogeneous principal polarization $\bar{\lambda}$ of the $\mathcal{L}$-set A.

(3) A K^p-level structure

$$\bar{\eta} : H_1(A, \mathbb{A}_f^p) \simeq W \otimes \mathbb{A}_f^p \mod K^p$$

which respects the bilinear forms on both sides up to a constant in $(\mathbb{A}_f^p)^\times$; see [loc. cit.] for details.

The set A should satisfy the determinant condition (i) of [loc. cit.]

For the definitions of the terms employed here we refer to [loc. cit., 6.3–6.8] and [Pappas 2000b, Section 3]. The functor $\mathcal{A}_{\boldsymbol{K}}^{\text{naive}}$ is representable by a quasiprojective scheme over $\mathcal{O}$. The moduli scheme $\mathcal{A}_{\boldsymbol{K}}^{\text{naive}}$ is connected to the *naive local model* $\mathrm{M}^{\text{naive}}$, see Section 2 for the explicit definition of $\mathrm{M}^{\text{naive}}$, via the local model diagram

$$\begin{array}{ccc} & \widetilde{\mathcal{A}}_{\boldsymbol{K}}^{\text{naive}} & \\ {}^{\pi_{\boldsymbol{K}}}\swarrow & & \searrow^{q_{\boldsymbol{K}}} \\ \mathcal{A}_{\boldsymbol{K}}^{\text{naive}} & & \mathrm{M}^{\text{naive}} \end{array} \tag{1B.1}$$

where the morphism $\pi_{\boldsymbol{K}}$ is a $\mathcal{G}$-torsor and $q_{\boldsymbol{K}}$ is a smooth and $\mathcal{G}$-equivariant morphism (see Section 6). Equivalently, using the language of algebraic stacks, there is a relatively representable smooth morphism

$$\mathcal{A}_{\boldsymbol{K}}^{\text{naive}} \to [\mathcal{G} \backslash \mathrm{M}^{\text{naive}}]$$

where the target is the quotient algebraic stack. In particular, since $\mathcal{G}$ is smooth, the above imply that $\mathcal{A}_{\boldsymbol{K}}^{\text{naive}}$ is étale locally isomorphic to $\mathrm{M}^{\text{naive}}$.

One can now consider a variation of the moduli of abelian schemes $\mathcal{A}_{\boldsymbol{K}}^{\text{spl}}$ over $\operatorname{Spec} O_{F_1}$ where we add in the moduli problem an additional subspace in the Hodge filtration $\mathrm{Fil}^0(A) \subset H^1_{\text{dR}}(A)$ of the universal abelian variety A with certain conditions to imitate the definition of the splitting local model $\mathcal{M}$; see Section 6B for the explicit definition of $\mathcal{A}_{\boldsymbol{K}}^{\text{spl}}$ and Section 2 where we define $\mathcal{M}$ for general signature (r, s). (Actually, $\mathcal{M}$ is a generalization of Krämer's local models [Krämer 2003, Definition 4.1]). There is a forgetful morphism

$$\tau : \mathcal{A}_{\boldsymbol{K}}^{\text{spl}} \to \mathcal{A}_{\boldsymbol{K}}^{\text{naive}} \otimes_{\mathcal{O}} O_{F_1}$$

defined by forgetting the extra subspace. Moreover, $\mathcal{A}_{\boldsymbol{K}}^{\text{spl}}$ has the same étale local structure as $\mathcal{M}$ and is a linear modification of $\mathcal{A}_{\boldsymbol{K}}^{\text{naive}} \otimes_{\mathcal{O}} O_{F_1}$ in the sense of [Pappas 2000b, Section 2] (see also [Pappas and Rapoport 2005, Section 15]). Therefore, there is a local model diagram for $\mathcal{A}_{\boldsymbol{K}}^{\text{spl}}$ similar to (1B.1) but with $\mathrm{M}^{\text{naive}}$ replaced by $\mathcal{M}$. Note, that there is also a corresponding forgetful morphism

$$\tau_1 : \mathcal{M} \to \mathrm{M}^{\text{naive}} \otimes_{\mathcal{O}} O_{F_1}.$$

In Section 2, we show that $\tau_1^{-1}(*)$ is isomorphic to the Grassmannian $\mathrm{Gr}(2, n)_k$. Here, $*$ is the "worst point" of $\mathrm{M}^{\text{naive}}$, i.e., the unique closed $\mathcal{G}$-orbit supported in the special fiber; see [Pappas 2000b, Section 4] for more details. Under the local model diagram, (see Section 6), $\tau_1^{-1}(*)$ corresponds to the locus where the Hodge filtration $\mathrm{Fil}^0(A)$ of the universal abelian scheme A is annihilated by the action of the uniformizer π. Consider the blow-up $\mathcal{A}^{\mathrm{bl}}_{\boldsymbol{K}}$ of $\mathcal{A}^{\mathrm{spl}}_{\boldsymbol{K}}$ along this locus.

1C. The main result of the paper is the following theorem.

Theorem 1.1. *$\mathcal{A}^{\mathrm{bl}}_{\boldsymbol{K}}$ is a semistable integral model for the Shimura variety* $\mathrm{Sh}_{\boldsymbol{K}}(G, X)$.

Since blowing-up commutes with étale localization and the étale local structure of the moduli scheme $\mathcal{A}^{\mathrm{spl}}_{\boldsymbol{K}}$ is controlled by the local structure of the local model $\mathcal{M}$, it is enough to show the above statement for the corresponding local models. In particular, it suffices to prove:

Theorem 1.2. *The blow-up* M^{bl} *of $\mathcal{M}$ along the smooth irreducible component $\tau_1^{-1}(*)$ of its special fiber is regular and has special fiber a divisor with normal crossings.*

To show the above theorem, we explicitly calculate an affine chart $\mathcal{U}$ of $\mathcal{M}$ in a neighborhood of $\tau_1^{-1}(*)$. In fact, we consider a more general situation where we calculate $\mathcal{U}$ for a general signature (r, s) and we show that $\mathcal{G}$-translates of $\mathcal{U}$ cover $\mathcal{M}$.

Proposition 1.3. *An affine chart $\mathcal{U} \subset \mathcal{M}$ containing a preimage of the worst point is isomorphic to*

$$\mathrm{Spec}\, O_{F_1}[X, Y]/(X - X^t, X \cdot (I_s + Y^t \cdot Y) - 2\pi I_s)$$

where X, Y are of sizes $s \times s$ and $(n - s) \times s$ respectively.

When $(r, s) = (n - 1, 1)$, Krämer [2003] shows that $\mathcal{U}$, and so $\mathcal{M}$, has semistable reduction. Therefore, she obtains a semistable integral model for the corresponding Shimura variety.

When $(r, s) = (n - 2, 2)$, $\mathcal{U}$ does not have semistable reduction anymore and so $\mathcal{M}$ does not give us a resolution. However, we use the explicit description of $\mathcal{U}$ above to calculate the blow-up of $\mathcal{M}$ along the $\mathcal{G}$-invariant smooth subscheme $\tau_1^{-1}(*)$. The blow-up gives a $\mathcal{G}$-birational projective morphism

$$r^{\mathrm{bl}} : \mathrm{M}^{\mathrm{bl}} \to \mathcal{M}$$

such that M^{bl} is regular and has special fiber a reduced divisor with normal crossings. We quickly see that the corresponding blow-up $\mathcal{A}^{\mathrm{bl}}_{\boldsymbol{K}}$ of the integral model $\mathcal{A}^{\mathrm{spl}}_{\boldsymbol{K}}$ inherits the same nice properties as M^{bl}. In fact, there is a local model diagram for $\mathcal{A}^{\mathrm{bl}}_{\boldsymbol{K}}$ similar to (1B.1) but with $\mathrm{M}^{\mathrm{naive}}$ replaced by M^{bl}. See Theorem 6.1 for the precise statement about the model $\mathcal{A}^{\mathrm{bl}}_{\boldsymbol{K}}$.

Let us mention here that we can obtain similar results for the Shimura varieties $\mathrm{Sh}_{\boldsymbol{K}'}(G, X)$ where $\boldsymbol{K}' = K^p K^\circ$ (see Section 6). (Recall that K° is the parahoric connected component of the stabilizer K.) Also, we can apply these results to obtain regular (formal) models of the corresponding Rapoport–Zink spaces.

Let us now explain the lay-out of the paper. In Section 2, we recall the definitions of certain variants of local models for ramified unitary groups. In Section 3, we give explicit equations that describe the affine chart $\mathcal{U}$ of the *splitting model* $\mathcal{M}$ for a general signature (r, s) and we also show that $\mathcal{G}$-translates of $\mathcal{U}$ cover $\mathcal{M}$. For the rest of the paper we assume $(r, s) = (n - 2, 2)$. In Section 4, we construct the semistable resolution $\rho : \mathcal{U}^{\mathrm{bl}} \to \mathcal{U}$ of the affine chart $\mathcal{U}$. In Section 5, we show that M^{bl} has semistable reduction by using the results of Section 4 and the structure of local models. In Section 6, we apply the above results to construct regular integral models for the corresponding Shimura varieties.

2. Preliminaries: local models and variants

We use the notation of [Pappas 2000b]. We take $F = \mathbb{Q}_p[t]/(t^2 - pu)$ and $O_F = \mathbb{Z}_p[t]/(t^2 - pu)$, where p is an odd prime and u is a unit in $\mathbb{Z}_p$. For $n > 3$, we set $V = F^n$ and denote by e_i, $1 \le i \le n$, the standard O_F-generators of the standard lattice $\Lambda = O_F^n \subset V$. Fix a uniformizer π of O_F with $\pi^2 = p\delta$. Also, since $p \neq 2$, $\delta = \pi^2/p$ has a square root in a finite étale extension of $\mathrm{Spec}(\mathbb{Z}_p)$. After such a base extension there is a uniformizer π such that $\pi^2 = p$. We will assume that we have such a uniformizer and suppress the notation of the étale base extension.

Set $k = \bar{\mathbb{F}}_p$. The uniformizing element π induces a $\mathbb{Z}_p$- linear mapping on Λ which we denote by t. We define a nondegenerate alternating $\mathbb{Q}_p$-bilinear form $\langle\,,\rangle : V \times V \to \mathbb{Q}_p$ given by

$$\langle e_i, te_j\rangle = \delta_{i,j}, \quad \langle e_i, e_j\rangle = 0, \quad \langle te_i, te_j\rangle = 0.$$

The restriction $\langle\,,\rangle : O_F^n \times O_F^n \to \mathbb{Z}_p$ is a perfect $\mathbb{Z}_p$-bilinear form. Using the duality isomorphism $\mathrm{Hom}_F(V, F) \cong \mathrm{Hom}_{\mathbb{Q}_p}(V, \mathbb{Q}_p)$ given by composing with the trace $\mathrm{Tr}_{F/\mathbb{Q}_p} : F \to \mathbb{Q}_p$ we see, as in [Pappas 2000b, Section 3], that there exists a unique nondegenerate hermitian form $\phi : V \times V \to F$ such that

$$\langle x, y\rangle = \mathrm{Tr}_{F/\mathbb{Q}_p}(\pi^{-1}\phi(x, y)), \quad x, y \in V.$$

We take $G := GU_n := GU(\phi)$ and we choose a partition $n = r + s$; we refer to the pair (r, s) as the signature. By replacing ϕ by $-\phi$ if needed, we can make sure that $s \le r$ and so we assume that $s \le r$ (see [Pappas and Rapoport 2009, Section 1.1] for more details). Identifying $G \otimes F \simeq \mathrm{GL}_{n,F} \times \mathbb{G}_{m,F}$, we define the cocharacter $\mu_{r,s}$ as $(1^{(s)}, 0^{(r)}, 1)$ of $D \times \mathbb{G}_m$, where D is the standard maximal torus of diagonal matrices in GL_n; for more details we refer the reader to [Smithling 2011, Section 3.2]. We denote by E the reflex field of $\{\mu_{r,s}\}$; then $E = \mathbb{Q}_p$ if $r = s$ and $E = F$ otherwise; see [Pappas and Rapoport 2009, Section 1.1]. We set $O := O_E$.

Next, we denote by K the stabilizer of Λ in $G(\mathbb{Q}_p)$. We also let $\mathcal{L}$ be the self-dual multichain consisting of $\{\pi^k\Lambda\}_{k\in\mathbb{Z}}$. Here $\mathcal{G} = \underline{\mathrm{Aut}}(\mathcal{L})$ is the group scheme over $\mathbb{Z}_p$ with $K = \mathcal{G}(\mathbb{Z}_p)$ the subgroup of $G(\mathbb{Q}_p)$ fixing the lattice chain $\mathcal{L}$. When n is odd, the stabilizer K is a parahoric subgroup but when n is even, K is not a parahoric subgroup since it contains a parahoric subgroup with index 2. The corresponding parahoric group scheme is its connected component K°; see [Pappas and Rapoport 2009, Section 1.2] for more details.

We briefly recall the definition of certain variants of local models for ramified unitary groups.

2A. ***Rapoport–Zink local models and variants.*** Let $\mathrm{M}^{\mathrm{naive}}$ be the functor which associates to each scheme S over $\mathrm{Spec}\, O$ the set of subsheaves $\mathcal{F}$ of $O \otimes \mathcal{O}_S$-modules of $\Lambda \otimes \mathcal{O}_S$ such that:

(1) $\mathcal{F}$ as an $\mathcal{O}_S$-module is Zariski locally on S a direct summand of rank n.

(2) $\mathcal{F}$ is totally isotropic for $\langle\,,\rangle \otimes \mathcal{O}_S$. item (Kottwitz condition) $\mathrm{char}_{t|\mathcal{F}}(X) = (X + \pi)^r(X - \pi)^s$.

The functor $\mathrm{M}^{\mathrm{naive}}$ is represented by a closed subscheme, which we again denote $\mathrm{M}^{\mathrm{naive}}$, of $\mathrm{Gr}(n, 2n) \otimes \operatorname{Spec} O$; hence $\mathrm{M}^{\mathrm{naive}}$ is a projective O-scheme. (Here we denote by $\mathrm{Gr}(n, d)$ the Grassmannian scheme parametrizing locally direct summands of rank n of a free module of rank d.) $\mathrm{M}^{\mathrm{naive}}$ is the *naive local model* of Rapoport and Zink [1996]. Also, $\mathrm{M}^{\mathrm{naive}}$ supports an action of $\mathcal{G}$.

Proposition 2.1. (a) *We have*

$$\mathrm{M}^{\mathrm{naive}} \otimes_O E \cong \mathrm{Gr}(s, n) \otimes E.$$

In particular, the generic fiber of $\mathrm{M}^{\mathrm{naive}}$ *is smooth and geometrically irreducible of dimension* rs.

(b) *We have*

$$\dim \mathrm{M}^{\mathrm{naive}} \otimes_O k \geq \begin{cases} n^2/4 & \text{if } n \text{ is even,} \\ (n^2-1)/4 & \text{if } n \text{ is odd.} \end{cases}$$

In particular, $\mathrm{M}^{\mathrm{naive}}$ *is not flat if* $|r - s| > 1$.

Proof. See [Pappas 2000b, Proposition 3.8; Krämer 2003, Proposition 2.2; 2003, Corollary 2.3]. □

The flat closure of $\mathrm{M}^{\mathrm{naive}} \otimes_O E$ in $\mathrm{M}^{\mathrm{naive}}$ is by definition the *local model* $\mathrm{M}^{\mathrm{loc}}$.

In [Pappas 2000b, Section 4], Pappas introduces the *wedge local model* $\mathrm{M}^{\wedge}$, in order to correct the nonflatness problem, by imposing the following additional condition:

$$\wedge^{r+1}(t - \pi \mid \mathcal{F}) = (0) \quad \text{and} \quad \wedge^{s+1}(t + \pi \mid \mathcal{F}) = (0).$$

More precisely, $\mathrm{M}^{\wedge}$ is the closed subscheme of $\mathrm{M}^{\mathrm{naive}}$ that classifies points given by $\mathcal{F}$ which satisfy the wedge condition. The scheme $\mathrm{M}^{\wedge}$ supports an action of $\mathcal{G}$ and the immersion $i : \mathrm{M}^{\wedge} \to \mathrm{M}^{\mathrm{naive}}$ is $\mathcal{G}$-equivariant. It is easy to see that:

Proposition 2.2. *The generic fibers of* $\mathrm{M}^{\mathrm{naive}}$ *and* $\mathrm{M}^{\wedge}$ *coincide, in particular the generic fiber of* $\mathrm{M}^{\wedge}$ *is a smooth, geometrically irreducible variety of dimension rs.*

Proof. See [Krämer 2003, Proposition 3.4] and [Arzdorf 2009, Lemma 1.1]. □

Also, $\mathrm{M}^{\mathrm{loc}} \subset \mathrm{M}^{\wedge}$ and $\mathrm{M}^{\mathrm{loc}} \otimes E = \mathrm{M}^{\wedge} \otimes E$. As in [Pappas 2000b, Section 4] and [Pappas and Rapoport 2009, Section 2.4.2, Section 5.5], the worst point of $\mathrm{M}^{\wedge}$, i.e., the unique closed $\mathcal{G}$-orbit which lies in the closure of any other orbit, is given by the k-valued point $\mathcal{F} = t\Lambda \subset \Lambda \otimes k \cong (k[t]/(t^2))^n$.

It is conjectured in [Pappas 2000b] that $\mathrm{M}^{\wedge}$ is flat for $n \geq 2$ and any signature and so $\mathrm{M}^{\mathrm{loc}} = \mathrm{M}^{\wedge}$. It has been shown, see [Pappas 2000b, Theorem 4.5], that this is true for the signature $(n-1, 1)$. For more general lattice chains, the wedge condition turns out to be insufficient, see [Pappas and Rapoport 2009, Remarks 5.3, 7.4]. In [loc. cit.], the authors introduced a further refinement of the moduli problem by also adding the so-called "spin condition" (for more information we refer the reader to [loc. cit.]); this will play no role in this paper.

Next, we consider the moduli scheme $\mathcal{M}$ over O_F, the *splitting* (*or Krämer*) *local model* as in [Pappas and Rapoport 2005, Remark 14.2] and [Krämer 2003, Definition 4.1], whose points for an O_F-scheme S are Zariski locally $\mathcal{O}_S$-direct summands $\mathcal{F}_0, \mathcal{F}_1$ of ranks s, n respectively, such that:

(1) $(0) \subset \mathcal{F}_0 \subset \mathcal{F}_1 \subset \Lambda \otimes \mathcal{O}_S$.

(2) $\mathcal{F}_1 = \mathcal{F}_1^\perp$, i.e., $\mathcal{F}_1$ is totally isotropic for $\langle\ ,\ \rangle \otimes \mathcal{O}_S$.

(3) $(t+\pi)\mathcal{F}_1 \subset \mathcal{F}_0$.

(4) $(t-\pi)\mathcal{F}_0 = (0)$.

The functor is represented by a projective O_F-scheme $\mathcal{M}$. The scheme $\mathcal{M}$ supports an action of $\mathcal{G}$ and there is a $\mathcal{G}$-equivariant morphism

$$\tau : \mathcal{M} \to \mathrm{M}^\wedge \otimes_O O_F$$

which is given by $(\mathcal{F}_0, \mathcal{F}_1) \mapsto \mathcal{F}_1$ on S-valued points. (Indeed, we can easily see, as in [Krämer 2003, Definition 4.1], that τ is well defined.)

Proposition 2.3. *The morphism $\tau : \mathcal{M} \to \mathrm{M}^\wedge \otimes_O O_F$ induces an isomorphism on the generic fibers.*

Proof. It follows by [Krämer 2003, Remark 4.2] and the proof of [Pappas 2000b, Proposition 3.8]. □

The following discussion appears in [Pappas 2000a]. Over the special fiber, the condition (4) amounts to $t\mathcal{F}_0 = (0)$. Thus, we have

$$(0) \subset \mathcal{F}_0 \subset t\Lambda \otimes k \subset \mathcal{F}_0^\perp \subset \Lambda \otimes k.$$

Also, we have

$$(0) \subset (t^{-1}(\mathcal{F}_0))^\perp \subset t\Lambda \otimes k \subset t^{-1}(\mathcal{F}_0) \subset \Lambda \otimes k.$$

The spaces $t^{-1}(\mathcal{F}_0)$, $\mathcal{F}_0^\perp$ have rank $n+s$, $2n-s = n+r$ respectively. Fixing $\mathcal{F}_0$, the rank of $t^{-1}(\mathcal{F}_0) \cap \mathcal{F}_0^\perp$ influences the dimension of the space of allowable $\mathcal{F}_1$ since

$$\mathcal{F}_0 + (t^{-1}(\mathcal{F}_0))^\perp \subset \mathcal{F}_1 \subset t^{-1}(\mathcal{F}_0) \cap \mathcal{F}_0^\perp.$$

Note that $\mathcal{F}_0 \subset t\Lambda \otimes k \simeq k^n \otimes \mathcal{O}_S$. Hence, we consider the morphism

$$\pi : \mathcal{M} \otimes k \to \mathrm{Gr}(s, n) \otimes k$$

given by $(\mathcal{F}_0, \mathcal{F}_1) \mapsto \mathcal{F}_0$. This has a section

$$\phi : \mathrm{Gr}(s, n) \otimes k \to \mathcal{M} \otimes k$$

given by $\mathcal{F}_0 \mapsto (\mathcal{F}_0, \mathcal{F}_1)$ with $\mathcal{F}_1 = t\Lambda \otimes k$. The image of the section ϕ is an irreducible component of $\mathcal{M} \otimes_O k$ which is the fiber $\tau^{-1}(t\Lambda)$ over the worst point. Hence, $\tau^{-1}(t\Lambda)$ is isomorphic to the Grassmannian $\mathrm{Gr}(s, n) \otimes k$ of dimension rs. Also, observe that $\{(\mathcal{F}_0, \mathcal{F}_1) \mid \mathrm{rank}(t^{-1}(\mathcal{F}_0) \cap \mathcal{F}_0^\perp) = n\} \subset \tau^{-1}(t\Lambda)$ since when $t^{-1}(\mathcal{F}_0) \cap \mathcal{F}_0^\perp$ has rank n then $t^{-1}(\mathcal{F}_0) \cap \mathcal{F}_0^\perp = t\Lambda$ which gives $\mathcal{F}_1 = t\Lambda$.

However, the morphism π has fibers of positive dimension over points of $\mathrm{Gr}(s, n) \otimes k$ which correspond to subspaces of $\mathrm{Gr}(s, n) \otimes k$ on which the dimension $t^{-1}(\mathcal{F}_0) \cap \mathcal{F}_0^\perp$ is more than n. Actually, $t^{-1}(\mathcal{F}_0) \cap \mathcal{F}_0^\perp$ has maximal dimension, i.e., $t^{-1}(\mathcal{F}_0) \subset \mathcal{F}_0^\perp$, if and only if $\mathcal{F}_0 \subset t\Lambda \otimes k \simeq k^n \otimes \mathcal{O}_S$ is totally isotropic for the (nondegenerate) symmetric form on $t\Lambda \otimes k$ which is defined as $\{tv, tw\} := \langle tv, w\rangle$; see the proof of [Krämer 2003, Theorem 4.5] for more details. Denote by $\mathcal{Q}(s, n)$ the smooth closed subscheme of

$\mathrm{Gr}(s,n)\otimes k$ of dimension $s(2n-3s-1)/2$ which parametrizes all these isotropic s-subspaces $\mathcal{F}_0$ in the n-space $k^n\otimes\mathcal{O}_S$. For such $\mathcal{F}_0\in\mathcal{Q}(s,n)$ we have that $t^{-1}(\mathcal{F}_0)\subset\mathcal{F}_0^{\perp}$ and thus the fiber $\pi^{-1}(\mathcal{F}_0)$ is given by $\mathcal{F}_1$ with $\mathcal{F}_1=\mathcal{F}_1^{\perp}$ such that

$$\mathcal{F}_0\subset(t^{-1}(\mathcal{F}_0))^{\perp}\subset\mathcal{F}_1\subset t^{-1}(\mathcal{F}_0).$$

We can see that these $\{\mathcal{F}_1\}$ correspond to Lagrangian subspaces in $t^{-1}(\mathcal{F}_0)/(t^{-1}(\mathcal{F}_0))^{\perp}$ which have dimension $2s$. This is a smooth $s(s+1)/2$-dimensional scheme which we denote by $L(s,2s)$. From the above discussion we see that $\pi^{-1}(\mathcal{Q}(s,n))$ is a $L(s,2s)$-bundle over $\mathcal{Q}(s,n)$ with dimension rs. Thus, $\pi^{-1}(\mathcal{Q}(s,n))$ is an irreducible component of $\mathcal{M}\otimes k$ which intersects with $\tau^{-1}(t\Lambda)$ over $\mathcal{Q}(s,n)$.

