Algebra & Number Theory

Volume 14
2020
No. 4

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2020 is US $415/year for the electronic version, and $620/year (+$60, 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. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

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 14:4 (2020)
https://doi.org/10.2140/ant.2020.14.815

The distribution of p-torsion in degree p cyclic fields

Jack Klys

We compute all the moments of the p-torsion in the first step of a filtration of the class group defined by Gerth (1987) for cyclic fields of degree p, unconditionally for $p = 3$ and under GRH in general. We show that it satisfies a distribution which Gerth conjectured as an extension of the Cohen–Lenstra–Martinet conjectures. In the $p = 3$ case this gives the distribution of the 3-torsion of the class group modulo the Galois invariant part. We follow the strategy used by Fouvry and Klüners (2007) in their proof of the distribution of the 4-torsion in quadratic fields.

1. Introduction

Let K be a number field of degree n. Let Cl_K denote the class group and $\mathrm{Cl}_{K,p}$ denote the p-part. Let S be the set of finite abelian p-groups. We are interested in the question: what is the probability of any $A \in S$ occurring as $\mathrm{Cl}_{K,p}$ for K of degree n? The Cohen–Lenstra heuristics [Cohen and Lenstra 1984] propose an answer to this question for quadratic fields.

We make the question more precise as follows. Let D_K denote the discriminant of K. Let $\mathcal{D}_X^{\pm}$ be the set of real (resp. complex) quadratic fields with $|D_K| < X$. For any X define

$$S_X^{\pm}(A) = \frac{|\{K \in \mathcal{D}_X^{\pm} \mid \mathrm{Cl}_{K,p} \cong A\}|}{|\mathcal{D}_X^{\pm}|}.$$

The probability of A occurring as $\mathrm{Cl}_{K,p}$ in the family of real (resp. complex) quadratic fields is $\lim_{X\to\infty} S_X^{\pm}(A)$. In general this is not known to exist. Cohen and Lenstra conjectured that it does and proposed a distribution on S which should equal this quantity.

For $s \in \mathbb{Z}_{\geq 1} \cup \{\infty\}$ let

$$\eta_s(p) = \prod_{i=1}^{s}\left(1 - \frac{1}{p^i}\right).$$

One can show [Cohen and Lenstra 1984; Hall 1938] that

$$\sum_{G \in S} \frac{1}{|\mathrm{Aut}\, G|} = \frac{1}{\eta_\infty(p)} < \infty.$$

MSC2010: primary 11R29; secondary 11R37, 11R45.
Keywords: Cohen–Lenstra heuristics, arithmetic statistics, class groups, cyclic fields.

Then for any $A \in S$ and $u \geq 0$ define

$$\mu_u(A) = \frac{\eta_\infty(p)}{|\mathrm{Aut}\, A||A|^u}.$$

This defines a probability measure on S, called the Cohen–Lenstra distribution. They originally considered the case $n = 2$, $p \neq 2$ and considered complex quadratic and real quadratic fields separately.

Conjecture 1.1 (Cohen–Lenstra). *For $A \in S$*

$$\mu_0(A) = \lim_{X\to\infty} S_X^-(A), \quad \mu_1(A) = \lim_{X\to\infty} S_X^+(A).$$

These conjectures were extended to higher degree number fields by Cohen and Martinet [1987] again for $p \nmid n$.

No cases of these conjectures are known in full strength, though there has been much recent work on the subject. In the setting of number fields there are results giving the average size of the class group or subgroup thereof. There is the classical result of Davenport and Heilbronn [1971] and Datskovsky and Wright [1988] for the average size of 3-torsion of quadratic fields. There are also partial results for 8 and 16 torsion of quadratic fields due to Milovic [2017; 2018].

Below we will discuss in more detail the work of Fouvry and Klüners [2007] on 4 torsion of quadratic fields. Recently Smith [2017] has proven the distribution of the whole 2^∞-torsion of complex quadratic fields, thus generalizing their work.

There are also nonabelian versions which have been studied by Alberts [2016] and Bhargava [2014]. The conjectures have also been studied in the setting of function fields which provides additional tools such as moduli schemes. Some results here are the work of Ellenberg, Venkatesh and Westerland [Ellenberg et al. 2016], Boston and Wood [2017] and Wood [2019].

The original conjectures ignored the case when p divides the degree of the number fields. Gerth proposed a way of extending them to p-torsion in degree p cyclic fields by considering a certain subgroup of $\mathrm{Cl}_K[p]$ (which he calls the narrow principal genus — see Section 4 of [Gerth 1987]). He proved theorems providing compelling evidence for these conjectures (Theorems 4.3 and 5.11 in [Gerth 1984] and Theorem 2 in [Gerth 1987]) For the case $p = 2$ and $n = 2$ Gerth's extension implies the conjectures should hold in their original form, but with Cl_K^2 instead of Cl_K. This was proved by Fouvry and Klüners [2007]. To state their result, let $\mathrm{rk}_4(\mathrm{Cl}_K) = \mathrm{rk}_2(\mathrm{Cl}_K^2)$ and for any $k \in \mathbb{Z}_{\geq 1}$ let

$$M_k^\pm(2) = \lim_{X\to\infty} \frac{\sum_{K, 0<\pm D_K<X} 2^{k\,\mathrm{rk}_4(\mathrm{Cl}_K)}}{\sum_{K,0<\pm D_K<X} 1}.$$

Define $\mathcal{N}(k, p)$ to be the number of subspaces of $\mathbb{F}_p^k$.

Theorem 1.2 (Fouvry–Klüners). *For every $k \in \mathbb{Z}_{\geq 1}$*

$$M_k^-(2) = \mathcal{N}(k, 2), \quad M_k^+(2) = \mathcal{N}(k+1, 2) - \mathcal{N}(k, 2).$$

By a separate result Fouvry and Klüners [2006] deduce that these moments are enough to determine a distribution.

Theorem 1.3 (Fouvry–Klüners). *The density of complex quadratic fields K with $\mathrm{rk}_4(\mathrm{Cl}_K) = s$ is*

$$\frac{\eta_\infty(2)}{\eta_s^2(2)2^{s^2}}$$

and the density of real quadratic fields with $\mathrm{rk}_4(\mathrm{Cl}_K) = s$ is

$$\frac{\eta_\infty(2)}{\eta_s(2)\eta_{s+1}(2)2^{s(s+1)}}.$$

Gerth conjectured a distribution for a certain subgroup of $\mathrm{Cl}_K[p]$ of cyclic p fields for all p. To state it we first define some notation. Throughout the paper p will denote an odd prime. Let K be a cyclic field of degree p with Galois group $G = \langle \sigma_K \rangle$. Let $\varphi_K = 1 - \sigma_K$ act on $\mathrm{Cl}_K[p]$. It can be shown (see Section 3) there is a filtration

$$\mathrm{Cl}_K[p]^G = \ker \varphi_K \subseteq \ker \varphi_K^2 \subseteq \cdots \subseteq \ker \varphi_K^{p-1} = \mathrm{Cl}_K[p]. \tag{1-1}$$

Then Gerth conjectured a distribution for the p-rank of $\varphi_K(\ker \varphi_K^2)$. Notice that for $p = 3$ we have $\ker \varphi_K^2 = \mathrm{Cl}_K[3]$ and so the above filtration implies $\varphi_K(\ker \varphi_K^2) \cong \mathrm{Cl}_K[3]/\mathrm{Cl}_K[3]^G$. We prove the following theorem which verifies Gerth's conjecture for $p = 3$:

Theorem 1.4. *The density of cyclic cubic fields with $\mathrm{rk}_3(\mathrm{Cl}_K[3]/\mathrm{Cl}_K[3]^G) = s$ is*

$$\frac{\eta_\infty(3)}{\eta_s(3)\eta_{s+1}(3)3^{s(s+1)}}.$$

We can extend this to all odd p under the assumption of GRH for Artin L-functions (we remark the L-functions we will consider are all known to be entire, and as such we do not need to assume Artin's holomorphy conjecture).

Theorem 1.5. *Assume GRH for Artin L-functions. Let p be odd. The density of degree p cyclic fields with $\mathrm{rk}_p(\varphi_K(\ker \varphi_K^2)) = s$ is*

$$\frac{\eta_\infty(p)}{\eta_s(p)\eta_{s+1}(p)p^{s(s+1)}}.$$

The above filtration (1-1) is analogous to the filtration

$$\mathrm{Cl}_{K,2}^G = \mathrm{Cl}_K[2] \subseteq \mathrm{Cl}_K[4] \subseteq \cdots \subseteq \mathrm{Cl}_{K,2}$$

when $p = 2$ and the object $\varphi_K(\ker \varphi_K^2)$ is hence analogous to $\mathrm{Cl}_K[4]^2 \cong \mathrm{Cl}_K[4]/\mathrm{Cl}_K[2]$ from Theorem 1.2. Since the completion of this paper Koymans and Pagano [2018] have extended the methods of Smith [2017] to determine the distribution of $\mathrm{Cl}_K[p^\infty]/\mathrm{Cl}_K[p^\infty]^G$ for odd p (conditional on the generalized Riemann hypothesis). They in fact prove a refined result which implies the distribution of $\varphi_K(\ker \varphi_K^{j+1})/\varphi_K(\ker \varphi_K^j)$ for all j.

Before continuing we make some remarks about $\mathrm{Cl}_K[p]^G$. It is the part of $\mathrm{Cl}_K[p]$ corresponding by class field theory to the genus field of K, that is the maximal unramified extension of K which is abelian over $\mathbb{Q}$. It can be shown $|\mathrm{Cl}_K[p]^G| = p^{r-1}$ where r is the number of primes ramified in K and that the average of $\mathrm{rk}_p(\mathrm{Cl}_K[p]^G)$ is ∞.

In the case $p = 2$ this quantity is $\mathrm{Cl}_K[4]^G = \mathrm{Cl}_K[2]$ and hence

$$\mathrm{rk}_2 \mathrm{Cl}_K^2 = \mathrm{rk}_2(\mathrm{Cl}_K[4]/\mathrm{Cl}_K[2]),$$

that is removing this part corresponds to replacing the 2-rank of Cl_K by 4-rank as defined above.

We deduce Theorems 1.4 and 1.5 from the following theorem together with [Fouvry and Klüners 2006]. Define

$$M_k(p) = \lim_{X\to\infty} \frac{\sum_{K, D_K < X} p^{k\,\mathrm{rk}_p(\varphi_K(\ker \varphi_K^2))}}{\sum_{K, D_K<X} 1}.$$

Theorem 1.6. *Let $k \in \mathbb{Z}_{\geq 1}$. Then unconditionally for $p = 3$ and under the assumption of GRH for Artin L-functions for $p > 3$ we have*

$$M_k(p) = \mathcal{N}(k+1, p) - \mathcal{N}(k, p).$$

The proof of Theorem 1.6 follows the strategy of Fouvry and Klüners. For any degree p cyclic field K we express $|\mathrm{Cl}_K[p]|$ using a sum of idele class characters, and then sum over all degree p cyclic fields of discriminant up to X. We then study the asymptotics of this expression using techniques from analytic number theory.

In the $p = 3$ case we require several versions of a large sieve inequality for cubic characters to bound the error term. We prove one such version as well as applying several others from the literature, due to Heath-Brown [2000], Baier and Young [2010] and Iwaniec and Kowalski [2004]. The reason for assuming the generalized Riemann hypothesis in the general case is that certain versions of the large sieve are not yet available for order p characters. In particular we lack analogs of Propositions 6.3 and 6.4. This is the only obstacle to an unconditional proof for all p.

Finally we remark briefly about an equivalent formulation of the Cohen–Lenstra conjectures which is commonly used. The distribution μ_u is characterized by the fact (see for instance [Ellenberg et al. 2016, Lemma 8.2]) that for all $A \in S$

$$\mathbb{E}_{G\sim\mu_u}(|\mathrm{Sur}(G, A)|) = \sum_{G\in S} \mu_u(G)\cdot|\mathrm{Sur}(G, A)| = \frac{1}{|A|^u}. \tag{1-2}$$

This is often called the A-moment of μ_u and computing it only for certain A can still provide information about the distribution of elements in Cl_K.

It is clear that $|\mathrm{Hom}(G, (\mathbb{Z}/p\mathbb{Z})^k)| = p^{k\,\mathrm{rk}_p(G)}$. Furthermore

$$|\mathrm{Hom}(G, (\mathbb{Z}/p\mathbb{Z})^k)| = \sum_{i=0}^{k} n(k, i, p)|\mathrm{Sur}(G, (\mathbb{Z}/p\mathbb{Z})^i)|$$

where $n(k, i, p)$ is the number of i-dimensional subspaces of $\mathbb{F}_p^k$. Hence Theorem 1.6 can be rephrased as computing the A moments in the above sense for all the groups $A = (\mathbb{Z}/p\mathbb{Z})^k$.

2. Preliminaries

2A. ***Class field theory.*** For any number field K Galois over $\mathbb{Q}$ and rational prime l let $K_\mathfrak{p}$ be the completion of K at the prime $\mathfrak{p} \mid l$. Let $N_\mathfrak{p} : K_\mathfrak{p} \to \mathbb{Q}_l$ be the norm map. Denote $K_l = K \otimes_\mathbb{Q} \mathbb{Q}_l$ and $N_l = \prod_{\mathfrak{p} \mid l} N_\mathfrak{p}$. Note the isomorphism $K_l \cong \prod_{\mathfrak{p} \mid l} K_\mathfrak{p}$ which lets us view N_l as a function on K_l.

Let C_K denote the idele class group of K. Let $N_{C_K} : C_K \to C_\mathbb{Q}$ denote the norm map defined by $(N_{C_K}\alpha)_l = N_l\big(\prod_{\mathfrak{p} \mid l} \alpha_\mathfrak{p}\big) = \prod_{\mathfrak{p} \mid l} N_\mathfrak{p}\alpha_\mathfrak{p}$.

In several places we will use the following isomorphism of $C_\mathbb{Q}$ with $\mathbb{R}_+ \times \prod_l \mathbb{Z}_l^\times$. For $x \in C_\mathbb{Q}$ there exists a unique $a_x \in \mathbb{Q}^\times$ such that $a_x \cdot x \in \mathbb{R}_+ \times \prod_l \mathbb{Z}_l^\times$. It is not hard to see that $x \mapsto a_x \cdot x$ is well defined and bijective.

Define the morphism $\langle\cdot\rangle_\mathfrak{p} : K_\mathfrak{p} \to C_K$ by $\langle b\rangle_\mathfrak{p} = (\dots, 1, b, 1, \dots)$ the class of the element with b in the $\mathfrak{p}$-th coordinate and 1 elsewhere. We will need the following lemma in Section 4.

Lemma 2.1. *Let $b \in \mathbb{Q}^*$ and $\langle b\rangle_l \in C_\mathbb{Q}$. Then $b \in N_l K_l$ if and only if $\langle b\rangle_l \in N_{C_K} C_K$.*

Proof. If $b = N_l\alpha_l$ for some $\alpha_l \in K_l$ then clearly $\langle b\rangle_l = N_{C_K}(\dots, 1, \alpha_l, 1, \dots)$.

For the converse note that, under the natural embedding of $\mathbb{Q}_l^*$ into $C_\mathbb{Q}$, we have $N_{C_K} C_K \cap \mathbb{Q}_l^* = N_l K_l^*$ by Corollary 5.8 from [Neukirch 1999, Section VI.5, page 394]. Hence if $\langle b\rangle_l \in N_{C_K} C_K$ then it follows immediately from the definition of N_{C_K} that $b \in N_l K_l^*$. □

2B. ***Cyclic degree p fields.*** Let $K/\mathbb{Q}$ be a degree p cyclic extension (p an odd prime). Then the discriminant is of the form $D_K = (p_1 \cdots p_r)^{p-1}$ where each p_i is either a prime congruent to 1 mod p or equal to p^2 and they are distinct. Conversely every integer of this form is a discriminant of a degree p cyclic field [Mayer 1992].

By class field theory each such extension corresponds to a character $\hat{\chi}$ of $C_\mathbb{Q}$ with $\ker \hat{\chi} = N_{C_K}(C_K)$ an index p subgroup of $C_\mathbb{Q}$. Through the identification $C_\mathbb{Q} \cong \mathbb{R}_+ \times \prod_l \mathbb{Z}_l^\times$ $\hat{\chi}$ descends to a character

$$\chi : (1 + p\mathbb{Z}_p) \times \prod_{l \mid D_K, l \neq p} \mathbb{F}_l^\times \to \mu_p \tag{2-1}$$

where the $(1 + p\mathbb{Z}_p)$ factor appears if and only if $p \mid D_K$. The character χ is nontrivial on each factor.

By an order p character we will mean a character $\chi_l : \mathbb{F}_l^\times \to \mu_p$ for any prime $l \neq p$, or $\chi_p : (1 + p\mathbb{Z}_p) \to \mu_p$. For each prime l (including $l = p$) there are $p - 1$ such distinct nontrivial characters. Thus χ factors into a product of order p characters $\chi = \prod_{l \mid D_K} \chi_l$. Hence there are $(p-1)^{\omega(n)}$ distinct nontrivial characters defined on the domain in (2-1) where $\omega(n)$ denotes the number of distinct prime divisors of n. Furthermore two distinct characters χ, ψ have the same kernel if and only if $\chi = \psi^i$ for some $1 \le i \le p - 1$.

It follows from these facts that for a fixed integer n which is a discriminant of a degree p cyclic extensions of $\mathbb{Q}$, the number of such extensions $K/\mathbb{Q}$ with $D_K = n$ is $(p-1)^{\omega(n)-1}$.

The following asymptotic formula for the number of degree p cyclic fields with discriminant up to X^{p-1} is well-known [Wright 1989]

$$\sum_{K, D_K < X^{p-1}} 1 \sim cX.$$

2C. ***The field $\mathbb{Q}(\zeta_3)$ and cubic reciprocity.*** Let ζ_3 be a cube root of unity. The following facts about the field $\mathbb{Q}(\zeta_3)$ and the cubic residue symbol can be found for instance in [Baier and Young 2010, Section 2.1].

The extension $\mathbb{Q}(\zeta_3)/\mathbb{Q}$ is quadratic, and for $x \in \mathbb{Q}(\zeta_3)$ we denote its Galois conjugate by $\bar{x}$. The ring of integers of $\mathbb{Q}(\zeta_3)$ is $\mathcal{O} = \mathbb{Z}[\zeta_3]$. It is a principal ideal domain and every ideal $(n) \subset \mathcal{O}$ with $(n, 3) = 1$ has a unique generator n which satisfies $n \equiv 1 \bmod 3\mathcal{O}$. The only prime which ramifies is $(3) = ((1-\zeta_3)^2)$. The primes of $\mathbb{Z}$ which split in $\mathcal{O}$ are exactly the ones congruent to $1 \bmod 3\mathbb{Z}$.

The set $\{1, \zeta_3\}$ is a basis for $\mathcal{O}$, so that every element of $\mathcal{O}$ can be written as $a + b\zeta_3$. Letting $N : \mathbb{Q}(\zeta_3) \to \mathbb{Q}$ denote the norm map we have the formula $N(a + b\zeta_3) = a^2 + b^2 - ab$. Using this it can be shown that $|\{a + b\zeta_3 \in \mathcal{O} \mid N(a + b\zeta_3) \le X\}| = O(X)$ and for a fixed $b \in \mathbb{Z}$ that

$$|\{a \in \mathbb{Z} \mid N(a + b\zeta_3) \le X\}| = O(X^{1/2}).$$

For $n, m \in \mathcal{O}$ coprime and $(m, 3) = 1$ denote by $\left(\frac{n}{m}\right)_3$ the cubic residue symbol. It satisfies

$$\left(\frac{\bar{n}}{\bar{m}}\right)_3 = \overline{\left(\frac{n}{m}\right)_3}, \quad \left(\frac{n}{m}\right)_3^2 = \overline{\left(\frac{n}{m}\right)_3} \tag{2-2}$$

and, if additionally $n, m \equiv 1 \bmod 3\mathcal{O}$ then there is the law of cubic reciprocity

$$\left(\frac{n}{m}\right)_3 = \left(\frac{m}{n}\right)_3. \tag{2-3}$$

3. Counting p-torsion in degree p cyclic fields

The first goal is to describe the subgroup of class group whose distribution we will be computing. See [Stevenhagen 1995] for a slightly different treatment of some of the material found in this section.

We recall that p will always denote an odd prime.

Let K be a degree p cyclic extension of $\mathbb{Q}$ with Galois group $G = \langle \sigma_K \rangle$. Let $\varphi_K = 1 - \sigma_K$. We view σ_K as a morphism acting on $\mathrm{Cl}_K[p^\infty]$. Let $NG = \sum_{i=0}^{p-1} \sigma_K^i$. Then since $NG : \mathrm{Cl}_K \to \mathrm{Cl}_{\mathbb{Q}}$ and the latter is trivial we have $\ker NG = \mathrm{Cl}_K$. Thus we can view $\mathrm{Cl}_K[p^\infty]$ as a module over the ring $\mathbb{Z}_p[\sigma_K]/\langle NG \rangle$.

It can be shown that in $\mathbb{Z}_p[\sigma_K]/\langle NG \rangle$ there is the relation $(\varphi_K^{p-1}) = (p)$. Thus there is a filtration

$$\mathrm{Cl}_K[p]^G = \ker \varphi_K \subseteq \ker \varphi_K^2 \subseteq \cdots \subseteq \ker \varphi_K^{p-1} = \mathrm{Cl}_K[p].$$

From this we can write down the exact sequence

$$1 \to \mathrm{Cl}_K[p]^G \to \ker \varphi_K^2 \to \varphi_K(\ker \varphi_K^2) \to 1$$

so that $|\ker \varphi_K^2| = |\mathrm{Cl}_K[p]|^G |\varphi_K(\ker \varphi_K^2)|$.