Krämer [2003] shows that τ defines a resolution of $\mathrm{M}^{\wedge}$ in the case that the signature is $(n-1,1)$. In particular, she proves that $\mathcal{M}$ is regular with special fiber a reduced divisor with simple normal crossings. Also she shows that the special fiber of $\mathcal{M}$ consists of two smooth irreducible components of dimension $n-1$ — one of which being isomorphic to $\mathbb{P}_k^{n-1}$ (this corresponds to $\tau^{-1}(t\Lambda)$), and the other one being a $\mathbb{P}_k^1$-bundle over a smooth quadric (this corresponds to $\pi^{-1}(\mathcal{Q}(1,n))$) — which intersect transversely in a smooth irreducible variety of dimension $n-2$ (this corresponds to $\mathcal{Q}(1,n)$).

3. An affine chart

The goal of this section is to write down the equations that define $\mathcal{M}$ in a neighborhood $\mathcal{U}$ of $(\mathcal{F}_0,t\Lambda)$ for a general signature (r,s); see Proposition 3.1. We also deduce, see Proposition 3.7, that $\mathcal{G}$-translates of $\mathcal{U}$ cover $\mathcal{M}$. (Recall from Section 2 that $(\mathcal{F}_0,t\Lambda)$ is a point in the fiber of $\tau:\mathcal{M}\to\mathrm{M}^{\wedge}\otimes_O O_F$ over the worst point.) In order to find the explicit equations that describe $\mathcal{U}$, we use similar arguments as in the proof of [Krämer 2003, Theorem 4.5]. In our case we consider:

$$\mathcal{F}_1=\left[\frac{A}{I_n}\right],\qquad \mathcal{F}_0=X=\left[\frac{X_1}{X_2}\right]$$

where A is of size $n\times n$, X is of size $2n\times s$ and X_1, X_2 are of size $n\times s$; with the additional condition that $\mathcal{F}_0$ has rank s and so X has a nonvanishing $s\times s$-minor. We also ask that $(\mathcal{F}_0,\mathcal{F}_1)$ satisfy the following four conditions:

(1) $\mathcal{F}_1^{\perp}=\mathcal{F}_1$.

(2) $\mathcal{F}_0\subset\mathcal{F}_1$.

(3) $(t-\pi)\mathcal{F}_0=(0)$.

(4) $(t+\pi)\mathcal{F}_1\subset\mathcal{F}_0$.

Observe that

$$M_t=\left[\begin{array}{c|c}0_n & pI_n\\\hline I_n & 0_n\end{array}\right]$$

of size $2n\times 2n$ is the matrix giving multiplication by t:

(1) The condition that $\mathcal{F}_1$ is isotropic translates to

$$A^t = A.$$

(2) The condition $\mathcal{F}_0 \subset \mathcal{F}_1$ translates to

$$\exists Y : X = \left[\frac{A}{I_n}\right] \cdot Y$$

where Y is of size $n \times s$. Thus, we have

$$\left[\frac{X_1}{X_2}\right] = \left[\frac{A}{I_n}\right] \cdot Y = \left[\frac{AY}{I_n Y}\right]$$

and so $X_1 = AY$ and $X_2 = Y$.

(3) The condition $(t - \pi)\mathcal{F}_0 = (0)$ is equivalent to

$$M_t \cdot X = \left[\frac{\pi X_1}{\pi X_2}\right],$$

which amounts to

$$\left[\frac{pX_2}{X_1}\right] = \left[\frac{\pi X_1}{\pi X_2}\right].$$

Thus, $X_1 = \pi X_2$ which translates to $AY = \pi Y$ by condition (2).

(4) The last condition $(t + \pi)\mathcal{F}_1 \subset \mathcal{F}_0$ translates to

$$\exists Z : M_t \cdot \left[\frac{A}{I_n}\right] + \left[\frac{\pi A}{\pi I_n Y}\right] = X \cdot Z^t$$

where Z is of size $n \times s$. This amounts to

$$\left[\frac{pI_n + \pi A}{A + \pi I_n}\right] = \left[\frac{X_1 Z^t}{X_2 Z^t}\right].$$

From the above we get $A + \pi I_n = X_2 Z^t$ which by condition (2) translates to $A = Y Z^t - \pi I_n$. Thus, A can be expressed in terms of Y, Z. In addition, by condition (2) and in particular by the relations $X_1 = AY$ and $X_2 = Y$ we deduce that the matrix X is given in terms of Y, Z. Also from $Y = X_2$ we get that the matrix Y is given in terms of A, X. (Below we will also show that Z can be expressed in terms of A, X.)

For later use, we break up the matrices Y, Z into blocks as follows. We write

$$Y = \left[\frac{Y_1}{Y_2}\right], \quad Z = \left[\frac{Z_1}{Z_2}\right]$$

where Y_1, Z_1 are of size $s \times s$ and Y_2, Z_2 are of size $(n - s) \times s$. Observe from $X_1 = \pi X_2$ that all the entries of X_1 are in the maximal ideal and thus a minor involving entries of X_1 cannot be a unit. Recall that the matrix X has a nonvanishing $s \times s$-minor and $X_2 = Y$ from condition (2). Therefore, we can assume that $Y_1 = I_s$ up to a change of basis.

We replace A by $YZ^t - \pi I_n$. Hence, conditions (1) and (3) are equivalent to

$$ZY^t = YZ^t, \tag{3.1}$$

$$YZ^tY = 2\pi Y. \tag{3.2}$$

Here, we want to mention how we can express Z in terms of A, X. From the above we have $YZ^t = A + \pi I_n$ and $Y = \left[\frac{I_s}{Y_2}\right]$ which gives $\left[\frac{Z^t}{Y_2 Z^t}\right] = A + \pi I_n$. Next, we break the matrices A, I_n into blocks: $A = \left[\frac{A_1}{A_2}\right]$, $I_n = \left[\frac{I_1}{I_2}\right]$ where A_1, I_1 are of size $s \times n$ and A_2, I_2 are of size $(n-s) \times n$. Hence, from $\left[\frac{Z^t}{Y_2 Z^t}\right] = A + \pi I_n$ we obtain $Z^t = A_1 + \pi I_1$ and thus $Z = A_1^t + \pi I_1^t$.

From the above we deduce that an affine neighborhood of $\mathcal{M}$ around $(\mathcal{F}_0, t\Lambda)$ is given by $\mathcal{U} = \operatorname{Spec} O_F[Y, Z]/(Y_1 - I_s, ZY^t - YZ^t, YZ^tY - 2\pi Y)$.

Our goal in this section is to prove the simplification of equations given by the following proposition.

Proposition 3.1. *The affine chart $\mathcal{U} \subset \mathcal{M}$ is isomorphic to*

$$\operatorname{Spec} O_F[Y_2, Z_1]/(Z_1 - Z_1^t, Z_1(I_s + Y_2^t Y_2) - 2\pi I_s).$$

Proof. From (3.1) we get

$$\left[\frac{Z_1}{Z_2}\right] \cdot [I_s \mid Y_2^t] = \left[\frac{I_s}{Y_2^t}\right] \cdot [Z_1^t \mid Z_2^t],$$

which gives

$$\left[\begin{array}{c|c} Z_1 & Z_1 Y_2^t \\ \hline Z_2 & Z_2 Y_2^t \end{array}\right] = \left[\begin{array}{c|c} Z_1^t & Z_2^t \\ \hline Y_2 Z_1 & Y_2 Z_2^t \end{array}\right].$$

From the above relation we obtain the relations $Z_1 = Z_1^t$ and $Z_2 = Y_2 Z_1^t$. Thus, Z_1 is symmetric and Z_2 can be expressed in terms of Y_2, Z_1.

Next, the relation (3.2) amounts to

$$\left[\frac{I_s}{Y_2}\right] \cdot [Z_1^t \mid Z_2^t] \cdot \left[\frac{I_s}{Y_2}\right] = \left[\frac{2\pi I_s}{2\pi Y_2}\right]$$

which is equivalent to

$$\left[\frac{Z_1^t + Z_2^t Y_2}{Y_2 Z_1^t + Y_2 Z_2^t Y_2}\right] = \left[\frac{2\pi I_s}{2\pi Y_2}\right].$$

From this we get $Z_1^t + Z_2^t Y_2 = 2\pi I_s$ which translates to $Z_1(I_s + Y_2^t Y_2) = 2\pi I_s$ as $Z_1^t = Z_1$ and $Z_2 = Y_2 Z_1^t$. From all the above the proof of the proposition follows. □

As corollaries of the above result we have:

Corollary 3.2. *For $(r, s) = (n-1, 1)$, the corresponding affine chart $\mathcal{U}$ will be isomorphic to*:

$$\mathcal{U} \cong \operatorname{Spec}\Bigl(O_F[(y_i)_{1 \le i \le n}, a] \Big/ \Bigl(a \sum_{c=1}^{n} y_c^2 - 2\pi,\, y_1 - 1\Bigr)\Bigr). \qquad \square$$

Remark 3.3. For the above signature, Krämer [2003] shows that $\mathcal{U}$ is regular with special fiber a divisor with simple normal crossings.

Corollary 3.4. *For* $(r, s) = (n-2, 2)$ *the corresponding affine chart* $\mathcal{U}$ *will be isomorphic to*

$$\mathcal{U} \cong \operatorname{Spec}(O_F[(x_i)_{3\le i\le n}, (y_i)_{3\le i\le n}, a, b, c]/(Z_1 N - 2\pi I_2))$$

where

$$Z_1 = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad N = \begin{pmatrix} 1+\sum_{i=3}^n x_i^2 & \sum_{i=3}^n x_i y_i \\ \sum_{i=3}^n x_i y_i & 1+\sum_{i=3}^n y_i^2 \end{pmatrix}. \qquad \square$$

Remark 3.5. In this case $s = 2$, $\mathcal{U}$ does not have semistable reduction as one of the irreducible components of the special fiber of $\mathcal{U}$ is not smooth. More precisely, over the special fiber $(\pi = 0)$ $\mathcal{U}$ has three irreducible components given by

$$T_i = \{(Z_1, N) \mid Z_1 N = 0, \operatorname{rank} Z_1 \le i, \operatorname{rank} N \le 2 - i\}, \quad \text{for } i = 0, 1, 2.$$

We can easily see that T_0, T_2 are smooth but T_1 is singular. We resolve the singularities of $\mathcal{U}$ in Section 3 by blowing up the irreducible component T_0 or in other words by blowing up the ideal (Z_1) generated by the entries of Z_1. Observe from the above and the proof of Proposition 3.1 that $A = Y \cdot \left[\frac{Z_1}{Z_2}\right]^t$ and $Z_2 = Y_2 Z_1^t$ over the special fiber. Hence, $Z_1 = 0$, i.e., $a = b = c = 0$, gives $A = 0$ which corresponds to $\mathcal{F}_1 = t\Lambda$. Thus $T_0 = \mathcal{U} \cap \tau^{-1}(t\Lambda)$ where $\tau : \mathcal{M} \to \mathrm{M}^\wedge \otimes_O O_F$; see Section 2A.

Remark 3.6. For a general signature (r, s), over the special fiber of $\mathcal{U}$ we have that $Z_1 = Z_1^t$ and $Z_1(I_s + Y_2^t Y_2) = 0$. As in Remark 3.5, $Z_1 = 0$ gives $A = 0$ which corresponds to $\mathcal{F}_1 = t\Lambda$.

Moreover, from the above and the definition of the (nondegenerate) symmetric form $\{\,,\}$ on $t\Lambda \otimes k$ (see Section 2A) we have $\{\mathcal{F}_0, \mathcal{F}_0\} = Y^t \cdot Y = I_s + Y_2^t Y_2$ since $\mathcal{F}_0 = \left[\frac{X_1}{X_2}\right]$ where $X_1 = \pi X_2$, $X_2 = Y$ and $Y = \left[\frac{I_s}{Y_2}\right]$. Thus, from the rank of $I_s + Y_2^t Y_2$ we read how isotropic $\mathcal{F}_0$ is with respect to $\{\,,\}$. When the rank of the matrix $I_s + Y_2^t Y_2$ is zero, which actually occurs, we have $\{\mathcal{F}_0, \mathcal{F}_0\} = 0$.

From all the above, we can easily see that $\mathcal{U}$ contains points $(\mathcal{F}_0, \mathcal{F}_1)$ where $\mathcal{F}_0 \in \mathcal{Q}(s, n)$ and $\mathcal{F}_1 = t\Lambda$. (Recall from Section 2A that $\mathcal{Q}(s, n)$ is the closed subscheme of $\operatorname{Gr}(s, n) \otimes k$ which contains all the totally isotropic s-subspaces $\mathcal{F}_0$ with respect to the symmetric form $\{\,,\}$.)

Proposition 3.7. *When* $s \ge 1$, $\mathcal{G}$*-translates of* $\mathcal{U}$ *cover* $\mathcal{M}$.

Proof. From Section 2A, we have the forgetful $\mathcal{G}$-equivariant morphism $\tau : \mathcal{M} \to \mathrm{M}^\wedge \otimes_O O_F$ given by $(\mathcal{F}_0, \mathcal{F}_1) \mapsto \mathcal{F}_1$. As in [Pappas 2000b, Section 4] and [Pappas and Rapoport 2009, Sections 2.4.2, 5.5], the worst point of $\mathrm{M}^\wedge \otimes_O O_F$ is given by the k-valued point $t\Lambda$. The reason for this terminology is that the geometric special fiber of $\mathrm{M}^\wedge \otimes_O O_F$ embeds into an appropriate affine flag variety, where it decomposes into unions of finitely many Schubert cells, and the worst point is the unique closed Schubert cell. This one point stratum lies in the closure of any other stratum and the inclusion relations between the Schubert varieties are given by the Bruhat order. From the construction of splitting local models and the above, in order to show that $\mathcal{G}$-translates of $\mathcal{U}$ cover $\mathcal{M}$ it is enough to prove that $\mathcal{G}$-translates of $\mathcal{U}$ cover $\tau^{-1}(t\Lambda)$.

Recall that K is the stabilizer of Λ in $G(\mathbb{Q}_p)$ and K° is the neutral component of K, i.e., the parahoric stabilizer of Λ. When n is odd $K = K^\circ$ and when n is even $K/K^\circ \simeq \mathbb{Z}/2\mathbb{Z}$; see §2. Also, $\mathcal{G}$ is the smooth group scheme of automorphisms of the polarized chain $\mathcal{L}$ over $\mathbb{Z}_p$ with $\mathcal{G}(\mathbb{Z}_p) = K$ and $\mathcal{G}^\circ$ is the neutral component of $\mathcal{G}$ with $\mathcal{G}^\circ(\mathbb{Z}_p) = K^\circ$.

From Section 2A, we have that $\tau^{-1}(t\Lambda) \cong \mathrm{Gr}(s,n)\otimes k$ and $\mathcal{G}_k$ acts via its action by reduction to $t\Lambda/t^2\Lambda$. This action factors through the orthogonal group $O(n)_k$ of the symmetric form $\{\,,\}$ on $t\Lambda\otimes k$ and gives the map $\mathcal{G}_k \to O(n)_k$. As in [Pappas and Rapoport 2008, Section 4] (see also [Tits 1979, Section 3.11]), $\mathcal{G}_k^{\circ}$ has $SO(n)_k$ as its maximal reductive quotient if n is even and $O(n)_k$ if n is odd via its action by reduction to $t\Lambda/t^2\Lambda$. The maps $\mathcal{G}_k^{\circ} \to O(n)_k$ and $\mathcal{G}_k^{\circ} \to SO(n)_k$ are surjective when n is odd and even respectively. Therefore, the map $\mathcal{G}_k \to O(n)_k$ is always surjective.

Next, the $O(n)$-action on $\mathrm{Gr}(s,n)$ has a finite number of orbits. More precisely, there are $s+1$ orbits $Q(0), Q(1), \dots, Q(s)$ where $Q(i)=\{\mathcal{F}_0 \in \mathrm{Gr}(s,n) \mid \dim(\mathrm{rad}(\mathcal{F}_0)) = i\}$; see [Barbasch and Evens 1994, Section 4]. For example in the case $s=2$ there are three $O(n)$-orbits: $\mathcal{F}_0$ can either contain no isotropic vectors at all or one isotropic vector or be totally isotropic. Observe that $Q(j)$ is contained in the (Zariski) closure of $Q(i)$ if and only if $j \geq i$ and $Q(s)=\mathcal{Q}(s,n)$ is the unique closed orbit; see for example [Barbasch and Evens 1994, Section 3.1] and [Arbarello et al. 1985]. Thus, $\mathcal{Q}(s,n)$ is contained in the closure of each orbit $Q(i)$.

Lastly, from Remark 3.6 we have that $\mathcal{U}$ contains points $(\mathcal{F}_0, t\Lambda)$ with $\mathcal{F}_0 \in \mathcal{Q}(s,n)$ and so $\mathcal{U}$ contains points from all the orbits. Therefore, from all the above we deduce that the $\mathcal{G}$-translates of $\mathcal{U}$ cover $\tau^{-1}(t\Lambda)$. □

Conjecture 3.8. *When $s \geq 1$, the scheme $\mathcal{M}$ is flat over* Spec O_F, *Cohen–Macaulay and normal. It's special fiber is reduced.*

Remark 3.9. (a) By Proposition 3.7, to prove the conjecture, it is enough to show that the affine chart $\mathcal{U}$ has the above properties. More precisely, the hard part of the conjecture is to prove that the special fiber of $\mathcal{U}$ is reduced and Cohen–Macaulay.

(b) For $(r,s)=(n-1,1)$, the conjecture is true as we can see from Remark 3.3.

(c) The above conjecture is supported by some computer calculations that we obtained with the help of Macaulay 2. In particular, we verified the conjecture when $(r,s)=(n-2,2)$ where $n=5,6,7,8,9,10$ for various primes $p>2$.

4. A blow-up

In what follows, we assume $(r,s)=(n-2,2)$. The goal of this section is to find a semistable resolution of the affine chart $\mathcal{U}$ (see Corollary 4.2).

From Corollary 3.4 we have that $\mathcal{U} \cong \mathrm{Spec}\, B$ where B is the quotient ring

$$B = O_F[(x_i)_{3\leq i\leq n}, (y_i)_{3\leq i\leq n}, a, b, c]/J$$

and J is the ideal generated by the entries of the relation:

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} Q(\underline{x}) & P(\underline{x},\underline{y}) \\ P(\underline{x},\underline{y}) & Q(\underline{y}) \end{pmatrix} = 2\pi \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

where $Q(\underline{x}) = 1+\sum_{i=3}^n x_i^2$, $Q(\underline{y}) = 1+\sum_{i=3}^n y_i^2$ and $P(\underline{x},\underline{y}) = \sum_{i=3}^n x_i y_i$.

Now, we will explicitly calculate the blow-up $\mathcal{U}^{\mathrm{bl}}$ of $\mathrm{Spec}(B)$ along the ideal (a, b, c). By Remark 3.5, $\mathcal{U}^{\mathrm{bl}}$ is the blow-up of $\mathcal{U}$ along the smooth subscheme $\mathcal{U} \cap \tau^{-1}(t\Lambda)$. Let $\rho : \mathcal{U}^{\mathrm{bl}} \to \mathcal{U}$ be the blow-up morphism. Define

$$\mathcal{U}' := \mathrm{Proj}(B[t_1, t_2, t_3]/J')$$

where

$$J' = (t_1 Q(\underline{x}) - t_3 Q(\underline{y}), t_2 Q(\underline{y}) + t_1 P(\underline{x}, \underline{y}), t_2 Q(\underline{x}) + t_3 P(\underline{x}, \underline{y}), at_2 - bt_1, at_3 - ct_1, bt_3 - ct_2).$$

By definition, $\mathcal{U}^{\mathrm{bl}}$ is a closed subscheme of the projective $\mathrm{Spec}(B)$-scheme $\mathcal{U}'$ (as the blow-up $\mathcal{U}^{\mathrm{bl}}$ may be, a priori, cut out by more equations). In fact, as a result of our analysis we will see that $\mathcal{U}^{\mathrm{bl}} = \mathcal{U}'$.

Proposition 4.1. (a) *$\mathcal{U}'$ has semistable reduction over O_F.*

(b) *The closed immersion $\mathcal{U}^{\mathrm{bl}} \to \mathcal{U}'$ is an isomorphism.*

Proof. There are three affine patches that cover $\mathcal{U}'$: For $t_1 = 1$ the affine open chart is given by $V(J_1) = \mathrm{Spec}\, R_1/J_1$ where $R_1 = O_F[(x_i)_{3 \le i \le n}, (y_i)_{3 \le i \le n}, a, t_2, t_3]$ and

$$J_1 = (t_2 Q(\underline{y}) + P(\underline{x}, \underline{y}), Q(\underline{x}) - t_3 Q(\underline{y}), a(Q(\underline{x}) + t_2 P(\underline{x}, \underline{y})) - 2\pi).$$

We will show that the scheme $V(J_1)$ has semistable reduction over O_F. It suffices to prove that $V(J_1)$ is regular and its special fiber is reduced with smooth irreducible components that have smooth intersections with correct dimensions. First we observe that

$$J_1 = (t_2 Q(\underline{y}) + P(\underline{x}, \underline{y}), Q(\underline{x}) - t_3 Q(\underline{y}), a(t_3 - t_2^2) Q(\underline{y}) - 2\pi).$$

Over the special fiber ($\pi = 0$) we have $V(\bar{J}_1) = \mathrm{Spec}\, \bar{R}_1/\bar{J}_1$ where

$$\bar{J}_1 = (t_2 Q(\underline{y}) + P(\underline{x}, \underline{y}), \quad Q(\underline{x}) - t_3 Q(\underline{y}), \quad a(t_3 - t_2^2) Q(\underline{y}))$$

and $\bar{R}_1 = k[(x_i)_{3 \le i \le n}, (y_i)_{3 \le i \le n}, a, t_2, t_3]$. Let $V(I_i) = \mathrm{Spec}\, \bar{R}_1/I_i$ of dimension $2(n-2)$, where

$$\begin{aligned} I_1 &= (a, t_2 Q(\underline{y}) + P(\underline{x}, \underline{y}), Q(\underline{x}) - t_3 Q(\underline{y})), \\ I_2 &= (t_3 - t_2^2, t_2 Q(\underline{y}) + P(\underline{x}, \underline{y}), Q(\underline{x}) - t_2^2 Q(\underline{y})), \\ I_3 &= (Q(\underline{y}), P(\underline{x}, \underline{y}), Q(\underline{x})). \end{aligned}$$

We can easily see that

$$V(\bar{J}_1) = V(I_1) \cup V(I_2) \cup V(I_3).$$

Using the Jacobi criterion we see that $V(I_i)$ are smooth and that their intersections

$$\begin{aligned} I_1 + I_2 &= (a, t_3 - t_2^2, t_2 Q(\underline{y}) + P(\underline{x}, \underline{y}), Q(\underline{x}) - t_2^2 Q(\underline{y})), \\ I_1 + I_3 &= (a, Q(\underline{y}), P(\underline{x}, \underline{y}), Q(\underline{x})), \\ I_2 + I_3 &= (t_3 - t_2^2, Q(\underline{y}), P(\underline{x}, \underline{y}), Q(\underline{x})), \\ I_1 + I_2 + I_3 &= (a, t_3 - t_2^2, Q(\underline{y}), P(\underline{x}, \underline{y}), Q(\underline{x})), \end{aligned}$$

are also smooth and with the correct dimensions. By the above, we get that $V(I_i)$ are the smooth irreducible components of $V(\bar{J}_1)$.

Now, we prove that the special fiber of $V(J_1)$ is reduced by showing that

$$\bar{J}_1 = I_1 \cap I_2 \cap I_3.$$

Recall that $\bar{J}_1 = (m_1, m_2, a(t_3 - t_2^2)Q(\underline{y}))$ where $m_1 := t_2 Q(\underline{y}) + P(\underline{x}, \underline{y})$ and $m_2 := Q(\underline{x}) - t_3 Q(\underline{y})$. Clearly, $\bar{J}_1 \subset I_1 \cap I_2 \cap I_3$. Let $g \in I_1 \cap I_2 \cap I_3$. Thus, $g \in I_1$ and

$$g = f_1 a + f_2 m_1 + f_3 m_2 \equiv f_1 a \mod \bar{J}_1$$

for $f_i \in \bar{R}_1$. Also, $g \in I_2$ and so $f_1 a \in I_2$. I_2 is a prime ideal and $a \notin I_2$. Thus, $f_1 \in I_2$ and

$$f_1 = h_1(t_3 - t_2^2) + h_2 m_1 + h_3 m_2 \equiv h_1(t_3 - t_2^2) \mod \bar{J}_1$$

for $h_i \in \bar{R}_1$. Lastly, $g \in I_3$ and from the above we obtain $h_1 \in I_3$ as a, $(t_3 - t_2^2) \notin I_3$. Thus,

$$h_1 = k_1 Q(\underline{y}) + k_2 P(\underline{x}, \underline{y}) + k_3 Q(\underline{x}) \equiv Q(\underline{y})(k_1 - k_2 t_2 + k_3 t_3) \mod \bar{J}_1$$

for $k_i \in \bar{R}_1$. Therefore, $g \equiv a(t_3 - t_2^2)Q(\underline{y})(k_1 - k_2 t_2 + k_3 t_3) \equiv 0 \mod \bar{J}_1$ and so $g \in \bar{J}_1$. Hence, $\bar{J}_1 = I_1 \cap I_2 \cap I_3$.