Note that as a special case of the ambiguous class number formula (see [Lemmermeyer 2013, Theorem 1]) we have $|\mathrm{Cl}_K| = p^{r-1}$. This implies $\mathrm{Cl}_K^G \subset \mathrm{Cl}_K[p^\infty]$.

Denote by N the norm map $N_{K/\mathbb{Q}}$ (both on ideals and elements of K). Let $\mathcal{J}$ be the group of fractional ideals of K. Furthermore let $P_1, \dots, P_r$ be the ramified primes of K, and let $\mathcal{B} = \{P_1^{e_k} \cdots P_r^{e_r} \mid e_i = 0, 1, \dots, p-1\}$. For any $I \in \mathcal{J}$ let $\bar{I}$ denote the natural projection to Cl_K.

Lemma 3.1. *Let $\overline{\mathcal{B}}$ be the projection of $\mathcal{B}$ to Cl_K. Then $|\overline{\mathcal{B}}| = p^{r-1}$.*

Proof. Clearly $\overline{\mathcal{B}} \subset \mathrm{Cl}_K[p]^G$. We will show $\overline{\mathcal{B}}$ generates Cl_K^G and the lemma will follow from $|\mathrm{Cl}_K^G| = p^{r-1}$.

Let $I \in \mathcal{J}$ such that $\bar{I} \in \mathrm{Cl}_K^G$, so that $I^{\sigma_K} = (\alpha)I$ for some $\alpha \in K$. Applying N to both sides gives $N(\alpha) = 1$, hence multiplying by -1 if necessary, we have $N\alpha = 1$ in K. By Hilbert's Theorem 90 there exists $\beta \in K$ such that $\alpha = \beta^{1-\sigma_K}$.

Thus $I^{\sigma_K} = (\beta)^{1-\sigma_K} I$ and rearranging $((\beta)I)^{\sigma_K} = (\beta)I$. So $(\beta)I$ is fixed by σ_K in $\mathcal{J}$. This implies gI is divisible only by ramified and rational primes in K. Thus $\overline{gI} \in \overline{\mathcal{B}}$. This completes the proof. □

Next we give another description of $\varphi_K(\ker \varphi_K^2)$.

Lemma 3.2. *Consider N acting on $\mathcal{J}$ the group of fractional ideals of K. Then*

$$\ker N = \varphi_K(\mathcal{J}).$$

Proof. It is clear that $\varphi_K(\mathcal{J}) \subset \ker N$. Suppose $NI = 1$ for some ideal $I \in \mathcal{J}$. Then I can only be divisible by split primes. Let $q \in \mathbb{Z}$ be a prime above which I is supported and let $Q_1, \dots, Q_p$ be all the prime ideals in K lying above q such that $Q_i^{\sigma_K} = Q_{i+1}$. Then $N(Q_1^{a_1} Q_2^{a_2} \cdots Q_p^{a_p}) = q^{\sum a_i}$ which implies that $\sum a_i = 0$. Then $Q_1^{a_1} Q_2^{a_2} \cdots Q_p^{a_p} = (Q_1^{a_1} Q_2^{a_1+a_2} \cdots Q_{p-1}^{a_1+\cdots+a_{p-1}})^{1-\sigma_K}$. Applying this to all primes q below I shows $I \in \varphi_K(\mathcal{J})$. □

Lemma 3.3. *For any $I \in \mathcal{J}$ such that $\bar{I} \in \mathrm{Cl}_K[p]^G$ we have*

$$\bar{I} \in \varphi_K(\ker \varphi_K^2) \iff NI = N(\alpha) \textit{ in } \mathcal{J} \textit{ for some } \alpha \in K$$

(*note this condition is independent of the ideal representing $\bar{I}$*).

Proof. Suppose first that $\bar{I} \in \varphi_K(\ker \varphi_K^2)$. So for some $\bar{J} \in \mathrm{Cl}_K$, we have $\bar{I} = \bar{J}^{1-\sigma_K}$. We have for some $\alpha \in K^\times$ that $(\alpha)I = J^{1-\sigma_K}$ in $\mathcal{J}$. Taking norm of this gives $NI = N(\alpha^{-1})$ which proves one direction.

Now suppose that $NI = N(\alpha)$ for some $\alpha \in K$. Hence $I = (\alpha)J$ for some ideal $J \in \ker N$. By Lemma 3.2 we have $J = J_0^{\varphi_K}$ for some ideal $J_0 \in \mathcal{J}$. So $\bar{I} = \bar{J}_0^{\varphi_K}$. Then $\bar{J}_0^{\varphi_K^2} = \bar{I}^{\varphi_K} = 1$ in Cl_K since $\bar{I} \in \mathrm{Cl}_K^G$. Thus $\bar{J}_0 \in \ker \varphi_K^2$. □

Let

$$\tilde{D}_K = \begin{cases} D_K & \text{if } p \nmid D_K, \\ D_K/p^{p-1} & \text{if } p \mid D_K. \end{cases} \tag{3-1}$$

Proposition 3.4. *With the above notation let $\Omega(\tilde{D}_K)$ be the set of positive integers dividing $\tilde{D}_K$. Then*

$$|\varphi_K(\ker \varphi_K^2)| = \tfrac{1}{p} |\{b \in \Omega(\tilde{D}_K) \mid b = N\alpha \textit{ for some } \alpha \in K^\times\}|.$$

Proof. Firstly it is clear that the map $N : \mathcal{B} \to \Omega(\tilde{D}_K)$ is a bijection. Since $|\bar{\mathcal{B}}| = p^{r-1}$ by Lemma 3.1, let $\{(1), (l_2) \dots, (l_p)\} \subset \mathcal{B}$ be the principal ideals. To each $J \in \mathrm{Cl}_K[p]^G = \bar{\mathcal{B}}$ we associate the set $\omega_J = \{N(I_J), N(I_J l_2), \dots, N(I_J l_p)\} \subset \mathcal{J}$ where I_J is a choice of representative of J supported on the ramified primes. For every element of ω_J we may choose the unique generator which is a positive integer. With this identification we may assume $\omega_J \subset \Omega(\tilde{D}_K)$.

We claim that if $J_1 \neq J_2$ in $\bar{\mathcal{B}}$ then $\omega_{J_1} \cap \omega_{J_2} = \varnothing$. Suppose to the contrary that $x \in \omega_{J_1} \cap \omega_{J_2}$ so $N(I_{J_1} l_i) = N(I_{J_2} l_j)$ for some i, j, and hence $N(I_{J_1}) = N(I_{J_2} l_k)$ for some l_k since the principal ideals of $\mathcal{B}$ form a subgroup. Since N is injective on $\mathcal{B}$ we get $I_{J_1} = I_{J_2} l_k$, a contradiction. Thus we can write as a disjoint union

$$\Omega(\tilde{D}_K) = \bigcup_{J \in \bar{\mathcal{B}}} \omega_J.$$

Then also

$$\{b \in \Omega(\tilde{D}_K) \mid b = N\alpha \text{ for some } \alpha \in K^\times\} = \bigcup_J \omega_J$$

where the union is over all J such that some element of ω_J is of the form $N\alpha$ for some $\alpha \in K$. Note this condition is equivalent to every element of ω_J is of the form $N\alpha$ for some $\alpha \in K$ since if $N(I_J l_i) = N(\alpha)$ then $N(I_J l_j) = N(\alpha l_j l_i^{-1})$.

By Lemma 3.3 $|\varphi_K(\ker \varphi_K^2)|$ is the number of classes $J \in \mathrm{Cl}_K[p]^G$ such that $NI = N(\alpha)$ for some (any) representative I of J and some $\alpha \in K$. In the above notation this is the set of $J \in \bar{\mathcal{B}}$ such that $N\alpha \in \omega_J$ for some $\alpha \in K$. Thus

$$|\varphi_K(\ker \varphi_K^2)| = \frac{1}{p} \bigcup_J |\omega_J|$$

where the union is over all J such that every element of ω_J is of the form $N\alpha$ for some $\alpha \in K$. This completes the proof. □

4. The *p*-torsion as a character sum

The goal of this section will be to prove a formula for the size of $\varphi_K(\ker \varphi_K^2)$ defined in the previous section.

Let $K/\mathbb{Q}$ be a cyclic degree p extension with discriminant D_K and Galois group $\mathrm{Gal}(K/\mathbb{Q}) = \langle \sigma_K \rangle$. Let $\hat{\chi}$ be a character of C_K corresponding to K and χ its quotient (see Section 2B). Recall that χ factors into a product of order p characters $\chi = \prod_{l \mid D_K} \chi_l$.

For simplicity we will henceforth write $\mathrm{im}(\varphi_K)$ for $\varphi_K(\ker \varphi_K^2)$.

Recall the definition of $\tilde{D}_K$ in (3-1) and that $\omega(\tilde{D}_K)$ denotes the number of distinct prime divisors of $\tilde{D}_K$. For any prime l of $\mathbb{Q}$ we defined the morphism $\langle \cdot \rangle_l : K_l \to C_K$ by $\langle b \rangle_l = (\dots, 1, b, 1, \dots)$ the element with b in the l-th coordinate.

Proposition 4.1. *For each degree p cyclic field K let σ_K denote a generator of the Galois group and D_K the discriminant and $\hat{\chi}$ a corresponding character. Then*

$$|\mathrm{im}(\varphi_K)| = \tfrac{1}{p^{\omega(\tilde{D}_K)+1}} \sum_{(b_0,\ldots,b_{p-1})} \prod_{i=0}^{p-1} \prod_{\substack{l \mid b_i \\ l \text{ prime}}} (1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle B\rangle_l)$$

where the sum is over p-tuples of coprime positive integers (b_i) satisfying

$$\left(\prod_{i=0}^{p-1} b_i\right)^{p-1} = \tilde{D}_K \quad \textit{and} \quad B = b_1 b_2^2 \cdots b_{p-1}^{p-1}.$$

Proof. Since $K/\mathbb{Q}$ is cyclic by the Hasse norm theorem $b \in \mathbb{Q}^\times$ is a global norm if and only if b is a local norm everywhere:

$$b = N\alpha \text{ for some } \alpha \in K \iff b = N_l\alpha_l \text{ for some } \alpha_l \in K_l, \text{ for all } l.$$

Recall $K_l = K \otimes_{\mathbb{Q}} \mathbb{Q}_l$. Hence by Proposition 3.4 we want to detect when $b \mid \tilde{D}_K$ satisfies $b \in N_l K_l$ for all l. If $l \nmid \tilde{D}_K$ this condition is trivial since b is then a local unit in $\mathbb{Q}_l$, and it is a standard fact from local class field theory that if l is unramified in $K/\mathbb{Q}$ then $N_l : K_l \to \mathbb{Q}_l$ surjects onto the local units. Hence we need to check the condition only for $l \mid \tilde{D}_K$.

By Lemma 2.1 $b \in N_l K_l$ if and only if the idele $\langle b\rangle_l \in C_{\mathbb{Q}}$ satisfies $\langle b\rangle_l \in N_{C_K} C_K = \ker\hat{\chi}$. Since $\hat{\chi}$ has order p this implies

$$(1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle b\rangle_l) = \begin{cases} p & \text{if } \langle b\rangle_l \in \ker\hat{\chi}, \\ 0 & \text{else.} \end{cases}$$

Hence we arrive at the following expression which detects when b is a norm at l:

$$\left(\frac{1+\hat{\chi}+\cdots+\hat{\chi}^{p-1}}{p}\right)(\langle b\rangle_l) = \begin{cases} 1 & \text{if } b \in N_l K_l, \\ 0 & \text{else.} \end{cases} \tag{4-1}$$

Note that $\tilde{D}_K$ is a $p-1$ power. Write a divisor of $\tilde{D}_K$ as $b_1 b_2^2 \cdots b_{p-1}^{p-1}$ where the b_i are square-free and coprime. Let

$$G(b) = \begin{cases} 1 & \text{if } b \in NK^\times, \\ 0 & \text{else.} \end{cases}$$

Thus by Proposition 3.4 we have

$$p|\mathrm{im}(\varphi_K)| = \sum_{b_1 b_2^2 \cdots b_{p-1}^{p-1} \mid \tilde{D}_K} G(b_1 b_2^2 \cdots b_{p-1}^{p-1})$$

where the sum is over positive integers dividing $\tilde{D}_K$. In the following we will let $B = b_1 b_2^2 \cdots b_{p-1}^{p-1}$. By (4-1) we get

$$p|\mathrm{im}(\varphi_K)| = \sum_{(b_0,\dots,b_{p-1})} \prod_{l \mid \tilde{D}_K} \left(\frac{1+\hat{\chi}+\cdots+\hat{\chi}^{p-1}}{p} \right)(\langle B \rangle_l)$$
$$= \frac{1}{p^{\omega(\tilde{D}_K)}} \sum_{(b_0,\dots,b_{p-1})} \prod_i \prod_{l \mid b_i} (1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle B \rangle_l)$$

where the sum is over p-tuples of positive integers satisfying $(\prod_i b_i)^{p-1} = \tilde{D}_K$. This completes the proof. □

We define some notation to state the next proposition. Define the function $\Phi : \mathbb{F}_p^2 \to \mathbb{F}_p$ by

$$\Phi(u, v) = \Phi((u_1, u_2), (v_1, v_2)) = u_1(v_2 - u_2). \tag{4-2}$$

Under the identification $C_\mathbb{Q} \cong \mathbb{R}_+ \times \prod_l \mathbb{Z}_l^\times$ (see Section 2A) for any integer b the class of $\langle b \rangle_l$ maps to

$$\left(\dots, \frac{1}{l^i}, \frac{b}{l^i}, \frac{1}{l^i}, \dots \right)$$

where $i = \mathrm{ord}_l\, b$. Hence if χ decomposes as $\prod_{l \mid D_K} \chi_l$ acting on $(1 + p\mathbb{Z}_p) \times \prod_{l \mid D_K, l \neq p} \mathbb{F}_l^\times$ (see Section 2B) then

$$\hat{\chi}(\langle b \rangle_l) = \chi_l\left(\frac{b}{l^i}\right) \prod_{q \neq l} \chi_q\left(\frac{1}{l^i}\right). \tag{4-3}$$

Theorem 4.2. *For each degree p cyclic field K let σ_K denote a generator of the Galois group and D_K the discriminant and $\chi = \prod_{l \mid D_K} \chi_l$ a corresponding character. Then*

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(\tilde{D}_K)+1}} \sum_{(D_u)} \prod_{v \in \mathbb{F}_p^2} \prod_{\substack{l \mid D_v \\ l \text{ prime}}} \chi_l \left(\prod_{u \in \mathbb{F}_p^2} D_u^{\Phi(u,v)} \right) \tag{4-4}$$

where the sum is over p^2-tuples of coprime positive integers (D_u) indexed by $u \in \mathbb{F}_p^2$ and satisfying $\left(\prod_{u \in \mathbb{F}_p^2} D_u\right)^{p-1} = \tilde{D}_K$.

Proof. For this proof we will denote $\tilde{D}_K = D^{p-1}$. By Proposition 4.1 we have

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(b_0,\dots,b_{p-1})} \prod_{i=0}^{p-1} \prod_{l \mid b_i} (1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle B \rangle_l) \tag{4-5}$$

where the sum is over p-tuples of positive integers satisfying $\prod_i b_i = D$ and $B = b_1 b_2^2 \cdots b_{p-1}^{p-1}$. We fix i and focus on the innermost product in (4-5). Expanding it we get

$$\prod_{l \mid b_i} (1+\hat{\chi}+\cdots+\hat{\chi}^{p-1})(\langle B \rangle_l) = \sum_{j_1,\dots,j_{\omega(b_i)}} \prod_{l \mid b_i} \hat{\chi}^{j_l}(\langle B \rangle_l) \tag{4-6}$$

where the sum is over tuples of integers with each $0 \le j_k \le p-1$. For any such fixed tuple $j_1, \ldots, j_{\omega(b_i)}$ and for each $0 \le j \le p-1$ define D_{ip+j} to be the product of all primes $l \mid b_i$ such that $j_l = j$. Then we can instead write (4-6) as

$$\sum_{j_1,\ldots,j_{\omega(b_i)}} \prod_{l \mid b_i} \hat{\chi}^{j_l}(\langle B\rangle_l) = \sum_{(D_{ip},D_{ip+1},\ldots,D_{ip+p-1})} \prod_{j=0}^{p-1} \prod_{l \mid D_{ip+j}} \hat{\chi}^{j}(\langle B\rangle_l)$$

where the sum on the right is over all p-tuples of positive integers satisfying $\prod_{j=0}^{p-1} D_{ip+j} = b_i$. Thus we get

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(b_0,\ldots,b_{p-1})} \prod_{i=0}^{p-1} \Bigg(\sum_{(D_{ip},D_{ip+1},\ldots,D_{ip+p-1})} \prod_{j=0}^{p-1} \prod_{l \mid D_{ip+j}} \hat{\chi}^{j}(\langle B\rangle_l) \Bigg)$$

where the inner sum is over all p-tuples of positive integers satisfying $\prod_{j=0}^{p-1} D_{ip+j} = b_i$. If $l \mid b_i$ then $\mathrm{ord}_l\, B = i$ hence by (4-3) we have

$$\hat{\chi}^{j}(\langle B\rangle_l) = \hat{\chi}(\langle B^j\rangle_l) = \chi_l\left(\frac{B^j}{l^{ij}}\right) \prod_{q \mid D, q \ne l} \chi_q\left(\frac{1}{l^{ij}}\right).$$

Plugging this in and rearranging summations we get

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\ldots,D_{p^2-1})} \prod_{i,j=0}^{p-1} \prod_{l \mid D_{ip+j}} \left[\chi_l\left(\frac{B^j}{l^{ij}}\right) \prod_{q \mid D, q \ne l} \chi_q\left(\frac{1}{l^{ij}}\right)\right] \tag{4-7}$$

where the first sum is over all p^2-tuples of positive integers satisfying $\prod_{i=0}^{p^2-1} D_i = D$. Taking the last two products from (4-7) and rearranging them gives

$$\prod_{l \mid D_{ip+j}} \left[\chi_l\left(\frac{B^j}{l^{ij}}\right) \prod_{q \mid D, q \ne l} \chi_q\left(\frac{1}{l^{ij}}\right)\right]$$
$$= \left[\prod_{q \mid D/D_{ip+j}} \chi_q\left(\prod_{l \mid D_{ip+j}} \frac{1}{l^{ij}} \right)\right] i \left[\prod_{q \mid D_{ip+j}} \chi_q\left(\prod_{l \mid D_{ip+j}/q} \frac{1}{l^{ij}} \right)\right] \times \left[\prod_{l \mid D_{ip+j}} \chi_l\left(\frac{B^j}{l^{ij}}\right)\right]$$

and grouping products and renaming the variable l to q in the last term gives

$$= \left[\prod_{q \mid D/D_{ip+j}} \chi_q\left(\frac{1}{D_{ip+j}^{ij}}\right)\right] \left[\prod_{q \mid D_{ip+j}} \chi_q\left(\frac{q^{ij}}{D_{ip+j}^{ij}}\right)\right] \left[\prod_{q \mid D_{ip+j}} \chi_q\left(\frac{B^j}{q^{ij}}\right)\right]. \tag{4-8}$$

Define $A_{i,j}, B_{i,j}, C_{i,j}$ to be respectively the first, second, and third factors in (4-8).

Let $\tilde{D} = \prod_{i,j=0}^{p-1} D_{ip+j}^{ij}$, $\tilde{D}_j = \prod_{i=0}^{p-1} D_{ip+j}^{ij}$ and $\overline{D}_j = \prod_{i=0}^{p-1} D_{ip+j}$. Then the last two terms in (4-8) can be combined:

$$B_{i,j} \cdot C_{i,j} = \prod_{q \mid D_{ip+j}} \chi_q\left(\frac{q^{ij}}{D_{ip+j}^{ij}}\right) \prod_{q \mid D_{ip+j}} \chi_q\left(\frac{B^j}{q^{ij}}\right) = \prod_{q \mid D_{ip+j}} \chi_q\left(\frac{B^j}{D_{ip+j}^{ij}}\right). \tag{4-9}$$

Let $B_j = \prod_{i=0}^{p-1} B_{i,j} C_{i,j}$.

Taking the product over $i=0,\dots,p-1$ of the first factor in (4-8) gives

$$\prod_{i=0}^{p-1} A_{i,j} = \Bigg[\prod_{q\,|\,D/\overline{D}_j} \chi_q\Big(\frac{1}{\tilde{D}_j}\Big)\Bigg]\Bigg[\prod_{i=0}^{p-1}\prod_{q\,|\,D_{ip+j}} \chi_q\Big(\frac{D^{ij}_{ip+j}}{\tilde{D}_j}\Big)\Bigg]. \tag{4-10}$$

Define $E_{1,j}$, $E_{2,j}$ to be respectively the first and second factor in (4-10).

So far we have shown

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\dots,D_{p^2-1})} \prod_{j=0}^{p-1} (E_{1,j}\cdot E_{2,j}\cdot B_j). \tag{4-11}$$

Now we can compute

$$\prod_{j=0}^{p-1} E_{1,j} = \prod_{j=0}^{p-1} \prod_{q\,|\,\overline{D}_j} \chi_q\Big(\frac{\tilde{D}_j}{\tilde{D}}\Big) \tag{4-12}$$

and

$$\prod_{j=0}^{p-1}(E_{2,j}B_j) = \prod_{i,j=0}^{p-1}\prod_{q\,|\,D_{ip+j}} \chi_q\Big(\frac{D^{ij}_{ip+j}}{\tilde{D}_j}\Big) \prod_{q\,|\,D_{ip+j}} \chi_q\Big(\frac{B^j}{D^{ij}_{ip+j}}\Big) = \prod_{j=0}^{p-1}\prod_{q\,|\,\overline{D}_j} \chi_q\Big(\frac{B^j}{\tilde{D}_j}\Big). \tag{4-13}$$

Plugging (4-12) and (4-13) into (4-11) we get

$$\begin{aligned} |\mathrm{im}(\varphi_K)| &= \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\dots,D_{p^2-1})} \prod_{j=0}^{p-1}\prod_{q\,|\,\overline{D}_j} \chi_q\Big(\frac{B^j}{\tilde{D}_j}\Big) \prod_{q\,|\,\overline{D}_j} \chi_q\Big(\frac{\tilde{D}_j}{\tilde{D}}\Big) \\ &= \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\dots,D_{p^2-1})} \prod_{j=0}^{p-1}\prod_{l\,|\,\overline{D}_j} \chi_l\Big(\frac{B^j}{\tilde{D}}\Big) \\ &= \frac{1}{p^{\omega(D)+1}} \sum_{(D_0,\dots,D_{p^2-1})} \prod_{i,j=0}^{p-1}\prod_{l\,|\,D_{ip+j}} \chi_l\Big(\frac{B^j}{\tilde{D}}\Big). \end{aligned}$$

From the definition of B we have

$$B = \prod_{i=0}^{p-1} (D_{ip}D_{ip+1}\cdots D_{ip+p-1})^i$$

hence for $0\le u_1,u_2,v_1,v_2\le p-1$ the exponent of $D_{u_1p+u_2}$ in $B^{v_2}/\tilde{D}$ is

$$u_1v_2 - u_1u_2 = u_1(v_2-u_2).$$

Let $u=(u_1,u_2)\in\mathbb{F}_p^2$ and $v=(v_1,v_2)\in\mathbb{F}_p^2$. Recall we defined $\Phi:\mathbb{F}_p^2\to\mathbb{F}_p$ by

$$\Phi(u,v) = u_1(v_2-u_2).$$

Thus relabelling $D_{ip+j} \mapsto D_{(i,j)}$ we conclude that

$$|\mathrm{im}(\varphi_K)| = \frac{1}{p^{\omega(D)+1}} \sum_{(D_{(0,0)},\dots,D_{(p-1,p-1)})} \prod_{v\in\mathbb{F}_p^2} \prod_{l \mid D_v} \chi_l\Bigl(\prod_{u\in\mathbb{F}_p^2} D_u^{\Phi(u,v)}\Bigr). \tag{4-14}$$

as required. □

5. An expression for the k-th moment

Define

$$S_k(X) = \sum_{K,\,D_K<X^{p-1}} |\mathrm{im}(\varphi_K)|^k.$$

Computing $S_k(X)$ will allow us to determine the k-th moment

$$M(k) = \lim_{X\to\infty} \frac{S_k(X)}{\sum_{K,D_K<X^{p-1}} 1}$$

of the function $|\mathrm{im}(\varphi_K)|$. We will then show that knowing $M(k)$ for all $k \in \mathbb{Z}_{\geq 1}$ will be enough to determine the distribution of the values of $|\mathrm{im}(\varphi_K)|$. Our goal for the remainder of the paper will thus be computing $S_k(X)$.

We want to use Theorem 4.2 to obtain a formula for $|\mathrm{im}(\varphi_K)|^k$.

Proposition 5.1. *For each degree p cyclic field K let σ_K denote a generator of the Galois group and D_K the discriminant and $\chi = \prod_{l \mid D_K} \chi_l$ a corresponding character. Then for any $k \in \mathbb{Z}_{k\geq 1}$*

$$|\mathrm{im}(\varphi_K)|^k = \frac{1}{p^k \cdot p^{k\omega(\tilde{D}_K)}} \sum_{(D_u)} \prod_{v\in(\mathbb{F}_p^2)^k} \prod_{\substack{l \text{ prime} \\ l \mid D_v}} \chi_l\Bigl(\prod_{u\in(\mathbb{F}_p^2)^k} D_u^{\Phi_k(u,v)}\Bigr)$$

where the sum is over p^{2k}-tuples of coprime positive integers (D_u) indexed by $u \in (\mathbb{F}_p^2)^k$ satisfying $\prod_{u\in(\mathbb{F}_p^2)^k} D_u = \tilde{D}_K$.

Proof. From Theorem 4.2 we see $|\mathrm{im}(\varphi_K)|^k$ involves a k-fold product of summations over factorizations of $\tilde{D}_K$, so we want to simultaneously consider k different factorizations of $\tilde{D}_K$. We follow the same method as in [Fouvry and Klüners 2007, pages 471–472], and denote any k factorizations of $\tilde{D}_K$ as

$$\tilde{D}_K = \prod_{u_1\in\mathbb{F}_p^2} D_{u_1}^g = \cdots = \prod_{u_k\in\mathbb{F}_p^2} D_{u_k}^{(k)}$$

where each index $u_i \in \mathbb{F}_p^2$ (note this differs from the notation in the previous section). Define $D_{u_1,\dots,u_k} = \gcd(D_{u_1}^g, \dots, D_{u_k}^{(k)})$. From this we obtain a further factorization of each $D_{u_l}^{(l)}$ by

$$D_{u_l}^{(l)} = [\prod_{\substack{1\leq j\leq k \\ j\neq l}} [\prod_{u_j\in\mathbb{F}_p^2} D_{u_1,\dots,u_k}.$$

Hence taking (4-4) in Theorem 4.2 to the k-th power we get

$$\frac{1}{p^k \cdot p^{k\omega(\tilde{D}_K)}} \sum_{(D^g_{u_1})} \cdots \sum_{(D^{(k)}_{u_k})} \prod_{i=1}^{k} \prod_{v_i \in \mathbb{F}_p^2} \prod_{\substack{l \text{ prime} \\ l \mid D^{(i)}_{v_i}}} \chi_l\Bigl(\prod_{u_i \in \mathbb{F}_p^2} (D^{(i)}_{u_i})^{\Phi(u_i,v_i)} \Bigr)$$

where the summations are over p^2-tuples of positive integers $(D^{(i)}_{u_i})$ such that $\prod_{u_i \in \mathbb{F}_p^2} D^{(i)}_{u_i} = \tilde{D}_K$. By multiplicativity of the χ_l we can simplify this as

$$\frac{1}{p^k \cdot p^{k\omega(\tilde{D}_K)}} \sum_{(D_{u_1,\dots,u_k})} \prod_{(v_1,\dots,v_k) \in (\mathbb{F}_p^2)^k} \prod_{\substack{l \text{ prime} \\ l \mid D_{v_1,\dots,v_k}}} \chi_l\Bigl(\prod_{(u_1,\dots,u_k) \in (\mathbb{F}_p^2)^k} D_{u_1,\dots,u_k}^{\Phi(u_1,v_1)+\cdots+\Phi(u_k,v_k)} \Bigr)$$

where the sum is over all p^{2k}-tuples of positive integers $(D_{u_1,\dots,u_k})$ such that

$$\prod_{(u_1,\dots,u_k) \in (\mathbb{F}_p^2)^k} D_{u_1,\dots,u_k} = \tilde{D}_K.$$

To simplify notation we let $u = (u_1, \dots, u_k)$, $v = (v_1, \dots, v_k)$, and $\Phi_k(u, v) = \sum_{i=1}^k \Phi(u_i, v_i)$. Then the expression becomes

$$\frac{1}{p^k \cdot p^{k\omega(\tilde{D}_K)}} \sum_{(D_u)} \prod_{v \in (\mathbb{F}_p^2)^k} \prod_{\substack{l \text{ prime} \\ l \mid D_v}} \chi_l\Bigl(\prod_{u \in (\mathbb{F}_p^2)^k} D_u^{\Phi_k(u,v)} \Bigr)$$

where the sum is over p^{2k}-tuples of coprime positive integers with $\prod_{u \in (\mathbb{F}_p^2)^k} D_u = \tilde{D}_K$. □

We now sum the expression from Proposition 5.1 over all degree p cyclic fields with discriminant up to X. To this end we define the following notation:

Let $\mathcal{P}(X)$ denote the set of p^{2k}-tuples of coprime positive integers (D_u) indexed by $u = (u_1, \dots, u_k) \in \mathbb{F}_p^{2k}$ (with $u_i \in \mathbb{F}_p^2$) whose prime factors are congruent to 1 mod p or equal to p and $p^{\mathrm{ord}_p(D)} D < X$ where we denote $D = \prod_{u \in (\mathbb{F}_p^2)^k} D_u$.

Let $\mathcal{C}(D)$ be the set of tuples of nontrivial order p characters $(\chi_{l'})_{l' \mid D}$ (see Section 2B).

Theorem 5.2. *For each degree p cyclic field K let σ_K denote a generator of the Galois group and D_K the discriminant. Then for any $k \in \mathbb{Z}_{k \geq 1}$*

$$\sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k = \frac{1}{(p-1) \cdot p^k} \sum_{(D_u) \in \mathcal{P}(X)} \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \frac{\mu^2(D)}{p^{k\omega(D)}} \times \prod_{v \in \mathbb{F}_p^{2k}} \prod_{\substack{l \text{ prime} \\ l \mid D_v}} \chi_l\Bigl(\prod_{u \in (\mathbb{F}_p^2)^k} D_u^{\Phi_k(u,v)} \Bigr)$$

where on the right hand side we denote $D = \prod_{u \in (\mathbb{F}_p^2)^k} D_u$ and $\Phi_k(u, v) = \sum_{i=1}^k \Phi(u_i, v_i)$ with Φ defined in (4-2).

Note that we are summing over cyclic degree p fields satisfying $D_K < X^{p-1}$ but on the right-hand side the condition is $p^{\mathrm{ord}_p(D)} D < X$.

Proof. Fix a degree p cyclic field K. Summing ((4-4) over all tuples of order p characters $(\chi_{l'})_{l' \mid \tilde{D}_K}$ characters corresponds to summing over all degree p cyclic fields of discriminant D_K but overcounts by a factor of $p-1$ since for a fixed discriminant D_K the characters $\prod_{l \mid \tilde{D}_K} \chi_l$ and $\prod_{l \mid \tilde{D}_K} \chi_l^j$ for $1 \le j \le p-1$ correspond to the same field (see Section 2B). Thus for any $D \in \mathbb{Z}$ which is a discriminant of a degree p cyclic field, by Proposition 5.1 we get

$$\sum_{K, D_K = D} |\mathrm{im}(\varphi_K)|^k = \frac{1}{(p-1)} \frac{1}{p^k p^{k\omega(D)}} \sum_{(D_u)} \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \prod_{v \in \mathbb{F}_p^{2k}} \prod_{\substack{l \mid D_v \\ l \text{ prime}}} \chi_l \Bigl(\prod_{u \in \mathbb{F}_p^{2k}} D_u^{\Phi_k(u,v)} \Bigr)$$

where the sum is over p^{2k}-tuples of coprime positive integers (D_u) indexed by $u \in (\mathbb{F}_p^2)^k$ satisfying $\prod_{u \in (\mathbb{F}_p^2)^k} D_u = \tilde{D}$ ($\tilde{D}$ defined as in (3-1)). Since we are interested in computing the average over all degree p Galois fields we sum over these to get

$$\sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k = \frac{1}{(p-1)p^k} \sum_{\substack{D \in \mathbb{Z} \\ 0 < p^{\mathrm{ord}_p(D)} D < X}} \mu^2(D) \frac{1}{p^{k\omega(D)}} \sum_{(D_u)} \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \prod_{v \in \mathbb{F}_p^{2k}} \prod_{\substack{l \mid D_v \\ l \text{ prime}}} \chi_l \Bigl(\prod_{u \in \mathbb{F}_p^{2k}} D_u^{\Phi_k(u,v)} \Bigr)$$

and $\prod_{u \in (\mathbb{F}_p^2)^k} D_u = D$. Then by definition of $\mathcal{P}(X)$ we obtain

$$\sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k = \frac{1}{(p-1)p^k} \sum_{(D_u) \in \mathcal{P}(X)} \mu^2(D) \frac{1}{p^{k\omega(D)}} \times \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \prod_{v \in \mathbb{F}_p^{2k}} \prod_{\substack{l \mid D_v \\ l \text{ prime}}} \chi_l \Bigl(\prod_{u \in \mathbb{F}_p^{2k}} D_u^{\Phi_k(u,v)} \Bigr).$$

This proves the theorem. □

The goal will now be an asymptotic analysis of this formula.

6. Analytic tools

We list the analytic results that will be needed in the sequel. The first two we take directly from [Fouvry and Klüners 2007].

Lemma 6.1. *There exists an absolute constant B_0, such that for every $X \ge 3$ and every $l \ge 0$ we have*

$$|\{n \le X \mid \omega(n) = l, \mu^2(n) = 1\}| \le B_0 \frac{X}{\log X} \frac{(\log\log X + B_0)^l}{l!}.$$

Lemma 6.2. *Let $\gamma \in \mathbb{R}$ with $\gamma > 0$. Then we have*

$$\sum_{X - Y < n < X} \gamma^{\omega(n)} \ll Y (\log X)^{\gamma - 1}$$

for $2 \le X \exp(-\sqrt{\log X}) \le Y \le X$.

Let $\mathcal{O} = \mathbb{Z}[\zeta_3]$, the ring of integers of the quadratic extension $\mathbb{Q}(\zeta_3)$. Let $\left(\frac{x}{y}\right)_3$ denote the cubic residue symbol for $x, y \in \mathcal{O}$ coprime. Let $N(\cdot) : \mathbb{Q}(\zeta_3) \to \mathbb{Q}$ denote the norm function. In the following A, B, Q will denote positive integers.

We will need the following results for estimating bilinear sums. They are all versions of the large sieve inequality. The first two containing the $(AB)^\epsilon$-type factor will be used when A and B are close together, and the latter two which do not contain this factor will be used when A and B are far apart. The first is Theorem 2 from [Heath-Brown 2000].

Proposition 6.3. *Let c_n be a sequence of complex numbers indexed by elements of $\mathcal{O}$. Then for any $\epsilon > 0$*

$$\sum_{\substack{m\in\mathcal{O},\\ N(m)\le A}} \left| \sum_{\substack{n\in\mathcal{O},\\ N(n)\le B}} \mu^2(N(n)N(m))c_n\left(\frac{n}{m}\right)_3\right|^2 \ll_\epsilon (A+B+(AB)^{2/3})(AB)^\epsilon \sum_{n\in\mathcal{O}}|c_n|^2$$

where each of the sums are over square-free elements $m, n \in \mathcal{O}$ congruent to $1 \bmod 3$.

Next we have a version for cubic Dirichlet characters and sums over integers, which is Theorem 1.4 from [Baier and Young 2010].

Proposition 6.4. *Let c_m be any sequence of complex numbers indexed by $\mathbb{N}$. Then for any $\epsilon > 0$*

$$\sum_{\substack{q\in\mathbb{N}\\ Q<q<2Q}} \sum_{\chi \bmod q} \left| \sum_{\substack{m\in\mathbb{N}\\ A<m<2A}} c_m\mu^2(m)\chi(m)\right|^2 \ll_\epsilon (Q^{11/9}+Q^{2/3}A)(QA)^\epsilon \sum_{m\in\mathbb{N}} \mu^2(m)|c_m|^2$$

where the second sum is over χ which are primitive Dirichlet characters satisfying $\chi^3 = 1$.

The next version is Theorem 7.13 from [Iwaniec and Kowalski 2004, Section 7.5, page 179] (due to Bombieri and Davenport) and applies to all Dirichlet characters (not necessarily cubic).

Proposition 6.5. *Let c_n be a sequence of complex numbers indexed by $\mathbb{N}$. Then*

$$\sum_{\substack{q\in\mathbb{N}\\ q<Q}} \sum_{\chi \bmod q} \left| \sum_{\substack{m\in\mathbb{N}\\ A<m<2A}} c_m\mu^2(m)\chi(m)\right|^2 \ll (Q^2+A) \sum_{m\in\mathbb{N}} |c_m|^2$$

where the second sum is over χ which are primitive Dirichlet characters.

Finally we will need a generalized version of Siegel–Walfisz for character sums, stated as Main Theorem in [Goldstein 1970]. We state a slightly weaker simplified version here.

Proposition 6.6. *Let $\epsilon > 0$. Let $K/\mathbb{Q}$ be Galois of degree n and let χ be a nontrivial finite Hecke character of K with conductor f_χ. Then there exists a positive constant $c = c(\epsilon)$, not depending on K or χ such that*

$$\sum_{\substack{\mathfrak{p}\subset\mathcal{O}_K \text{ prime}\\ N(\mathfrak{p})\le x\\ (\mathfrak{p},f_\chi)=1}} \chi(\mathfrak{p}) = O(dx\log^2 x\exp(-cn(\log x)^{1/2}/d))$$

where $d = n^3|D_K N(f_\chi)|^\epsilon c^{-n}$. The implied constant does not depend on K or χ.

We now prove another version of a large sieve bound for cubic characters.

We require a preliminary lemma. The next result is Exercise 2 from [Iwaniec and Kowalski 2004, Section 7.4, page 178]. In their terminology a set of $\alpha_r = (\alpha_{r,1}, \ldots, \alpha_{r,k}) \in \mathbb{R}^k$ is δ-spaced if $\max_i |\alpha_{r,i} - \alpha_{r',i}| \geq \delta$ for all $r \neq r'$. The definition extends to elements of $\mathbb{R}^k/\mathbb{Z}^k$ by choosing representatives in $\mathbb{R}^k$ for which $|\alpha_{r,i} - \alpha_{r',i}|$ is minimal, for all i (that is, which make the spacing minimal).

Lemma 6.7. *Let $d \geq 1$ and $\delta > 0$ and let $\alpha_r = (\alpha_{r,1}, \ldots, \alpha_{r,d})$ be δ-spaced points in $\mathbb{R}^d/\mathbb{Z}^d$ and a_n a sequence in $\mathbb{C}$ indexed by $n = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ with $1 \leq n_i \leq N$. Then*

$$\sum_r \Big| \sum_n a_n \exp(2\pi i (n \cdot \alpha_r)) \Big|^2 \ll_d (\delta^{-d} + N^d) \sum_n |a_n|^2.$$

Proposition 6.8. *For each $n \in \mathcal{O}$ let ψ_n be a primitive cubic Hecke character of modulus $(n) \subset \mathcal{O}$.*

For any $d \in \mathcal{O}$ let z_d be the smallest positive integer such that $d^{-1} = z/z_d$ for some $z \in \mathcal{O}$ (clearly $z_d \leq |N(d)|$). Let $\mathcal{P}(B) \subset \mathcal{O}$ be a set of elements d satisfying $z_d < B$.

For all $d \in \mathcal{O}$ let $a_d \in \mathbb{C}$ such that $|a_d| \leq 1$. Then

$$\sum_{n \in \mathcal{P}(B)} \Big| \sum_{\substack{m \in \mathcal{O} \\ N(m) \leq A}} a_m \psi_n(m) \Big|^2 \ll (B^2 + A)A.$$

Proof. For $r, n \in \mathcal{O}$ define the generalized Gauss sum for the character ψ as

$$g(r, n) = \sum_{d \in (\mathcal{O}/n)^\times} \psi_n(d) \check{e}(rd/n)$$

where $\check{e}(z) = \exp(2\pi i (z + \bar{z}))$. It satisfies the property (see [Baier and Young 2010, Section 2.2])

$$g(rs, n) = \overline{\psi_n(s)} g(r, n)$$

for $s \in \mathcal{O}$ coprime to n. We will write $g(n) = g(1, n)$.

We can write the sum in the statement as

$$\sum_n \Big| \sum_m a_m \psi_n(m) \Big|^2 = \sum_n \frac{1}{|g(n)|^2} \Big| \sum_m a_m g(m, n) \Big|^2 = \sum_n \frac{1}{|g(n)|^2} \Big| \sum_{d \in (\mathcal{O}/n)^\times} \psi_n(d) \sum_m a_m \check{e}(dm/n) \Big|^2.$$

Since ψ_n is a primitive character of $(\mathcal{O}/n)^\times$ we can sum over all such characters to get the bound

$$\sum_n \Big| \sum_m a_m \psi_n(m) \Big|^2 \leq \sum_n \frac{1}{|g(n)|^2} \sum_\chi \Big| \sum_{d \in (\mathcal{O}/n)^\times} \chi(d) \sum_m a_m \check{e}(dm/n) \Big|^2$$

where the summation $\sum_\chi$ is over primitive characters of $(\mathcal{O}/n)^\times$. Then expanding the square and using orthogonality of characters we obtain, with the notation $b_d = \sum_m a_m \check{e}(dm/n)$,

$$\begin{aligned}\sum_n \Big|\sum_m a_m \psi_n(m)\Big|^2 &\le \sum_n \frac{1}{|g(n)|^2} \sum_\chi \sum_{d,d' \in (\mathcal{O}/n)^\times} \chi(d)\bar{\chi}(d') b_d \bar{b}_{d'} \\ &\le \sum_n \frac{1}{|g(n)|^2} \sum_{d,d' \in (\mathcal{O}/n)^\times} b_d \bar{b}_{d'} \sum_\chi \chi(dd'^{-1}) \\ &\le \sum_n \frac{|(\mathcal{O}/n)^\times|}{|g(n)|^2} \sum_{d \in (\mathcal{O}/n)^\times} \Big|\sum_m a_m \check{e}(dm/n)\Big|^2 \\ &\le \sum_n \sum_{d \in (\mathcal{O}/n)^\times} \Big|\sum_m a_m \check{e}(dm/n)\Big|^2. \end{aligned} \tag{6-1}$$

We now want to apply the multivariable large sieve inequality of Lemma 6.7 so we will rewrite the summation accordingly.