Next, we can easily see that the ideals I_1, I_2, I_3 are principal over $V(J_1)$. In particular, for I_1 we have $I_1 = (a)$, for I_2 we have $I_2 = (t_3 - t^2)$ and for I_3 we get $I_3 = (Q(\underline{y}))$. From the above we deduce that $V(J_1)$ is regular; see [Hartl 2001, Remark 1.1.1].

From all the above discussion we deduce that $V(J_1)$ has semistable reduction over O. By symmetry, we get similar results for $t_3 = 1$.

For $t_2 = 1$, the affine open chart is given by $V(J_2) = \operatorname{Spec} R_2/J_2$ where

$$R_2 = O_F[(x_i)_{3 \le i \le n}, (y_i)_{3 \le i \le n}, b, t_1, t_3]$$

and

$$J_2 = (Q(\underline{y}) + t_1 P(\underline{x}, \underline{y}), Q(\underline{x}) + t_3 P(\underline{x}, \underline{y}), b(1 - t_1 t_3)P(\underline{x}, \underline{y}) - 2\pi).$$

To show that $V(J_2)$ has semistable reduction one proceeds exactly as above. In this case, the special fiber of $V(J_2)$ is isomorphic to $\operatorname{Spec} \bar{R}_2/\bar{J}_2$ where

$$\bar{J}_2 = (Q(\underline{y}) + t_1 P(\underline{x}, \underline{y}), \quad Q(\underline{x}) + t_3 P(\underline{x}, \underline{y}), \quad b(1 - t_1 t_3)P(\underline{x}, \underline{y}))$$

and $\bar{R}_2 = k[(x_i)_{3 \le i \le n}, (y_i)_{3 \le i \le n}, b, t_1, t_3]$. Let $V(I_i') = \operatorname{Spec} \bar{R}_2/I_i'$ of dimension $2(n-2)$, where

$$I_1' = (b, Q(\underline{y}) + t_1 P(\underline{x}, \underline{y}), Q(\underline{x}) + t_3 P(\underline{x}, \underline{y})),$$
$$I_2' = (1 - t_1 t_3, Q(\underline{y}) + t_1 P(\underline{x}, \underline{y}), Q(\underline{x}) + t_3 P(\underline{x}, \underline{y})),$$
$$I_3' = (P(\underline{x}, \underline{y}), Q(\underline{y}), Q(\underline{x})).$$

and their intersections

$$
\begin{aligned}
I'_1 + I'_2 &= (b,\, 1 - t_1 t_3,\, Q(\underline{y}) + t_1 P(\underline{x}, \underline{y}),\, Q(\underline{x}) + t_3 P(\underline{x}, \underline{y})),\\
I'_1 + I'_3 &= (b,\, Q(\underline{y}),\, P(\underline{x}, \underline{y}),\, Q(\underline{x})),\\
I'_2 + I'_3 &= (1 - t_1 t_3,\, Q(\underline{y}),\, P(\underline{x}, \underline{y}),\, Q(\underline{x})),\\
I'_1 + I'_2 + I'_3 &= (b,\, 1 - t_1 t_3,\, Q(\underline{y}),\, P(\underline{x}, \underline{y}),\, Q(\underline{x})).
\end{aligned}
$$

As in the case $t_1 = 1$, by using the Jacobi criterion we see that the irreducible components $V(I'_i)$ are smooth and they intersect transversely. Also, by a similar argument as above we can easily see that $V(J_2)$ is regular and its special fiber is reduced. Now, the semistability of $V(J_2)$ follows.

By the above, we conclude that $\mathcal{U}'$ is regular, of relative dimension $2(n-2)$, that $\mathcal{U}'$ is O_F-flat and that its special fiber is a reduced divisor with normal crossings. This shows part (a). Let us show part (b). The blow-up $\mathcal{U}^{\mathrm{bl}}$ is a closed subscheme of $\mathcal{U}'$. By the above, $\mathcal{U}'$ is integral of dimension $2(n-2)$. However, the dimension of the blow-up $\mathcal{U}^{\mathrm{bl}}$ is also $2(n-2)$. Indeed, on one hand $\mathcal{U}^{\mathrm{bl}}$ is a closed subscheme of $\mathcal{U}'$ while on the other hand it is birational to $\mathrm{Spec}(B)$. We deduce that $\mathcal{U}^{\mathrm{bl}} = \mathcal{U}'$ which is the claim in (b). □

As a consequence of the above proposition we obtain:

Corollary 4.2. *The morphism $\rho : \mathcal{U}^{\mathrm{bl}} \to \mathcal{U}$ is a semistable resolution, i.e., $\mathcal{U}^{\mathrm{bl}}$ has semistable reduction over O_F.*

Proof. It follows from part (a) and (b) of Proposition 4.1. □

Remark 4.3. From the proof of Proposition 4.1 we obtain that the special fiber of $\mathcal{U}^{\mathrm{bl}}$ has three irreducible components. In fact, we explicitly describe the equations defining these irreducible components over the three affine patches that cover $\mathcal{U}^{\mathrm{bl}}$. It is then easy to see that the exceptional locus of $\rho : \mathcal{U}^{\mathrm{bl}} \to \mathcal{U}$ is the irreducible component of the special fiber of $\mathcal{U}^{\mathrm{bl}}$

$$
\mathrm{Proj}\left(\frac{k[(x_i)_{3 \le i \le n}, (y_i)_{3 \le i \le n}][t_1, t_2, t_3]}{(t_1 Q(\underline{x}) - t_3 Q(\underline{y}),\, t_2 Q(\underline{y}) + t_1 P(\underline{x}, \underline{y}),\, t_2 Q(\underline{x}) + t_3 P(\underline{x}, \underline{y}))} \right)
$$

that corresponds to $V(I_1)$ and $V(I'_1)$ for the affine patches $t_1 = 1$ and $t_2 = 1$ respectively.

5. A resolution for the local model

We use the notation from Section 2. In particular, recall the morphism

$$
\tau : \mathcal{M} \to \mathrm{M}^{\wedge} \otimes_O O_F
$$

and the following isomorphisms over the generic fiber

$$
\mathcal{M} \otimes F \cong \mathrm{M}^{\wedge} \otimes F \cong \mathrm{M}^{\mathrm{loc}} \otimes F. \tag{5.1}
$$

Let $\mathcal{Z} = \tau^{-1}(t\Lambda)$ be the smooth $\mathcal{G}$-invariant subscheme of dimension $2(n-2)$, which is supported in the special fiber. (Recall from Section 2 that $t\Lambda$ is the worst point of $\mathrm{M}^{\wedge}$ and $\tau^{-1}(t\Lambda) \cong \mathrm{Gr}(2, n) \otimes k$.)

We consider the blow-up of $\mathcal{M}$ along the subscheme $\mathcal{Z}$. This gives a $\mathcal{G}$-birational projective morphism

$$r^{\mathrm{bl}} : \mathrm{M}^{\mathrm{bl}} \to \mathcal{M}$$

which induces an isomorphism on the generic fibers.

Theorem 5.1. *The scheme* M^{bl} *is regular and has special fiber a reduced divisor with normal crossings.*

Proof. From Proposition 3.7 we have that the $\mathcal{G}$-translates of $\mathcal{U}$ cover $\mathcal{M}$ and since r^{bl} is $\mathcal{G}$-equivariant we obtain that the $\mathcal{G}$-translates of the open $\mathcal{U}^{\mathrm{bl}} = (r^{\mathrm{bl}})^{-1}(\mathcal{U}) \subset \mathrm{M}^{\mathrm{bl}}$ cover M^{bl}. Therefore, it is enough to show the conclusion of the theorem for the blow-up $\mathcal{U}^{\mathrm{bl}}$ of $\mathcal{U}$ at the ideal (a, b, c) and by Corollary 4.2 the proof of the theorem follows. □

Remark 5.2. It would be useful to have a simple moduli-theoretic description of the blow-up M^{bl} similar in spirit to the description of $\mathcal{M}$ given in Section 2.

We just proved that M^{bl} has semistable reduction, and is therefore flat over O_F. Combining all the above we have

$$\mathrm{M}^{\mathrm{bl}} \xrightarrow{r^{\mathrm{bl}}} \mathcal{M} \xrightarrow{\tau} \mathrm{M}^{\wedge} \otimes_O O_F$$

which factors through $\mathrm{M}^{\mathrm{loc}} \otimes_O O_F \subset \mathrm{M}^{\wedge} \otimes_O O_F$ because of flatness; the generic fiber of all of these is the same as we can see from (5.1). Then, we obtain that $\mathrm{M}^{\mathrm{bl}} \to \mathrm{M}^{\mathrm{loc}} \otimes_O O_F$ is a $\mathcal{G}$-equivariant birational projective morphism.

6. Application to Shimura varieties

6A. ***Unitary Shimura data.*** We now discuss some Shimura varieties to which we can apply these results. We follow [Pappas and Rapoport 2009, Section 1.1] for the description of the unitary Shimura varieties; see also [Pappas 2000b, Section 3].

Let F_0 be an imaginary quadratic field and fix an embedding $\epsilon : F_0 \hookrightarrow \mathbb{C}$. Let O be the ring of integers of F_0 and denote by $a \mapsto \bar{a}$ the nontrivial automorphism of F_0. Assuming $n > 3$, we let $W = F_0^n$ be a n-dimensional F_0-vector space, and we suppose that $\phi : W \times W \to F_0$ is a nondegenerate hermitian form. Set $W_{\mathbb{C}} = W \otimes_{F_0,\epsilon} \mathbb{C}$. Choosing a suitable isomorphism $W_{\mathbb{C}} \cong \mathbb{C}^n$ we may write ϕ on $W_{\mathbb{C}}$ in a normal form $\phi(w_1, w_2) = {}^t\bar{w}_1 H w_2$ where

$$H = \mathrm{diag}(-1, \dots, -1, 1, \dots, 1).$$

We denote by s (resp. r) the number of places, where -1, (resp. 1) appears in H. We will say that ϕ has signature (r, s). By replacing ϕ by $-\phi$ if needed, we can make sure that $s \le r$ and so we assume that $s \le r$. Let $J : W_{\mathbb{C}} \to W_{\mathbb{C}}$ be the endomorphism given by the matrix $-\sqrt{-1}H$. We have $J^2 = -\mathrm{id}$ and so the endomorphism J gives an $\mathbb{R}$-algebra homomorphism $h_0 : \mathbb{C} \to \mathrm{End}_{\mathbb{R}}(W \otimes_{\mathbb{Q}} \mathbb{R})$ with $h_0(\sqrt{-1}) = J$ and hence a complex structure on $W \otimes_{\mathbb{Q}} \mathbb{R} = W_{\mathbb{C}}$. For this complex structure we have

$$\mathrm{Tr}_{\mathbb{C}}(a; W \otimes_{\mathbb{Q}} \mathbb{R}) = s \cdot \epsilon(a) + r \cdot \bar{\epsilon}(a), \quad a \in F_0.$$

Denote by E the subfield of $\mathbb{C}$ which is generated by the traces above (the "reflex field"). We have that $E=\mathbb{Q}$ if $r=s$ and $E=F_0$ otherwise. The representation of F_0 on $W\otimes_{\mathbb{Q}}\mathbb{R}$ with the above trace is defined over E, i.e., there is an n-dimensional E-vector space W_0 on which F_0 acts such that

$$\mathrm{Tr}_E(a;\, W_0)=s\cdot a+r\cdot\bar{a}$$

and such that $W_0\otimes_E\mathbb{C}$ together with the above F_0-action is isomorphic to $W\otimes_{\mathbb{Q}}\mathbb{R}$ with the F_0-action induced by $\epsilon : F_0\hookrightarrow\mathbb{C}$ and the above complex structure.

Next, fix a nonzero element $a\in F_0$ with $\bar{a}=-a$ and set

$$\psi(x,y)=\mathrm{Tr}_{F_0/\mathbb{Q}}(a^{-1}\phi(x,y))$$

which is a nondegenerate alternating form $W\otimes_{\mathbb{Q}}W\to\mathbb{Q}$. This satisfies

$$\psi(av,w)=\psi(v,\bar{a}w),\quad \text{for all } a\in F_0,\ v,w\in W.$$

By replacing a by $-a$, we can make sure that the symmetric $\mathbb{R}$-bilinear form on $W_{\mathbb{C}}$ given by $\psi(x,Jy)$ for $x,y\in W_{\mathbb{C}}$ is positive definite. Let G be the reductive group over $\mathbb{Q}$ which is given by

$$G(\mathbb{Q})=\{g\in\mathrm{GL}_{F_0}(W)\mid \psi(gv,gw)=c(g)\psi(v,w),\ c(g)\in\mathbb{Q}^{\times}\}.$$

The group G can be identified with the unitary similitude group of the form ϕ. Set

$$GU(r,s):=\{A\in\mathrm{GL}_n(\mathbb{C})\mid {}^t\bar{A}HA^{-}=c(A)H,\, c(A)\in\mathbb{R}^{\times}\}.$$

By the above, the embedding $\epsilon : F_0\hookrightarrow\mathbb{C}$ induces an isomorphism $G(\mathbb{R})\cong GU(r,s)$. We define a homomorphism $h:\mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_{m,\mathbb{C}}\to G_{\mathbb{R}}$ by restricting h_0 to $\mathbb{C}^{\times}$. Then $h(a)$ for $a\in\mathbb{R}^{\times}$ acts on $W\otimes_{\mathbb{Q}}\mathbb{R}$ by multiplication by a and $h(\sqrt{-1})$ acts as J. Consider $h_{\mathbb{C}}(z,1):\mathbb{C}^{\times}\to G(\mathbb{C})\cong\mathrm{GL}_n(\mathbb{C})\times\mathbb{C}^{\times}$. Up to conjugation $h_{\mathbb{C}}(z,1)$ is given by

$$\mu_{r,s}(z)=(\mathrm{diag}(z^{(s)},1^{(r)}),z);$$

this is a cocharacter of G defined over the number field E. Denote by X_h the conjugation orbit of $h(i)$ under $G(\mathbb{R})$. The pair (G,h) gives rise to a Shimura variety $\mathrm{Sh}(G,h)$ which is defined over the reflex field E.

6B. ***Unitary integral models.*** We continue with the notations and assumptions of the previous paragraph. In particular, we take $G=GU_n$ and $X=X_h$ above that define the unitary similitude Shimura datum (G,X). Assume that $(r,s)=(n-2,2)$.

Assume that p is an odd prime number and is ramified in F_0. Let $F_1=F_0\otimes\mathbb{Q}_p$ and $V=W\otimes_{\mathbb{Q}}\mathbb{Q}_p$. We fix a square root π of p and we set $k=\bar{\mathbb{F}}_p$. In addition, we assume that the hermitian form ϕ on V is split. This means that there exists a basis $e_1,\dots,e_n$ of V such that $\phi(e_i,e_{n+1-j})=\delta_{ij}$ for $i,j\in\{1,\dots,n\}$. We denote by Λ the standard lattice $O^n\otimes_{\mathbb{Z}}\mathbb{Z}_p$ in V. Denote by K the stabilizer of Λ in $G(\mathbb{Q}_p)$.

We let $\mathcal{L}$ be the self-dual multichain consisting of $\{\pi^k\Lambda\}_{k\in\mathbb{Z}}$. Here $\mathcal{G} = \underline{\mathrm{Aut}}(\mathcal{L})$ is the group scheme over $\mathbb{Z}_p$ with $K = \mathcal{G}(\mathbb{Z}_p)$ the subgroup of $G(\mathbb{Q}_p)$ fixing the lattice chain $\mathcal{L}$. Denote by K° the neutral component of K. As in Section 2, when n is odd $K = K^\circ$ and when n is even $K/K^\circ \simeq \mathbb{Z}/2\mathbb{Z}$.

Choose also a sufficiently small compact open subgroup K^p of the prime-to-p finite adelic points $G(\mathbb{A}_f^p)$ of G and set $\boldsymbol{K} = K^pK$ and $\boldsymbol{K}' = K^pK^\circ$. As was observed in [Pappas and Rapoport 2009, Section 1.3], the Shimura varieties $\mathrm{Sh}_{\boldsymbol{K}'}(G, X)$ and $\mathrm{Sh}_{\boldsymbol{K}}(G, X)$ have isomorphic geometric connected components. Therefore, from the point of view of constructing reasonable integral models, we may restrict our attention to $\mathrm{Sh}_{\boldsymbol{K}}(G, X)$; since K corresponds to a lattice set stabilizer, this Shimura variety is given by a simpler moduli problem. The Shimura variety $\mathrm{Sh}_{\boldsymbol{K}}(G, X)$ with complex points

$$\mathrm{Sh}_{\boldsymbol{K}}(G, X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/\boldsymbol{K}$$

is of PEL type and has a canonical model over the reflex field E. We set $\mathcal{O} = O_{E_v}$ where v the unique prime ideal of E above (p).

We consider the moduli functor $\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}}$ over $\operatorname{Spec}\mathcal{O}$ given in [Rapoport and Zink 1996, Definition 6.9]: A point of $\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}}$ with values in the $\operatorname{Spec}\mathcal{O}$-scheme S is the isomorphism class of the following set of data $(A, \bar{\lambda}, \bar{\eta})$:

(1) An $\mathcal{L}$-set of abelian varieties $A = \{A_\Lambda\}$.

(2) A $\mathbb{Q}$-homogeneous principal polarization $\bar{\lambda}$ of the $\mathcal{L}$-set A.

(3) A K^p-level structure

$$\bar{\eta} : H_1(A, \mathbb{A}_f^p) \simeq W \otimes \mathbb{A}_f^p \bmod K^p$$

which respects the bilinear forms on both sides up to a constant in $(\mathbb{A}_f^p)^\times$; see [loc. cit.] for details.

The set A should satisfy the determinant condition (i) of [loc. cit.].

For the definitions of the terms employed here we refer to [6.3–6.8] and [Pappas 2000b, Section 3]. The functor $\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}}$ is representable by a quasiprojective scheme over $\mathcal{O}$. Since the Hasse principle is satisfied for the unitary group, we can see as in [loc. cit.] that there is a natural isomorphism

$$\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}} \otimes_{\mathcal{O}} E_v = \mathrm{Sh}_{\boldsymbol{K}}(G, X) \otimes_E E_v.$$

As is explained in [Rapoport and Zink 1996] and [Pappas 2000b] the naive local model $\mathrm{M}^{\mathrm{naive}}$ is connected to the moduli scheme $\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}}$ via the local model diagram

$$\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}} \xleftarrow{\psi_1} \tilde{\mathcal{A}}^{\mathrm{naive}}_{\boldsymbol{K}} \xrightarrow{\psi_2} \mathrm{M}^{\mathrm{naive}}$$

where the morphism ψ_1 is a $\mathcal{G}$-torsor and ψ_2 is a smooth and $\mathcal{G}$-equivariant morphism. Therefore, there is a relatively representable smooth morphism

$$\mathcal{A}^{\mathrm{naive}}_{\boldsymbol{K}} \to [\mathcal{G}\backslash\mathrm{M}^{\mathrm{naive}}]$$

where the target is the quotient algebraic stack.

As we mentioned in Section 2, the scheme $\mathrm{M}^{\mathrm{naive}}$ is never flat and by the above, the same is true for $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{naive}}$. Denote by $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{flat}}$ the flat closure of $\mathrm{Sh}_{\boldsymbol{K}}(G, X) \otimes_E E_v$ in $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{naive}}$. Recall from Section 2 that the flat closure of $\mathrm{M}^{\mathrm{naive}} \otimes_{\mathcal{O}} E_v$ in $\mathrm{M}^{\mathrm{naive}}$ is by definition the local model $\mathrm{M}^{\mathrm{loc}}$. By the above we can see, as in [Pappas and Rapoport 2009], that there is a relatively representable smooth morphism of relative dimension $\dim(G)$,

$$\mathcal{A}_{\boldsymbol{K}}^{\mathrm{flat}} \to [\mathcal{G} \backslash \mathrm{M}^{\mathrm{loc}}].$$

This of course implies imply that $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{flat}}$ is étale locally isomorphic to the local model $\mathrm{M}^{\mathrm{loc}}$.

One can now consider a variation of the moduli of abelian schemes $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{spl}}$ where we add in the moduli problem an additional subspace in the Hodge filtration $\mathrm{Fil}^0(A) \subset H^1_{\mathrm{dR}}(A)$ of the universal abelian variety A (see [Haines 2005, Section 6.3] for more details) with certain conditions to imitate the definition of the splitting local model $\mathcal{M}$. $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{spl}}$ associates to an O_{F_1}-scheme S the set of isomorphism classes of objects $(A, \bar{\lambda}, \bar{\eta}, \mathscr{F}_0)$. Here $(A, \bar{\lambda}, \bar{\eta})$ is an object of $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{naive}}(S)$. Set $\mathscr{F}_1 := \mathrm{Fil}^0(A)$. The final ingredient $\mathscr{F}_0$ of an object of $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{spl}}$ is the subspace $\mathscr{F}_0 \subset \mathscr{F}_1 \subset H^1_{\mathrm{dR}}(A)$ of rank s which satisfies the following conditions:

$$(t+\pi)\mathscr{F}_1 \subset \mathscr{F}_0, \quad (t-\pi)\mathscr{F}_0 = (0).$$

There is a forgetful morphism

$$\tau : \mathcal{A}_{\boldsymbol{K}}^{\mathrm{spl}} \to \mathcal{A}_{\boldsymbol{K}}^{\mathrm{naive}} \otimes_{\mathcal{O}} O_{F_1}$$

defined by $(A, \bar{\lambda}, \bar{\eta}, \mathscr{F}_0) \mapsto (A, \bar{\lambda}, \bar{\eta})$. Moreover, $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{spl}}$ has the same étale local structure as $\mathcal{M}$; it is a "linear modification" of $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{naive}} \otimes_{\mathcal{O}} O_{F_1}$ in the sense of [Pappas 2000b, Section 2]; see also [Pappas and Rapoport 2005, Section 15]. Also we want to mention that under the local model diagram the subspace $\mathscr{F}_1$ corresponds to $\mathcal{F}_1$ of $(\mathcal{F}_0, \mathcal{F}_1) \in \mathcal{M}$.

Theorem 6.1. *For every K^p as above, there is a scheme $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{bl}}$, flat over* $\mathrm{Spec}(O_{F_1})$, *with*

$$\mathcal{A}_{\boldsymbol{K}}^{\mathrm{bl}} \otimes_{O_{F_1}} F_1 = \mathrm{Sh}_{\boldsymbol{K}}(G, X) \otimes_E F_1,$$

and which supports a local model diagram

$$\begin{array}{ccc} & \widetilde{\mathcal{A}}_{\boldsymbol{K}}^{\mathrm{bl}}(G, X) & \\ {\scriptstyle \pi_K^{\mathrm{reg}}} \swarrow & & \searrow {\scriptstyle q_K^{\mathrm{reg}}} \\ \mathcal{A}_{\boldsymbol{K}}^{\mathrm{bl}} & & \mathrm{M}^{\mathrm{bl}} \end{array} \tag{6B.1}$$

such that:

(a) $\pi_{\boldsymbol{K}}^{\mathrm{reg}}$ *is a $\mathcal{G}$-torsor for the parahoric group scheme $\mathcal{G}$ that corresponds to K_p.*

(b) $q_{\boldsymbol{K}}^{\mathrm{reg}}$ *is smooth and $\mathcal{G}$-equivariant.*

(c) $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{bl}}$ *is regular and has special fiber which is a reduced divisor with normal crossings.*

Proof. By the above, we have

$$\begin{array}{ccc} & \widetilde{\mathcal{A}}_{\boldsymbol{K}}^{\mathrm{spl}} & \\ {}^{\pi_K}\swarrow & & \searrow^{q_K} \\ \mathcal{A}_{\boldsymbol{K}}^{\mathrm{spl}} & & \mathcal{M} \end{array} \tag{6B.2}$$

with $\pi_{\boldsymbol{K}}$ a $\mathcal{G}$-torsor and $q_{\boldsymbol{K}}$ smooth and $\mathcal{G}$-equivariant. We set

$$\widetilde{\mathcal{A}}_{\boldsymbol{K}}^{\mathrm{bl}} = \widetilde{\mathcal{A}}_{\boldsymbol{K}}^{\mathrm{spl}} \times_{\mathcal{M}} \mathrm{M}^{\mathrm{bl}}$$

which carries a diagonal $\mathcal{G}$-action. Since $\mathrm{M}^{\mathrm{bl}} \to \mathcal{M}$ is given by a blow-up, is projective, and we can see [Pappas 2000b, Section 2] that the quotient

$$\pi_K^{\mathrm{reg}} : \widetilde{\mathcal{A}}_{\boldsymbol{K}}^{\mathrm{bl}} \to \mathcal{A}_{\boldsymbol{K}}^{\mathrm{bl}} := \mathcal{G}\backslash\widetilde{\mathcal{A}}_K^{\mathrm{bl}}(G, X)$$

is represented by a scheme and gives a $\mathcal{G}$-torsor. (This is an example of a linear modification, see [Pappas 2000b, Section 2].) In fact, since blowing-up commutes with étale localization, $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{bl}}$ is the blow-up of $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{spl}}$ along the locus of its special fiber where $t\mathscr{F}_1 = 0$. The projection gives a smooth $\mathcal{G}$-morphism

$$q_{\boldsymbol{K}}^{\mathrm{reg}} : \widetilde{\mathcal{A}}_{\boldsymbol{K}}^{\mathrm{bl}} \to \mathrm{M}^{\mathrm{bl}}$$

which completes the local model diagram. Property (c) follows from Theorem 5.1 and properties (a) and (b) which imply that $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{bl}}$ and M^{bl} are locally isomorphic for the étale topology. □

Corollary 6.2. *$\mathcal{A}_{\boldsymbol{K}}^{\mathrm{bl}}$ is the blow-up of $\mathcal{A}_{\boldsymbol{K}}^{\mathrm{spl}}$ along the locus of its special fiber where the deRham filtration $\mathscr{F}_1 = \mathrm{Fil}^0(A)$ is annihilated by the action of the uniformizer π.*

Proof. It follows from the proof of the above theorem. □

Remarks 6.3. (1) From the above discussion, we can obtain a semistable integral model for the Shimura variety $\mathrm{Sh}_{\boldsymbol{K}'}(G, X)$ where $\boldsymbol{K}' = K^p K^\circ$. In this case, the corresponding local models $\mathrm{M}^{\mathrm{loc}}$ of $\mathrm{Sh}_{\boldsymbol{K}'}(G, X)$ agree with the Pappas–Zhu local models $\mathbb{M}_{K^\circ}(G, \mu_{r,s})$ for the local model triples $(G, \{\mu_{r,s}\}, K^\circ)$; see [Pappas and Zhu 2013, Theorem 1.2] and [Pappas and Zhu 2013, Section 8] for more details.