Let $R = \{(n, d) \mid n \in \mathcal{P}(B), d \in (\mathcal{O}/n)^\times\}$. For any $(n, d) \in R$, using $d \in \mathcal{O}$ to also denote any choice of representative of $d \in (\mathcal{O}/n)^\times$ (everything that follows will be independent of such a choice), write $d/n = d_1 + \zeta_3 d_2$ in the basis $\{1, \zeta_3\}$ with $d_i \in \mathbb{Q}$ and similarly write $m = s_1 + \zeta_3 s_2$ with $s_i \in \mathbb{Z}$. Then a computation shows $dm/n = (d_1 s_1 - d_2 s_2, d_1 s_2 + d_2 s_1 - d_2 s_2)$ in coordinates in $\{1, \zeta_3\}$, and

$$\operatorname{tr}(dm/n) = dm/n + \overline{dm}/\bar{n} = s_1(2d_1 - d_2) + s_2(-d_1 - d_2) = (s_1, s_2) \cdot (2d_1 - d_2, d_1 - d_2).$$

So given $r = (n, d) \in R$ define $\alpha_r = (2d_1 - d_2, d_1 - d_2) \in \mathbb{Q}^2$. Then

$$\check{e}(dm/n) = \exp(2\pi i \operatorname{tr}(dm/n)) = \exp(2\pi i (s_1, s_2) \cdot \alpha_r)$$

Hence we can rewrite (6-1) as

$$\sum_n \Big|\sum_m a_m \psi_n(m)\Big|^2 \le \sum_{r \in R} \Big|\sum_{(s_1,s_2) \in \mathbb{Z}^2} a_{s_1 + \zeta s_2} \exp(2\pi i (s_1, s_2) \cdot \alpha_r)\Big|^2 \tag{6-2}$$

where $s_1, s_2 \ll A^{1/2}$ since $N(m) \le A$ (see Section 2C).

We claim the sequence $S = \{\alpha_r\}_{r \in R}$ is $1/B$-spaced (see definition before Lemma 6.7). For any $(n, d) \in R$, d is coprime to n so the map $R \to \mathbb{Q}(\zeta_3)$ defined by $(n, d) \mapsto d/n$ is injective. Furthermore $(d_1, d_2) \to (2d_1 - d_2, d_1 - d_2)$ is an invertible linear map, hence the elements of S are all distinct. Note that for any distinct $a/c, b/c \in \mathbb{Q}$ we have $|a/c - b/c| \ge 1/c$. Hence the spacing of a set in $\mathbb{Q}^2$ is bounded below in terms of the denominators of the coordinates of its elements.

Since $n \in \mathcal{P}(B)$ there exists $z \in \mathcal{O}$ such that $nz = z_n \in \mathbb{Z}$ and $z_n < B$. We can write

$$\frac{d}{n} = \frac{dz}{nz} = \frac{a}{z_n} + \frac{b}{z_n}\zeta$$

for some $a, b \in \mathbb{Z}$. Since $z_n \le B$ it follows that S is $1/B$-spaced as required.

Thus by Lemma 6.7 and (6-2) we get

$$\sum_n \Big|\sum_m a_m \psi_n(m)\Big|^2 \ll \sum_{r\in R}\Big|\sum_{s_1,s_2} a_{s_1+\zeta s_2}\exp(2\pi i(s_1,s_2)\cdot\alpha_r)\Big|^2 \ll (B^2+A)A. \qquad \square$$

7. Determining the main term

We start with the expression for $\sum_{K,D_K<X^{p-1}}|\mathrm{im}(\varphi_K)|^k$ which we derived in Theorem 5.2,

$$S_k(X)=\frac{1}{(p-1)\cdot p^k}\sum_{(D_u)\in\mathcal{P}(X)}\ \sum_{(\chi_{l'})\in\mathcal{C}(\prod D_u)}\frac{\mu^2(\prod D_u)}{p^{k\omega(\prod D_u)}}\prod_v\prod_{l\mid D_v}\chi_l\Big(\prod_u D_u^{\Phi_k(u,v)}\Big) \tag{7-1}$$

and recall the notation:

- $\mathcal{P}(X)$ denotes the set of p^{2k}-tuples of coprime positive integers (D_u) indexed by $u=(u_1,\dots,u_k)\in\mathbb{F}_p^{2k}$ whose prime factors are congruent to 1 mod p or equal to p and $p^{\mathrm{ord}_p(D)}D<X$ (with $D=\prod_{u\in\mathbb{F}_p^{2k}}D_u$).
- $\mathcal{C}(D)$ denotes the set of tuples of nontrivial order p characters $(\chi_{l'})_{l'\mid D}$.
- $\Phi_k(u,v)=\sum_{i=1}^k\Phi(u_i,v_i)$ with Φ defined in (4-2).

For the remainder of the paper we will use the convention that the implied constants in any big-O notation which appears are allowed to depend on p and k, but not X.

Fix $k\in\mathbb{Z}_{\geq 1}$ and let $\Delta=1+\log^{-(p-1)\cdot p^k}X$. Define $\boldsymbol{A}$ to be a p^{2k}-tuple of variables $(A_u)_{u\in\mathbb{F}_p^{2k}}$ with each $A_u=\Delta^j$ for some $j\geq 0$. We can partition $S_k(X)$ according to the various $\boldsymbol{A}$.

Let $\mathcal{P}(X,\boldsymbol{A})\subset\mathcal{P}(X)$ be the subset of tuples (D_u) satisfying $A_u\leq D_u<\Delta A_u$ for all $u\in\mathbb{F}_p^{2k}$. Let $S_k(X,\boldsymbol{A})$ be the above sum (7-1) but now restricted in the first summation to tuples $(D_u)\in\mathcal{P}(X,\boldsymbol{A})$. Thus we have

$$S_k(X)=\sum_{\boldsymbol{A}}S_k(X,\boldsymbol{A})$$

summing over all $\boldsymbol{A}$ with $\prod_{u\in\mathbb{F}_p^{2k}}A_u<X$.

Note that since $\Delta=1+\log^{-(p-1)\cdot p^k}X$ there are $O((\log X)^{p^{2k}(1+(p-1)\cdot p^k)})$ possible $\boldsymbol{A}$ with $S_k(X,\boldsymbol{A})$ not empty. This is since there are $O((\log X)^{(1+(p-1)\cdot p^k)})$ choices for each $1<A_u\leq X$.

We now consider certain families of tuples $\boldsymbol{A}$ for which $S_k(X,\boldsymbol{A})$ makes a negligible contribution to the sum. These will be the same as the four families from Section 5.4 in [Fouvry and Klüners 2007]. In the case of the first the argument is identical but for completeness we reproduce it here. The proofs in the remaining three cases must be generalized.

First we reduce the sum $S_k(X)$ to terms where all of the D_u satisfy $\omega(D_u)\leq\Omega$, where we define $\Omega=e(p-1)p^{2k}(\log\log X+B_0)$ with B_0 the constant given by Lemma 6.1.

Let $\mathcal{P}(X,\boldsymbol{A},\Omega)\subset\mathcal{P}(X,\boldsymbol{A})$ be the subset of tuples (D_u) additionally satisfying $\omega(D_u)\leq\Omega$ for all $u\in\mathbb{F}_p^{2k}$. Let $S_k(X,\boldsymbol{A},\Omega)$ be the above sum (7-1) but now restricted in the first summation to tuples $(D_u)\in\mathcal{P}(X,\boldsymbol{A},\Omega)$.

Lemma 7.1. *With the above notation, for all tuples* $\boldsymbol{A}$

$$\sum_{\boldsymbol{A}} S_k(X, \boldsymbol{A}) = \sum_{\boldsymbol{A}} S_k(X, \boldsymbol{A}, \Omega) + O\left(\frac{X}{\log X}\right).$$

Proof. Let S_0 be the sum of the terms in (7-1) where not all of the D_u satisfy $\omega(D_u) \le \Omega$. We will bound S_0. Let $n = \prod_{u \in \mathbb{F}_p^{2k}} D_u$. We can trivially bound (7-1) by setting all $\chi_l = 1$. For any positive square-free $n \in \mathbb{Z}$ we have $\left|\left\{(D_u) \in \mathcal{P}(X) \mid \prod_{u \in \mathbb{F}_p^{2k}} D_u = n\right\}\right| = p^{2k\omega(n)}$ (this is just the number of ways of writing n as a product of p^{2k} positive integers) and $|\mathcal{C}(n)| = (p-1)^{\omega(n)}$. Thus applying the trivial bound to S_0 gives

$$S_0 \ll \sum_{n < X, \omega(n) > \Omega} \mu^2(n)(p^k(p-1))^{\omega(n)}.$$

Then splitting the sum up by the number of prime factors and applying Lemma 6.1 we get the bound

$$\begin{aligned}\sum_{n<X,\omega(n)>\Omega} \mu^2(n)(p^k(p-1))^{\omega(n)} &\ll \sum_{l \ge \Omega} \frac{X}{\log X}(p^k(p-1))^l \frac{(\log\log X + B_0)^l}{l!} \\ &\ll \frac{X}{\log X} \sum_{l \ge \Omega} \left(\frac{p^k(p-1)(\log\log X + B_0)}{l/e}\right)^l \\ &\ll \frac{X}{\log X} \sum_{l \ge \Omega} \left(\frac{1}{p^k}\right)^l \\ &\ll \frac{X}{\log X}\end{aligned}$$

where in the second-last inequality we are using $l/e \ge (p-1)p^{2k}(\log\log X + B_0)$ by definition of Ω. □

Thus we can assume in the remainder that all variables D_u satisfy $\omega(D_u) \le \Omega$ (we will only need this fact to bound family 4 in Section 7D).

7A. ***The first family.*** Note that it is possible there exists an $\boldsymbol{A}$ for which $S_k(X, \boldsymbol{A}, \Omega)$ is not empty, but

$$\prod_{u \in \mathbb{F}_p^{2k}} \Delta A_u > X. \tag{7-2}$$

Thus for any tuple $(D_u) \in \mathcal{P}(X, \boldsymbol{A}, \Omega)$ the condition $\prod_u D_u < X$ imposes dependencies between the D_u. We wish to remove this dependency to allow application of subsequent analytic results in which we will sum over each D_u independently.

Let $\mathcal{F}_1$ denote the set of $\boldsymbol{A}$ such that (7-2) is satisfied.

Lemma 7.2. *With the above notation*

$$\sum_{\boldsymbol{A} \in \mathcal{F}_1} S_k(X, \boldsymbol{A}, \Omega) \ll X/\log X.$$

Proof. Applying the trivial bound as in Lemma 7.1 and applying Lemma 6.2 we have

$$\sum_{\boldsymbol{A}\in\mathcal{F}_1} S_k(X, \boldsymbol{A}, \Omega) \ll \sum_{\Delta^{-p^{2k}} X\le D\le X} (p^k(p-1))^{\omega(D)} \ll (1-\Delta^{-p^{2k}})X(\log X)^{(p-1)\cdot p^k-1}.$$

Using that $(1+x)^\alpha = 1+\alpha x + O(x^2)$ for $x\to 0$, setting $\alpha = -p^{2k}$ and $x = \log^{-(p-1)\cdot p^k} X$ we get

$$\Delta^{-p^{2k}} = (1+\log^{-(p-1)\cdot p^k} X)^{-p^{2k}} = 1 - p^{2k}\log^{-(p-1)\cdot p^k} X + O(\log^{-2((p-1)\cdot p^k)} X).$$

This gives the bound

$$\sum_{\boldsymbol{A}\in\mathcal{F}_1} S_k(X, \boldsymbol{A}, \Omega) \ll (p^{2k}\log^{-(p-1)\cdot p^k} X + O(\log^{-2((p-1)\cdot p^k)} X))X(\log X)^{(p-1)\cdot p^k-1} \ll X/\log X. \qquad \square$$

Thus if $\boldsymbol{A}\notin\mathcal{F}_1$ then any $(D_u)\in\mathcal{P}(X, \boldsymbol{A}, \Omega)$ automatically satisfies $\prod_{u\in\mathbb{F}_p^{2k}} D_u < X$ so this condition can be dropped from the definition of $\mathcal{P}(X, \boldsymbol{A}, \Omega)$ for $\boldsymbol{A}\notin\mathcal{F}_1$.

7B. ***The second family.*** We now bound the terms in which the range of summation is too short for too many variables D_u. Let $X^{\ddagger} = \exp(\log^\eta X)$ for some small $\eta > 0$ which we will specify later.

Let $\mathcal{F}_2$ be the set of $\boldsymbol{A}$ which satisfy

$$\text{at most } p^{k-1} \text{ variables satisfy } A_u > X^{\ddagger}. \tag{7-3}$$

Lemma 7.3. *In the above notation*

$$\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X, \boldsymbol{A}, \Omega) \ll X(\log X)^{\eta(p-1)\cdot p^k-1/p}.$$

Proof. Let r be the number of variables greater than $X^{\ddagger}$. We factor the sum $\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X, \boldsymbol{A}, \Omega)$ into two parts corresponding to terms with all variables $D_u\le X^{\ddagger}$ and terms with all $D_u > X^{\ddagger}$ and then apply the trivial bound as in the proof of Lemma 7.1. This results in

$$\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X, \boldsymbol{A}, \Omega) \ll \sum_{r=0}^{p^{k-1}} \sum_{m<(X^{\ddagger})^{p^{2k}-r}} \mu^2(m)(p^{2k}-r)^{\omega(m)}\left(\frac{p-1}{p^k}\right)^{\omega(m)} \sum_{n<X/m} \mu^2(n)\left(\frac{(p-1)r}{p^k}\right)^{\omega(n)}.$$

Then applying Lemma 6.2 to the second term above we get

$$\begin{aligned}\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X, \boldsymbol{A}, \Omega) &\ll \sum_{r=0}^{p^{k-1}} \sum_{m<(X^{\ddagger})^{p^{2k}-r}} \frac{((p-1)\cdot p^k)^{\omega(m)}}{m} X(\log X)^{(p-1)r/p^k-1} \\ &\ll X\left(\sum_{r=0}^{p^{k-1}} (\log X)^{(p-1)r/p^k-1}\right)\left(\sum_{m<(X^{\ddagger})^{p^{2k}-r}} \frac{((p-1)\cdot p^k)^{\omega(m)}}{m}\right).\end{aligned}$$

We trivially bound

$$\sum_{r=0}^{p^{k-1}} (\log X)^{(p-1)r/p^k - 1} \ll (\log X)^{(p-1)p^{k-1}/p^k - 1} = (\log X)^{-1/p}. \tag{7-4}$$

We will apply Mertens' formula $\prod_{q<x}\bigl(1-\frac{1}{q}\bigr)^{-1} \ll \log x$ (where the product is over q prime), to the second term above. Recall $X^{\ddagger} = \exp(\log^{\eta} X)$. We have

$$\begin{aligned}\sum_{m<(X^{\ddagger})^{p^{2k-r}}} \frac{((p-1)\cdot p^k)^{\omega(m)}}{m} &\ll \prod_{q<(X^{\ddagger})^{p^{2k-r}}} \left(1+\frac{(p-1)\cdot p^k}{q}+\cdots\right) \\ &\ll \left[\prod_{q<(X^{\ddagger})^{p^{2k-r}}} \left(1-\frac{1}{q}\right)^{-1}\right]^{(p-1)\cdot p^k} \\ &\ll (\log X^{\ddagger})^{p^k(p-1)} = (\log X)^{\eta p^k (p-1)}. \end{aligned} \tag{7-5}$$

Putting together (7-4) and (7-5) we get

$$\sum_{\boldsymbol{A}\in\mathcal{F}_2} S_k(X, \boldsymbol{A}, \Omega) \ll X(\log X)^{\eta(p-1)\cdot p^k - 1/p}. \qquad \square$$

Clearly for any $\epsilon > 0$, for small enough η we get $X(\log X)^{\eta(p-1)\cdot p^k - 1/p} \ll X/(\log X)^{1/p-\epsilon}$.

7C. *The third family, the case $p = 3$.* For the third and fourth families we will first let $p = 3$ and bound the error term unconditionally. In Section 7E we will handle the case of general p under the assumption of GRH.

We define some terminology which will be used in the remainder of the paper.

Definition 7.4. We say indices $u, v \in \mathbb{F}_p^{2k}$ are linked if $\Phi_k(u, v) \neq 0$ or $\Phi_k(v, u) \neq 0$. Otherwise we say they are unlinked. We say a set $\mathcal{U} \subset \mathbb{F}_p^{2k}$ is unlinked if u and v are unlinked for all $u, v \in \mathcal{U}$.

Let $X^{\dagger} = \log^{8(1+9^k(1+2\cdot 3^k))} X$.

Let $\mathcal{F}_3$ denote the set of $\boldsymbol{A}$ such that there are two linked indices u and v with

$$A_u, A_v > X^{\dagger}. \tag{7-6}$$

Fix such an $\boldsymbol{A}$ and two linked indices u, v. We split

$$S_k(X, \boldsymbol{A}, \Omega) = \sum_{i=1}^{4} S_{k,i}(X, \boldsymbol{A}, \Omega) \tag{7-7}$$

into four terms depending on whether $3 \mid D_u$, $3 \mid D_v$, $3 \mid D_w$ for some $w \neq u, v$, or $3 \nmid D_w$ for all w. For simplicity we only present the proof of bounding $S_{k,1}(X, \boldsymbol{A}, \Omega)$, the arguments in the other cases being almost identical.

We now consider two cases: case 1 occurs when both $\Phi_k(u, v)$ and $\Phi_k(v, u)$ are nonzero in (7-1) and case 2 occurs when only one of these is nonzero.

Case 1: Both $\Phi_k(u,v)$ and $\Phi_k(v,u)$ are nonzero.

Define $\mathcal{Q}(X) \subset \mathcal{O}$ to be the set of elements $d \in \mathcal{O}$ congruent to $1 \bmod 3\mathcal{O}$ which are products of split primes, $\mu^2(N(d)) = 1$ and $N(d) \le X$.

Note $\mathcal{Q}(X)$ is closed under conjugation in $\mathbb{Q}(\zeta_3)$.

Lemma 7.5. *If $\Phi_k(u,v)$ and $\Phi_k(v,u)$ are both nonzero then*

$$S_{k,1}(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)_{w \neq u,v} \in \mathcal{P}(X,\boldsymbol{A},\Omega)} \left| \sum_{\substack{d_u \in \mathcal{Q}(\Delta A_u/3) \\ d_v \in \mathcal{Q}(\Delta A_v)}} \mu^2(D_{uv}) a(d_u) a(d_v) \left(\frac{d_u}{d_v}\right)_3 \right| \tag{7-8}$$

with $|a(d_u)|, |a(d_u)| \le 1$, and we denote $D_{uv} = N(d_u)N(d_v)$.

Proof. In the following equation to simplify notation we write $D = \prod_{w \in \mathbb{F}_3^{2k}} D_w$ and $D' = D_u D_v$. Let $\boldsymbol{A}_1$ denote the $(3^{2k}-2)$-tuple $(A_w)_{w \neq u,v}$ and $\boldsymbol{A}_2$ denote the pair (A_u, A_v).

Then from (7-1) we get by splitting up the summations and bounding

$$\begin{aligned} S_{k,1}(X,\boldsymbol{A},\Omega) &= \frac{1}{2 \cdot 3^k} \sum_{\substack{(D_w) \in \mathcal{P}(X,\boldsymbol{A},\Omega) \\ 3 \mid D_u}} \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \frac{\mu^2(D)}{3^{k\omega(D)}} \prod_{y \in \mathbb{F}_3^{2k}} \prod_{l \mid D_y} \chi_l \Bigl(\prod_{z \in \mathbb{F}_3^{2k}} D_z^{\Phi_k(z,y)} \Bigr) \\ &\le \frac{1}{2 \cdot 3^k} \sum_{(D_w) \in \mathcal{P}(X,\boldsymbol{A}_1,\Omega)} \sum_{(\chi_{l'}) \in \mathcal{C}(D/D')} \frac{1}{3^{k\omega(D/D')}} \\ &\qquad \times \left| \sum_{\substack{(D_u,D_v) \in \mathcal{P}(X,\boldsymbol{A}_2,\Omega) \\ 3 \mid D_u}} \sum_{(\chi_{l'}) \in \mathcal{C}(D')} \frac{\mu^2(D)}{3^{k\omega(D')}} \prod_{\substack{y=u,v \\ l \mid D_y}} \chi_l \Bigl(\prod_{z \in \mathbb{F}_3^{2k}} D_z^{\Phi_k(z,y)} \Bigr) \right|. \end{aligned}$$

Suppose $z \in \mathbb{Z}$ is square free and a product of primes congruent to $1 \bmod 3\mathbb{Z}$. This implies z factors into split primes in $\mathbb{Q}(\zeta_3)$, hence there are exactly $2^{\omega(z)}$ ideals $I \subset \mathcal{O}$ such that $N(I) = z$. Furthermore for any ideal $I \subset \mathcal{O}$ such that $N(I) = z$ there exists a unique element with $(d) = I$ such that $d \in \mathcal{Q}(z)$.