(2) Similar results can be obtained for corresponding Rapoport–Zink formal schemes; see [He et al. 2020, Section 4] for an example of this parallel treatment.

Acknowledgements

I thank G. Pappas for his valuable comments, insights and corrections on a preliminary version of this article. I also thank B. Howard and K. Madapusi Pera for useful suggestions and the referee for their careful reading that lead to several corrections.

References

[Arbarello et al. 1985] E. Arbarello, M. Cornalba, P. A. Griffiths, and J. Harris, *Geometry of algebraic curves, I*, Grundl. Math. Wissen. **267**, Springer, 1985. MR Zbl

[Arzdorf 2009] K. Arzdorf, "On local models with special parahoric level structure", *Michigan Math. J.* **58**:3 (2009), 683–710. MR Zbl

[Barbasch and Evens 1994] D. Barbasch and S. Evens, "K-orbits on Grassmannians and a PRV conjecture for real groups", *J. Algebra* **167**:2 (1994), 258–283. MR Zbl

[Bruinier et al. 2020] J. H. Bruinier, B. Howard, S. S. Kudla, M. Rapoport, and T. Yang, "Modularity of generating series of divisors on unitary Shimura varieties", pp. 7–125 in *Arithmetic divisors on orthogonal and unitary Shimura varieties*, Astérisque **421**, Soc. Math. France, Paris, 2020. MR Zbl

[Faltings 1997] G. Faltings, "Explicit resolution of local singularities of moduli-spaces", *J. Reine Angew. Math.* **483** (1997), 183–196. MR Zbl

[Genestier 2000] A. Genestier, "Un modèle semi-stable de la variété de Siegel de genre 3 avec structures de niveau de type $\Gamma_0(p)$", *Compos. Math.* **123**:3 (2000), 303–328. MR Zbl

[Haines 2005] T. J. Haines, "Introduction to Shimura varieties with bad reduction of parahoric type", pp. 583–642 in *Harmonic analysis*, *the trace formula*, *and Shimura varieties* (Toronto, ON, 2003), edited by J. Arthur et al., Clay Math. Proc. **4**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Hartl 2001] U. T. Hartl, "Semi-stability and base change", *Arch. Math.* (*Basel*) **77**:3 (2001), 215–221. MR Zbl

[He et al. 2020] X. He, G. Pappas, and M. Rapoport, "Good and semi-stable reductions of Shimura varieties", *J. Éc. Polytech. Math.* **7** (2020), 497–571. MR Zbl

[He et al. 2023] Q. He, C. Li, Y. Shi, and T. Yang, "A proof of the Kudla–Rapoport conjecture for Krämer models", *Invent. Math.* **234**:2 (2023), 721–817. MR Zbl

[de Jong 1993] A. J. de Jong, "The moduli spaces of principally polarized abelian varieties with $\Gamma_0(p)$-level structure", *J. Algebraic Geom.* **2**:4 (1993), 667–688. MR Zbl

[Krämer 2003] N. Krämer, "Local models for ramified unitary groups", *Abh. Math. Sem. Univ. Hamburg* **73** (2003), 67–80. MR Zbl

[Pappas 2000a] G. Pappas, "Letter to M. Rapoport", September 2000.

[Pappas 2000b] G. Pappas, "On the arithmetic moduli schemes of PEL Shimura varieties", *J. Algebraic Geom.* **9**:3 (2000), 577–605. MR Zbl

[Pappas and Rapoport 2005] G. Pappas and M. Rapoport, "Local models in the ramified case, II: Splitting models", *Duke Math. J.* **127**:2 (2005), 193–250. MR Zbl

[Pappas and Rapoport 2008] G. Pappas and M. Rapoport, "Twisted loop groups and their affine flag varieties", *Adv. Math.* **219**:1 (2008), 118–198. MR Zbl

[Pappas and Rapoport 2009] G. Pappas and M. Rapoport, "Local models in the ramified case, III: Unitary groups", *J. Inst. Math. Jussieu* **8**:3 (2009), 507–564. MR Zbl

[Pappas and Zachos 2022] G. Pappas and I. Zachos, "Regular integral models for Shimura varieties of orthogonal type", *Compos. Math.* **158**:4 (2022), 831–867. MR Zbl

[Pappas and Zhu 2013] G. Pappas and X. Zhu, "Local models of Shimura varieties and a conjecture of Kottwitz", *Invent. Math.* **194**:1 (2013), 147–254. MR Zbl

[Rapoport and Zink 1996] M. Rapoport and T. Zink, *Period spaces for p-divisible groups*, Ann. of Math. Stud. **141**, Princeton Univ. Press, 1996. MR Zbl

[Smithling 2011] B. D. Smithling, "Topological flatness of local models for ramified unitary groups, I: The odd dimensional case", *Adv. Math.* **226**:4 (2011), 3160–3190. MR Zbl

[Tits 1979] J. Tits, "Reductive groups over local fields", pp. 29–69 in *Automorphic forms*, *representations and L-functions*, *I* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979. MR Zbl

[Zhang 2021] W. Zhang, "Weil representation and arithmetic fundamental lemma", *Ann. of Math.* (2) **193**:3 (2021), 863–978. MR Zbl

Communicated by Shou-Wu Zhang
Received 2022-11-02 Revised 2023-05-09 Accepted 2023-10-12

zachosi@bc.edu *Deptartment of Mathematics, Boston College, Boston, MA, United States*

ALGEBRA AND NUMBER THEORY 18:9 (2024)
https://doi.org/10.2140/ant.2024.18.1737

A unipotent realization of the chromatic quasisymmetric function

Lucas Gagnon

We realize two families of combinatorial symmetric functions via the complex character theory of the finite general linear group $\mathrm{GL}_n(\mathbb{F}_q)$: chromatic quasisymmetric functions and vertical strip LLT polynomials. The associated $\mathrm{GL}_n(\mathbb{F}_q)$ characters are elementary in nature and can be obtained by induction from certain well-behaved characters of the unipotent upper triangular groups $\mathrm{UT}_n(\mathbb{F}_q)$. The proof of these results also gives a general Hopf algebraic approach to computing the induction map. Additional results include a connection between the relevant $\mathrm{GL}_n(\mathbb{F}_q)$ characters and Hessenberg varieties and a reinterpretation of known theorems and conjectures about the relevant symmetric functions in terms of $\mathrm{GL}_n(\mathbb{F}_q)$.

1. Introduction

The chromatic symmetric function sits at a nexus of disparate areas of mathematics. At face value, this symmetric function encodes the coloring problem of a graph as an analogue of the chromatic polynomial [40]. However, through a well-known equivalence between the ring of symmetric functions and the representation theory of the symmetric groups (see, e.g., [35]), some chromatic symmetric functions are also complex characters of the symmetric group [21]. By way of a t-analogue known as the chromatic quasisymmetric function, Brosnan and Chow [10] and Guay-Paquet [25] independently proved that the characters corresponding to indifference graphs are afforded by symmetric group representations on the cohomology rings of regular semisimple Hessenberg varieties, as predicted by a conjecture of Shareshian and Wachs [39]. Thus, certain questions about graphs, representation theory, and algebraic geometry coincide in the combinatorics of these symmetric functions, and vice versa.

At about the same time, a sequence of superficially unrelated developments occurred in the character theory of the group of unipotent upper triangular matrices UT_n over a finite field $\mathbb{F}_q$. Unlike the symmetric group, the conjugacy classes and irreducible characters of UT_n are exceptionally complicated and cannot be described with modern combinatorial tools [26]. However, beginning with the work of André [8], a theory of well-behaved reducible characters — known a supercharacters — has developed, leading to a combinatorial representation theory of UT_n without irreducible characters, as in [3; 4]. A recent example given by Aliniaeifard and Thiem [7] constructs supercharacters which are imbued with Catalan combinatorics coming from a family of normal subgroups of UT_n. These same subgroups and

MSC2020: primary 05E05, 05E10, 16T30; secondary 05E14, 20C33.
Keywords: chromatic quasisymmetric function, LLT polynomial, unipotent group, combinatorial Hopf algebra, supercharacter.

supercharacters will appear in this paper, where they will be indexed by indifference graphs using a canonical bijection between Catalan-enumerated objects.

This paper uses the representation theory of the general linear group GL_n over $\mathbb{F}_q$ to establish a connection between the supercharacter theory of UT_n and the chromatic (quasi)symmetric function. Both UT_n and its subgroups are contained in GL_n. The main result, Theorem 3.1, shows that up to a factor of $(q-1)^n$, inducing the trivial character from each of these subgroups gives a map

$$\left\{\begin{matrix}\text{indifference graph}\\ \text{indexed subgroups}\end{matrix}\right\} \xrightarrow{\mathrm{Ind}_{(-)}^{\mathrm{GL}_n}(\mathbb{1})} \left\{\begin{matrix}\text{chromatic quasisymmetric functions for}\\ \text{indifference graphs evaluated at } t=q\end{matrix}\right\},$$

using an implicit identification between characters of GL_n with unipotent support and symmetric functions coming from the Hall algebra; more details can be found in Section 3. This result is a $\mathrm{GL}_n(\mathbb{F}_q)$-analogue of the Brosnan–Chow–Guay-Paquet theorem, in which cohomology rings are replaced by a permutation representation on the cosets of certain unipotent subgroups.

The remaining sections of the paper explore the implications of the main result for the theory of chromatic quasisymmetric functions. Many of these consequences are reminiscent of consequences of the Brosnan–Chow–Guay-Paquet theorem. Along with Theorem 3.1 itself, these similarities come as a surprise, especially since the association between characters of GL_n and symmetric functions used above is markedly different from the classical association for the symmetric groups. Intuition notwithstanding, each result appears to be straightforward, or even inevitable once the right perspective is achieved.

Section 4 relates Theorem 3.1 to the study of Hessenberg varieties, but not the ones appearing in the Brosnan–Chow–Guay-Paquet theorem. Instead, the values of the GL_n characters in Theorem 3.1 count the points of a nilpotent Hessenberg variety over $\mathbb{F}_q$ associated to an ad-nilpotent ideal. The analogous complex Hessenberg varieties have been studied by Precup and Sommers [38], who found an independent connection to the chromatic quasisymmetric function via Poincaré polynomials. Corollary 4.5 links these results by showing that the Poincaré polynomials for certain complex Hessenberg varieties also count the points of the corresponding Hessenberg variety over $\mathbb{F}_q$.

The chromatic quasisymmetric functions of indifference graphs are also closely related to another family of symmetric functions known as unicellular LLT polynomials [11] (see also [28]), and Section 5 reframes this relationship as a GL_n representation theoretic one. There is a second, more standard realization of symmetric functions as *unipotent* characters of GL_n, and up to a twist by the involution ω, Theorem 5.1 gives a map

$$\left\{\begin{matrix}\text{indifference graph}\\ \text{indexed subgroups}\end{matrix}\right\} \xrightarrow{\mathrm{proj}_{\mathrm{unipotent}} \circ \mathrm{Ind}_{(-)}^{\mathrm{GL}_n}(\mathbb{1})} \left\{\begin{matrix}\text{unicellular LLT polynomials}\\ \text{evaluated at } t=q\end{matrix}\right\},$$

where $\mathrm{proj}_{\mathrm{unipotent}}$ is the operation which replaces a character of GL_n with the sum of its irreducible unipotent constituents. In fact, by applying the composite map to additional characters of UT_n — including supercharacters — Theorem 5.1 finds the larger family of vertical strip LLT polynomials as unipotent characters of GL_n. These symmetric functions are known to appear in the representation theory of quantum groups [34], affine Hecke algebras [23], and the symmetric groups [25; 28], but this is their first appearance in the representation theory of GL_n.

Finally, both chromatic quasisymmetric functions and LLT polynomials are the subject of "positivity conjectures" which are at least partially open. Such a conjecture postulates that when a particular symmetric function is expressed in a chosen basis, the coefficient of each basis element will be a polynomial in t with nonnegative coefficients. For chromatic quasisymmetric functions, the modified Stanley–Stembridge conjecture [39, Conjecture 1.3] (see also [42]) concerns the elementary basis, and is almost entirely open. For LLT polynomials, positivity in the Schur basis has been established by Grojnowski and Haiman [23], but no "positive" combinatorial formula is known in general [27]. Section 6 describes the meaning of these conjectures — and one more, recently resolved by D'Adderio [13] and Alexandersson and Sulzgruber [6] — in GL_n representation theory. This does not lead to immediate progress on any conjecture, but it may be a useful guide for future work.

The method of proof for Theorems 3.1 and 5.1 may also be of independent interest. At a high level, I am able to translate Guay-Paquet's proof in [25] into the (super)character theory of UT_n and GL_n in such a way that both results follow immediately. However, this translation also gives a more general Hopf algebraic conduit from the combinatorial representation theory of UT_n to that of GL_n. Since matters of UT_n character theory are usually very difficult, the tractability of this approach alone is a significant development. These results begin to answer lingering questions from [3] about the Hopf algebraic enumerative invariants of certain supercharacters of UT_n.

A short summary of the aforementioned framework and the machinery of [25] is given in this paragraph. In [2], Aguiar, Bergeron, and Sottile constructively classify all Hopf algebra homomorphisms from an arbitrary Hopf algebra to the Hopf algebra of symmetric functions Sym using linear functionals of the domain. This generalizes Zelevinsky's theory of PSH algebras, which completely describes the character theory of GL_n by constructing a collection of homomorphisms from a Hopf algebra $\mathrm{cf}(\mathrm{GL}_\bullet)$ of GL_n-class functions to Sym. In [19], I construct an analogous Hopf algebra $\mathrm{cf}(\mathrm{UT}_\bullet)$ on the class functions of UT_n, and show that induction $\mathrm{Ind}_{\mathrm{UT}_n}^{\mathrm{GL}_n}$ induces a Hopf algebra homomorphism to $\mathrm{cf}(\mathrm{GL}_\bullet)$. By composing induction with any of Zelevinsky's maps to Sym, the classification of [2] can be used to describe the induction map itself, and Theorems 3.8 and 5.11 do so. The classification of [2] was also used in [25] to construct the chromatic quasisymmetric function using a Hopf algebra structure on Hessenberg varieties, and I show that this coincides with induction of Catalan supercharacters and related objects.

This Hopf algebraic approach builds on the previously understood relationship between the combinatorics of unipotent subgroups and of finite groups of Lie type, including GL_n [9; 22; 32; 45]. Future work should continue to push this connection: it may be possible to transplant some of the framework in this paper and [19] into other Lie types. In doing so, one might find the generalized LLT polynomials defined in [23], yet-to-be-discovered variants of the chromatic quasisymmetric function, and more nilpotent Hessenberg varieties.

The remainder of the paper is organized as follows. Section 2 describes the general background material for the paper. Section 3 concerns Theorem 3.1 and the chromatic quasisymmetric function, and Section 4 relates these results to Hessenberg varieties. Section 5 concerns Theorem 5.1 and the vertical strip LLT polynomial, and is essentially independent of Sections 3 and 4. Finally, Section 6 connects my results to various positivity conjectures.

2. Preliminaries

This section gives the shared preliminary material for Sections 3 and 5. This includes definitions of each of the relevant Hopf algebras, background material from representation theory and combinatorics, and a short review of the theory of combinatorial Hopf algebras.

2A. *Hopf algebras and (quasi)symmetric functions.* This section will describe the Hopf algebras of quasisymmetric and symmetric functions, and their role as universal objects in the theory of combinatorial Hopf algebras. Throughout this paper, the term "Hopf algebra" will refer to a graded connected Hopf algebra over the field of complex numbers $\mathbb{C}$, and all homomorphisms and sub-Hopf algebras are graded.

A *composition* of $n \in \mathbb{Z}_{\geq 0}$ is a finite (possibly empty) sequence of positive integers $\alpha = (\alpha_1, \ldots, \alpha_k)$ with $\alpha_1 + \cdots + \alpha_k = n$. Call each α_i a *part* of α, and write $\ell(\alpha) = k$ for the number of parts of α. The *monomial quasisymmetric function* associated to the composition α is

$$M_\alpha = \sum_{i_1 < \cdots < i_{\ell(\alpha)}} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_{\ell(\alpha)}}^{\alpha_{\ell(\alpha)}} \in \mathbb{C}[\![\boldsymbol{x}]\!].$$

where $\boldsymbol{x} = \{x_1, x_2, \ldots\}$ is an infinite, totally ordered set of commuting indeterminates. The *Hopf algebra of quasisymmetric functions* is the graded commutative, noncocommutative Hopf algebra

$$\mathcal{Q}\mathrm{Sym} = \mathbb{C}\text{-span}\{M_\alpha \mid \alpha \text{ is a composition}\}.$$

The product of $\mathcal{Q}\mathrm{Sym}$ is inherited from $\mathbb{C}[\![\boldsymbol{x}]\!]$ and the coproduct is given by deconcatenation:

$$\Delta(M_\alpha) = \sum_{\ell(\alpha) \geq k \geq 0} M_{(\alpha_1, \ldots, \alpha_k)} \otimes M_{(\alpha_k, \ldots, \alpha_{\ell(\alpha)})}.$$

A *partition of n* is a composition of n is with nonincreasing parts. Let

$$\mathcal{P} = \bigsqcup_{n \geq 0} \mathcal{P}(n) \quad \text{with } \mathcal{P}(n) = \{\text{partitions of } n\}.$$

The *Hopf algebra of symmetric functions* is the cocommutative sub-Hopf algebra

$$\mathrm{Sym} = \mathbb{C}\text{-span}\{m_\lambda \mid \lambda \in \mathcal{P}\} \subseteq \mathcal{Q}\mathrm{Sym} \quad \text{with } m_\lambda = \sum_{\mathrm{sort}(\alpha) = \lambda} M_\alpha,$$

where $\mathrm{sort}(\alpha)$ denotes the partition obtained by listing the parts of α in nonincreasing order.

Three additional bases of Sym will be used in later sections. The first basis consists of the *elementary symmetric functions* $\{e_\lambda \mid \lambda \in \mathcal{P}\}$ defined by

$$e_\lambda = e_{\lambda_1} \cdots e_{\lambda_\ell} \quad \text{where } e_k = m_{(1^k)}.$$

The second basis comprises the *Schur functions* $\{s_\lambda \mid \lambda \in \mathcal{P}\}$, which I will not define; see [35, I.3]. The final basis consists of the Hall–Littlewood elements $P_\lambda(\boldsymbol{x}; t)$ [35, III.2], which are discussed further in Section 3A.

The antipode of Sym acts as $(-1)^n\omega$ on the n-th graded component, where ω is the involutive automorphism of Sym defined in [35, I.4], given by $\omega(s_\lambda) = s_{\lambda'}$, where λ' denotes the transpose partition of λ: $(\lambda')_i = \#\{j \in [\ell(\lambda)] \mid \lambda_j \geq i\}$ for $1 \leq i \leq \lambda_1$.

2A1. *Combinatorial Hopf algebras.* This section will give an abridged description of the framework for classifying Hopf algebra homomorphisms to $\mathcal{Q}$Sym established in [2]. The original result also includes an explicit formula for any such map, which is omitted from this paper as the relevant maps are already known.

A *combinatorial Hopf algebra* (*CHA*) is a pair (H, ζ) where H is a Hopf algebra and $\zeta : H \to \mathbb{C}$ is an algebra homomorphism, which will be called a *zeta function* in order to avoid confusion with group characters. An important example of a CHA is $\mathcal{Q}$Sym with the *first principal specialization*,

$$(\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1) \qquad \text{with } \ \mathrm{ps}_1 : \mathcal{Q}\mathrm{Sym} \to \mathbb{C}, \quad M_\alpha \mapsto \begin{cases} 1 & \text{if } \ell(\alpha) \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 2.1. The name "first principal specialization" comes from the fact that ps_1 is equivalent to specializing $x_1 = 1$ and $x_i = 0$ for $i > 1$ in any quasisymmetric function.

A *CHA morphism* between combinatorial Hopf algebras (H, ζ) and (H', ζ') is a graded Hopf algebra homomorphism $\Psi : H \to H'$ for which $\zeta = \zeta' \circ \Psi$. For example, the inclusion of Sym into $\mathcal{Q}$Sym gives a CHA morphism $(\mathrm{Sym}, \mathrm{ps}_1) \to (\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1)$.

Theorem 2.2 [2, Theorem 4.1]. *Let (H, ζ) be a combinatorial Hopf algebra. There is a unique CHA morphism*

$$\mathrm{cano} : (H, \zeta) \to (\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1).$$

A consequence of Theorem 2.2 is that for every Hopf algebra H, there is a bijection

$$\{\text{homomorphisms } H \to \mathcal{Q}\mathrm{Sym}\} \longleftrightarrow \{\text{combinatorial Hopf algebras } (H, \zeta)\},$$
$$\Psi \mapsto (H, \mathrm{ps}_1 \circ \Psi), \qquad \mathrm{cano} \leftarrow\!\!\dashv (H, \zeta),$$

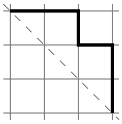
where cano refers to the Hopf algebra homomorphisms underlying the CHA morphism in Theorem 2.2. This paper will frequently appeal to this bijective interpretation.

2B. ***Dyck paths and related objects.*** The results of this paper build on the combinatorics of Dyck paths, indifference graphs, and Schröder paths, each of which are described in this section.

A *Dyck path* of size $n \geq 0$ is a lattice path consisting of $2n$ steps east $E = (1, 0)$ and south $S = (0, -1)$ from $(0, 0)$ to $(n, -n)$ which does not go below the main diagonal $y = -x$. Let

$$\mathcal{D} = \bigsqcup_{n \geq 0} \mathcal{D}_n \quad \text{ with } \mathcal{D}_n = \{\text{Dyck paths of size } n\}.$$

For example,



$$= (EESESS) \in \mathcal{D}_3. \tag{2.3}$$

It is well known that the size of $\mathcal{D}_n$ is the n-th Catalan number, $\frac{1}{n+1}\binom{2n}{n}$; see, for instance, [41].

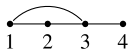
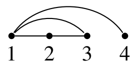
An *indifference graph* of size $n \geq 0$ is a simple, undirected graph γ with vertex set $[n] = \{1, \dots, n\}$ and edge set $E(\gamma)$ satisfying

$$\text{for each } \{i < l\} \in E(\gamma), \quad \{\{j, k\} \mid i \leq j < k \leq l\} \subseteq E(\gamma).$$

The empty graph on $\varnothing$ is the unique indifference graph of size zero. Let

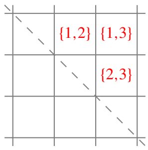
$$\mathcal{IG} = \bigsqcup_{n \geq 0} \mathcal{IG}_n \quad \text{with } \mathcal{IG}_n = \{\text{indifference graphs on } [n]\}.$$

For example,

$\gamma =$  $\in \mathcal{IG}_4$ but $\sigma =$  $\notin \mathcal{IG}_4$,

as $\{1, 4\} \in E(\sigma)$ but $\{3, 4\} \notin E(\sigma)$.

There is a size-preserving bijection between Dyck paths and indifference graphs. Label the unit squares above $y = -x$ in the fourth quadrant of $\mathbb{Z} \times \mathbb{Z}$ by edges so that the square with lower right corner $(j, -i)$ is labeled by $\{i, j\}$; for example,



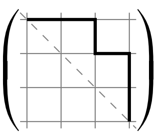
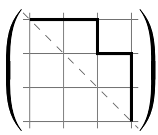
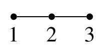
shows the first three of these unit squares with their labels. For any Dyck path π, let

$$\text{Area}(\pi) = \{\{i, j\} \mid \text{the unit square } \{i, j\} \text{ is below } \pi\}$$

and if π has size n, define the *graph of* π to be

$$\text{Graph}(\pi) = ([n], \text{Area}(\pi)).$$

For example, taking the Dyck path in (2.3),

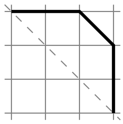
Area  $= \{\{1, 2\}, \{2, 3\}\}$ and Graph  $=$  . (2.4)

Proposition 2.5 [41, Solution 187]. *For $n \geq 0$, the map $\pi \mapsto \text{Graph}(\pi)$ is a bijection from $\mathcal{D}_n$ to $\mathcal{IG}_n$.*

Remark 2.6. Both $\mathcal{D}_n$ and $\mathcal{IG}_n$ also correspond to the family of integer partitions bounded termwise by $(n-1, \dots, 2, 1)$ [41, Item 167]. Reflecting a Dyck path π across $y = -x$ gives the Ferrer's shape (in

French notation) of such a partition, and the edges of $\mathrm{Graph}(\pi)$ are the excluded squares. For example, the objects in (2.4) correspond to $\lambda = (1)$.

A common generalization of Dyck paths will appear in Sections 5 and 6. A *Schröder path* of size $n \geq 0$ is a lattice path from $(0, 0)$ to $(n, -n)$ consisting of steps E, S, and $D = (1, -1)$ that never goes below the main diagonal. Thus, every Dyck path is a Schröder path, but there are more Schröder paths, for example



$$= (EEDSS). \tag{2.7}$$

Say that a Schröder path σ is *tall* if σ has no D steps along the main diagonal. Let

$$\mathcal{TS} = \bigsqcup_{n \geq 0} \mathcal{TS}_n \quad \text{with } \mathcal{TS}_n = \{\text{tall Schröder paths of size } n\}.$$

The Schröder path in (2.7) above is tall, as is any Dyck path, taken as a Schröder path. The number of tall Schröder paths by size is given by the small Schröder numbers, [37, A001003].

Finally, for any tall Schröder path $\sigma \in \mathcal{TS}$, define

$$\mathrm{Area}(\sigma) = \{\{i, j\} \mid \text{the unit square } \{i, j\} \text{ is completely below } \sigma\}$$

and

$$\mathrm{Diag}(\sigma) = \{\{i, j\} \mid \sigma \text{ has a diagonal step through the unit square } \{i, j\}\},$$

so that taking σ as in (2.7) gives $\mathrm{Area}(\sigma) = \{\{1, 2\}, \{2, 3\}\}$ and $\mathrm{Diag}(\sigma) = \{\{1, 3\}\}$.

2C. ***Supercharacter theory.*** Let G be a finite group, let $\mathrm{Irr}(G)$ denote the irreducible complex characters of G, and let $\mathsf{cf}(G)$ denote the space of complex-valued class functions on G. The set $\mathrm{Irr}(G)$ is an orthonormal basis for $\mathsf{cf}(G)$ under the inner product $\langle \cdot, \cdot \rangle : \mathsf{cf}(G) \otimes \mathsf{cf}(G) \to \mathbb{C}$ defined by

$$\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \chi(g) \overline{\psi(g)},$$

where $\overline{\psi(g)}$ denotes the complex conjugate of $\psi(g)$.

Following Diaconis and Isaacs [15], a *supercharacter theory* $(\mathtt{Cl}, \mathtt{Ch})$ of G comprises a set partition $\mathtt{Cl}$ of G and a basis of orthogonal characters $\mathtt{Ch}$ for the space

$$\mathsf{scf}(G) = \{\phi : G \to \mathbb{C} \mid \phi \text{ is constant on each part of } \mathtt{Cl}\},$$

such that $\mathsf{scf}(G)$ contains the regular character reg_G. Since

$$\mathrm{reg}_G(g) = \begin{cases} |G| & \text{if } g = 1_G, \\ 0 & \text{otherwise}, \end{cases}$$

the final condition above is equivalent to $\{1_G\} \in \mathtt{Cl}$.

The elements of Cl and Ch are respectively called *superclasses* and *supercharacters*. Every group has at least one supercharacter theory, with superclasses given by conjugacy classes and supercharacters given by irreducible characters, and in this case $\mathsf{scf}(G) = \mathsf{cf}(G)$.

Each supercharacter theory of G comes with two canonical bases: the supercharacters in Ch and the set of *superclass identifier functions*

$$\{\delta_K \mid K \in \mathtt{Cl}\} \quad \text{with} \quad \delta_K(g) = \begin{cases} 1 & \text{if } g \in K, \\ 0 & \text{otherwise.} \end{cases} \tag{2.8}$$

These bases are each orthogonal. For any $\chi \in \mathsf{scf}(G)$, define an element $\chi\rangle \in \mathsf{scf}(G)^*$ by

$$\chi\rangle : \mathsf{scf}(G) \to \mathbb{C}, \qquad \psi \mapsto \langle \psi, \chi \rangle, \tag{2.9}$$

so that $\mathsf{scf}(G)^* = \{\chi\rangle \mid \chi \in \mathsf{scf}(G)\}$.

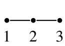
The rest of the section describes a particular collection of supercharacter theories originating in the work of Aliniaeifard and Thiem [7]. Fix a prime power q, let $\mathbb{F}_q$ denote the field with q elements, and let $\mathrm{GL}_n = \mathrm{GL}_n(\mathbb{F}_q)$. The *unipotent upper triangular group* is the subgroup

$$\mathrm{UT}_n = \{g \in \mathrm{GL}_n \mid (g - 1_n)_{i,j} \neq 0 \text{ only if } i < j\}$$

where 1_n denotes the $n \times n$ identity matrix. This group has a family of normal subgroups — called *normal pattern subgroups* — indexed by indifference graphs [36, Lemma 4.1]: for $\gamma \in \mathcal{IG}_n$, let

$$\mathrm{UT}_\gamma = \{g \in \mathrm{UT}_n \mid g_{i,j} = 0 \text{ if } \{i, j\} \in E(\gamma)\}$$

where $E(\gamma)$ denotes the edge set of γ. If $\pi \in \mathcal{D}_n$ is the Dyck path for which $\gamma = \mathrm{Graph}(\pi)$, UT_γ can be visualized in terms of π: UT_γ is the subset of elements of UT_n with nonzero entries occurring only on the diagonal or above the path π. For example, using the graph and Dyck path from (2.4),

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

The paper [7] also shows that the set $\{\mathrm{UT}_\gamma \mid \gamma \in \mathcal{IG}_n\}$ is a lattice under containment. This order is dual to the spanning subgraph relation on $\mathcal{IG}_n$, in that the containment $\mathrm{UT}_\gamma \subseteq \mathrm{UT}_\sigma$ holds if and only if σ is a spanning subgraph of γ. The top element of this lattice is UT_n, corresponding to the edgeless graph $([n], \varnothing)$, and $|\mathrm{UT}_n : \mathrm{UT}_\gamma| = q^{|E(\gamma)|}$ for all $\gamma \in \mathcal{IG}_n$.

The lattice structure on normal pattern subgroups partitions the set UT_n into parts

$$\mathrm{UT}^\circ_\gamma = \{g \in \mathrm{UT}_\gamma \mid g \notin \mathrm{UT}_\sigma \text{ for any } \sigma \supsetneq \gamma\}$$

for each $\gamma \in \mathcal{IG}_n$. Similarly, $\mathcal{IG}_n$ indexes the parts of a partition of the set of irreducible characters $\mathrm{Irr}(\mathrm{UT}_n)$ of UT_n: let

$$\widehat{\mathrm{UT}}^\circ_\gamma = \{\psi \in \mathrm{Irr}(\mathrm{UT}_n) \mid \mathrm{UT}_\gamma \subseteq \ker(\psi) \text{ and } \mathrm{UT}_\sigma \not\subseteq \ker(\psi) \text{ for each } \sigma \supsetneq \gamma\}.$$

for each $\gamma \in \mathcal{IG}_n$, and further define

$$\chi^\gamma = \sum_{\psi \in \widehat{\mathrm{UT}}^\circ_\gamma} \psi(1)\psi.$$

Proposition 2.10 [7, Section 3.2]. *With*

$$\mathtt{Ch} = \{\mathrm{UT}^\circ_\gamma \mid \gamma \in \mathcal{IG}_n\} \quad \textit{and} \quad \mathtt{Cl} = \{\chi^\gamma \mid \gamma \in \mathcal{IG}_n\},$$

the pair $(\mathtt{Cl}, \mathtt{Ch})$ *is a supercharacter theory of* UT_n.

For the remainder of the paper, write $\delta_\gamma = \delta_{\mathrm{UT}^\circ_\gamma}$ for the superclass identifier functions in this supercharacter theory. In addition to these functions and the supercharacters, the space $\mathsf{scf}(\mathrm{UT}_n)$ has two interesting bases: $\{\bar{\delta}_\gamma \mid \gamma \in \mathcal{IG}_n\}$ and $\{\bar{\chi}^\gamma \mid \gamma \in \mathcal{IG}_n\}$, with

$$\bar{\delta}_\gamma = \sum_{\sigma \supseteq \gamma} \delta_\sigma \quad \text{and} \quad \bar{\chi}^\gamma = \sum_{\sigma \subseteq \gamma} \chi^\sigma.$$

Remarkably, if $\mathbb{1} \in \mathsf{cf}(\mathrm{UT}_\gamma)$ denotes the character of the trivial representation then

$$\bar{\chi}^\gamma = \mathrm{Ind}_{\mathrm{UT}_\gamma}^{\mathrm{UT}_n}(\mathbb{1}) = q^{|E(\gamma)|}\bar{\delta}_\gamma, \tag{2.11}$$

the character of the UT_n-module $\mathbb{C}[\mathrm{UT}_n/\mathrm{UT}_\gamma]$.

2D. ***Homomorphisms between Hopf algebras of class functions.*** In [45, III], Zelevinsky defines a graded connected Hopf algebra on the space

$$\mathsf{cf}(\mathrm{GL}_\bullet) = \bigoplus_{n \geq 0} \mathsf{cf}(\mathrm{GL}_n),$$

with structure maps coming from the parabolic induction and restriction functors. The paper [19] defines a similar Hopf structure on the spaces

$$\mathsf{scf}(\mathrm{UT}_\bullet) = \bigoplus_{n \geq 0} \mathsf{scf}(\mathrm{UT}_n) \quad \text{and} \quad \mathsf{cf}(\mathrm{UT}_\bullet) = \bigoplus_{n \geq 0} \mathsf{cf}(\mathrm{UT}_n),$$

in which $\mathsf{scf}(\mathrm{UT}_n)$ is the subspace of class functions defined in Section 2C, with $\mathsf{scf}(\mathrm{UT}_\bullet)$ a sub-Hopf algebra of $\mathsf{cf}(\mathrm{UT}_\bullet)$. This section will describe several homomorphisms involving these Hopf algebras.

In [25, Section 6], Guay-Paquet defines a $\mathbb{C}[t]$-Hopf algebra on the free $\mathbb{C}[t]$-module $\mathbb{C}[t][\mathcal{IG}]$, and specializing $t \mapsto q^{-1}$ gives a Hopf algebra over $\mathbb{C}$; see [19, Section 7]. Recall the basis $\{\bar{\delta}_\gamma \mid \gamma \in \mathcal{IG}\}$ of $\mathsf{scf}(\mathrm{UT}_\bullet)$ defined in Section 2C.

Theorem 2.12 [19, Corollary 7.2]. *The map* $\gamma \mapsto \bar{\delta}_\gamma$ *is an isomorphism from Guay-Paquet's specialized Hopf algebra to* $\mathsf{scf}(\mathrm{UT}_\bullet)$.

A second map comes from the induction functors $\mathrm{Ind}_{\mathrm{UT}_n}^{\mathrm{GL}_n} : \mathsf{cf}(\mathrm{UT}_n) \to \mathsf{cf}(\mathrm{GL}_n)$: let

$$\mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}} = \bigoplus_{n \geq 0} \mathrm{Ind}_{\mathrm{UT}_n}^{\mathrm{GL}_n} : \mathsf{cf}(\mathrm{UT}_\bullet) \to \mathsf{cf}(\mathrm{GL}_\bullet). \tag{2.13}$$

Theorem 2.14 [19, Theorem 6.1]. *The map* $\mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}$ *is a Hopf algebra homomorphism.*

The homomorphism $\mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}$ also induces a linear map on dual spaces. Using the identification in (2.8), the dual of the direct sum $\mathsf{cf}(\mathrm{GL}_\bullet)$ becomes a product

$$\mathsf{cf}(\mathrm{GL}_\bullet)^* = \prod_{n\geq 0} \mathsf{cf}(\mathrm{GL}_n)^* = \{(\chi_n\rangle)_{n\geq 0} \mid \chi_n \in \mathsf{cf}(\mathrm{GL}_n)\}.$$

Making the analogous identification for $\mathsf{cf}(\mathrm{UT}_\bullet)^*$ and $\mathsf{scf}(\mathrm{UT}_\bullet)^*$, Frobenius reciprocity gives that

$$(\chi_n\rangle)_{n\geq 0} \circ \mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}} = \left(\mathrm{Res}_{\mathrm{UT}_n}^{\mathrm{GL}_n}(\chi_n)\rangle\right)_{n\geq 0}.$$

If $\mathrm{Res}_{\mathrm{UT}_n}^{\mathrm{GL}_n}(\chi_n) \in \mathsf{scf}(\mathrm{UT}_\bullet)$ for each $n \geq 0$, the same equation applies when considering each side as an element of $\mathsf{scf}(\mathrm{UT}_\bullet)^*$.

3. The chromatic quasisymmetric function as a GL_n character

This section will state and prove Theorem 3.1, following some initial context. Recall the Hopf algebras $\mathsf{scf}(\mathrm{UT}_\bullet)$ and $\mathsf{cf}(\mathrm{UT}_\bullet)$ from Section 2D. The Hopf algebra of GL-*class functions with unipotent support* is the image

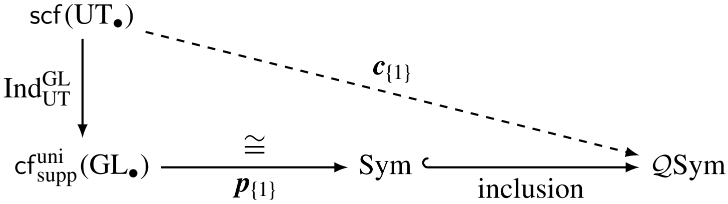
$$\mathsf{cf}_{\mathrm{supp}}^{\mathrm{uni}}(\mathrm{GL}_\bullet) = \mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}(\mathsf{cf}(\mathrm{UT}_\bullet)) \subseteq \mathsf{cf}(\mathrm{GL}_\bullet).$$

Zelevinsky [45] has defined a Hopf algebra isomorphism $\boldsymbol{p}_{\{1\}} : \mathsf{cf}_{\mathrm{supp}}^{\mathrm{uni}}(\mathrm{GL}_\bullet) \to \mathrm{Sym}$ which will be used in the theorem; see Section 3A. Finally, for each indifference graph γ, recall the subgroup UT_γ defined in Section 2C, and let $X_\gamma(\boldsymbol{x}; t)$ denote the chromatic quasisymmetric function of γ in an indeterminate 't', which will be formally defined in Section 3B.

Theorem 3.1. *For* $n \geq 0$ *and* $\gamma \in \mathcal{IG}_n$,

$$\mathrm{Ind}_{\mathrm{UT}_\gamma}^{\mathrm{GL}_n}(\mathbb{1}) = (q-1)^n\, \boldsymbol{p}_{\{1\}}^{-1}(X_\gamma(\boldsymbol{x}; q)).$$

I will describe briefly how the results in this section prove Theorem 3.1. Define a Hopf algebra homomorphism $\boldsymbol{c}_{\{1\}} : \mathsf{scf}(\mathrm{UT}_\bullet) \to \mathcal{Q}\mathrm{Sym}$ as the composite map in the diagram



(3.2)

of Hopf algebra homomorphisms. By the transitivity of induction, the theorem is equivalent to computing the image of the character $\bar{\chi}^\gamma = \mathrm{Ind}_{\mathrm{UT}_\gamma}^{\mathrm{UT}_n}(\mathbb{1}) \in \mathsf{scf}(\mathrm{UT}_\bullet)$ under $\boldsymbol{c}_{\{1\}}$.

Recalling the theory of combinatorial Hopf algebras from Section 2A1, there is a unique combinatorial Hopf algebra structure on $\mathsf{scf}(\mathrm{UT}_\bullet)$ for which $\boldsymbol{c}_{\{1\}}$ is a CHA morphism to $(\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1)$, and this structure is given by a zeta function of the Hopf algebra $\mathsf{scf}(\mathrm{UT}_\bullet)$. Theorem 3.8 computes this zeta function, and

Proposition 3.13 shows that it is essentially the same as one defined by Guay-Paquet [25]. This leads to a formula for $\boldsymbol{c}_{\{1\}}$ on the basis $\{\bar{\delta}_\gamma \mid \gamma \in \mathcal{IG}\}$ of $\mathsf{scf}(\mathrm{UT}_\bullet)$ from Section 2C, stated formally in Corollary 3.14:

$$\boldsymbol{c}_{\{1\}}(\bar{\delta}_\gamma) = (q-1)^n X_\gamma(\boldsymbol{x}; q^{-1}) \quad \text{for } \gamma \in \mathcal{IG}_n. \tag{3.3}$$

From here, the theorem follows from an identity of Shareshian–Wachs [39]. Recalling from Section 2C that $\bar{\chi}^\gamma = q^{|E(\gamma)|}\bar{\delta}_\gamma$, [39, Proposition 2.6] reformulates (3.3) as

$$\boldsymbol{c}_{\{1\}}(\bar{\chi}^\gamma) = (q-1)^n q^{|E(\gamma)|} X_\gamma(\boldsymbol{x}; q^{-1}) = (q-1)^n X_\gamma(\boldsymbol{x}; q).$$

The results used in the proof are given in the remainder of this section, which comprises two subsections. Section 3A describes the zeta functions of the Hopf algebras $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$ and $\mathsf{scf}(\mathrm{UT}_\bullet)$ needed to make Diagram (3.2) a diagram of combinatorial Hopf algebras. Then, Section 3B uses results from [25] and Section 2 to describe the chromatic quasisymmetric function as the image of a CHA morphism from $\mathsf{scf}(\mathrm{UT}_\bullet)$ and subsequently shows that up to a power of $(q-1)$ this map coincides with $\boldsymbol{c}_{\{1\}}$.

3A. ***Factoring $\boldsymbol{c}_{\{1\}}$ through $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$.*** This section describes the Hopf algebra $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$ and its isomorphism with Sym.

An element $g \in \mathrm{GL}_n$ is called *unipotent* if $g - 1_n$ is nilpotent. There is a canonical indexing of the unipotent GL_n-conjugacy classes by partitions; this is stated without proof in [45, 10.1] so a bit more detail has been included here. The Jordan canonical form of an element $g \in \mathrm{GL}_n$ is defined over any field that contains every root of the characteristic polynomial of g. Assuming that g is unipotent, the characteristic polynomial is $(t-1)^n$, so the Jordan canonical form of g is defined over $\mathbb{F}_q$. The Jordan matrices corresponding to $(t-1)^n$ are naturally indexed by partitions of n: $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_\ell)$ corresponds to

$$J_\lambda = \begin{bmatrix} J_{\lambda_1} & 0 & \cdots & 0 \\ 0 & J_{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{\lambda_\ell} \end{bmatrix} \quad \text{with } J_k = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Thus, if we write O_λ for the GL_n conjugacy class of J_λ, the set of unipotent elements of GL_n is partitioned by the conjugacy classes $\{O_\lambda \mid \lambda \in \mathcal{P}_n\}$.

This shows that an element of GL_n is unipotent if and only if it is conjugate to an element of $\mathrm{UT}_n(\mathbb{F}_q)$, so that the sub-Hopf algebra $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$ from Section 3 is exactly

$$\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet) = \bigoplus_{n\geq 0} \{\psi \in \mathsf{cf}(\mathrm{GL}_n) \mid \psi(h) = 0 \text{ for } h \in \mathrm{GL}_n \text{ not unipotent}\},$$

the space of GL-class functions with support only on unipotent elements. This fact is the source of the notational choice "$\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}$".

The preceding paragraphs demonstrate that $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$ has a $\mathcal{P}$-indexed basis of identifier functions $\delta_\lambda = \delta_{O_\lambda}$ for unipotent conjugacy classes,

$$\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet) = \mathbb{C}\text{-span}\{\delta_\lambda \mid \lambda \in \mathcal{P}\}.$$

Zelevinsky [45, 10.13] (see also [35, IV.4.1]) constructs a graded Hopf algebra isomorphism

$$\boldsymbol{p}_{\{1\}} : \mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet) \to \mathrm{Sym}, \quad \delta_\lambda \mapsto \widetilde{P_\lambda}(\boldsymbol{x}; q) = q^{-n(\lambda)} P_\lambda(\boldsymbol{x}; q^{-1}), \tag{3.4}$$

where $n(\lambda) = \sum_{i=1}^{\lambda_1} \binom{\lambda'_i}{2}$, $P_\lambda(\boldsymbol{x}; t)$ is an element of the Hall–Littlewood 'P' basis of $\mathrm{Sym}[t]$ defined in [35, III.2], and we have specialized $t = q^{-1}$.

In the framework of Theorem 2.2, the isomorphism $\boldsymbol{p}_{\{1\}}$ is equivalent to a zeta function of $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$. This datum was also determined by Zelevinsky in [45]. The *regular* unipotent elements of GL_n are the members of the conjugacy class $O_{(n)}$. Using the notation of Section 2D, define a linear functional

$$\delta^*_{(\bullet)}\rangle = (\delta^*_{(n)}\rangle)_{n\geq 0} \in \mathsf{cf}(\mathrm{GL}_\bullet)^* \quad \text{with } \delta^*_{(n)}\rangle = \frac{\delta_{(n)}\rangle}{\langle \delta_{(n)}, \delta_{(n)}\rangle},$$

so that for $\psi \in \mathsf{cf}(\mathrm{GL}_n)$, the value of $\delta^*_{(\bullet)}\rangle(\psi)$ is the value of ψ at any regular unipotent element, $\psi(J_{(n)})$. By embedding $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$ into $\mathsf{cf}(\mathrm{GL}_\bullet)$, $\delta^*_{(\bullet)}\rangle$ is also a linear functional on $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$.

Proposition 3.5 [45, 10.8]. *The map $\delta^*_{(\bullet)}\rangle$ is a zeta function of the Hopf algebra $\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$ and $\boldsymbol{p}_{\{1\}}$ is the unique CHA morphism $(\mathsf{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet), \delta^*_{(\bullet)}\rangle) \to (\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1)$.*

Remark 3.6. In [45], this result is stated in terms of symmetric functions, since the language of CHAs was not yet available. However, the underlying theory naturally extends to this context, essentially because the inclusion $(\mathrm{Sym}, \mathrm{ps}_1) \hookrightarrow (\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1)$ is a CHA morphism.

Now consider the Hopf algebra $\mathsf{scf}(\mathrm{UT}_\bullet)$. Recall that $([n], \varnothing)$ is the minimal indifference graph on n vertices and define a linear functional

$$(q-1)^\bullet \delta^*_{([\bullet],\varnothing)}\rangle = ((q-1)^n \delta^*_{([n],\varnothing)}\rangle)_{n\geq 0} \in \mathsf{scf}(\mathrm{UT}_\bullet)^* \quad \text{with } \delta^*_{([n],\varnothing)}\rangle = \frac{\delta_{([n],\varnothing)}\rangle}{\langle \delta_{([n],\varnothing)}, \delta_{([n],\varnothing)}\rangle}.$$

Remark 3.7. There is an unfortunate coincidence of notation between the class functions $\delta^*_{(n)}$ and $\delta^*_{([n],\varnothing)}$, and care should be taken to distinguish between the two: up to normalization $\delta^*_{(n)}$ is the GL_n-class function which identifies the conjugacy class $O_{(n)}$ of regular unipotent elements, and $\delta^*_{([n],\varnothing)}$ is the UT_n-class function which identifies the superclass

$$\mathrm{UT}^\circ_{([n],\varnothing)} = \{X \in \mathrm{UT}_n \mid X_{i,i+1} \neq 0 \text{ for } 1 \leq i < n\}.$$

However, the two are closely related, as described in the proof of Theorem 3.8 below.