Since $p = 3$ there are 2 nontrivial cubic characters χ_l corresponding to any prime l. They are $\left(\frac{n}{\pi}\right)_3$ and $\left(\frac{n}{\bar{\pi}}\right)_3$ where $\pi, \bar{\pi} \equiv 1 \bmod 3\mathcal{O}$ and $N(\pi) = N(\bar{\pi}) = l$. Thus there is a bijection between $\mathcal{C}(z)$ and ideals $I \subset \mathcal{O}$ such that $N(I) = z$, defined by

$$(\chi_{l'})_{l' \mid z} \mapsto \Bigl(\prod_{\substack{l' \mid z \\ \chi_{l'} = (\frac{\cdot}{\pi})}} \pi \Bigr). \tag{7-9}$$

This map is injective by unique factorization of ideals and hence surjective since the size of both sets is equal. In particular this implies

$$\sum_{(\chi_{l'}) \in \mathcal{C}(D/D')} \frac{1}{3^{k\omega(D/D')}} \le \frac{2^{\omega(D/D')}}{3^{k\omega(D/D')}} < 1.$$

So far we have shown

$$S_{k,1}(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)\in\mathcal{P}(X,\boldsymbol{A}_1,\Omega)} \Bigg| \sum_{\substack{(D_u,D_v)\in\mathcal{P}(X,\boldsymbol{A}_2,\Omega)\\ 3\,|\,D_u}} \sum_{(\chi_{l'})\in\mathcal{C}(D')} \frac{\mu^2(D)}{3^{k\omega(D')}} \prod_{\substack{y=u,v\\ l\,|\,D_y}} \chi_l\Bigg(\prod_{z\in\mathbb{F}_3^{2k}} D_z^{\Phi_k(z,y)}\Bigg)\Bigg|.$$

It also follows from the above bijection (7-9) that in the sum over D_u, D_v we can replace the χ_l with cubic residue symbols

$$\sum_{\substack{(D_u,D_v)\in\mathcal{P}(X,\boldsymbol{A}_2,\Omega)\\ 3\,|\,D_u}} \sum_{(\chi_{l'})\in\mathcal{C}(D')} \frac{\mu^2(D)}{3^{k\omega(D')}} \prod_{\substack{y=u,v\\ l\,|\,D_y}} \chi_l\Bigg(2\prod_{z\in\mathbb{F}_3^{2k}} D_z^{\Phi_k(z,y)}\Bigg)$$

$$\ll \sum_{\substack{(D_u,D_v)\in\mathcal{P}(X,\boldsymbol{A}_2,\Omega)\\ 3\,|\,D_u}} \sum_{\substack{d_u\in\mathcal{Q}(\Delta A_u/3)\\ N(d_u)=D_u/3}} \sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ N(d_v)=D_v}} \frac{\mu^2(D)}{3^{k\omega(D')}} \prod_{y=u,v} \left(\frac{\prod_{z\in\mathbb{F}_3^{2k}} D_z^{\Phi_k(z,y)}}{d_y}\right)_3 \chi_3\Bigg(\prod_{z\in\mathbb{F}_3^{2k}} D_z^{\Phi_k(z,u)}\Bigg) \tag{7-10}$$

For any d_u in the above sum let

$$b(d_u) = \frac{\mu^2\big(\prod_{w\neq v} D_w\big)}{3^{k\omega(N(d_u))}} \prod_{y\neq u,v} \prod_{l\,|\,D_y} \chi_l((3N(d_u))^{\Phi_k(u,y)}) \left(\frac{\prod_{y\neq u,v} D_v^{\Phi_k(y,u)}}{d_u}\right)_3$$

and similarly for $b(d_v)$ $\big($which will additionally contain the factor of $\chi_3\big(\prod_{z\in\mathbb{F}_3^{2k}} D_z^{\Phi_k(z,u)}\big)\big)$. Note also that $\mu^2(D)=\mu^2\big(\prod_{w\neq v} D_w\big)\mu^2\big(\prod_{w\neq u} D_w\big)\mu^2(D_uD_v)$. Changing notation to $D_u=N(d_u)$ and $D_v=N(d_v)$ and plugging in $b(d_u)$ and $b(d_v)$ we can rewrite (7-10) as

$$\sum_{\substack{d_u\in\mathcal{Q}(\Delta A_u/3)\\ A_u/3\le N(d_u)<\Delta A_u/3}} \sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ A_v\le N(d_v)<\Delta A_v}} \mu^2(D_uD_v)b(d_u)b(d_v)\left(\frac{D_u}{d_v}\right)_3^{\Phi_k(u,v)}\left(\frac{D_v}{d_u}\right)_3^{\Phi_k(v,u)}.$$

Note Φ_k is either 1 or 2, and squaring a cubic character is the same as conjugating it. Since $\mathcal{Q}(X)$ is closed under conjugation, removing Φ_k from the exponent permutes the coefficients. As a result of this procedure rename $b(d_u)$ to $a(d_u)$ and $b(d_v)$ to $a(d_v)$ if necessary. Letting $a(d_u)=0$ for $N(d_u)<A_u/3$ we can extend the range of summation to $N(d_u)\le\Delta A_u/3$, and similarly to $N(d_v)\le\Delta A_v$.

For any $D_u=N(d_u)=d_u\bar{d}_u$ and $D_v=N(d_v)=d_v\bar{d}_v$ not divisible by 3 by the properties (2-2) and the law of cubic reciprocity (2-3) we have

$$\left(\frac{D_u}{d_v}\right)_3\left(\frac{D_v}{d_u}\right)_3 = \left(\frac{d_u}{d_v}\right)_3\left(\frac{\bar{d}_u}{d_v}\right)_3\left(\frac{d_v}{d_u}\right)_3\left(\frac{\bar{d}_v}{d_u}\right)_3 = \left(\frac{d_u}{d_v}\right)_3^2\left(\frac{\bar{d}_u}{d_v}\right)_3\left(\frac{\bar{d}_u}{d_v}\right)_3^2 = \left(\frac{\bar{d}_u}{\bar{d}_v}\right)_3. \tag{7-11}$$

This proves the lemma. □

Proposition 7.6. *For all $\boldsymbol{A}\in\mathcal{F}_3$*

$$S_{k,1}(X, \boldsymbol{A}, \Omega) \ll X/\log^{1+9^k(1+2\cdot 3^k)} X.$$

Proof. We will apply the standard strategy of bounding bilinear sums using Cauchy–Schwarz followed by a large sieve type bound. By Cauchy–Schwarz applied to the summand on the right-hand side in Lemma 7.5 we have, for any fixed tuple $(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)$,

$$\Biggl|\sum_{\substack{d_u\in\mathcal{Q}(\Delta A_u/3)\\ d_v\in\mathcal{Q}(\Delta A_v)}}\mu^2(D_{uv})a(d_u)a(d_v)\biggl(\frac{d_u}{d_v}\biggr)_3\Biggr| \ll A_v^{1/2}\Biggl(\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\Biggl|\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\mu^2(D_{uv})a(d_u)\biggl(\frac{d_u}{d_v}\biggr)_3\Biggr|^2\Biggr)^{1/2} \tag{7-12}$$

where $D_{uv}=N(d_u)N(d_v)$.

Note that $d_u\mapsto\mu^2(D_{uv})\bigl(\frac{d_u}{d_v}\bigr)_3$ is a primitive cubic Hecke character of modulus (d_v). Note also that $\mathcal{Q}(\Delta A_v)$ satisfies the conditions of the set $\mathcal{P}(\Delta A_v)$ in Proposition 6.8 since by definition for any $d\in\mathcal{Q}(\Delta A_v)$, $N(d)=d\bar{d}<\Delta A_v$. Thus by Proposition 6.8

$$\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\Biggl|\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\mu^2(D_{uv})a(d_u)\biggl(\frac{d_u}{d_v}\biggr)_3\Biggr|^2\ll(A_v^2+A_u)A_u.$$

Plugging this into the bound (7-12) we get

$$\Biggl|\sum_{d_v}\sum_{d_u}\mu^2(D_{uv})a(d_u)a(d_v)\biggl(\frac{d_u}{d_v}\biggr)_3\Biggr|\ll A_v^{1/2}((A_v^2+A_u)A_u)^{1/2}=A_vA_u\biggl(\frac{A_v}{A_u}+\frac{1}{A_v}\biggr)^{1/2}. \tag{7-13}$$

By symmetry (recall by cubic reciprocity $\bigl(\frac{d_u}{d_v}\bigr)_3=\bigl(\frac{d_v}{d_u}\bigr)_3$) we can also bound this by $A_vA_u\bigl(\bigl(\frac{A_u}{A_v}+\frac{1}{A_u}\bigr)\bigr)^{1/2}$.

Now by symmetry we can assume without loss of generality that $A_v\le A_u$.

First suppose $A_v^2<A_u$. Recall $\boldsymbol{A}\in\mathcal{F}_3$ implies $A_v,A_u>X^\dagger=\log^{8(1+9^k(1+2\cdot3^k))}X$. Plugging the bound (7-13) into Lemma 7.5 we get

$$S_{k,1}(X,\boldsymbol{A},\Omega)\ll\sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)}A_vA_u\biggl(\frac{A_v}{A_u}+\frac{1}{A_v}\biggr)^{1/2}\ll X\biggl(\frac{1}{A_u^{1/2}}+\frac{1}{A_v}\biggr)^{1/2}\ll X/\log^{1+9^k(1+2\cdot3^k)}X.$$

Now suppose $A_u\le A_v^2$. Then by Proposition 6.3 we directly get the bound, for any $\epsilon>0$,

$$\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\Biggl|\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\mu^2(D_{uv})a(d_u)\biggl(\frac{d_u}{d_v}\biggr)_3\Biggr|^2\ll(A_u+A_v+(A_uA_v)^{2/3})(A_uA_v)^\epsilon A_u. \tag{7-14}$$

Plugging (7-14) into Lemma 7.5 we get

$$\begin{aligned}S_{k,1}(X,\boldsymbol{A},\Omega)&\ll\sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)}A_v^{1/2}((A_u+A_v+(A_uA_v)^{2/3})(A_uA_v)^\epsilon A_u)^{1/2}\\&\ll X\biggl(\biggl(\frac{1}{A_v}+\frac{1}{A_u}+\frac{1}{(A_uA_v)^{1/3}}\biggr)(A_uA_v)^\epsilon\biggr)^{1/2}\\&\ll X\biggl(\frac{1}{X^{\dagger1/2}}\biggr)^{1/2}\\&\ll X/\log^{1+9^k(1+2\cdot3^k)}X.\end{aligned}$$

□

This proves the desired bound in Case 1.

Case 2: Only one of $\Phi_k(u,v)$ and $\Phi_k(v,u)$ is nonzero. Without loss of generality assume $\Phi_k(u,v)$ is nonzero.

For any $X>0$ define $\mathcal{R}(X)$ to be the set of positive $d\in\mathbb{Z}$ which are a product of primes congruent to 1 mod $3\mathbb{Z}$ and $d<X$.

Lemma 7.7. *For any linked indices u and v with $\Phi_k(u,v)\neq 0$ and $\Phi_k(v,u)=0$*

$$S_{k,1}(X,\boldsymbol{A},\Omega)\ll\sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)}\left|\sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ D_u\in\mathcal{R}(\Delta A_u/3)}}\mu^2(D_{uv})a(D_u)a(d_v)\left(\frac{D_u}{d_v}\right)_3\right|,\tag{7-15}$$

with $|a(d_u)|,|a(d_u)|\leq 1$, $D_{uv}=D_uN(d_v)$.

Proof. This is a simpler version of the proof of Lemma 7.5. Since $\Phi_k(v,u)=0$ the symbol $\left(\frac{D_v}{d_u}\right)_3$ does not appear and hence we do not apply cubic reciprocity unlike in that proof. Hence we are left with $\left(\frac{D_u}{d_v}\right)_3$ which is what appears in the statement above. □

Proposition 7.8. *For all $\boldsymbol{A}\in\mathcal{F}_3$*

$$S_{k,1}(X,\boldsymbol{A},\Omega)\ll X/\log^{1+9^k(1+2\cdot 3^k)}X.$$

Proof. The above expression is no longer symmetric in u and v hence we must consider several subcases.

By Cauchy–Schwarz applied to the summand on the right-hand side in Lemma 7.7 we have, for any fixed tuple $(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)$,

$$\left|\sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ D_u\in\mathcal{R}(\Delta A_u/3)}}\mu^2(D_{uv})a(D_u)a(d_v)\left(\frac{D_u}{d_v}\right)_3\right|$$

$$\ll A_u^{1/2}\left(\sum_{D_u\in\mathcal{R}(\Delta A_u/3)}\left|\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\mu^2(D_{uv})a(d_v)\left(\frac{D_u}{d_v}\right)_3\right|^2\right)^{1/2}\tag{7-16}$$

where $D_{uv}=D_uN(d_v)$.

For fixed $D_u\in\mathbb{Z}$ the map $d_v\mapsto\mu^2(D_{uv})\left(\frac{D_u}{d_v}\right)_3$ is a primitive cubic Hecke character on $\mathcal{O}$ with modulus $(9D_u)$ (see [Baier and Young 2010, Section 2.1]). Note also that $\mathcal{R}(\Delta A_u/3)$ satisfies the conditions of the set $\mathcal{P}(\Delta A_u/3)$ in Proposition 6.8, since by definition for any $d\in\mathcal{R}(\Delta A_u/3)$, $d\in\mathbb{Z}$ and $d<\Delta A_u/3$. Thus by Proposition 6.8

$$\sum_{D_u\in\mathcal{R}(\Delta A_u/3)}\left|\sum_{d_v\in\mathcal{Q}(\Delta A_v)}\mu^2(D_{uv})a(d_v)\left(\frac{D_u}{d_v}\right)_3\right|^2\ll(A_u^2+A_v)A_v.$$

Plugging this into (7-16) we get

$$\left|\sum_{\substack{d_v\in\mathcal{Q}(\Delta A_v)\\ D_u\in\mathcal{R}(\Delta A_u/3)}}\mu^2(D_{uv})a(D_u)a(d_v)\left(\frac{D_u}{d_v}\right)_3\right|\ll A_u^{1/2}((A_u^2+A_v)A_v)^{1/2}=A_uA_v\left(\frac{A_u}{A_v}+\frac{1}{A_u}\right)^{1/2}.\tag{7-17}$$

First suppose $A_u^2 < A_v$. Plugging (7-17) into Lemma 7.7 we get

$$\begin{aligned} S_{k,1}(X, \boldsymbol{A}, \Omega) &\ll \sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)} A_u A_v \left(\frac{A_u}{A_v} + \frac{1}{A_u}\right)^{1/2} \\ &\ll X\left(\frac{1}{A_v^{1/2}} + \frac{1}{A_u}\right)^{1/2} \\ &\ll X/\log^{1+9^k(1+2\cdot 3^k)} X. \end{aligned}$$

Next suppose $A_v^2 < A_u$. We again apply Cauchy–Schwarz as in (7-16) with summations reversed. Note $\chi(D_u) = \left(\frac{D_u}{d_v}\right)_3$ is a primitive Dirichlet character of modulus $N(d_v)$ for all $d_v \in \mathcal{Q}(\Delta A_v)$. Then by Proposition 6.5 we have

$$\sum_{d_v\in\mathcal{Q}(\Delta A_v)} \left| \sum_{D_u\in\mathcal{R}(\Delta A_u/3)} a(D_u)\mu^2(D_{uv})\left(\frac{D_u}{d_v}\right)_3 \right|^2 \ll (A_v^2 + A_u)A_u$$

and hence

$$S_{k,1}(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)} A_v^{1/2}((A_v^2 + A_u)A_u)^{1/2} \ll X\left(\frac{1}{A_u^{1/2}} + \frac{1}{A_v}\right)^{1/2} \ll X/\log^{1+9^k(1+2\cdot 3^k)} X.$$

In the case when the variables A_u and A_v are close together, specifically $A_u < A_v < A_u^2$ or $A_v < A_u < A_v^2$ we again apply Cauchy–Schwarz, followed by Proposition 6.4. We obtain

$$\sum_{d_v\in\mathcal{Q}(\Delta A_v)} \left| \sum_{D_u\in\mathcal{R}(\Delta A_u/3)} a(D_u)\mu^2(D_{uv})\left(\frac{D_u}{d_v}\right)_3 \right|^2 \ll (A_v^{11/9} + A_v^{2/3}A_u)(A_u A_v)^{\epsilon} A_u.$$

Then

$$S_k(X, \boldsymbol{A}, \Omega) \ll \sum_{(D_w)_{w\neq u,v}\in\mathcal{P}(X,\boldsymbol{A},\Omega)} A_v^{1/2}((A_v^{11/9} + A_v^{2/3}A_u)(A_u A_v)^{\epsilon} A_u)^{1/2} \ll X\left(\left(\frac{A_v^{2/9}}{A_u} + \frac{1}{A_v^{1/3}}\right)(A_u A_v)^{\epsilon}\right)^{1/2}.$$

Then using that $A_u < A_v < A_u^2$ we get

$$S_k(X, \boldsymbol{A}, \Omega) \ll X\left(\left(\frac{1}{A_u^{5/9}} + \frac{1}{A_u^{1/3}}\right)(A_u A_v)^{\epsilon}\right)^{1/2} \ll X\left(\frac{1}{A_u^{1/4}}\right)^{1/2} \ll X/\log^{1+9^k(1+2\cdot 3^k)} X.$$

The case $A_v < A_u < A_v^2$ is similar. □

This proves the desired bound in Case 2.

Finally summing over all $\boldsymbol{A} \in \mathcal{F}_3$ and recalling that there are $O((\log X)^{9^k(1+2\cdot 3^k)})$ possible $\boldsymbol{A}$ with $S_k(X, \boldsymbol{A}, \Omega)$ not empty, we have proven

$$\sum_{\boldsymbol{A}\in\mathcal{F}_3} S_k(X, \boldsymbol{A}, \Omega) \ll X/\log X.$$

7D. ***The fourth family, the case $p=3$.*** Recall we previously defined$X^{\ddagger}=\exp(\log^{\eta} X)$ and $X^{\dagger}=\log^{8(1+9^k(1+2\cdot 3^k))} X$.

Now consider the fourth family $\mathcal{F}_4$ which consists of those $\boldsymbol{A}$ such that $\boldsymbol{A}\notin\mathcal{F}_3$ and there are two linked indices u, v and

$$A_u > X^{\ddagger}, \quad 2/\Delta \le A_v < X^{\dagger}. \tag{7-18}$$

Note that given $A_u > X^{\ddagger}$ the condition $A_v < X^{\dagger}$ is forced by the assumption that $\boldsymbol{A}\notin\mathcal{F}_3$. For fixed u we in fact consider the collection of all indices $v\in\mathbb{F}_3^{2k}$ which satisfy the above condition. Of the set of $v\in\mathbb{F}_3^{2k}$ linked with u satisfying (7-18), let S_1 be the subset of v such that $\Phi_k(u,v)\neq 0$ and let S_2 be the subset of v such that $\Phi_k(v,u)\neq 0$. We assume $S_1\cup S_2$ is not empty.

As in Section 7C we split $S_k(X,\boldsymbol{A},\Omega)=\sum_{i=1}^4 S_{k,i}(X,\boldsymbol{A},\Omega)$ into four terms depending on whether $3\mid D_u$, $3\mid D_v$ for some $v\in S$, $3\mid D_w$ for some $w\neq u$ and $w\notin S$, or $3\nmid D_w$ for all w. In the following we bound $S_{k,1}(X,\boldsymbol{A},\Omega)$, the arguments in the other cases being almost identical.

Lemma 7.9. *For each prime $l\in\mathbb{Z}$ congruent to* 1 mod 3 *fix a nontrivial order* 3 *character of modulus l, denoted χ_l.*

For $\boldsymbol{A}$, u and S_1, S_2 as defined above we have

$$S_{k,1}(X,\boldsymbol{A},\Omega)\ll \sum_{(D_w)_{w\neq u}\in\mathcal{P}(X,\boldsymbol{A},\Omega)}\left|\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\frac{\mu^2(D)}{3^{\omega(N(d_u))}}\psi(d_u)\right|$$

where we denote $D=\prod_{w\in\mathbb{F}_3^{2k}} D_w$, $D'=\prod_{w\neq u} D_w$, and ψ is a cubic Hecke character defined by

$$\psi(d_u)=\left(\frac{\prod_{v\in S_2} D_v}{d_u}\right)_3 \prod_{v\in S_1}\prod_{l\mid D_v}\chi_l(D_u).$$

Proof. The proof is similar to Lemma 7.5. □

We remark that the $v\in S$ by assumption satisfy $A_v<X^{\dagger}$. The modulus of ψ is $f_\psi=9\prod_{v\in S_1\cup S_2} D_v$ and $N(f_\psi)\le (X^{\dagger})^{2\cdot 3^{2k}}$. Also note ψ is a nontrivial character since $D_v\ge A_v\ge 2/\Delta>1$.

Proposition 7.10. *With the above notation we have, for some $t\in\mathbb{R}$*

$$\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\frac{\mu^2(D)}{3^{\omega(d_u)}}\psi(d_u)\ll\frac{(\log X)^t}{\exp(c^4(\log X)^{\eta/4-2\cdot 9^k\epsilon}/3^{2+\epsilon})}.$$

Proof. Partitioning the sum according to the number of prime factors in $\mathcal{O}$ we get

$$\sum_{d_u}\frac{\mu^2(D)}{3^{\omega(d_u)}}\psi(d_u)=\sum_{l=0}^{\Omega}\frac{1}{3^l}\sum_{\pi_1,\dots,\pi_l}\mu^2\biggl(\prod_{v\neq u} D_v N(\pi_1\cdots\pi_l)\biggr)\psi(\pi_1\cdots\pi_l)$$

where (π_i) are prime ideals in $\mathcal{O}$ with $\pi_i \equiv 1 \bmod 3$ and $N(\pi_1 \cdots \pi_l) \le \Delta A_u$. We can relabel the π_i so that $N(\pi_1) \le N(\pi_2) \le \cdots \le N(\pi_l)$ and split the sum on the right-hand side up to get

$$\sum_{l=0}^{\Omega} \frac{1}{3^l} \sum_{\pi_1,\dots,\pi_{l-1}} \psi(\pi_1 \cdots \pi_{l-1}) \sum_{\pi_l} \mu^2\Bigl(\prod_{v\neq u} D_v N(\pi_1 \cdots \pi_l)\Bigr)\psi(\pi_l) \tag{7-19}$$

where $N(\pi_1 \cdots \pi_{l-1}) \le \Delta A_u$ and $A_u^{1/l} \le N(\pi_l) \le \Delta A_u / N(\pi_1 \cdots \pi_{l-1})$.

Note $\omega\bigl(\prod_{v\neq u} D_v N(\pi_1 \cdots \pi_{l-1})\bigr) \le 3^{2k}\Omega$ factoring in $\mathbb{Z}$, hence the number of prime factors in $\mathcal{O}$ is at most $2\cdot 3^{2k}\Omega$. Hence removing μ^2 from (7-19) adds at most an additional $O(\Omega)$ terms of absolute value 1. Furthermore notice the summations in (7-19) are only over primes split in $\mathcal{O}$. The number of inert primes (π) in $\mathcal{O}$ with $\pi^2 = N\pi < A_u$ is $O(A_u^{1/2})$. Then looking at the inner sum of (7-19) we obtain the bound

$$\sum_{\pi_l} \mu^2\Bigl(\prod_{v\neq u} D_v N(\pi_1 \cdots \pi_l)\Bigr)\psi(\pi_l) \ll \sum_{\pi_l,(\pi_l, f_\psi)=1} \psi(\pi_l) + \Omega + A_u^{1/2}$$

where the summation on the right-hand side is now over all prime ideals in $\mathcal{O}$ with $A_u^{1/l} \le N(\pi_l) \le \Delta A_u / N(\pi_1 \cdots \pi_{l-1})$.

Now we apply Proposition 6.6 with $f_\chi = f_\psi$ and $x = \Delta A_u / N(\pi_1 \cdots \pi_{l-1})$ to get, for some constant c,

$$\sum_{\pi_l,(\pi_l, f_\psi)=1} \psi(\pi_l) \ll \frac{N(f_\psi)^\epsilon x (\log x)^2}{\exp(c^4 (\log x)^{1/2} / 3^{2+\epsilon} N(f_\psi)^\epsilon)}.$$

Using that $N(f_\psi)^\epsilon \le (X^\dagger)^{2\cdot 9^k \epsilon}$ which implies $\exp(-1/N(f_\psi)^\epsilon) \ll \exp(-1/(X^\dagger)^{2\cdot 9^k\epsilon})$, we get

$$\sum_{\pi_l,(\pi_l, f_\psi)=1} \psi(\pi_l) \ll \frac{(X^\dagger)^{2\cdot 9^k\epsilon} x (\log x)^2}{\exp(c^4 (\log x)^{1/2} / 3^{2+\epsilon} (X^\dagger)^{2\cdot 9^k \epsilon})}. \tag{7-20}$$

Now $x \ge N(\pi_l) \ge A_u^{1/l}$. We claim $A_u^{1/l} \gg \exp(\log^{\eta/2} X)$. Note $\Omega \ll \log\log X$ by definition. Let $\theta = \log X$. Then

$$-\log l + \eta \log\theta \ge -\log\Omega + \eta\log\theta \gg -\log\log\theta + \eta\log\theta \ge \frac{\eta}{2}\log\theta.$$

Taking exp of this inequality gives $(1/l)\log^\eta X \gg \log^{\eta/2} X$. Thus

$$\log A_u^{1/l} \ge \log(X^\ddagger)^{1/l} = (1/l)\log^\eta X \gg \log^{\eta/2} X.$$

Combining these facts we have $(\log x)^{1/2} \gg \log^{\eta/4} X$ so we can write $c_0(\log x)^{1/2} \ge \log^{\eta/4} X$ for some constant c_0. Noting that $X^\dagger$ is some fixed power of $\log X$ we get the bound

$$\sum_{\pi_l,(\pi_l, f_\psi)=1} \psi(\pi_l) \ll X^{\dagger 2\cdot 9^k\epsilon} \frac{A_u}{N(\pi_1 \cdots \pi_{l-1})} \frac{(\log x)^2}{\exp(c^4 (\log X)^{\eta/4 - 2\cdot 9^k\epsilon} / c_0 3^{2+\epsilon})}. \tag{7-21}$$

We plug (7-21) back into (7-19) to get

$$\sum_{d_u}\frac{\mu^2(D)}{3^{\omega(d_u)}}\psi(d_u)\ll\sum_{l=0}^{\Omega}\frac{1}{3^l}\sum_{\pi_1,\dots,\pi_{l-1}}X^{\dagger 2\cdot 9^k\epsilon}\frac{A_u}{N(\pi_1\cdots\pi_{l-1})}\frac{(\log x)^2}{\exp(c^4(\log X)^{\eta/4-2\cdot 9^k\epsilon}/c_0 3^{2+\epsilon})}$$
$$\ll A_u\frac{X^{\dagger 2\cdot 9^k\epsilon}(\log x)^2}{\exp(c^4(\log X)^{\eta/4-2\cdot 9^k\epsilon}/c_0 3^{2+\epsilon})}\sum_{l=0}^{\Omega}\sum_{\pi_1,\dots,\pi_{l-1}}\frac{1}{N(\pi_1\cdots\pi_{l-1})}.$$

Noting the bounds

$$\sum_{\substack{\pi_1,\dots,\pi_{l-1}\\ N(\pi_1\cdots\pi_{l-1})\le\Delta A_u}}\frac{1}{N(\pi_1\cdots\pi_{l-1})}\ll\log A_u\Omega\ll\log\log X\log x\ll\log X$$

we get, for some $t\in\mathbb{R}$

$$\sum_{d_u}\frac{\mu^2(D)}{3^{\omega(d_u)}}\psi(d_u)\ll A_u\frac{(\log X)^t}{\exp(c^4(\log X)^{\eta/4-2\cdot 9^k\epsilon}/c_0 3^{2+\epsilon})}.$$ □

Combining Lemma 7.9 and Proposition 7.10 we get

$$S_{k,1}(X,\boldsymbol{A})\ll\sum_{(D_w)_{w\ne u}\in\mathcal{P}(X,\boldsymbol{A},\Omega)}\left|\sum_{d_u\in\mathcal{Q}(\Delta A_u/3)}\frac{\mu^2(D)}{3^{\omega(N(d_u))}}\psi(d_u)\right|\ll X\frac{(\log X)^t}{\exp(c^4(\log X)^{\eta/4-2\cdot 9^k\epsilon}/c_0 3^{2+\epsilon})}.$$

Then summing over all $\boldsymbol{A}$ and using that there are $O((\log X)^{9^k(1+2\cdot 3^k)})$ possible $\boldsymbol{A}$ with $S_k(X,\boldsymbol{A},\Omega)$ not empty

$$\sum_{\boldsymbol{A}\in\mathcal{F}_4}S_k(X,\boldsymbol{A})\ll X\frac{(\log X)^{t+9^k(1+2\cdot 3^k)}}{\exp(c^4(\log X)^{\eta/4-2\cdot 9^k\epsilon}/c_0 3^{2+\epsilon})}=o(X).$$

7E. *The third and fourth families for all p.* Assume GRH for Artin L-functions. The missing ingredients required to extend our result to general p unconditionally are analogs of Proposition 6.3 and 6.4. That is, we cannot deal with the case in family 3 when A_u and A_v are close together. We will instead give a proof assuming GRH. The following argument replaces the sections containing families 3 and 4 for $p=3$ above.

Suppose $\boldsymbol{A}\notin\mathcal{F}_1\cup\mathcal{F}_2$. In particular there are at least $p^{k-1}+1$ indices $w\in\mathbb{F}_p^{2k}$ which satisfy $A_w>X^{\ddagger}$. Let A_u be the largest of these. Let S_1 be the set of indices v linked with u such that $\Phi_k(u,v)\ne 0$ and let S_2 be the set of v such that $\Phi_k(v,u)\ne 0$ and suppose $S_1\cup S_2$ is not empty.

Let $\zeta=e^{2\pi i/p}$ and let $\mathcal{O}=\mathbb{Z}[\zeta]$ the ring of integers of $\mathbb{Q}(\zeta)$ which is a degree $p-1$ extension of $\mathbb{Q}$. For each prime $l\in\mathbb{Z}$ congruent to 1 mod p fix a nontrivial order p character of modulus l, denoted χ_l.

For $A,B\in\mathbb{Z}$ with $(A,B)=1$ define

$$\left[\frac{A}{B}\right]_p=\prod_{l\mid B}\frac{(\chi_l+\cdots+\chi_l^{p-1})}{p}(A)$$

which does not depend on the choice of χ_l above.

Recall we defined $\mathcal{R}(X)$ to be the set of positive $d \in \mathbb{Z}$ which are a product of primes congruent to 1 mod $p\mathbb{Z}$ and $d < X$. Then we have, by an argument similar to the proof of Lemma 7.5,

$$S_k(X, \boldsymbol{A}) \ll \sum_{(D_w)_{w\neq u} \in \mathcal{P}(X, \boldsymbol{A}, \Omega)} \left| \sum_{D_u \in \mathcal{R}(\Delta A_u)} \mu^2(D) \left(\frac{D_u}{f_u}\right)_p \left[\frac{C_u}{D_u}\right]_p \right| \tag{7-22}$$

where $D = \prod_{w \in \mathbb{F}_p^{2k}} D_w$, $C_u = \prod_{w \in S_2} D_w$, and any choice of $f_u \in \mathcal{O}$ with $N(f_u) = \prod_{w \in S_1} D_w$. Divisibility by p of the D_w is handled as in the $p = 3$ cases in the previous section. For simplicity we will assume $p \nmid D_u C_u N(f_u)$.

Let $K = \mathbb{Q}(\zeta, \sqrt[p]{C_u})$ and let F_q denote the Frobenius of q in K. Let $K_1 = \mathbb{Q}(\zeta)$. Define a character $\rho : \mathrm{Gal}(K/K_1) \to \mu_p$ by $\sigma(\sqrt[p]{C_u}) = \rho(\sigma)\sqrt[p]{C_u}$ (viewing $\mathrm{Gal}(K/K_1)$ as a subgroup of $\mathrm{Gal}(K/\mathbb{Q})$). Denote by $\rho' = \mathrm{Ind}_{\mathrm{Gal}(K/K_1)}^{\mathrm{Gal}(K/\mathbb{Q})} \rho$ the induction to $\mathrm{Gal}(K/\mathbb{Q})$.

Lemma 7.11. *With the above notation, for any $D_u \in \mathcal{R}(\Delta A_u)$ and any prime $q \equiv 1 \bmod p$*

$$\left[\frac{C_u}{D_u}\right]_p = \frac{\prod_{q \mid D_u} \mathrm{tr}\,\rho'(F_q)}{p^{\omega(D_u)}}.$$

Proof. Note $F_q \in \mathrm{Gal}(K/K_1)$. We claim that for any $q \equiv 1 \bmod p$ we have $\rho(F_q)$ is trivial if and only $\chi_q(C_u)$ is trivial. Indeed $\chi_q : \mathbb{F}_q^\times \to \mu_p$ has kernel equal to the p-th powers in $\mathbb{F}_q^\times$. For any prime $\hat{q}$ in $\mathbb{Q}(\zeta, \sqrt[p]{C_u})$ lying above q we have, using $q = 1 + pn$ for some n, that

$$F_q(\sqrt[p]{C_u}) \bmod \hat{q} = (\sqrt[p]{C_u})^q \bmod \hat{q} = C_u^n \sqrt[p]{C_u} \bmod \hat{q}.$$

Let $\mathbb{F}_{\hat{q}}$ denote the residue field of the prime $\hat{q}$. Note $C_u \in \mathbb{F}_q^\times \subset \mathbb{F}_{\hat{q}}^\times$ and $C_u^n \bmod q$ is trivial exactly when C_u is a p-th power. Thus we have shown $\rho(F_q)$ is trivial if and only $\chi_q(C_u)$ is trivial. In particular $\sum_{i=1}^{p-1} \rho^i(F_q) = \sum_{i=1}^{p-1} \chi_q^i(C_u)$.

By properties of induced representations and since $F_q \in \mathrm{Gal}(K/K_1)$, we have

$$\mathrm{tr}\,\rho'(F_q) = \sum_{g \in \mathrm{G}(K/\mathbb{Q})/\mathrm{G}(K/K_1)} \hat{\mathrm{tr}}(g^{-1} F_q g) = \sum_{i=1}^{p-1} \rho(F_q^i)$$

where $\hat{\mathrm{tr}}(h) = \rho(h)$ if $h \in \mathrm{Gal}(K/K_1)$ and 0 otherwise. The result follows. □

Let M be the degree p cyclic field corresponding to the character $\chi_{N(f_u)}$ and let $L = KM$. Define the representation $\sigma = \chi_{N(f_u)} \otimes \rho'$ of $\mathrm{Gal}(L/\mathbb{Q})$.

Lemma 7.12. *The L-function*

$$L_0(s, \sigma) = \prod_{\substack{q \equiv 1 (p) \\ q \nmid D}} \det\left(I - \frac{\sigma(F_q)}{q^s}\right)^{-1}$$

is convergent for $\mathrm{Re}\, s > \frac{1}{2}$.

Proof. Consider the regular representation $\chi' = \oplus_{i=0}^{p-2} \chi^i$ where $\chi : \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \to \mu_{p-1}$ is a fixed character. Let τ be the representation of $\mathrm{Gal}(L/\mathbb{Q})$ given by $\tau = \sigma \otimes \chi'$.

Note the following facts. By standard properties of Artin L-functions we can factor $L(s, \tau) = \prod_{i=0}^{p-1} L(s, \chi_{N(f_u)} \cdot \chi^i \otimes \rho'(F_q))$. If $q \equiv 1 \bmod p$ then $\chi^i(F_q) = 1$ for all i. If $q \not\equiv 1 \bmod p$ then $\mathrm{tr}[\chi_{N(f_u)} \cdot \chi' \otimes \rho'(F_q)] = 0$. Furthermore for any q we have

$$\det\left(I - \frac{\chi_{N(f_u)} \cdot \chi' \otimes \rho'(F_q)}{q^s}\right) = 1 - \frac{\mathrm{tr}[\chi_{N(f_u)} \cdot \chi' \otimes \rho'(F_q)]}{q^s} + O\left(\frac{1}{q^{2s}}\right).$$

Putting these together we get

$$\begin{aligned} L(s, \tau) &= \prod_{q \nmid D} \det\left(I - \frac{\chi_{N(f_u)} \cdot \chi' \otimes \rho'(F_q)}{q^s}\right)^{-1} \\ &= \prod_{\substack{q \equiv 1 (p) \\ q \nmid D}} \det\left(I - \frac{\chi_{N(f_u)} \otimes \rho'(F_q)}{q^s}\right)^{-p} \prod_{\substack{q \not\equiv 1 (p) \\ q \nmid D}} \left(1 + O\left(\frac{1}{q^{2s}}\right)\right)^{-1}. \end{aligned}$$

The last product above is absolutely convergent for $\mathrm{Re}\, s > \frac{1}{2}$ and hence has no zeros. Note $L(s, \tau)$ is entire since it can be factored as a product of 1-dimensional L-functions or ones which are induced from 1-dimensional L-functions, which are all known to be entire. By assumption of GRH $L(s, \tau)$ has no zeros to the right of $s = \frac{1}{2}$. Thus there exists a branch of $\log(L_0(s, \sigma)^p)$ (and hence of $(L_0(s, \sigma)^p)^{1/p}$) on $\mathrm{Re}\, s > \frac{1}{2}$ and the result follows. □

Define the function

$$L_0(s) = \prod_{\substack{q \equiv 1 (p) \\ q \nmid D}} \left(1 + \frac{\mathrm{tr}\, \sigma(F_q)}{p q^s}\right).$$

Lemma 7.13. *There exists a function* $F(s) = \prod_q (1 + O(1/q^{2s}))$ *which is absolutely convergent for* $\mathrm{Re}\, s > \frac{1}{2}$ *and a branch of* $(L_0(s, \sigma) F(s))^{1/p}$ *defined on* $\mathrm{Re}\, s > \frac{1}{2}$ *such that*

$$L_0(s) = (L_0(s, \sigma) F(s))^{1/p}.$$

Proof. By definition we have

$$L_0(s, \sigma) = \prod_{\substack{q \equiv 1 (p) \\ q \nmid D}} \det(I - \frac{\sigma(F_q)}{q^s})^{-1} = \prod_{\substack{q \equiv 1 (p) \\ q \nmid D}} \left(1 - \frac{\mathrm{tr}\, \sigma(F_q)}{q^s} + O(1/q^{2s})\right)^{-1}.$$

By a similar computation we have

$$L_0(s)^p = \prod_{\substack{q \equiv 1 (p) \\ q \nmid D}} \left(1 + \frac{\mathrm{tr}\, \sigma(F_q)}{q^s} + O(1/q^{2s})\right).$$

Now we have

$$L_0(s,\sigma)L_0(s)^{-p} = \prod_{\substack{q\equiv 1(p)\\ q\nmid D}} (1+O(1/q^{2s}))^{-1}.$$

By assumption of GRH $L(s,\sigma)$ (and hence $L_0(s,\sigma)$) has no zeros to the right of $s=\frac{1}{2}$ and neither does $F(s)$ since it is a convergent product. Thus there exists a branch of $\log L_0(s,\sigma)F(s)$ (and hence of $(L_0(s,\sigma)F(s))^{1/p}$) on $\operatorname{Re} s > \frac{1}{2}$ and the result follows. □

In particular it follows from the above lemma that $L_0(s,\sigma)^{1/p}$ has no poles for $\operatorname{Re} s > \frac{1}{2}$.

Let $\psi(d) = \mu^2(D)\prod_{q\mid d} \operatorname{tr}\sigma(F_q)/p^{\omega(d)}$. Then by Lemma 7.11 we have

$$L_0(s) = \sum_{d\in\mathcal{R}(\infty)} \frac{\psi(d)}{d^s} = \sum_{D_u\in\mathcal{R}(\infty)} \mu^2(D)\left(\frac{D_u}{f_u}\right)_p \left[\frac{C_u}{D_u}\right]_p \cdot d^{-s}.$$

We now apply a standard argument for bounding sums of L-series coefficients (see for instance [Davenport 2000, pages 105–106]). We have

$$\sum_{d\in\mathcal{R}(x)} \psi(d) = \int_{2-iT}^{2+iT} L_0(s)x^s\,\frac{ds}{s} + O\left(\frac{x^2}{T\log x}\right). \tag{7-23}$$

Consider the integral of $L_0(s)x^s/s$ over the rectangle with vertices $\left(\frac{1}{2}+\epsilon, \pm iT\right)$, $(2,\pm iT)$. By Lemma 7.13 we have $|L_0(s)| \ll |L_0(s,\sigma)^{1/p}|$ on this rectangle. Furthermore by the Lindelof conjecture $|L_0(s,\sigma)| \ll (TD)^\epsilon$ on this rectangle. In addition for s on the lines $y\pm iT$ with $y\in\left[\frac{1}{2}+\epsilon, 2\right]$ we have the bounds $|x^s| \ll x^2$ and $1/|s| \ll 1/T$. Thus shifting the above integral to the $\frac{1}{2}+\epsilon$ line we get the bound

$$\begin{aligned}\left|\int_{2-iT}^{2+iT} L_0(s)x^s\,\frac{ds}{s}\right| &\ll \int_{2-iT}^{2+iT} |L_0(s,\sigma)^{1/p}|\left|\frac{x^s}{s}\right| ds\\ &\ll \int_{1/2+\epsilon-iT}^{1/2+\epsilon+iT} |L_0(s,\sigma)^{1/p}||x^s|\,\frac{ds}{|s|} + O\left(\frac{x^2(TD)^\epsilon}{T}\right)\\ &\ll x^{1/2+\epsilon}(TD)^\epsilon \int_{1/2+\epsilon-iT}^{1/2+\epsilon+iT} \frac{1}{|s|}\,ds + O\left(\frac{x^2(TD)^\epsilon}{T}\right)\\ &\ll x^{1/2+\epsilon}(TD)^\epsilon + O\left(\frac{x^2(TD)^\epsilon}{T}\right).\end{aligned} \tag{7-24}$$

Combining (7-23) and (7-24) and setting $T=x^3$ we get

$$\sum_{d\in\mathcal{R}(x)} \psi(d) \ll x^{1/2+\epsilon'}D^\epsilon.$$

Then we bound the inner sum in (7-22) as

$$\sum_{D_u\in\mathcal{R}(\Delta A_u)} \mu^2(D)\left(\frac{D_u}{f_u}\right)_p \left[\frac{C_u}{D_u}\right]_p = \sum_{d\in\mathcal{R}(\Delta A_u)} \psi(d) \ll A_u^{1/2+\epsilon}D^\epsilon.$$

Note that $D^\epsilon \le A_u^{p^{2k}\epsilon}$ since $A_w < A_u$ for all $w \in \mathbb{F}_p^{2k}$. Then summing over all the remaining D_w we get

$$S_k(X, \boldsymbol{A}) \ll X/A_u^{1/4} \ll X/X^{\ddagger 1/4} = o(X).$$

This argument shows that we can remove all $\boldsymbol{A}$ in which there is a variable larger than $X^\ddagger$ and linked with any other $A_w > 1$. This is equivalent to removing the $\boldsymbol{A}$ which belong to families 3 or 4.

We summarize the results of this section in the following theorem:

Theorem 7.14. *Let $\sum_{\boldsymbol{A}}' S_k(X, \boldsymbol{A})$ denote a summation over all tuples $\mathbf{A}$ which do not belong to any of the* 4 *families*, *that is they do not satisfy any of* (7-2), (7-3), (7-6), (7-18). *Then*

$$S_k(X) = \sum\nolimits_{\boldsymbol{A}}' S_k(X, \boldsymbol{A}) + o(X).$$

8. Computing the k-th moment

We now want to prove Theorem 1.6.

For this section we define the following notation. Let $\mathcal{N}(k) = \mathcal{N}(k, p)$ which we recall is the number of vector subspaces of $\mathbb{F}_p^k$. Let $\mathcal{S}(X)$ be the set of positive square-free integers of the form $n = p_1 \cdots p_r$ and each p_i is either a prime congruent to 1 mod p or equal to p^2, and such that $n < X$.