Theorem 3.8. *The linear functional $(q-1)^\bullet \delta^*_{([\bullet],\varnothing)}\rangle$ is a zeta function of $\mathsf{scf}(\mathrm{UT}_\bullet)$ and*

$$(q-1)^\bullet \delta^*_{([\bullet],\varnothing)}\rangle = \delta^*_{(\bullet)}\rangle \circ \mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}},$$

so $\mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}$ is a CHA morphism

$$(\mathsf{scf}(\mathrm{UT}_\bullet), (q-1)^\bullet \delta^*_{([\bullet],\varnothing)}\rangle) \xrightarrow{\mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}} (\mathsf{cf}(\mathrm{GL}_\bullet), \delta^*_{(\bullet)}\rangle).$$

Proof. The first and third assertions follow from the second. The proof of the second assertion will make use of the fact that the superclass $\mathrm{UT}^{\circ}_{([n],\varnothing)}$ is also the set of all regular unipotent elements in UT_n, so that $\delta_{([n],\varnothing)} = \mathrm{Res}^{\mathrm{GL}_n}_{\mathrm{UT}_n}(\delta_{(n)})$.

For $\gamma \in \mathcal{IG}_n$, Frobenius reciprocity (as described in Section 2D) gives

$$\delta^*_{(\bullet)}\rangle \circ \mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}(\bar{\delta}_\gamma) = \frac{\mathrm{Res}^{\mathrm{GL}_n}_{\mathrm{UT}_n}(\delta_{(n)})\rangle(\bar{\delta}_\gamma)}{\langle \delta_{(n)}, \delta_{(n)}\rangle} = \frac{\langle \bar{\delta}_\gamma, \delta_{([n],\varnothing)}\rangle}{\langle \delta_{(n)}, \delta_{(n)}\rangle} = \begin{cases} \dfrac{\langle \delta_{([n],\varnothing)}, \delta_{([n],\varnothing)}\rangle}{\langle \delta_{(n)}, \delta_{(n)}\rangle} & \text{if } \gamma = ([n],\varnothing), \\ 0 & \text{otherwise,} \end{cases}$$

with the last equation following from the definition of $\bar{\delta}_\gamma$, the minimality of $([n],\varnothing)$, and the orthogonality of the superclass identifiers; see Section 2C. Direct computation then gives that

$$\frac{\langle \delta_{([n],\varnothing)}, \delta_{([n],\varnothing)}\rangle}{\langle \delta_{(n)}, \delta_{(n)}\rangle} = \frac{|\mathrm{GL}_n|}{|O_{(n)}|}\frac{|\mathrm{UT}^{\circ}_{([n],\varnothing)}|}{|\mathrm{UT}_n|} = (q-1)^n,$$

where the final equality comes from the order formulas

$$O_{(n)} = \frac{|\mathrm{GL}_n|}{q^{n-1}(q-1)} \quad \text{and} \quad \mathrm{UT}^{\circ}_{([n],\varnothing)} = (q-1)^{n-1}\frac{|\mathrm{UT}_n|}{q^{n-1}}. \qquad \square$$

Now recall the map $\boldsymbol{c}_{\{1\}}$ defined in Diagram (3.2). Theorem 3.8 and Proposition 3.5 give the following.

Corollary 3.9. *The map $\boldsymbol{c}_{\{1\}}$ is the unique CHA morphism*

$$\boldsymbol{c}_{\{1\}} : (\mathrm{scf}(\mathrm{UT}_\bullet), (q-1)^\bullet \delta^*_{([\bullet],\varnothing)}\rangle) \to (\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1).$$

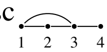
Remark 3.10. Theorem 3.8 actually establishes the stronger result that $(q-1)^\bullet\delta^*_{([\bullet],\varnothing)}\rangle$ is a zeta function of $\mathrm{cf}(\mathrm{UT}_\bullet)$, and that we may extend the domain of the CHA morphisms $\mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}$ and $\boldsymbol{c}_{\{1\}}$ to the combinatorial Hopf algebra $\big(\mathrm{cf}(\mathrm{UT}_\bullet), (q-1)^\bullet\delta^*_{([\bullet],\varnothing)}\rangle\big)$. While this level of generality is unnecessary for the scope of this work, it may be of general interest.

3B. ***The chromatic quasisymmetric function.*** This section defines the chromatic quasisymmetric function of a graph and describes how it can be realized as the image of a character of $\mathrm{GL}_n(\mathbb{F}_q)$ under a particular a CHA morphism, leading to a proof of Theorem 3.1.

Let γ be a simple, undirected graph with vertex set $[n]$ and edge set $E(\gamma)$. A *coloring* of γ is a function $\kappa : [n] \to \mathbb{Z}_{>0}$. A coloring κ of γ is *proper* if $\kappa(i) \neq \kappa(j)$ for all $\{i, j\} \in E(\gamma)$. The *γ-ascent number* of a coloring κ is

$$\mathrm{asc}_\gamma(\kappa) = \big|\{\{i, j\} \in E(\gamma) \mid i < j \text{ and } \kappa(i) < \kappa(j)\}\big|. \tag{3.11}$$

For example, if $\kappa : [5] \to \mathbb{Z}_{>0}$ is given by $\kappa(1) = 2$, $\kappa(2) = 5$, $\kappa(3) = 1$, and $\kappa(4) = 5$, then



$$\mathrm{asc}(\kappa) = \big|\{\{1, 2\}, \{3, 4\}\}\big| = 2.$$

In this example, κ is a proper coloring of the given graph.

The *chromatic quasisymmetric function* of γ is

$$X_\gamma(\boldsymbol{x};t) = \sum_{\substack{\kappa:[n]\to\mathbb{Z}_{>0}\\ \text{proper}}} t^{\mathrm{asc}_\gamma(\kappa)} x_{\kappa(1)}x_{\kappa(2)}\cdots x_{\kappa(n)} \in \mathcal{Q}\mathrm{Sym}[t],$$

so that $X_\gamma(\boldsymbol{x};t)$ is a polynomial in an indeterminate t whose coefficients — by properties of the ascent statistic — are quasisymmetric functions. For an indifference graph $\gamma \in \mathcal{IG}_n$, it is known that these coefficients are elements of Sym [39, Theorem 4.5]. For example,

$$X_{\underset{1\ \ 2\ \ 3}{\bullet\!-\!\bullet\!-\!\bullet}}(\boldsymbol{x};t) = tm_{(2,1)} + (t^2+4t+1)m_{(1^3)}.$$

However, this property is not used, and a novel proof of it follows from Corollary 3.14 below; see Remarks 3.15 (R1).

Evaluating the indeterminate t in $X_\gamma(\boldsymbol{x};t)$ at a complex number gives an actual (quasi)symmetric function. For example, $X_\gamma(\boldsymbol{x};1)$ is the ordinary chromatic symmetric function of the graph γ, as defined by Stanley in [40]. In Theorem 3.1 the chromatic quasisymmetric functions are evaluated at q, the order of the finite field $\mathbb{F}_q$.

In [25], Guay-Paquet constructs the chromatic quasisymmetric by way of a homomorphism of $\mathbb{C}[t]$-Hopf algebras. By evaluating at $t=q^{-1}$ as in Theorem 2.12, this result descends to a Hopf algebra homomorphism $\mathsf{scf}(\mathrm{UT}_\bullet) \to \mathcal{Q}\mathrm{Sym}$. Define a linear functional

$$\zeta_0 : \mathsf{scf}(\mathrm{UT}_\bullet) \to \mathbb{C}, \qquad \bar{\delta}_\gamma \mapsto \begin{cases} 1 & \text{if } \gamma = ([n],\varnothing),\\ 0 & \text{otherwise.}\end{cases}$$

The following theorem is translated from its original context in [25] to that of the Hopf algebra $\mathsf{scf}(\mathrm{UT}_\bullet)$ using the Hopf algebra isomorphism in Theorem 2.12.

Theorem 3.12 [25, Theorem 57]. *The map ζ_0 is a zeta function of $\mathsf{scf}(\mathrm{UT}_\bullet)$, and the unique CHA morphism*

$$(\mathsf{scf}(\mathrm{UT}_\bullet),\zeta_0) \to (\mathcal{Q}\mathrm{Sym},\mathrm{ps}_1)$$

is given by

$$\bar{\delta}_\gamma \mapsto X_\gamma(\boldsymbol{x};q^{-1}).$$

Along with Theorem 2.2, this result is the key to compute the image of the map $\boldsymbol{c}_{\{1\}}$ defined at the outset of Section 3. Recall the zeta function $(q-1)^\bullet\delta^*_{([\bullet],\varnothing)}\rangle$ of the Hopf algebra $\mathsf{scf}(\mathrm{UT}_\bullet)$ defined in Section 3A.

Proposition 3.13. *Let γ be an indifference graph of size $n \geq 0$. Then*

$$(q-1)^\bullet\delta^*_{([\bullet],\varnothing)}\rangle(\bar{\delta}_\gamma) = \begin{cases} (q-1)^n & \textit{if } \gamma = ([n],\varnothing),\\ 0 & \textit{otherwise.}\end{cases}$$

Proof. By definition, $\bar{\delta}_\gamma = \sum_{\sigma\supseteq\gamma}\delta_\sigma$. Explicit computation then gives

$$\frac{\langle (q-1)^n\delta_{([n],\varnothing)},\bar{\delta}_\gamma\rangle}{\langle \delta_{([n],\varnothing)},\delta_{([n],\varnothing)}\rangle} = (q-1)^n\frac{\sum_{\sigma\supseteq\gamma}\langle\delta_{([n],\varnothing)},\delta_\sigma\rangle}{\langle \delta_{([n],\varnothing)},\delta_{([n],\varnothing)}\rangle}.$$

Using the orthogonality of the basis $\{\delta_\gamma \mid \gamma \in \mathcal{IG}_n\}$ and the minimality of $([n], \varnothing)$ under the spanning subgraph order on $\mathcal{IG}_n$, the above expression reduces to the desired formula. □

Thus, on homogeneous elements of degree n, the zeta functions $(q-1)^\bullet \delta^*_{([\bullet],\varnothing)}\rangle$ and ζ_0 only differ by a factor of $(q-1)^n$. This leads to the following result, which is a restatement of (3.3) and accordingly the last step in the proof of Theorem 3.1.

Corollary 3.14. *Let γ be an indifference graph of size $n \geq 0$. Then*

$$\boldsymbol{c}_{\{1\}}(\bar{\delta}_\zeta) = (q-1)^n X_\gamma(\boldsymbol{x}; q^{-1}).$$

Proof. By comparison with the Hopf algebra homomorphism in Theorem 3.12, it is clear that the given map is a graded Hopf algebra homomorphism, and further, that

$$\mathrm{ps}_1((q-1)^n X_\gamma(\boldsymbol{x}; q^{-1})) = (q-1)^n \zeta_0(\bar{\delta}_\gamma) = (q-1)^\bullet \delta^*_{([\bullet],\varnothing)}\rangle(\bar{\delta}_\gamma).$$

Thus, the given map is a CHA morphism

$$(\mathsf{scf}(\mathrm{UT}_\bullet), (q-1)^\bullet \delta^*_{([\bullet],\varnothing)}\rangle) \to (\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1).$$

By Theorem 2.2, the above map must be equal to $\boldsymbol{c}_{\{1\}}$. □

Remarks 3.15. (R1) As the image of $\boldsymbol{c}_{\{1\}}$ is $\mathrm{Sym} \subseteq \mathcal{Q}\mathrm{Sym}$, Corollary 3.14 gives a novel proof that the coefficients of $X_\gamma(\boldsymbol{x}; t)$ are symmetric functions.

(R2) Proposition 3.13 also shows that $\zeta_0 = (\delta^*_{([n],\varnothing)}\rangle)_{n\geq 0}$; however this fact seems not to have any representation theoretic significance beyond its relation to the proof above.

4. Connections to Hessenberg varieties

This section will describe the relationship between the characters $\mathrm{Ind}_{\mathrm{UT}_\gamma}^{\mathrm{GL}_n}(\mathbb{1})$ in Theorem 3.1, certain Hessenberg varieties over $\mathbb{F}_q$, and the analogous Hessenberg varieties over $\mathbb{C}$. These results follow a short overview of Hessenberg varieties. Throughout, the algebraic groups defined over $\mathbb{F}_q$ in Section 2C and their analogues over $\mathbb{C}$ are used, so the underling field will be explicitly written for each such group to avoid confusion.

Take a field $\mathbb{K} \in \{\mathbb{F}_q, \mathbb{C}\}$, and for $n \geq 0$ let $B_n(\mathbb{K})$ denote the subgroup of upper triangular matrices in $\mathrm{GL}_n(\mathbb{K})$. For each subspace $M \subseteq \mathrm{Mat}_n(\mathbb{K})$ which is stable under conjugation by elements of $B_n(\mathbb{K})$ and each matrix $A \in \mathrm{Mat}_n(\mathbb{K})$, the *Hessenberg variety* associated to A and M is

$$\mathcal{B}_A^M = \{gB_n(\mathbb{K}) \in \mathrm{GL}_n(\mathbb{K})/B_n(\mathbb{K}) \mid g^{-1}Ag \in M\}.$$

This is a slight variation — apparently due to [44] — of the original definition in [14], which requires that M contain all upper triangular matrices. The generalization is crucial, since the following results exclusively concern Hessenberg varieties associated to strictly upper triangular subspaces known as *ad-nilpotent ideals*. For $\gamma \in \mathcal{IG}_n$, let

$$\begin{aligned}\mathrm{ut}_\gamma(\mathbb{K}) &= \{A \in \mathrm{Mat}_n(\mathbb{K}) \mid A_{i,j} \neq 0 \text{ only if } i < j \text{ and } (i,j) \notin \gamma\} \\ &= \mathrm{UT}_\gamma(\mathbb{K}) - 1.\end{aligned}$$

These sets are in fact ideals in the algebra (and Lie algebra) of upper triangular matrices. Key examples of the Hessenberg varieties of the form $\mathcal{B}_A^{\mathrm{ut}_\gamma(\mathbb{K})}$ have been known for some time, but a specific study of these varieties is quite recent; see [31; 38].

Proposition 4.1. *Let $n \geq 0$ and $\gamma \in \mathcal{IG}_n$. For any $A \in \mathrm{Mat}_n(\mathbb{F}_q)$ with $1 + A \in \mathrm{GL}_n(\mathbb{F}_q)$,*

$$\mathrm{Ind}_{\mathrm{UT}_\gamma(\mathbb{F}_q)}^{\mathrm{GL}_n(\mathbb{F}_q)}(\mathbb{1})(1+A) = (q-1)^n q^{|E(\gamma)|} |\mathcal{B}_A^{\mathrm{ut}_\gamma(\mathbb{F}_q)}|.$$

Proof. The proof will compute the left side of the equation directly. Equation (2.11) and the standard formula for induced character values give

$$\mathrm{Ind}_{\mathrm{UT}_\gamma(\mathbb{F}_q)}^{\mathrm{GL}_n(\mathbb{F}_q)}(\mathbb{1})(1+A) = \left|\{h\mathrm{UT}_\gamma(\mathbb{F}_q) \in \mathrm{GL}_n(\mathbb{F}_q)/\mathrm{UT}_\gamma(\mathbb{F}_q) \mid h^{-1}(1+A)h \in \mathrm{UT}_\gamma(\mathbb{F}_q)\}\right|.$$

Each left $B_n(\mathbb{F}_q)$ coset in $\mathrm{GL}_n(\mathbb{F}_q)$ comprises $q^{|E(\gamma)|}(q-1)^n$ left $\mathrm{UT}_\gamma(\mathbb{F}_q)$ cosets, and for each $h\mathrm{UT}_\gamma(\mathbb{F}_q) \subseteq gB_n(\mathbb{F}_q)$, it is the case that $h^{-1}(1+A)h \in \mathrm{UT}_\gamma(\mathbb{F}_q)$ if and only if $g^{-1}(1+A)g \in \mathrm{UT}_\gamma(\mathbb{F}_q)$, because $\mathrm{UT}_\gamma(\mathbb{F}_q)$ is normalized by $B_n(\mathbb{F}_q)$. Finally, $g^{-1}(1+A)g \in \mathrm{UT}_\gamma(\mathbb{F}_q)$ if and only if $g^{-1}Ag \in \mathrm{ut}_\gamma(\mathbb{F}_q)$, in which case $gB_n(\mathbb{F}_q)$ belongs to $\mathcal{B}_A^{\mathrm{ut}_\gamma(\mathbb{F}_q)}$. □

This result reveals a relationship between the chromatic quasisymmetric function and Hessenberg varieties for ad-nilpotent ideals over $\mathbb{F}_q$. To state this relationship precisely, recall the degree-shifted Hall–Littlewood elements $\widetilde{P}_\lambda(\boldsymbol{x}; t)$ from Section 3A, and define Laurent polynomials $d_\lambda^\gamma(t)$ by

$$X_\gamma(\boldsymbol{x}; t) = \sum_{\lambda \in \mathcal{P}_n} d_\lambda^\gamma(t) \widetilde{P}_\lambda(\boldsymbol{x}; t). \tag{4.2}$$

Each $\widetilde{P}_\lambda(\boldsymbol{x}; t)$ is a polynomial in t^{-1} rather than t, so there is some subtlety to this definition: one must first express $t^{-|E(\gamma)|}X_\gamma(\boldsymbol{x}; t)$ in the basis $P_\lambda(\boldsymbol{x}; t^{-1})$ of $\mathrm{Sym}[t^{-1}]$ and then multiply each term by appropriate powers of t to obtain (4.2).

Evaluating both sides of (4.2) at $t = q$ and applying the map $\boldsymbol{p}_{\{1\}}^{-1}$ defined in Section 3A gives

$$\frac{1}{(q-1)^n} \mathrm{Ind}_{\mathrm{UT}_\gamma(\mathbb{F}_q)}^{\mathrm{GL}n(\mathbb{F}_q)}(\mathbb{1}) = \sum_{\lambda \in \mathcal{P}_n} d_\lambda^\gamma(q) \delta_\lambda,$$

where Theorem 3.1 is used to evaluate the left side. Each side of the above equation is a class function, so for any partition $\lambda \in \mathcal{P}_n$ we can evaluate both sides at an element $1 + A \in O_\lambda$ for some fixed partition $\lambda \in \mathcal{P}_n$, Proposition 4.1 gives

$$q^{|E(\gamma)|} |\mathcal{B}_A^{\mathrm{ut}_\gamma(\mathbb{F}_q)}| = d_\lambda^\gamma(q). \tag{4.3}$$

The coefficients $d_\lambda^\gamma(t)$ also appear in the complex geometry of Hessenberg varieties for ad-nilpotent subspaces in a manner discovered by Precup and Sommers in [38]. For the following theorem, note that the discussion of Jordan canonical form in Section 3A shows that the similarity classes of nilpotent matrices over any field are indexed by partitions of n: the class indexed by $\lambda \in \mathcal{P}_n$ consists of all matrices similar to $J_\lambda - 1$.

Theorem 4.4 [38, Equation (4.7)]. *For $n \geq 0$, take $\gamma \in \mathcal{IG}_n$ and $\lambda \in \mathcal{P}_n$. Then*

$$\sum_{k \geq 0} \beta_k^\lambda t^{k/2} = t^{-|E(\gamma)|} d_\lambda^\gamma(t),$$

where β_k^λ denotes the k-th Betti number of $\mathcal{B}_A^{\mathrm{ut}_\gamma(\mathbb{C})}$ for any nilpotent matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ in the similarity class indexed by λ.

Thus, [38] shows that the $d_\lambda^\gamma(t)$ are in fact polynomials. Combining this result with (4.3) leads to the following corollary.

Corollary 4.5. *For $n \geq 0$, take $\gamma \in \mathcal{IG}_n$ and $\lambda \in \mathcal{P}_n$. Let $A \in \mathrm{Mat}_n(\mathbb{F}_q)$ be a nilpotent elements in similarity class indexed by λ. Then*

$$\sum_{k \geq 0} \beta_k^\lambda q^{k/2} = |\mathcal{B}_A^{\mathrm{ut}_\gamma(\mathbb{F}_q)}|,$$

where the numbers β_k^λ are as in Theorem 4.4.

Remarks 4.6. (R1) Aside from this paper, I am aware of two works about Hessenberg varieties over $\mathbb{F}_q$. The preprint [17] concerns the Hessenberg variety associated to a split regular element of $\mathrm{GL}_n(\mathbb{F}_q)$ and a subspace containing all upper triangular matrices; under some nontrivial assumptions on q a result similar to Corollary 4.5 is established. This generalizes Fulman's use of Weil conjecture machinery on a subset of smooth Hessenberg varieties in order prove some identities on q-Eulerian numbers [18].

(R2) In [31], Ji and Precup give a combinatorial formula for the polynomials $d_\lambda^\gamma(t)$ by constructing an affine paving of $\mathcal{B}_A^{\mathrm{ut}_\gamma(\mathbb{C})}$. Precup has also suggested that a second proof of Corollary 4.5 could be obtained from a careful study of this paving, which would independently reprove Theorem 3.1 (private communication, 2022).

5. The vertical strip LLT polynomial as a GL_n character

This section gives a second result of the same type as Theorem 3.1, in that it interprets a family of t-graded symmetric functions as the images of certain GL_n characters obtained by induction from UT_n under a particular isomorphism; see Table 1. Here, the initial UT_n characters come from a larger set $\{\psi^\sigma \mid \sigma \in \mathcal{TS}\}$ indexed by the set of tall Schröder paths $\mathcal{TS}$ from Section 2B, the map to Sym is a homomorphism $\boldsymbol{p}_{\mathbb{1}} : \mathrm{cf}(\mathrm{GL}_\bullet) \to \mathrm{Sym}$ which records the unipotent constituent of a character, and the symmetric functions are the vertical strip LLT polynomials $G_\sigma(\boldsymbol{x}; t)$, also indexed by the set $\mathcal{TS}$. Each object mentioned will be defined in this section.

	Theorem 3.1	Theorem 5.1
indexing set	indifference graphs $\gamma \in \mathcal{IG}_n$	tall Schröder paths $\sigma \in \mathcal{TS}_n$
UT_n-characters	permutation characters $\bar{\chi}^{\gamma}$	pseudosupercharacters ψ^{σ}
symmetric functions	chromatic quasisymmetric functions $X_{\gamma}(\boldsymbol{x};t)$	vertical strip LLT polynomials $G_{\sigma}(\boldsymbol{x};t)$
map to Sym	$\boldsymbol{p}_{\{1\}} : \mathrm{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_{\bullet}) \to \mathrm{Sym}$	$\boldsymbol{p}_{\mathbb{1}} : \mathrm{cf}(\mathrm{GL}_{\bullet}) \to \mathrm{Sym}$
meaning of map to Sym	records unipotently supported GL_n-class functions	records the irreducible unipotent constituents

Table 1. A comparison of the results of Theorems 3.1 and 5.1 in degree n.

Theorem 5.1. *Let σ be a tall Schröder path. Then*

$$\boldsymbol{p}_{\mathbb{1}} \circ \mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}(\psi^{\sigma}) = (q-1)^{|\mathrm{Diag}(\sigma)|}\omega G_{\sigma}(\boldsymbol{x};q),$$

where $\mathrm{Diag}(\sigma)$ *is the set of diagonal steps in σ.*

I will now describe the meaning of this result in greater depth and outline its proof. In the study of finite groups of Lie type, including GL_n, Deligne–Lusztig theory identifies an exemplary set of irreducible characters known as *unipotent characters*. For GL_n, the unipotent characters are relatively well understood and will be described in Section 5B. Here, the relevant fact is that Zelevinsky [45] has shown that the subspace

$$\mathrm{cf}^{\mathrm{uni}}_{\mathrm{char}}(\mathrm{GL}_{\bullet}) = \mathbb{C}\text{-span}\{\text{irreducible unipotent characters of } \mathrm{GL}_n,\ n \geq 0\}$$

is a sub-Hopf algebra of $\mathrm{cf}(\mathrm{GL}_{\bullet})$, and that $\mathrm{cf}^{\mathrm{uni}}_{\mathrm{char}}(\mathrm{GL}_{\bullet})$ is isomorphic to Sym. Furthermore, [45] shows that the orthogonal projection from $\mathrm{cf}(\mathrm{GL}_{\bullet})$ to $\mathrm{cf}^{\mathrm{uni}}_{\mathrm{char}}(\mathrm{GL}_{\bullet})$ (with respect to the inner product $\langle \cdot , \cdot \rangle$ in Section 2C) is a Hopf algebra homomorphism. Consequently, there is a homomorphism $\boldsymbol{p}_{\mathbb{1}} : \mathrm{cf}(\mathrm{GL}_{\bullet}) \to \mathrm{Sym}$ obtained by projecting onto $\mathrm{cf}^{\mathrm{uni}}_{\mathrm{char}}(\mathrm{GL}_{\bullet})$ and then applying the aforementioned isomorphism, as in the diagram

$$\begin{array}{ccc} \mathrm{cf}(\mathrm{GL}_{\bullet}) & \xrightarrow{\ \boldsymbol{p}_{\mathbb{1}}\ } & \mathrm{Sym} \\ \searrow\!\!\nwarrow & & \nearrow\!\!\swarrow\ \cong \\ & \mathrm{cf}^{\mathrm{uni}}_{\mathrm{char}}(\mathrm{GL}_{\bullet}) & \end{array} \tag{5.2}$$

of Hopf algebra homomorphisms. The map $\boldsymbol{p}_{\mathbb{1}}$ faithfully records the irreducible unipotent constituents of any class function of GL_n, which can be recovered by reversing the isomorphism $\mathrm{cf}^{\mathrm{uni}}_{\mathrm{char}}(\mathrm{GL}_{\bullet}) \cong \mathrm{Sym}$. Thus, Theorem 5.1 states that the vertical strip LLT polynomial $G_{\sigma}(\boldsymbol{x};q)$ determines the irreducible unipotent constituents of the character $\mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}(\psi^{\sigma})$.

An interesting connection arises from the interplay of Theorems 3.1 and 5.1. Carlsson and Mellit [11, Proposition 3.5] show that for a Dyck path $\pi \in \mathcal{D}_n$, the plethystic relationship

$$(t-1)^n X_{\mathrm{Graph}(\pi)}(\boldsymbol{x};t)\left[\frac{\boldsymbol{x}}{t-1}\right] = G_\pi(\boldsymbol{x};t)$$

holds, where $\mathrm{Graph}(\pi)$ is the indifference graph associated to π in Section 2B. It is also known [35, IV.4] that the composite map

$$\mathrm{Sym} \xrightarrow{\boldsymbol{p}_{\{1\}}^{-1}} \mathsf{cf}_{\mathrm{supp}}^{\mathrm{uni}}(\mathrm{GL}_\bullet) \hookrightarrow \mathsf{cf}(\mathrm{GL}_\bullet) \xrightarrow{\boldsymbol{p}_{\mathbb{1}}} \mathrm{Sym}$$

is an isomorphism which can be expressed in plethystic notation as $f[\boldsymbol{x}] \mapsto \omega f\big[\frac{\boldsymbol{x}}{t-1}\big]\big|_{t=q}$, so my results give a GL_n-representation theoretic interpretation of Carlsson and Mellit's result; at the same time, [11, Proposition 3.5] could be used to prove Theorem 5.1 via Theorem 3.1.

The proof of Theorem 3.1 will instead use the machinery of combinatorial Hopf algebras, which has the benefit of giving a new description of the map $\boldsymbol{p}_{\mathbb{1}} \circ \mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}$. Define a Hopf algebra homomorphism $\boldsymbol{c}_{\mathbb{1}} : \mathsf{scf}(\mathrm{UT}_\bullet) \to \mathcal{Q}\mathrm{Sym}$ as the composite map in the diagram

$$\begin{array}{ccccccc}
\mathsf{scf}(\mathrm{UT}_\bullet) & & & \overset{\boldsymbol{c}_{\mathbb{1}}}{\dashrightarrow} & & & \mathcal{Q}\mathrm{Sym} \\
\Big\downarrow \mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}} & & & & & & \\
\mathsf{cf}_{\mathrm{supp}}^{\mathrm{uni}}(\mathrm{GL}_\bullet) & \hookrightarrow & \mathsf{cf}(\mathrm{GL}_\bullet) & \xrightarrow[\boldsymbol{p}_{\mathbb{1}}]{} & \mathrm{Sym} & \hookrightarrow & \mathcal{Q}\mathrm{Sym}
\end{array} \tag{5.3}$$

of Hopf algebras, so that Theorem 5.1 describes $\boldsymbol{c}_{\mathbb{1}}$ implicitly. By definition, $\boldsymbol{c}_{\mathbb{1}}$ can be computed by inducing a character of UT_n to GL_n and recording its unipotent constituents as symmetric functions. However, Theorem 2.2 shows that $\boldsymbol{c}_{\mathbb{1}}$ is also determined by the zeta function $\mathrm{ps}_1 \circ \boldsymbol{c}_{\mathbb{1}}$ of the Hopf algebra $\mathsf{scf}(\mathrm{UT}_\bullet)$. It happens that this zeta function coincides exactly with one defined by Guay-Paquet, so that a result of [25] — restated in Corollary 5.16 — shows that

$$\boldsymbol{c}_{\mathbb{1}}(\bar{\delta}_{\mathrm{Graph}(\pi)}) = G_\pi(\boldsymbol{x};q^{-1}) \quad \text{for } \pi \in \mathcal{D}. \tag{5.4}$$

Several known identities for LLT polynomials complete the proof; these are given in Proposition 5.18.

The remainder of the section is divided into three parts. First, Section 5A describes the characters ψ^σ appearing in Theorem 5.1 and shows that this family includes both the permutation characters and supercharacters of $\mathsf{scf}(\mathrm{UT}_\bullet)$. Then, Section 5B describes the map $\boldsymbol{c}_{\mathbb{1}}$ as a CHA morphism to $(\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1)$, defining the necessary combinatorial Hopf algebra structures on $\mathsf{scf}(\mathrm{UT}_\bullet)$ and $\mathsf{cf}(\mathrm{GL}_\bullet)$ along the way. Finally, Section 5C formally defines the vertical strip LLT polynomial, shows how it can be realized as the image of a CHA morphism, and concludes with a proof of Theorem 5.1.

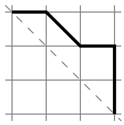
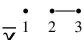
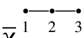
Remark 5.5. It is possible to "remove" the factors of $q-1$ in Theorem 5.1. With results in Section 5A, work of Andrews and Thiem [9, Remark on p. 490] and Aliniaeifard and Thiem [7, Remark (1) on p. 13] show that each ψ^σ is the sum of $(q-1)^{|\mathrm{Diag}(\sigma)|}$ distinct characters which each have the same image under $\boldsymbol{p}_{\mathbb{1}} \circ \mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}$; this image must be $\omega G_\sigma(\boldsymbol{x}; q)$.

5A. ***The pseudosupercharacters ψ^σ.*** This section will define the characters ψ^σ appearing in Theorem 5.1. Recall the terminology used for Schröder paths in Section 2B and the characters of UT_n defined in Section 2C.

For $\sigma \in \mathcal{TS}_n$, the *pseudosupercharacter* indexed by σ is the class function

$$\psi^\sigma = \sum_{S \subseteq \mathrm{Diag}(\sigma)} (-1)^{|\mathrm{Diag}(\sigma)-S|} \overline{\chi}^{([n],\, \mathrm{Area}(\sigma)\cup S)} \in \mathsf{scf}(\mathrm{UT}_\bullet).$$

The definition of $\mathrm{Diag}(\sigma)$ ensures that each graph $([n], \mathrm{Area}(\sigma) \cup S)$ above is in fact an indifference graph. For example, with

$\sigma =$  we have $\psi^\sigma = -\overline{\chi}$  $+\, \overline{\chi}$  . (5.6)

A noteworthy family of examples is the pseudosupercharacters indexed by Dyck paths: for $\pi \in \mathcal{D}$, $\mathrm{Diag}(\pi) = \varnothing$, from which it follows that

$$\psi^\pi = \overline{\chi}^{\mathrm{Graph}(\pi)}.$$

Proposition 5.7. *Let σ be a tall Schröder path of size $n \geq 0$. Then ψ^σ is a character, and in particular*

$$\psi^\sigma = \sum_{\substack{E(\gamma)\subseteq(\mathrm{Area}(\sigma)\cup\mathrm{Diag}(\sigma)) \\ \mathrm{Diag}(\sigma)\subseteq E(\gamma)}} \chi^\gamma,$$

where the sum is over indifference graphs $\gamma \in \mathcal{IG}_n$ satisfying the given conditions.

Proof. Using the definition of ψ^σ,

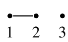
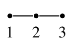
$$\psi^\sigma = \sum_{S \subseteq \mathrm{Diag}(\sigma)} (-1)^{|\mathrm{Diag}(\sigma)-S|} \sum_{E(\gamma)\subseteq \mathrm{Area}(\sigma)\cup S} \chi^\gamma,$$

where the sum is over indifference graphs γ as in the proposition. Reversing the order of summation above, we obtain

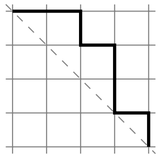
$$\psi^\sigma = \sum_{E(\gamma)\subseteq\mathrm{Area}(\sigma)\cup\mathrm{Diag}(\sigma)} \Biggl(\sum_{\substack{T\subseteq\mathrm{Diag}(\sigma) \\ T\supseteq E(\gamma)\cap\mathrm{Diag}(\sigma)}} (-1)^{|\mathrm{Diag}(\sigma)-T|} \Biggr) \chi^\gamma,$$

where the innermost sum is over subsets T of $\mathrm{Diag}(\sigma)$ that contain $E(\gamma) \cap \mathrm{Diag}(\sigma)$. Combining terms in this sum, the proposition follows from the binomial theorem. □

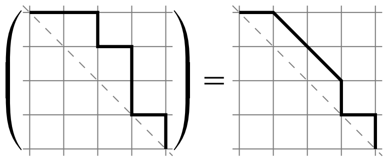
As an example of Proposition 5.7, the pseudosupercharacter in (5.6) expands as the sum of supercharacters

$\psi^\sigma = \chi$  $+\, \chi$  .

The final result in this section shows that every supercharacter of $\mathrm{scf}(\mathrm{UT}_n)$ occurs as a pseudosupercharacter. Given a Dyck path π, a *peak* of π is a sequence of steps ES; say that a peak is *tall* if the first step E does not begin on the diagonal $x = -y$. For example,

 $= (EESESSES)$

has three peaks, but only two tall peaks. Define the *Mesa path* of $\pi \in \mathcal{D}_n$ to be the tall Schröder path $\mathrm{Mesa}(\pi) \in \mathcal{TS}_n$ obtained by first constructing $\mathrm{Dyck}(\pi)$ and then replacing each tall peak ES with a diagonal step D; for example,

Mesa  $= (EDDSES)$.

Proposition 5.8. *Let π be a Dyck path. Then $\psi^{\mathrm{Mesa}(\pi)} = \chi^{\mathrm{Graph}(\pi)}$.*

Proof. By assumption,

$$\mathrm{Area}(\pi) = \mathrm{Area}(\mathrm{Mesa}(\pi)) \cup \mathrm{Diag}(\mathrm{Mesa}(\pi)),$$

so by Proposition 5.7,

$$\psi^{\mathrm{Mesa}(\pi)} = \sum_{\substack{\gamma \subseteq \mathrm{Graph}(\pi) \\ \mathrm{Diag}(\mathrm{Mesa}(\pi)) \subseteq E(\gamma)}} \chi^{\gamma}.$$

Now suppose that an indifference graph γ is a proper spanning subgraph of $\mathrm{Graph}(\pi)$. Then γ must be missing at least one edge $\{i, j\}$ such that the unit square indexed by $\{i, j\}$ is bordered directly by a tall peak of π, so that $\{i, j\} \in \mathrm{Diag}(\mathrm{Mesa}(\pi))$, and χ^{γ} does not appear in the sum above. Thus the only summand above is $\chi^{\mathrm{Graph}(\pi)}$. □

5B. ***Factoring $c_{\mathbb{1}}$ through* $\mathrm{cf}(\mathrm{GL}_\bullet)$.** This section will describe the unipotent characters of GL_n, and their relation to the Hopf algebra structure of $\mathrm{cf}(\mathrm{GL}_\bullet)$. As stated at the outset of Section 5, unipotent characters originate in Deligne–Lusztig theory, and are typically defined using cohomological induction. However, the unipotent characters of GL_n can also be described with much more elementary methods; see [16, Theorem 15.8 and proof] for the details. This paper will take this alternate description as a definition: an irreducible character of GL_n is *unipotent* if it is a constituent of $\mathrm{Ind}_{B_n}^{\mathrm{GL}_n}(\mathbb{1})$, where $B_n = B_n(\mathbb{F}_q)$ is the subgroup of upper triangular matrices in GL_n.

It is also known that irreducible unipotent characters of GL_n are indexed by the partitions of n [16, Theorem 15.8]; write χ^{λ} for the unipotent character corresponding to $\lambda \in \mathcal{P}(n)$. This paper follows the convention of [35] in which $\chi^{(1^n)}$ is the trivial character $\mathbb{1}$ of GL_n and $\chi^{(n)}$ is the *Steinberg character* St_n; this differs from the convention of [45] and others by the transposition of each partition.

The homomorphism $\boldsymbol{p}_{\mathbb{1}}$ was constructed by Zelevinsky [45, 9.4], and is given by

$$\boldsymbol{p}_{\mathbb{1}} : \mathsf{cf}(\mathrm{GL}_\bullet) \to \mathrm{Sym}, \quad \psi \mapsto \sum_\lambda \langle \psi, \chi^\lambda \rangle s_\lambda. \tag{5.9}$$

As a linear transformation, $\boldsymbol{p}_{\mathbb{1}}$ has a right inverse $s_\lambda \mapsto \chi^\lambda$, and [45] shows that this right inverse is also a Hopf algebra homomorphism. Thus, the image

$$\mathsf{cf}^{\mathrm{uni}}_{\mathrm{char}}(\mathrm{GL}_\bullet) = \mathbb{C}\text{-span}\{\chi^\lambda \mid \lambda \in \mathcal{P}\} \subseteq \mathsf{cf}(\mathrm{GL}_\bullet)$$

is a sub-Hopf algebra of $\mathsf{cf}(\mathrm{GL}_\bullet)$ through which $\boldsymbol{p}_{\mathbb{1}}$ factors, as shown in Diagram (5.2).

By Theorem 2.2, the map $\boldsymbol{p}_{\mathbb{1}}$ is equivalent to a zeta function of the Hopf algebra $\mathsf{cf}(\mathrm{GL}_\bullet)$. This zeta function is also given in [45], and is

$$\mathrm{St}_\bullet\rangle = (\mathrm{St}_n\rangle)_{n\geq 0} \in \mathsf{cf}(\mathrm{GL}_\bullet)^*.$$

Proposition 5.10 [45, 9.4–5]. *The map* $\mathrm{St}_\bullet\rangle$ *is a zeta function of* $\mathsf{cf}(\mathrm{GL}_\bullet)$ *and* $\boldsymbol{p}_{\mathbb{1}}$ *is the unique CHA morphism* $(\mathsf{cf}(\mathrm{GL}_\bullet), \mathrm{St}_\bullet\rangle) \to (\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1)$.

Now, for $n \geq 0$, write $\mathrm{reg}_{\mathrm{UT}_n}$ for the regular character of UT_n. Define a linear functional

$$\mathrm{reg}_\bullet\rangle = (\mathrm{reg}_{\mathrm{UT}_n}\rangle)_{n\geq 0} \in \mathsf{scf}(\mathrm{UT}_\bullet)^*.$$

Theorem 5.11. *The function* $\mathrm{reg}_\bullet\rangle$ *is a zeta function of* $\mathsf{scf}(\mathrm{UT}_\bullet)$ *and*

$$\mathrm{reg}_\bullet\rangle = \mathrm{St}_\bullet\rangle \circ \mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}},$$

so $\mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}$ *is a CHA morphism*

$$(\mathsf{scf}(\mathrm{UT}_\bullet), \mathrm{reg}_\bullet\rangle) \xrightarrow{\mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}} (\mathsf{cf}(\mathrm{GL}_\bullet), \mathrm{St}_\bullet\rangle).$$

Proof. It is sufficient to prove that $\mathrm{reg}_\bullet\rangle = \mathrm{St}_\bullet\rangle \circ \mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}$. Doing so requires the well-known fact (see, for example, [45, 10.3]) that for unipotent $X \in \mathrm{GL}_n$,

$$\mathrm{St}_n(X) = \begin{cases} |\mathrm{UT}_n| & \text{if } X = 1_n, \\ 0 & \text{for other unipotent } X. \end{cases}$$

As a consequence,

$$\mathrm{Res}^{\mathrm{GL}_n}_{\mathrm{UT}_n}(\mathrm{St}_n) = \mathrm{reg}_{\mathrm{UT}_n}.$$

With this, the claim follows from Frobenius reciprocity as described in Section 2D:

$$\mathrm{St}_\bullet\rangle \circ \mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}} = \big(\mathrm{Res}^{\mathrm{GL}_n}_{\mathrm{UT}_n}(\mathrm{St}_n)\rangle\big)_{n\geq 0} = \mathrm{reg}_\bullet\rangle. \qquad \square$$

Remark 5.12. Like Theorem 3.8, Theorem 5.11 actually shows that $\mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}$ is a CHA morphism from the larger combinatorial Hopf algebra $(\mathsf{cf}(\mathrm{UT}_\bullet), \mathrm{reg}_\bullet\rangle)$ to $(\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1)$.

5C. ***The vertical strip LLT polynomial.*** The *vertical strip LLT polynomial* indexed by a tall Schröder path σ is

$$G_\sigma(\boldsymbol{x};t) = \sum_{\kappa \in A(\sigma)} t^{\mathrm{asc}_{([n],\mathrm{Area}(\sigma))}(\kappa)} x_{\kappa(1)} x_{\kappa(2)} \cdots x_{\kappa(n)} \in \mathbb{C}[\![\boldsymbol{x}]\!][t],$$

where the sum is over the set $A(\sigma)$ of functions $\kappa : [n] \to \mathbb{Z}_{>0}$ which satisfy $\kappa(i) < \kappa(j)$ for each $i < j$ with $\{i, j\} \in \mathrm{Diag}(\sigma)$. Viewed as a polynomial in t, the coefficients of $G_\sigma(\boldsymbol{x};t)$ are actually symmetric functions [28, Lemma 10.2], though this is not obvious. For example,

G  $(\boldsymbol{x};t) = tm_{(2,1)} + (t^2 + 2t)m_{(1^3)}.$

Remark 5.13. There are several essentially equivalent definitions of LLT polynomials; the one above is due to [11] in the unicellular case and to [13] (see also [5]) in general.

If σ is a Dyck path, so that $\mathrm{Diag}(\sigma) = \varnothing$, then the sum in $G_\sigma(\boldsymbol{x};t)$ is over all possible colorings; this special case is know as a *unicellular* LLT polynomial. In [25], Guay-Paquet realizes the unicellular LLT polynomials by way of a homomorphism of Hopf algebras over $\mathbb{C}[t]$. By evaluating at $t = q^{-1}$ as in Theorem 2.12, this result descends to a Hopf algebra homomorphism $\mathsf{scf}(\mathrm{UT}_\bullet) \to \mathcal{Q}\mathrm{Sym}$. Define a linear functional

$$\zeta_1 : \mathsf{scf}(\mathrm{UT}_\bullet) \to \mathbb{C}, \quad \bar{\delta}_\gamma \mapsto 1.$$

Theorem 5.14 [25, Theorem 57]. *The map ζ_1 is a zeta function of* $\mathsf{scf}(\mathrm{UT}_\bullet)$, *and the unique CHA morphism*

$$(\mathsf{scf}(\mathrm{UT}_\bullet), \zeta_1) \to (\mathcal{Q}\mathrm{Sym}, \mathrm{ps}_1)$$

is given by

$$\bar{\delta}_{\mathrm{Graph}(\pi)} \mapsto G_\pi(\boldsymbol{x}; q^{-1}) \quad \textit{for } \pi \in \mathcal{D}.$$

Now recall the zeta function $\mathrm{reg}_\bullet\rangle$ defined in the previous section.

Proposition 5.15. *As a zeta function of the Hopf algebra* $\mathsf{scf}(\mathrm{UT}_\bullet)$, $\mathrm{reg}_\bullet\rangle$ *is equal to* ζ_1; *in particular*

$$\mathrm{reg}_\bullet\rangle(\bar{\delta}_\gamma) = 1 \quad \textit{for } \gamma \in \mathcal{IG}.$$

Proof. This follows from direct computation: if $\gamma \in \mathcal{IG}_n$,

$$\mathrm{reg}_\bullet\rangle(\bar{\delta}_\gamma) = \langle \bar{\delta}_\gamma, \mathrm{reg}_{\mathrm{UT}_n}\rangle = \bar{\delta}_\gamma(1_n) = 1. \qquad \square$$

The uniqueness result of Theorem 2.2 now gives the following, which restates (5.4).

Corollary 5.16. *The map $\boldsymbol{c}_{\mathbb{1}}$ is the CHA morphism described in Theorem 5.14. In particular,*

$$\boldsymbol{c}_{\mathbb{1}}(\bar{\delta}_{\mathrm{Graph}(\pi)}) = G_\pi(\boldsymbol{x}; q^{-1}) \quad \textit{for } \pi \in \mathcal{D}.$$

Remark 5.17 (cf. Remarks 3.15(R1)). Corollary 5.16 can be used to give a novel proof that the unicellular LLT polynomial $G_\pi(\boldsymbol{x};t)$ has symmetric coefficients.

The proof of Theorem 5.1 is given below following two identities for LLT polynomials.

Proposition 5.18 [6, Theorem 2.1; 11, Proposition 3.4]. *Let n be a positive integer.*

(1) *For any Dyck paths $\pi \in \mathcal{D}_n$,*

$$q^{|\mathrm{Area}(\pi)|} G_{\mathrm{Dyck}(\pi)}(\boldsymbol{x}; q^{-1}) = \omega G_{\mathrm{Dyck}(\pi)}(\boldsymbol{x}; q).$$

(2) *For any tall Schröder paths $\sigma \in \mathcal{TS}_n$,*

$$(q-1)^{|\mathrm{Diag}(\sigma)|} G_\sigma(\boldsymbol{x}; q) = \sum_{S \subseteq \mathrm{Diag}(\sigma)} (-1)^{|\mathrm{Diag}(\sigma) - S|} G_{\mathrm{Area}^{-1}(\mathrm{Area}(\sigma) \cup S)}(\boldsymbol{x}; q),$$

where $\mathrm{Area}^{-1}(\mathrm{Area}(\sigma) \cup S)$ *denotes the unique Dyck path with area* $\mathrm{Area}(\sigma) \cup S$.

Proof of Theorem 5.1. For $\pi \in \mathcal{D}$, (2.11) states that $\overline{\chi}^{\mathrm{Graph}(\pi)} = q^{|\mathrm{Area}(\pi)|} \overline{\delta}_{\mathrm{Graph}(\pi)}$, so by Proposition 5.18(i),

$$\boldsymbol{c}_{\mathbb{1}}(\overline{\chi}^{\mathrm{Graph}(\pi)}) = \omega G_\pi(\boldsymbol{x}; q).$$

Combining this with Proposition 5.18(ii) and the linearity of ω,

$$\boldsymbol{c}_{\mathbb{1}}(\psi^\sigma) = \sum_{S \subseteq \mathrm{Diag}(\sigma)} (-1)^{|\mathrm{Diag}(\sigma) - S|} \omega G_{\pi + S}(\boldsymbol{x}; q) = (q-1)^{|\mathrm{Diag}(\sigma)|} \omega G_\sigma(\boldsymbol{x}; q). \qquad \square$$

6. Positivity conjectures

Recall the bases of Sym given in Section 2A. An element $f(\boldsymbol{x}; t) \in \mathrm{Sym}[t]$ is said to be *e-positive* if the coefficients $a_\lambda(t)$ in

$$f(\boldsymbol{x}; t) = \sum_{\lambda \in \mathcal{P}} a_\lambda(t) e_\lambda$$

are polynomials in t with nonnegative coefficients: $a_\lambda(t) \in \mathbb{Z}_{\geq 0}[t]$. Likewise, if the coefficients of $f(\boldsymbol{x}; t)$ in any other basis of Sym have this property — for example, the Schur basis $\{s_\lambda \mid \lambda \in \mathcal{P}\}$ — say that $f(\boldsymbol{x}; t)$ is positive in that basis. The positivity of the symmetric functions in this paper are of some interest, and this section will describe the meaning of positivity in the context of $\mathrm{GL}_n(\mathbb{F}_q)$ representation theory.

For the chromatic quasisymmetric functions in Section 3B, e-positivity generalizes the Stanley–Stembridge conjecture [42, Conjecture 5.5], which by [24] is the $t = 1$ case below.

Conjecture 6.1 [39, Conjecture 1.3]. *For each $\gamma \in \mathcal{IG}$, $X_\gamma(\boldsymbol{x}; t)$ is e-positive.*

Special cases of Conjecture 6.1 have explicit solutions, as in [1; 12; 29; 30].