We will do this by proving the following:

Theorem 8.1. *For any $k \in \mathbb{Z}_{\ge 1}$*

$$S_k(X) = p^{-k}(\mathcal{N}(k+1) - \mathcal{N}(k)) \sum_{n \in \mathcal{S}(X)} (p-1)^{\omega(n)-1} + o(X).$$

Note $S_k(X) = \sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k$ is a sum over discriminants up to X^{p-1}. Recall from Section 2 that the number of degree p cyclic fields with discriminant up to X^{p-1} is

$$\sum_{n \in \mathcal{S}(X)} (p-1)^{\omega(n)-1} = cX + o(X).$$

Thus it follows immediately from combining these facts with the above theorem that

$$\lim_{X \to \infty} \frac{\sum_{K, D_K < X} |\mathrm{im}(\varphi_K)|^k}{\sum_{K, D_K < X} 1} = \frac{\mathcal{N}(k+1) - \mathcal{N}(k)}{p^k}.$$

We start by proving some facts about maximal unlinked sets of indices. Recall that for $u, v \in \mathbb{F}_p^2$ written as $u = (u_1, u_2)$ and $v = (v_1, v_2)$ we defined

$$\Phi(u, v) = (u_1)(v_2 - u_2).$$

If we represent each index $u \in \mathbb{F}_p^{2k}$ as $u = (u_{11}, u_{12}, u_{21}, u_{22}, \ldots, u_{k1}, u_{k2})$ then

$$\Phi_k(u, v) = \sum_{i=1}^{k} \Phi((u_{i1}, u_{i2}), (v_{i1}, v_{i2})) = \sum_{i=1}^{k} (u_{i1})(v_{i2} - u_{i2}).$$

We defined $u, v \in \mathbb{F}_p^{2k}$ to be unlinked if $\Phi_k(u, v) = 0$ and $\Phi_k(v, u) = 0$. We say a set $\mathcal{U} \subset \mathbb{F}_p^{2k}$ is unlinked if u and v are unlinked for all $u, v \in \mathcal{U}$.

We will show that for each $\boldsymbol{A}$ in the sum in Theorem 7.14 all the indices $u \in \mathbb{F}_p^{2k}$ with $\Delta A_u \geq 2$ form a maximal unlinked set.

Let $\pi : \mathbb{F}_p^{2k} \to \mathbb{F}_p^k$ be the projection onto the even coordinates, that is

$$\pi(u_{11}, u_{12}, u_{21}, u_{22}, \ldots, u_{k1}, u_{k2}) = (u_{12}, u_{22}, \ldots, u_{k2})$$

and let ρ be the projection onto the odd coordinates. Let $V_1 = \ker \pi$ and let $V_2 = \ker \rho$. Then $\mathbb{F}_p^{2k} \cong V_1 \oplus V_2$ given by $\xi : v \mapsto (\rho v, \pi v)$.

For any subset $V \subseteq \mathbb{F}_p^{2k}$ define $\pi(V)^{\perp} = \{v \in \mathbb{F}_p^k \mid v \cdot u = 0, \forall u \in \pi(V)\}$.

We will start by classifying the maximal unlinked subspaces of $\mathbb{F}_p^{2k}$.

Lemma 8.2. *Let V be a subspace of $\mathbb{F}_p^{2k}$. Then V is an unlinked set if and only if $\xi(V) \subseteq \pi(V)^{\perp} \oplus \pi(V)$ (here $\pi(V)^{\perp} \oplus \pi(V)$ is viewed as a subspace of $V_1 \oplus V_2$). Equality holds if and only if V is a maximal unlinked subspace.*

Proof. For any $v, w \in V$

$$\Phi_k(v, w) = \sum_{i=1}^{k} v_{i1}(w_{i2} - v_{i2}) = \rho(v) \cdot \pi(w - v).$$

Suppose V is unlinked. Fix any $u \in V$. Let $w = u + v \in V$ in the above equation. Then we get $\Phi_k(v, w) = \rho(v) \cdot \pi(u) = 0$ so $\rho(v) \in \pi(V)^{\perp}$. Since $v \in V$ was arbitrary this implies $\xi(V) \subseteq \pi(V)^{\perp} \oplus \pi(V)$. The converse is clear from the above equation.

For the second part note that $\dim \pi(V)^{\perp} \oplus \pi(V) = k$ for all subspaces V of $\mathbb{F}_p^{2k}$. Thus by the first part any unlinked subspace is contained in one of dimension k. This completes the proof. □

Next we determine when translation preserves the property of being unlinked.

Lemma 8.3. *Suppose $V \subset \mathbb{F}_p^{2k}$ is an unlinked subspace. Let $a \in \mathbb{F}_p^{2k}$. Then $V + a$ is unlinked if and only if $\rho(a) \in \pi(V)^{\perp}$.*

Proof. Let $v, w \in V$. Then we compute

$$\Phi_k(v + a, w + a) = \sum_{i=1}^{k} (v_{i1} + a_{i1})(w_{i2} - v_{i2}) = \sum_{i=1}^{k} (a_{i1})(w_{i2} - v_{i2}) = \rho(a) \cdot \pi(w) - \rho(a) \cdot \pi(v).$$

If $\rho(a) \in \pi(V)^{\perp}$ we see $V + a$ is unlinked. If $V + a$ is unlinked then setting $w = v + u$ for any $u \in V$ we see $\rho(a) \in \pi(V)^{\perp}$. □

Next we show that every maximal unlinked set is a coset of some unlinked subspace.

Lemma 8.4. *Let $\mathcal{U} \subset \mathbb{F}_p^{2k}$ be a maximal unlinked set and let $a \in \mathcal{U}$. Let $V = \mathcal{U} - a$. Then $V \subset \mathbb{F}_p^{2k}$ is an unlinked subspace.*

Proof. First we show V is a subspace. Let $u, v \in \mathcal{U}$. We need to show that $(u-a)+(v-a)+a = u+v-a \in \mathcal{U}$. Since $\mathcal{U}$ is maximal we show $u+v-a$ is unlinked with every element of $\mathcal{U}$. Let $w \in \mathcal{U}$. We have

$$\begin{aligned}\Phi_k(u+v-a, w) &= \sum_{i=1}^{k}(u_{i1}+v_{i1}-a_{i1})(w_{i2}-u_{i2}-v_{i2}+a_{i2}) \\ &= \sum_{i=1}^{k}(u_{i1}+v_{i1}-a_{i1})((a_{i2}-u_{i2})+(w_{i2}-v_{i2})) \\ &= \sum_{i=1}^{k} v_{i1}(a_{i2}-u_{i2})+u_{i1}(w_{i2}-v_{i2})-a_{i1}(w_{i2}-v_{i2}) \\ &= 0\end{aligned}$$

where the last two equalities follow since u, v, w, a are all unlinked and for instance $v_{i1}(-u_{i2}+a_{i2}) = -v_{i1}(u_{i2}-v_{i2})+v_{i1}(-v_{i2}+a_{i2})$. Similarly

$$\Phi_k(w, u+v-a) = \sum_{i=1}^{k}(w_{i1})(u_{i2}+v_{i2}-a_{i2}-w_{i2}) = 0.$$

Thus V is a subspace. Next we show V is unlinked.

For any $w \in V$ we have $\mathcal{U} = \mathcal{U}+w$. Let $u' \in V$ and let $u = u'+a$, so $u \in \mathcal{U}$. Note $a \in \mathcal{U}$. Then we have

$$\begin{aligned}0 &= \Phi_k(u+w, a+w) \\ &= \sum_{i=1}^{k}(u_{i1}+w_{i1})(a_{i2}+w_{i2}-u_{i2}-w_{i2}) \\ &= \Phi_k(u, a)+\sum_{i=1}^{k} w_{i1}(a_{i2}-u_{i2}) \\ &= -\sum_{i=1}^{k} w_{i1}u'_{i2} = \rho(w)\cdot\pi(u')\end{aligned}$$

Since $w, u' \in V$ were arbitrary this shows $\rho(V) \in \pi(V)^{\perp}$ so $\xi(V) \subset \pi(V)^{\perp} \oplus \pi(V)$. By Lemma 8.2 V is unlinked. □

With the above results we can classify all the maximal unlinked sets.

Proposition 8.5. *The maximal unlinked sets $\mathcal{U} \subset \mathbb{F}_p^{2k}$ are exactly the sets of the form $\mathcal{U} = V + a$ where V is a subspace which is a maximal unlinked set and $\rho(a) \in \pi(V)^{\perp}$.*

Proof. Suppose $\mathcal{U}$ is a maximal unlinked set. By Lemma 8.4 $\mathcal{U}-a = V$ for some subspace V which is unlinked and some $a \in \mathcal{U}$. Since $\mathcal{U} = V+a$ by Lemma 8.3 we see $\rho(a) \in \pi(V)^{\perp}$.

Let $\mathcal{W} = \xi^{-1}(\pi(V)^{\perp} \oplus \pi(V))$ so that $V \subset \mathcal{W}$. By Lemma 8.2 $\mathcal{W}$ is unlinked. Note $\rho(a) \in \pi(V)^{\perp} = \pi(\mathcal{W})^{\perp}$. Thus $\mathcal{W}+a$ is unlinked and $\mathcal{U} \subset \mathcal{W}+a$. By maximality we get $\mathcal{U} = \mathcal{W}+a$.

The converse is clear from the above lemmas. □

As mentioned in the above proofs $\dim \pi(V)^{\perp} \oplus \pi(V) = k$ for all subspaces V and hence, by Proposition 8.5, every maximal unlinked set has size p^k.

With this we can rewrite $S_k(X)$ in a form closer to Theorem 8.1.

Proposition 8.6. *Let $k \in \mathbb{Z}_{\geq 1}$. Let U be the number of maximal unlinked sets in $\mathbb{F}_p^{2k}$. Then*

$$\sum_{K, D_K < X^{p-1}} |\mathrm{im}(\varphi_K)|^k = \binom{U}{p^k} \sum_{n \in \mathcal{S}(X)} (p-1)^{\omega(n)-1} + o(X)$$

where $\mathcal{S}(X)$ is the set of square-free positive integers of the form $n = p_1 \cdots p_r$ where each p_i is either a prime congruent to 1 mod p *or equal to p^2, such that $n < X$.*

Proof. Given two maximal unlinked sets $\mathcal{U}_i$ for $i = 1, 2$, if $a \in \mathcal{U}_1 \cap \mathcal{U}_2$ then $V_i = \mathcal{U}_i - a$ is a vector space which is also a maximal unlinked set. If the $\mathcal{U}_i$ are distinct then so are the V_i and hence $V_1 \cap V_2$ is at most $k-1$ dimensional. Hence the largest possible intersection of two distinct maximal unlinked sets has size p^{k-1}. Thus a set of $p^{k-1}+1$ unlinked indices determines a unique maximal unlinked set.

Let $\boldsymbol{A}$ be a tuple as in the statement of Theorem 7.14, that is $\boldsymbol{A} \notin \mathcal{F}_i$ for $i = 1, 2, 3, 4$. Let $\mathcal{U}_0$ be the set of indices $u \in \mathbb{F}_p^{2k}$ such that $A_u > X^{\ddagger}$. Since $\boldsymbol{A} \notin \mathcal{F}_2$ this implies there are at least $p^{k-1}+1$ indices $u \in \mathbb{F}_p^{2k}$ with $A_u > X^{\ddagger}$ so $|\mathcal{U}_0| \geq p^{k-1}+1$. Then since $\boldsymbol{A} \notin \mathcal{F}_3$ and $X^{\ddagger} > X^{\dagger}$ the set $\mathcal{U}_0$ is unlinked, so by the above remark determines a unique maximal unlinked set $\mathcal{U} \supset \mathcal{U}_0$.

Hence any $u \in \mathbb{F}_p^{2k}$ such that $u \notin \mathcal{U}$ is linked with some $v \in \mathcal{U}_0$. Since $\boldsymbol{A} \notin \mathcal{F}_3 \cup \mathcal{F}_4$ this implies $A_u < 2/\Delta$ and hence $D_u = 1$.

Let $\boldsymbol{A}_1$ be the tuple consisting of the coordinates of $\boldsymbol{A}$ in $\mathcal{U}$. Thus in the expression $S_k(X, \boldsymbol{A})$ all of the characters χ_l evaluate to 1 so we can write

$$\begin{aligned} S_k(X, \boldsymbol{A}) &= \frac{1}{(p-1)\cdot p^k} \sum_{(D_u) \in \mathcal{P}(X, \boldsymbol{A})} \sum_{(\chi_{l'}) \in \mathcal{C}(D)} \frac{\mu^2(D)}{p^{k\omega(D)}} \prod_v \prod_{l \mid D_v} \chi_l \Bigl(\prod_u D_u^{\Phi_k(u,v)} \Bigr) + o(X) \\ &= \frac{1}{(p-1)\cdot p^k} \sum_{(D_u)_{u \in \mathcal{U}} \in \mathcal{P}(X, \boldsymbol{A}_1)} \sum_{(\chi_{l'}) \in \mathcal{C}(D')} \frac{\mu^2(D')}{p^{k\omega(D')}} + o(X) \\ &= \frac{1}{(p-1)\cdot p^k} \sum_{(D_u)_{u \in \mathcal{U}} \in \mathcal{P}(X, \boldsymbol{A}_1)} \frac{\mu^2(D')}{p^{k\omega(D')}} \cdot (p-1)^{\omega(D')} + o(X) \end{aligned}$$

where we denote $D = \prod_{u \in \mathbb{F}_p^{2k}} D_u$ and $D' = \prod_{u \in \mathcal{U}} D_u$.

Let $\mathcal{A}(\mathcal{U})$ be the set of tuples $\boldsymbol{A}$ which determine $\mathcal{U}$ by the above procedure, that is for which $A_u < 2/\Delta$ for all $u \notin \mathcal{U}$ and $A_u > X^{\ddagger}$ for all $u \in \mathcal{U}$.

We can partition $\sum'_{\boldsymbol{A}} S_k(X, \boldsymbol{A}) = \sum_{\mathcal{U}} S_k(X, \mathcal{U})$ where we define $S_k(X, \mathcal{U}) = \sum'_{\boldsymbol{A} \in \mathcal{A}(\mathcal{U})} \sum S_k(X, \boldsymbol{A})$. Notice that for each $u \in \mathcal{U}$ the range of summation of each D_u in $S_k(X, \mathcal{U})$ is $X^{\ddagger} < D_u < X$. It follows from Section 7B that we can extend this to $1 \leq D_u < X$ since it is proven there that the summation over these terms is contained in the error term.

Thus we have

$$S_k(X,\mathcal{U}) = \frac{1}{p^k(p-1)} \sum_{\prod_{j=1}^{p^k} n_j \in \mathcal{S}(X)} \mu^2\Big(\prod_{j=1}^{p^k} n_j\Big)\Big(\frac{p-1}{p^k}\Big)^{\omega\left(\prod_{j=1}^{p^k} n_j\right)} + o(X)$$
$$= \frac{1}{p^k} \sum_{n\in\mathcal{S}(X)} \mu^2(n)(p-1)^{\omega(n)-1} + o(X).$$

The last equality follows since there are $p^{k\omega(n)}$ ways of writing a positive integer n as a product of p^k positive integers. The proposition follows by summing $S_k(X,\mathcal{U})$ over all maximal unlinked sets $\mathcal{U}$. □

The final step of the proof will be the next proposition. Define $n(k,r)$ to be the number of r-dimensional subspaces of $\mathbb{F}_p^k$. We will need two properties of this function which can be found in Lemmas 1 and 3 from [Fouvry and Klüners 2007].

Lemma 8.7. *The function $n(k,r)$ satisfies*

$$n(k,r) = n(k,k-r), \quad \sum_{r=0}^{k} p^r n(k,r) = \mathcal{N}(k+1) - \mathcal{N}(k).$$

Lemma 8.8. *The number U of maximal unlinked sets $\mathcal{U} \subset \mathbb{F}_p^{2k}$ is*

$$U = \mathcal{N}(k+1) - \mathcal{N}(k).$$

Proof. By Lemma 8.3 if V is a maximal unlinked subspace then $V+a$ is maximal unlinked if and only if $\rho(a) \in \pi(V)^\perp$. Hence given any k-dimensional subspace $V \subset \mathbb{F}_p^{2k}$ which is a maximal unlinked set there are $p^k(p^{\dim \pi(V)^\perp})$ vectors which translate V to a maximal unlinked set. However since translating by a_1 and a_2 gives the same set if and only if a_1 and a_2 are in the same coset of V this implies that there are $p^{\dim \pi(V)^\perp}$ distinct maximal unlinked sets that can be obtained from V.

Now let S be the set of k-dimensional subspaces $V \subset \mathbb{F}_p^{2k}$ which satisfy Lemma 8.2. We compute the size of this set. Fix some subspace $V_0 \subset \mathbb{F}_p^k$ with $\dim V_0 = r$ and suppose V satisfies $\pi(V) = V_0$. So $\dim \pi(V)^\perp = k - r$. If $V \in S$ then $V = \pi(V)^\perp \oplus \pi(V) = V_0^\perp \oplus V_0$ and hence there is a unique $V \in S$ with $\pi(V) = V_0$. Hence the number of $V \in S$ with $\dim \pi(V) = r$ is $n(k,r)$.

Thus we have

$$U = \sum_{V\in S} p^{\dim \pi(V)^\perp} = \sum_{r=0}^{k} p^r n(k,r) = \mathcal{N}(k+1) - \mathcal{N}(k)$$

by Lemma 8.7. □

Thus combining Lemma 8.8 with Proposition 8.6 we have shown

$$S_k(X) = p^{-k}(\mathcal{N}(k+1) - \mathcal{N}(k)) \sum_{n<X} (p-1)^{\omega(n)-1} + o(X)$$

which proves Theorem 8.1. As remarked at the beginning of the section it follows that

$$\lim_{X\to\infty} \frac{\sum_{K,D_K<X} |\mathrm{im}(\varphi_K)|^k}{\sum_{K,D_K<X} 1} = p^{-k}(\mathcal{N}(k+1) - \mathcal{N}(k)). \tag{8-1}$$

Thus we have computed all the moments of the function $|\mathrm{im}(\varphi_K)|$ over degree p cyclic fields. We now refer to a result of Fouvry and Klüners which shows that these moments determine a distribution.

The combination of Proposition 1 and Theorem 2 from [Fouvry and Klüners 2006] can be summarized in the following form. This form is slightly more general than the original but follows by the same exact proof (see [Fouvry and Klüners 2006, pages 7–15]) which only uses properties of the function $\mathcal{N}(k, p)$ and $\eta_s(p)$ (defined in the introduction).

Let $\mathcal{F}$ be a family of number fields. Let f be a function on $\mathcal{F}$ valued in $\{1, p, p^2, \ldots\}$. Let

$$\mathcal{M}(k) = \lim_{X\to\infty} \frac{\sum_{K\in\mathcal{F},|D_K|<X} f^k(K)}{\sum_{K\in\mathcal{F},|D_K|<X} 1}.$$

Proposition 8.9. *Let p be a prime. Suppose that for every $k \in \mathbb{Z}_{\geq 1}$*

$$\mathcal{M}(k) = p^{-k}(\mathcal{N}(k+1) - \mathcal{N}(k)).$$

Then for every $s \in \mathbb{Z}_{\geq 0}$ the density of the set $\{K \in \mathcal{F} \mid f(K) = p^s\}$ is

$$\frac{\eta_\infty(p)}{\eta_s(p)\eta_{s+1}(p)p^{s(s+1)}}.$$

By letting $\mathcal{F}$ be the set of degree p cyclic fields and $f = |\mathrm{im}(\varphi_K)| = p^{\mathrm{rk}_p \mathrm{im}(\varphi_K)}$ we see that Proposition 8.9 combined with (8-1) immediately implies Theorems 1.4 and 1.5.

Acknowledgements

The author extends his thanks to several people for helpful and stimulating discussions while working on this problem: John Friedlander, Jürgen Klüners, Jacob Tsimerman, Asif Zaman. He would also like to thank Leo Goldmakher and Matthew Young for providing useful references on character sums.