For the vertical strip LLT polynomials in Section 5C, Schur positivity has implications for the study of Macdonald polynomials [28]. Adapting results from the case of general LLT polynomials, it is known [23, Corollary 6.9] that $G_\sigma(\boldsymbol{x}; t)$ is positive in the Schur basis for every $\sigma \in \mathcal{TS}$. However, their proof is algebraic and does not construct the Schur coefficients. In some special cases, explicit formulas are known, including the q-Kostka numbers [33] and the results of [30; 43], but in general these coefficients are a mystery.

Open Problem 6.2 [27, Open Problem 6.6]. *Find a (manifestly positive) combinatorial formula for the Schur coefficients of $G_\sigma(\boldsymbol{x}; t)$.*

The e-positivity of vertical strip LLT polynomials is also the subject of study; in this context, the paradigm is altered by considering the shifted polynomial $G_\sigma(\boldsymbol{x}; t+1)$. The e-positivity of shifted vertical strip LLT polynomials is proved in [13, Theorem 5.5], and the paper [6] gives an explicit combinatorial formula the e-coefficients, which will be restated in Section 6C. Using Theorem 5.1, this formula implies a result about the characters $\mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}(\psi^\sigma)$, inadvertently giving some representation theoretic intuition for the $t \leftrightarrow t+1$ shift.

Returning to the general discussion of positivity, if a polynomial $f(\boldsymbol{x}; t) \in \mathrm{Sym}[t]$ is positive with respect to a certain basis, then evaluating t at any positive integer will give a symmetric function with nonnegative integer coefficients in the chosen basis. Thus, evaluating $t = q$ above gives positivity results about the GL_n characters in this paper. Conversely, polynomial equations can be verified on any infinite set — like the set of prime powers — so GL_n characters offer a novel approach to some of the open problems above.

This section reinterprets each of the positivity statements above in the context of GL_n representation theory. Section 6A will discuss the e-positivity of the chromatic quasisymmetric function, Section 6B will discuss Schur positivity of the vertical strip LLT polynomials, and Section 6C will discuss the implications of the e-positivity of vertical strip LLT polynomials.

6A. *Interpreting the e-positivity of $X_\gamma(\boldsymbol{x}; t)$.* In light of Theorem 3.1, there should be a restatement of Conjecture 6.1 involving the characters $\mathrm{Ind}_{\mathrm{UT}_\gamma}^{\mathrm{GL}_n}(\mathbb{1})$. However, the isomorphism $\boldsymbol{p}_{\{1\}}$ in Theorem 3.1 does not associate e_λ to a character of GL_n, so some interpretation is required. My choice to use the particular restatement below is informed by ongoing work on the subject.

Recall the Steinberg character $\mathrm{St}_n \in \mathrm{cf}(\mathrm{GL}_n)$ defined in Section 5B. For any partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$, define $\mathrm{St}_\lambda \in \mathrm{cf}(\mathrm{GL}_\bullet)$ to be the product

$$\mathrm{St}_\lambda = \mathrm{St}_{\lambda_1} \mathrm{St}_{\lambda_2} \cdots \mathrm{St}_{\lambda_\ell}.$$

Conjecture 6.3. *Let $n \geq 0$ and $\gamma \in \mathcal{IG}_n$. There are polynomials $a_\lambda^\gamma(t) \in \mathbb{Z}_{\geq 0}[t]$ such that for each prime power q the character*

$$\eta_\gamma = \sum_{\lambda \in \mathcal{P}_n} a_\lambda^\gamma(q)\, \mathrm{St}_\lambda$$

satisfies $(q-1)^n \eta_\gamma(u) = \mathrm{Ind}_{\mathrm{UT}_\gamma}^{\mathrm{GL}_n}(\mathbb{1})(u)$ for every unipotent element $u \in \mathrm{GL}_n(\mathbb{F}_q)$.

Proposition 6.4. *Conjectures 6.1 and 6.3 are equivalent.*

Proof. For a class function $\psi \in \mathrm{cf}(\mathrm{GL}_n)$, write $\psi|_{\mathrm{uni}} \in \mathrm{cf}_{\mathrm{supp}}^{\mathrm{uni}}(\mathrm{GL}_\bullet)$ for the element defined by

$$\psi|_{\mathrm{uni}}(g) = \begin{cases} \psi(g) & \text{if } g \text{ is unipotent,} \\ 0 & \text{otherwise,} \end{cases}$$

so that Conjecture 6.3 states $\sum_{\lambda \in \mathcal{P}_n} a_\lambda^\gamma(q) \mathrm{St}_\lambda|_{\mathrm{uni}} = \frac{1}{(q-1)^n} \mathrm{Ind}_{\mathrm{UT}_\gamma}^{\mathrm{GL}_n}(\mathbb{1})$.

I now claim that $\boldsymbol{p}_{\{1\}}(\mathrm{St}_\lambda|_{\mathrm{uni}}) = e_\lambda$, so that with the preceding remarks and Theorem 3.1 the proof will be complete. The claim is relatively well-known to experts, but a proof sketch is included for the sake of completeness. Direct computation gives that $\mathrm{St}_n|_{\mathrm{uni}} = q^{\binom{n}{2}}\delta_{(1^n)}$ (see the proof of Theorem 5.11), and

$$\boldsymbol{p}_{\{1\}}(q^{\binom{n}{2}}\delta_{(1^n)}) = \tilde{P}_{(1^n)}(\boldsymbol{x}; q) = e_n,$$

with the second equality due to the definition of the Hall–Littlewood polynomial; see [35, III.2 (2.8)]. The claim then follows from the fact that the extension of $\psi \mapsto \psi_{\mathrm{uni}}$ to all of $\mathrm{cf}(\mathrm{GL}_\bullet)$ is a Hopf algebra homomorphism to $\mathrm{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_\bullet)$ [45, 10.1]. □

Remarks 6.5. (R1) A direct proof of Conjecture 6.3 would probably find an organic realization of the character η_γ using the representation theory of GL_n, and in a manner which does not depend on q. Ongoing work has identified a promising candidate for the character η_γ, but has not led to any progress on the conjecture itself.

(R2) It is not clear that Conjecture 6.3 offers an easier approach to Conjecture 6.1 than other equivalent statements. However, as the clearest restatement of Conjecture 6.1 in the $\mathrm{GL}_n(\mathbb{F}_q)$ context, the wide interest in e-positivity seems to justify its inclusion.

6B. ***Interpreting the Schur positivity of $G_\sigma(x; t)$.*** Let σ be a tall Schröder path and write

$$G_\sigma(\boldsymbol{x}; t) = \sum_{\lambda\in\mathcal{P}} b^\sigma_\lambda(t) s_\lambda.$$

It is immediate that each $b^\sigma_\lambda(t)$ is a polynomial in t with integral coefficients, and the content of [23, Corollary 6.9] is that the coefficients of this polynomial are nonnegative.

Recall from Section 5B that the irreducible unipotent characters of GL_n are $\{\chi^\lambda \mid \lambda \in \mathcal{P}_n\}$, and that $\boldsymbol{p}_{\mathbb{1}}(\chi^\lambda) = s_\lambda$ for each partition $\lambda \in \mathcal{P}$. Thus, for a tall Schröder path σ, Theorem 5.1 implies that

$$(q-1)^{|\mathrm{Diag}(\sigma)|} b^\sigma_\lambda(q) = \langle \chi^{\lambda'}, \mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}(\psi^\sigma)\rangle, \tag{6.6}$$

which is the multiplicity of the irreducible unipotent GL_n-module indexed by λ' in the GL_n-module affording $\mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}(\psi^\sigma)$. Thus, Theorem 5.1 implies the known fact that $b^\sigma_\lambda(q)$ is nonnegative for each prime power q, but falls short of giving a second proof of Schur positivity: a polynomial with negative coefficients can still take on infinitely many positive values. Nonetheless, progress on Open Problem 6.2 might be obtained through explicit representation theoretic formulas.

Open Problem 6.7. *For $n \geq 0$, $\sigma \in \mathcal{TS}_n$, and $\lambda \in \mathcal{P}_n$, find a combinatorial formula for $\langle \chi^{\lambda'}, \mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}(\psi^\sigma)\rangle$ as a function of q.*

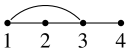
Such a formula would almost certainly be divisible by $(q-1)^{|\mathrm{Diag}(\sigma)|}$ in a straightforward manner; see Remark 5.5. This would give an answer to Open Problem 6.2.

6C. ***Interpreting the e-positivity of $G_\sigma(x; t)$.*** The final section of this paper will show how the explicit e-positivity formula for vertical strip LLT polynomials given in [6] leads to a deeper understanding of the characters $\mathrm{Ind}^{\mathrm{GL}}_{\mathrm{UT}}(\psi^\sigma)$ from Theorem 5.1; see Corollary 6.10. I will begin by recalling the main result of [6].

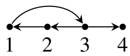
Fix a graph $\gamma = ([n], E(\gamma))$ on $[n]$. An *orientation* of γ is a collection of directed edges

$$\theta = \{(i, j) \mid \{i, j\} \in E(\gamma)\},$$

so that $([n], \theta)$ is a directed graph whose underlying undirected graph is γ. For example, with

$\gamma =$  and $\theta = \{(2, 1), (1, 3), (3, 2), (3, 4)\}$ (6.8)

we have

$([n], \theta) =$  .

Write $\mathcal{O}(\gamma)$ for the set of orientations of γ. For $\theta \in \mathcal{O}(\gamma)$ and $i \in [n]$, say that the *highest reachable vertex* from i under θ is

$$\mathrm{hrv}(\theta, i) = \max\{j \in [n] \mid \text{there is an } \textit{increasing} \text{ path in } ([n], \theta) \text{ from } i \text{ to } j\}.$$

For example, taking γ and θ as in (6.8)

$$\mathrm{hrv}(\theta, 1) = 4, \quad \mathrm{hrv}(\theta, 2) = 2, \quad \mathrm{hrv}(\theta, 3) = 4, \quad \text{and} \quad \mathrm{hrv}(\theta, 4) = 4.$$

Finally, for $\theta \in \mathcal{O}(\gamma)$, the *type* of θ is the partition $\mathrm{type}(\theta) \in \mathcal{P}_n$ obtained by truncating all zeros from the nonincreasing reordering of the sequence

$$\big(\big|\{i \in [n] \mid \mathrm{hrv}(\theta, i) = 1\}\big|, \ldots, \big|\{i \in [n] \mid \mathrm{hrv}(\theta, i) = n\}\big|\big).$$

For example, taking γ and θ as in (6.8), $\mathrm{type}(\theta) = (3, 1)$.

Theorem 6.9 [6, Theorem 2.9]. *For $n \geq 0$, let $\sigma \in \mathcal{TS}_n$ and let γ be the natural unit interval order on $[n]$ with edge set $E(\gamma) = \mathrm{Area}(\sigma) \cup \mathrm{Diag}(\sigma)$. Then*

$$G_\sigma(\boldsymbol{x}; t) = \sum_{\substack{\mathrm{Diag}(\sigma)\text{-ascending} \\ \theta \in \mathcal{O}(\gamma)}} (t-1)^{|\{\{i,j\} \in \mathrm{Area}(\sigma) \mid (i,j) \in \theta \text{ with } i<j\}|} e_{\mathrm{type}(\theta)},$$

where the sum is over orientations $\theta \in \mathcal{O}(\gamma)$ with $(i, j) \in \theta$ for each $i < j$ with $\{i, j\} \in \mathrm{Diag}(\sigma)$.

Evaluating the identity above at $t = q$, the expression $q - 1$ can be interpreted as $|\mathbb{F}_q^\times|$, the number of units in the field $\mathbb{F}_q$. As $|\mathbb{F}_q^\times|$ is a positive integer, it can be interpreted as the multiplicity of a submodule, as will be discussed at the end of this section.

The *Gelfand–Graev character* of GL_n [22] is the class function

$$\Gamma_n = \frac{1}{(q-1)^{n-1}} \mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}(\psi^{ED^{n-1}S}),$$

where $\psi^{ED^{n-1}S}$ is as defined in Section 5A; as the name suggests, Γ_n is actually a character of GL_n; see Remark 5.5. The *degenerate Gelfand–Graev character* [45, 12] indexed by a partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ is

$$\Gamma_\lambda = \Gamma_{\lambda_1} \cdots \Gamma_{\lambda_\ell}.$$

Corollary 6.10. *For $n \geq 0$, let $\sigma \in \mathcal{TS}_n$, and let γ be the natural unit interval order on $[n]$ with edge set $E(\gamma) = \mathrm{Area}(\sigma) \cup \mathrm{Diag}(\sigma)$. Then*

$$\mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}(\psi^{\sigma}) = \sum_{\substack{\mathrm{Diag}(\sigma)\text{-ascending} \\ \theta \in \mathcal{O}(\gamma)}} (q-1)^{|\{\{i,j\} \in E(\gamma) \mid (i,j) \in \theta \text{ with } i<j\}|} \Gamma_{\mathrm{type}(\theta)},$$

where the sum is over orientations $\theta \in \mathcal{O}(\gamma)$ with $(i, j) \in \theta$ for each $i < j$ with $\{i, j\} \in \mathrm{Diag}(\sigma)$.

Proof. Since the map $\boldsymbol{p}_{\mathbb{1}}$ restricts to an isomorphism from $\mathrm{cf}^{\mathrm{uni}}_{\mathrm{supp}}(\mathrm{GL}_{\bullet})$ to Sym (discussed in Section 5), and the involution ω is also an isomorphism, it is sufficient to establish that the above equation holds after the application of $\omega \circ \boldsymbol{p}_{\mathbb{1}}$ to both sides. By Theorems 5.1 and 6.9, the left side becomes

$$\omega \circ \boldsymbol{p}_{\mathbb{1}} \circ \mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}(\psi^{\sigma}) = \sum_{\substack{\mathrm{Diag}(\sigma)\text{-ascending} \\ \theta \in \mathcal{O}(\gamma)}} (q-1)^{|\{\{i,j\} \in E(\gamma) \mid (i,j) \in \theta \text{ with } i<j\}|} e_{\mathrm{type}(\theta)},$$

so the claim will follow from $\omega \circ \boldsymbol{p}_{\mathbb{1}}(\Gamma_n) = e_n$. This fact is known, but a short proof is included below for completeness.

Theorem 5.1 states that $\omega \circ \boldsymbol{p}_{\mathbb{1}}(\Gamma_n) = G_{ED^{n-1}S}(\boldsymbol{x}; q)$. With $\mathrm{Diag}(ND^{n-1}S) = \{\{i, i+1\} \mid 1 \leq i < n\}$, the definition of vertical strip LLT polynomials given in Section 5C becomes

$$G_{ED^{n-1}S}(\boldsymbol{x}; q) = \sum_{\substack{\kappa:[n] \to \mathbb{Z}_{>0} \\ \kappa(1) < \cdots < \kappa(n)}} x_{\kappa(1)} \cdots x_{\kappa(n)} = e_n. \qquad \square$$

This result implies that the GL_n-module affording $\mathrm{Ind}_{\mathrm{UT}}^{\mathrm{GL}}(\psi^{\sigma})$ decomposes into a direct sum of submodules that each afford some degenerate Gelfand–Graev character. Exhibiting this decomposition explicitly would give a new proof of Corollary 6.10 and Theorem 6.9.

Open Problem 6.11. *Find a module theoretic proof of Corollary 6.10.*

Acknowledgments

Along with [19], this paper is part of a Ph.D. thesis undertaken at the University of Colorado Boulder, and I am extremely thankful for the support and insights of my advisor Nathaniel Thiem throughout this process. I am also grateful to Martha Precup for a helpful conversation about Hessenberg varieties, and to an anonymous referee for suggesting many small improvements to this paper. An extended abstract of this work will appear in the proceedings of the FPSAC 2023 conference [20].

References

[1] A. Abreu and A. Nigro, "Chromatic symmetric functions from the modular law", *J. Combin. Theory Ser. A* **180** (2021), art. id. 105407. MR Zbl

[2] M. Aguiar, N. Bergeron, and F. Sottile, "Combinatorial Hopf algebras and generalized Dehn–Sommerville relations", *Compos. Math.* **142**:1 (2006), 1–30. MR Zbl

[3] M. Aguiar, C. André, C. Benedetti, N. Bergeron, Z. Chen, P. Diaconis, A. Hendrickson, S. Hsiao, I. M. Isaacs, A. Jedwab, K. Johnson, G. Karaali, A. Lauve, T. Le, S. Lewis, H. Li, K. Magaard, E. Marberg, J.-C. Novelli, A. Pang, F. Saliola, L. Tevlin, J.-Y. Thibon, N. Thiem, V. Venkateswaran, C. R. Vinroot, N. Yan, and M. Zabrocki, "Supercharacters, symmetric functions in noncommuting variables, and related Hopf algebras", *Adv. Math.* **229**:4 (2012), 2310–2337. MR Zbl

[4] M. Aguiar, N. Bergeron, and N. Thiem, "Hopf monoids from class functions on unitriangular matrices", *Algebra Number Theory* **7**:7 (2013), 1743–1779. MR Zbl

[5] P. Alexandersson and G. Panova, "LLT polynomials, chromatic quasisymmetric functions and graphs with cycles", *Discrete Math.* **341**:12 (2018), 3453–3482. MR Zbl

[6] P. Alexandersson and R. Sulzgruber, "A combinatorial expansion of vertical-strip LLT polynomials in the basis of elementary symmetric functions", *Adv. Math.* **400** (2022), art. id. 108256. MR Zbl

[7] F. Aliniaeifard and N. Thiem, "Pattern groups and a poset based Hopf monoid", *J. Combin. Theory Ser. A* **172** (2020), art. id. 105187. MR Zbl

[8] C. A. M. André, "Basic characters of the unitriangular group", *J. Algebra* **175**:1 (1995), 287–319. MR Zbl

[9] S. Andrews and N. Thiem, "The combinatorics of GL_n generalized Gelfand–Graev characters", *J. Lond. Math. Soc.* (2) **95**:2 (2017), 475–499. MR Zbl

[10] P. Brosnan and T. Y. Chow, "Unit interval orders and the dot action on the cohomology of regular semisimple Hessenberg varieties", *Adv. Math.* **329** (2018), 955–1001. MR Zbl

[11] E. Carlsson and A. Mellit, "A proof of the shuffle conjecture", *J. Amer. Math. Soc.* **31**:3 (2018), 661–697. MR Zbl

[12] L. Colmenarejo, A. H. Morales, and G. Panova, "Chromatic symmetric functions of Dyck paths and q-rook theory", *Eur. J. Combin.* **107** (2023), art. id. 103595. MR Zbl

[13] M. D'Adderio, "e-positivity of vertical strip LLT polynomials", *J. Combin. Theory Ser. A* **172** (2020), art. id. 105212. MR Zbl

[14] F. De Mari, C. Procesi, and M. A. Shayman, "Hessenberg varieties", *Trans. Amer. Math. Soc.* **332**:2 (1992), 529–534. MR Zbl

[15] P. Diaconis and I. M. Isaacs, "Supercharacters and superclasses for algebra groups", *Trans. Amer. Math. Soc.* **360**:5 (2008), 2359–2392. MR Zbl

[16] F. Digne and J. Michel, *Representations of finite groups of Lie type*, Lond. Math. Soc. Stud. Texts **21**, Cambridge Univ. Press, 1991. MR Zbl

[17] L. Escobar, M. Precup, and J. Shareshian, "Hessenberg varieties of codimension one in the flag variety", preprint, 2022. arXiv 2208.06299

[18] J. Fulman, "Descent identities, Hessenberg varieties, and the Weil conjectures", *J. Combin. Theory Ser. A* **87**:2 (1999), 390–397. MR Zbl

[19] L. Gagnon, "A $GL(\mathbb{F}_q)$-compatible Hopf algebra of unitriangular class functions", *J. Algebra* **632** (2023), 426–461. MR Zbl

[20] L. Gagnon, "A unipotent realization of the chromatic quasisymmetric function", *Sém. Lothar. Combin.* **89B** (2023), art. id. 72. MR Zbl

[21] V. Gasharov, "Incomparability graphs of $(3+1)$-free posets are s-positive", *Discrete Math.* **157**:1-3 (1996), 193–197. MR Zbl

[22] I. M. Gelfand and M. I. Graev, "Construction of irreducible representations of simple algebraic groups over a finite field", *Dokl. Akad. Nauk SSSR* **147** (1962), 529–532. In Russian. MR Zbl

[23] I. Grojnowski and M. Haiman, "Affine Hecke algebras and positivity of LLT and Macdonald polynomials", preprint, 2007, available at https://math.berkeley.edu/~mhaiman/ftp/llt-positivity/new-version.pdf.

[24] M. Guay-Paquet, "A modular relation for the chromatic symmetric functions of $(3+1)$-free posets", preprint, 2013. arXiv 1306.2400

[25] M. Guay-Paquet, "A second proof of the Shareshian–Wachs conjecture, by way of a new Hopf algebra", preprint, 2016. arXiv 1601.05498

[26] P. M. Gudivok, Y. V. Kapitonova, S. S. Polyak, V. P. Rudko, and A. I. Tsitkin, "Classes of conjugate elements of the unitriangular group", *Kibernetika* (*Kiev*) **1990**:1 (1990), 40–48. In Russian; translation in *Cybernetics* **26**:1 (1990), 47–57. MR Zbl

[27] J. Haglund, *The q,t-Catalan numbers and the space of diagonal harmonics*, Univ. Lect. Ser. **41**, Amer. Math. Soc., Providence, RI, 2008. MR Zbl

[28] J. Haglund, M. Haiman, and N. Loehr, "A combinatorial formula for Macdonald polynomials", *J. Amer. Math. Soc.* **18**:3 (2005), 735–761. MR Zbl

[29] M. Harada and M. E. Precup, "The cohomology of abelian Hessenberg varieties and the Stanley–Stembridge conjecture", *Algebr. Comb.* **2**:6 (2019), 1059–1108. MR Zbl

[30] J. Huh, S.-Y. Nam, and M. Yoo, "Melting lollipop chromatic quasisymmetric functions and Schur expansion of unicellular LLT polynomials", *Discrete Math.* **343**:3 (2020), art. id. 111728. MR Zbl

[31] C. Ji and M. Precup, "Hessenberg varieties associated to ad-nilpotent ideals", *Comm. Algebra* **50**:4 (2022), 1728–1749. MR Zbl

[32] N. Kawanaka, "Generalized Gelfand–Graev representations and Ennola duality", pp. 175–206 in *Algebraic groups and related topics* (Kyoto/Nagoya, 1983), edited by R. Hotta, Adv. Stud. Pure Math. **6**, North-Holland, Amsterdam, 1985. MR Zbl

[33] A. Lascoux and M.-P. Schützenberger, "Sur une conjecture de H. O. Foulkes", *C. R. Acad. Sci. Paris Sér. A-B* **286**:7 (1978), 323–324. MR Zbl

[34] A. Lascoux, B. Leclerc, and J.-Y. Thibon, "Ribbon tableaux, Hall–Littlewood functions, quantum affine algebras, and unipotent varieties", *J. Math. Phys.* **38**:2 (1997), 1041–1068. MR Zbl

[35] I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed., Oxford Univ. Press, 1995. MR Zbl

[36] E. Marberg, "A supercharacter analogue for normality", *J. Algebra* **332** (2011), 334–365. MR Zbl

[37] N. J. A. Sloane et al., "The on-line encyclopedia of integer sequences", available at http://oeis.org/.

[38] M. Precup and E. Sommers, "Perverse sheaves, nilpotent Hessenberg varieties, and the modular law", 2022. To appear in *Pure Appl. Math. Q.* arXiv 2201.13346

[39] J. Shareshian and M. L. Wachs, "Chromatic quasisymmetric functions", *Adv. Math.* **295** (2016), 497–551. MR Zbl

[40] R. P. Stanley, "A symmetric function generalization of the chromatic polynomial of a graph", *Adv. Math.* **111**:1 (1995), 166–194. MR Zbl

[41] R. P. Stanley, *Catalan numbers*, Cambridge Univ. Press, 2015. MR Zbl

[42] R. P. Stanley and J. R. Stembridge, "On immanants of Jacobi–Trudi matrices and permutations with restricted position", *J. Combin. Theory Ser. A* **62**:2 (1993), 261–279. MR Zbl

[43] F. Tom, "A combinatorial Schur expansion of triangle-free horizontal-strip LLT polynomials", *Comb. Theory* **1** (2021), art. id. 14. MR Zbl

[44] J. S. Tymoczko, "Hessenberg varieties are not pure dimensional", *Pure Appl. Math. Q.* **2**:3 (2006), 779–794. MR Zbl

[45] A. V. Zelevinsky, *Representations of finite classical groups: a Hopf algebra approach*, Lecture Notes in Math. **869**, Springer, 1981. MR Zbl

Communicated by Victor Reiner
Received 2023-02-21 Revised 2023-07-26 Accepted 2023-09-18

lgagnon@yorku.ca *Department of Mathematics, University of Colorado Boulder, Boulder, CO, 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 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 LaTeX but submissions in other varieties of TeX, 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 BibTeX 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 18 No. 9 2024