References

[Alberts 2016] B. Alberts, "Cohen–Lenstra moments for some nonabelian groups", preprint, 2016. arXiv

[Baier and Young 2010] S. Baier and M. P. Young, "Mean values with cubic characters", *J. Number Theory* **130**:4 (2010), 879–903. MR Zbl

[Bhargava 2014] M. Bhargava, "The geometric sieve and the density of squarefree values of invariant polynomials", preprint, 2014. arXiv

[Boston and Wood 2017] N. Boston and M. M. Wood, "Non-abelian Cohen–Lenstra heuristics over function fields", *Compos. Math.* **153**:7 (2017), 1372–1390. MR Zbl

[Cohen and Lenstra 1984] H. Cohen and H. W. Lenstra, Jr., "Heuristics on class groups of number fields", pp. 33–62 in *Number theory* (Noordwijkerhout, Netherlands, 1983), edited by H. Jager, Lecture Notes in Math. **1068**, Springer, 1984. MR Zbl

[Cohen and Martinet 1987] H. Cohen and J. Martinet, "Class groups of number fields: numerical heuristics", *Math. Comp.* **48**:177 (1987), 123–137. MR Zbl

[Datskovsky and Wright 1988] B. Datskovsky and D. J. Wright, "Density of discriminants of cubic extensions", *J. Reine Angew. Math.* **386** (1988), 116–138. MR Zbl

[Davenport 2000] H. Davenport, *Multiplicative number theory*, 3rd ed., Graduate Texts in Math. **74**, Springer, 2000. MR Zbl

[Davenport and Heilbronn 1971] H. Davenport and H. Heilbronn, "On the density of discriminants of cubic fields, II", *Proc. Roy. Soc. Lond. Ser. A* **322**:1551 (1971), 405–420. MR Zbl

[Ellenberg et al. 2016] J. S. Ellenberg, A. Venkatesh, and C. Westerland, "Homological stability for Hurwitz spaces and the Cohen–Lenstra conjecture over function fields", *Ann. of Math.* (2) **183**:3 (2016), 729–786. MR Zbl

[Fouvry and Klüners 2006] É. Fouvry and J. Klüners, "Cohen–Lenstra heuristics of quadratic number fields", pp. 40–55 in *Algorithmic number theory*, edited by F. Hess et al., Lecture Notes in Comput. Sci. **4076**, Springer, 2006. MR Zbl

[Fouvry and Klüners 2007] É. Fouvry and J. Klüners, "On the 4-rank of class groups of quadratic number fields", *Invent. Math.* **167**:3 (2007), 455–513. MR Zbl

[Gerth 1984] F. Gerth, III, "The 4-class ranks of quadratic fields", *Invent. Math.* **77**:3 (1984), 489–515. MR Zbl

[Gerth 1987] F. Gerth, III, "Densities for ranks of certain parts of p-class groups", *Proc. Amer. Math. Soc.* **99**:1 (1987), 1–8. MR Zbl

[Goldstein 1970] L. J. Goldstein, "A generalization of the Siegel–Walfisz theorem", *Trans. Amer. Math. Soc.* **149** (1970), 417–429. MR Zbl

[Hall 1938] P. Hall, "A partition formula connected with Abelian groups", *Comment. Math. Helv.* **11**:1 (1938), 126–129. MR Zbl

[Heath-Brown 2000] D. R. Heath-Brown, "Kummer's conjecture for cubic Gauss sums", *Israel J. Math.* **120**:part A (2000), 97–124. MR Zbl

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, Amer. Math. Soc. Colloq. Publ. **53**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Koymans and Pagano 2018] P. Koymans and C. Pagano, "On the distribution of $\mathrm{Cl}(K)[l^\infty]$ for degree l cyclic fields", preprint, 2018. arXiv

[Lemmermeyer 2013] F. Lemmermeyer, "The ambiguous class number formula revisited", *J. Ramanujan Math. Soc.* **28**:4 (2013), 415–421. MR Zbl

[Mayer 1992] D. C. Mayer, "Multiplicities of dihedral discriminants", *Math. Comp.* **58**:198 (1992), 831–847. MR Zbl

[Milovic 2017] D. Milovic, "On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-8p})$ for $p \equiv -1 \mod 4$", *Geom. Funct. Anal.* **27**:4 (2017), 973–1016. MR Zbl

[Milovic 2018] D. Z. Milovic, "On the 8-rank of narrow class groups of $\mathbb{Q}(\sqrt{-4pq})$, $\mathbb{Q}(\sqrt{-8pq})$, and $\mathbb{Q}(\sqrt{8pq})$", *Int. J. Number Theory* **14**:8 (2018), 2165–2193.

[Neukirch 1999] J. Neukirch, *Algebraic number theory*, Grundlehren der Math. Wissenschaften **322**, Springer, 1999. MR Zbl

[Smith 2017] A. Smith, "2^∞-Selmer groups, 2^∞-class groups, and Goldfeld's conjecture", preprint, 2017. arXiv

[Stevenhagen 1995] P. Stevenhagen, "Rédei-matrices and applications", pp. 245–259 in *Number theory* (Paris, 1992-93), edited by S. David, Lond. Math. Soc. Lecture Note Ser. **215**, Cambridge Univ. Press, 1995. MR Zbl

[Wood 2019] M. M. Wood, "Nonabelian Cohen–Lenstra moments", *Duke Math. J.* **168**:3 (2019), 377–427. MR Zbl

[Wright 1989] D. J. Wright, "Distribution of discriminants of abelian extensions", *Proc. Lond. Math. Soc.* (3) **58**:1 (1989), 17–50. MR Zbl

Communicated by Bjorn Poonen
Received 2017-06-30 Revised 2019-10-30 Accepted 2019-11-27

jack.klys@ucalgary.ca *Department of Mathematics and Statistics, University of Calgary, Canada*

ALGEBRA AND NUMBER THEORY 14:4 (2020)
https://doi.org/10.2140/ant.2020.14.855

On the motivic class of an algebraic group

Federico Scavia

Let F be a field of characteristic zero admitting a biquadratic field extension. We give an example of a torus G over F whose classifying stack BG is stably rational and such that $\{BG\} \neq \{G\}^{-1}$ in the Grothendieck ring of algebraic stacks over F. We also give an example of a finite étale group scheme A over F such that BA is stably rational and $\{BA\} \neq 1$.

1. Introduction

Let F be a field. The Grothendieck ring of algebraic stacks $K_0(\mathrm{Stacks}_F)$ was introduced by Ekedahl [2009b], following up on earlier works [Behrend and Dhillon 2007; Joyce 2007; Toën 2005]. It is a variant of the Grothendieck ring of varieties $K_0(\mathrm{Var}_F)$. By definition, $K_0(\mathrm{Stacks}_F)$ is generated as an abelian group by the equivalence classes $\{X\}$ of all algebraic stacks X of finite type over F with affine stabilizers. These classes are subject to the scissor relations $\{X\} = \{Y\} + \{X \setminus Y\}$ for every closed substack $Y \subseteq X$, and the relations $\{E\} = \{\mathbb{A}^n \times X\}$ for every vector bundle E of rank n over X. The product is defined by $\{X\} \cdot \{Y\} := \{X \times Y\}$, and extended by linearity.

Given a group scheme G over F, we may consider the class $\{BG\}$ of its classifying stack in $K_0(\mathrm{Stacks}_F)$. The problem of computing $\{BG\}$ appears to be related to the problem of the stable rationality of BG, although no direct implications are known. Recall that BG is stably rational if for one (equivalently, every) generically free representation V of G, the rational quotient V/G is stably rational. An equivalent terminology is that the Noether problem for stable rationality has a positive solution for G; see [Florence and Reichstein 2018, Section 3]. The case of a finite (constant) group G was considered in [Ekedahl 2009a]: it frequently happens that $\{BG\} = 1$ (notably for the symmetric groups, see [loc. cit., Theorem 4.3]), although there are examples of finite groups G for which $\{BG\} \neq 1$; see [loc. cit., Corollaries 5.2 and 5.8]. Further work on the triviality of $\{BG\}$ for finite groups G has been done in [Martino 2016; 2017]. So far, all the known examples of finite group schemes G for which $\{BG\} \neq 1$ are such that BG is not stably rational. This suggests the following question.

Question 1.1 (see [Ekedahl 2009a, Section 6]). Is it true that, for a finite group scheme G, the following two conditions are equivalent?

- BG is stably rational.
- $\{BG\} = 1$ in $K_0(\mathrm{Stacks}_F)$.

MSC2010: primary 14L15; secondary 14D23.
Keywords: motivic class, Grothendieck ring of stacks, classifying stack, algebraic torus.

We will answer Question 1.1 in the negative in Theorem 1.6.

Now let G be a connected linear algebraic group. Recall that G is special if every G-torsor is Zariski-locally trivial. For example, GL_n, SL_n and Sp_{2n} are special; see [Serre 1958]. It was shown by Ekedahl that if $P \to S$ is a torsor under the special group G, then $\{P\} = \{G\}\{S\}$. This is immediate if S is a scheme, but less obvious when S is a stack; see [Bergh 2016, Corollary 2.4]. Applying this to the universal G-torsor $\operatorname{Spec} F \to BG$, one obtains $\{BG\}\{G\} = 1$.

The equality $\{BG\} = \{G\}^{-1}$ appears to be the analogue for connected groups of the relation $\{BG\} = 1$ for finite group schemes. In [Bergh 2016], these equalities are referred to as *expected class formulas*, and there is a sense in which they are "almost" true. Ekedahl [2009b, Section 2] defines a generalized Euler characteristic

$$\chi_c : K_0(\mathrm{Stacks}_F) \to K_0(\mathrm{Coh}_F)$$

taking values in a Grothendieck ring $K_0(\mathrm{Coh}_F)$ of Galois representations over F. If G is a finite group scheme, the equality $\chi_c(\{BG\}) = 1$ always holds [Ekedahl 2009a, Proposition 3.1]. On the other hand, if G is connected, then $\chi_c(\{BG\}\{G\}) = 1$; see [Bergh 2016, Section 2.2]. Since $\{BG\} \neq 1$ for some finite groups G, the following question naturally arises:

Question 1.2. Let F be a field. Is it true that

$$\{BG\} = \{G\}^{-1} \tag{1.3}$$

in $K_0(\mathrm{Stacks}_F)$ for every connected group G?

In Theorem 1.5, we show that the answer to Question 1.2 is also negative. Computations for nonspecial G have been carried out for PGL_2 and PGL_3 in [Bergh 2016], for SO_n and n odd in [Dhillon and Young 2016], for SO_n and n even (and O_n for any n) in [Talpo and Vistoli 2017], and for Spin_7, Spin_8 and G_2 in [Pirisi and Talpo 2019]. In each of these cases, (1.3) was found to be true. The expectation was that, for a connected linear algebraic group G over a field F of characteristic 0, Question 1.4 below should have an affirmative answer. If F is an algebraically closed field, then there are no examples of connected G where BG is known not to be stably rational. If F is not assumed to be algebraically closed, then such examples exist. The following variant of Question 1.1 seems natural in this context.

Question 1.4 (see [Talpo and Vistoli 2017, Section 1] and [Pirisi and Talpo 2019, Remark 4.1]). Is it true that, for a connected linear algebraic group G, the following two conditions are equivalent?

- BG is stably rational.
- $\{BG\} = \{G\}^{-1}$ in $K_0(\mathrm{Stacks}_F)$.

Our first result gives a negative answer to Questions 1.2 and 1.4.

Theorem 1.5. *Let F be a field of characteristic zero which admits a biquadratic field extension K, let E_1 and E_2 be two distinct quadratic subextensions of K/F, and set $G := R^{(1)}_{E_1\times E_2/F}(\mathbb{G}_m)$. Then*

(a) *BG is stably rational, and*

(b) *$\{BG\} \neq \{G\}^{-1}$ in $K_0(\mathrm{Stacks}_F)$.*

The torus G is an example of a norm-one torus; see Section 2 for the definition. It follows from Theorem 1.5 that counterexamples H to (1.3) exist in any dimension $\dim H \geq 3$: consider for example $H := G \times \mathbb{G}_m^r$ for $r \geq 0$.

The key ingredient in the proof of Theorem 1.5 is the *refined Euler characteristic* of Ekedahl, introduced in [Ekedahl 2009b, Sections 6 and 3]; see Section 4.

Our second result gives a negative answer to Question 1.1.

Theorem 1.6. *Let F be a field of characteristic zero which admits a biquadratic field extension K, and let E_1 and E_2 be two distinct quadratic subextensions of K/F. Define $G := R^{(1)}_{E_1\times E_2/F}(\mathbb{G}_m)$, and let $A := G[2]$ be the 2-torsion subgroup of G. Then*

(a) *BA is stably rational, and*

(b) *$\{BA\} \neq 1$ in $K_0(\mathrm{Stacks}_F)$.*

Questions 1.1, 1.2 and 1.4 remain open in the case, where the base field F is assumed to be algebraically closed. Our arguments do not shed any new light in this setting.

The remainder of this paper is structured as follows. In Section 2 we review well known computations of motivic classes for nonsplit tori. In Section 3 we obtain explicit formulas for the motivic classes of G and BG, and in Section 4 we give the required background on the refined Euler characteristic. In Section 5 we prove Theorem 1.5, and in Section 6 we prove Theorem 1.6.

2. Preliminaries

Let F be a field. We will write $\mathbb{L}$ for the class $\{\mathbb{A}^1\}$ in $K_0(\mathrm{Var}_F)$ or $K_0(\mathrm{Stacks}_F)$. If E is an étale algebra over F, we will denote by $\{E\}$ the class $\{\mathrm{Spec}\, E\}$ in $K_0(\mathrm{Var}_F)$ or $K_0(\mathrm{Stacks}_F)$. If X is a quasiprojective scheme over E, we will denote by $R_{E/F}(X)$ the *Weil restriction* of X to F. By definition, for every F-scheme S one has $R_{E/F}(X)(S) = X(S_E)$. We refer the reader to [Voskresensky 1998, Section 3.12] for an account of the main properties of the Weil restriction.

Let G be a linear algebraic group over F, and $\alpha \in H^1(F, G)$ be represented by a G-torsor $P \to \mathrm{Spec}\, F$. For every quasiprojective F-scheme Z, we denote by ${}^{\alpha}Z$ the *twist* of Z by P, that is,

$${}^{\alpha}Z := (Y \times P)/G,$$

where G acts diagonally. We refer the reader to [Florence 2008, Section 2] for the definition and the basic properties of the twisting operation.

We will write C_2 for the cyclic group of two elements and S_n for the symmetric group on n symbols.

The following observations will be repeatedly used in the sequel.

Lemma 2.1. *Let X be a scheme over F, E an étale algebra of degree n over F, $\alpha \in H^1(F, S_n)$ the class corresponding to E/F.*

(a) *Let S_n act on the disjoint union $\coprod_{i=1}^n X$ by permuting the n copies of X. Then*

$$^{\alpha}\Big(\coprod_{i=1}^{n} X\Big) \cong X_E.$$

(b) *Let S_n act on X^n by permuting the n factors. Then*

$$^{\alpha}(X^n) \cong R_{E/F}(X).$$

Proof. (a) Let $Y := \coprod_{i=1}^n X$, and let S_n act on Y by permuting the copies of X. By definition,

$$^{\alpha}Y = (Y \times \operatorname{Spec} E)/S_n \cong (Y \times_X X_E)/S_n,$$

where S_n acts diagonally. This shows that $^{\alpha}Y$ is the twist of X_E by the trivial S_n-torsor $Y \to X$ in the category of X-schemes, which implies $^{\alpha}Y \cong X_E$.

(b) See the bottom of page 5 in [Florence et al. 2017]. □

Lemma 2.2. *Let*

$$1 \to N \to G \to H \to 1$$

be an exact sequence of group schemes over F, and assume that G is special. Then

$$\{BN\} = \{H\}/\{G\}.$$

Proof. See [Bergh 2016, Proposition 2.9]. □

Let F_s be a separable closure of F. Recall that a group scheme T over F is called a *torus* if $T_{F_s} \cong \mathbb{G}^n_{\mathrm{m},F_s}$ for some $n \geq 0$. The *character lattice* of T is the finitely generated $\mathbb{Z}$-free $\operatorname{Gal}(F)$-module $\operatorname{Hom}_{F_s}(T_{F_s}, \mathbb{G}_{\mathrm{m},F_s})$. The character lattice induces an antiequivalence between the category of F-tori and the category of $\operatorname{Gal}(F)$-lattices, i.e., finitely generated $\mathbb{Z}$-free continuous $\operatorname{Gal}(F)$-modules; see [Favi and Florence 2008, Section 2]. Similarly, for every separable finite extension L/F, we have an antiequivalence between $\operatorname{Gal}(L/F)$-lattices and F-tori T split by L, i.e., such that $T_L \cong \mathbb{G}^n_{\mathrm{m},L}$ for some $n \geq 0$. The *dual torus* of T is the torus T' whose character lattice is dual to that of T.

Let E be an étale algebra over F. If G is a group scheme over E, then $R_{E/F}(G)$ is a group scheme over F. The group $R_{E/F}(\mathbb{G}_\mathrm{m}) := R_{E/F}(\mathbb{G}_{\mathrm{m},E})$ is an F-torus. Tori of this kind are called *quasisplit*. They are special groups, and they correspond to permutation $\operatorname{Gal}(F)$-lattices, that is, lattices admitting a $\mathbb{Z}$-basis that is permuted by $\operatorname{Gal}(F)$; see [Voskresensky 1998, Section 3.12, Example 19].

Lemma 2.3. *Let T be an algebraic torus over F, and let T' be its dual. Assume that T is stably rational. Then*:

(a) *BT' is stably rational.*

(b) *$\{BT'\}\{T\} = 1$ in $K_0(\mathrm{Stacks}_F)$.*

Proof. Since T is stably rational, by [Voskresensky 1998, Section 4.7, Theorem 2] there is a short exact sequence

$$1 \to T_1 \to T_2 \to T \to 1 \tag{2.4}$$

where T_1 and T_2 are quasisplit. Since quasisplit tori are isomorphic to their dual, the sequence dual to (2.4),

$$1 \to T' \to T_2 \to T_1 \to 1, \tag{2.5}$$

shows that T' embeds in T_2. We may view T_2 as a maximal torus inside GL_n, where $n = \operatorname{rank} T_2$. This gives a faithful representation of T' with quotient birational to T_1. Since quasisplit tori are rational, it follows that BT' is stably rational.

Quasisplit tori are special, so we may apply Lemma 2.2 to (2.4) and (2.5). We obtain $\{T\} = \{T_2\}/\{T_1\}$ and $\{BT'\} = \{T_1\}/\{T_2\}$, so $\{BT'\}\{T\} = 1$. □

Let E/F be an étale algebra, and let $R_{E/F}(\mathbb{G}_\mathrm{m})$ be the associated quasisplit torus. The kernel of the norm homomorphism $R_{E/F}(\mathbb{G}_\mathrm{m}) \to \mathbb{G}_\mathrm{m}$ is called a *norm-one torus*, and is denoted by $R^{(1)}_{E/F}(\mathbb{G}_\mathrm{m})$. Its dual torus is isomorphic to $R_{E/F}(\mathbb{G}_\mathrm{m})/\,\mathbb{G}_\mathrm{m}$.

Lemma 2.6. *Assume that* char $F \neq 2$. *Let $E := F(\sqrt{m})$ be a separable quadratic field extension, and let α denote the class of E/F in $H^1(F, C_2)$. Then:*

(a) $R^{(1)}_{E/F}(\mathbb{G}_\mathrm{m}) \cong R_{E/F}(\mathbb{G}_\mathrm{m})/\,\mathbb{G}_\mathrm{m}$.

(b) *Let* $\mathrm{Gal}(E/F)$ *act on* $\mathbb{P}^1$ *via* $z \mapsto z^{-1}$. *Then* ${}^{\alpha}\mathbb{P}^1 \cong \mathbb{P}^1$.

(c) $R_{E/F}(\mathbb{G}_\mathrm{m})/\,\mathbb{G}_\mathrm{m}$ *is rational and*

$$\{R_{E/F}(\mathbb{G}_\mathrm{m})/\,\mathbb{G}_\mathrm{m}\} = \{B(R_{E/F}(\mathbb{G}_\mathrm{m})/\,\mathbb{G}_\mathrm{m})\}^{-1} = \mathbb{L} - \{E\} + 1.$$

(d) $\{R_{E/F}(\mathbb{G}_\mathrm{m})\} = \{BR_{E/F}(\mathbb{G}_\mathrm{m})\}^{-1} = (\mathbb{L} - 1)(\mathbb{L} - \{E\} + 1)$.

(e) $\{R_{E/F}(\mathbb{P}^1)\} = \mathbb{L}^2 + \{E\}\mathbb{L} + 1$.

Proof. (a) Both tori correspond to the unique nontrivial $\mathrm{Gal}(E/F)$-lattice of rank 1. Here $\mathrm{Gal}(E/F) \cong C_2$.

(b) The C_2-action on $\mathbb{P}^1$ has a fixed point $z = 1$, hence ${}^{\alpha}\mathbb{P}^1$ has an F-point. By Châtelet's theorem [Gille and Szamuely 2006, Theorem 5.1.3], a form of $\mathbb{P}^n$ which admits an F-point is trivial (the case $n = 1$ is particularly simple, see [loc. cit., Remark 1.3.5]). We conclude that ${}^{\alpha}\mathbb{P}^1 \cong \mathbb{P}^1$.

(c) Let $T := R^{(1)}_{E/F}(\mathbb{G}_\mathrm{m}) \cong R_{E/F}(\mathbb{G}_\mathrm{m})/\,\mathbb{G}_\mathrm{m}$. The open embedding $\mathbb{G}_\mathrm{m} \hookrightarrow \mathbb{P}^1$, as the complement of $Z := \{0, \infty\}$, is equivariant under the C_2-action on $\mathbb{G}_\mathrm{m}$ and $\mathbb{P}^1$ given by $z \mapsto z^{-1}$. Twisting by α, we obtain by (b) an open embedding of T in $\mathbb{P}^1$ as the complement of ${}^{\alpha}Z$. In particular, T is rational. By Lemma 2.1(a), ${}^{\alpha}Z \cong \operatorname{Spec} E$, so

$$\{T\} = \{\mathbb{P}^1\} - \{{}^{\alpha}Z\} = \mathbb{L} + 1 - \{E\}.$$

Now (c) follows from Lemma 2.3(b).

(d) The first equality holds because $R_{E/F}(\mathbb{G}_{\mathrm{m}})$ is special. Consider the short exact sequence

$$1 \to \mathbb{G}_{\mathrm{m}} \to R_{E/F}(\mathbb{G}_{\mathrm{m}}) \to T \to 1.$$

Since $R_{E/F}(\mathbb{G}_{\mathrm{m}})$ is special, Lemma 2.2 yields

$$\{R_{E/F}(\mathbb{G}_{\mathrm{m}})\} = (\mathbb{L}-1)\{BT\}^{-1},$$

thus (d) follows from (c).

(e) Write $\mathbb{P}^1 = \mathbb{A}^1 \cup \{\infty\}$, and consider the C_2-equivariant decomposition

$$(\mathbb{P}^1)^2 = (\mathbb{A}^1)^2 \amalg (\mathbb{A}^1 \times \{\infty\} \cup \{\infty\} \times \mathbb{A}^1) \amalg \{(\infty, \infty)\}.$$

By Hilbert's Theorem 90 and Lemma 2.1(a), twisting by α gives

$$R_{E/F}(\mathbb{P}^1) = \mathbb{A}^2 \amalg \mathbb{A}^1_E \amalg \operatorname{Spec} F,$$

thus $\{R_{E/F}(\mathbb{P}^1)\} = \mathbb{L}^2 + \{E\}\mathbb{L} + 1$. □

3. The classes of G and BG

Let F be a field of characteristic not 2, and assume that there exists a biquadratic extension

$$K := F(\sqrt{m_1}, \sqrt{m_2})$$

of F. Let

$$E_1 := F(\sqrt{m_1}), \quad E_2 := F(\sqrt{m_2}), \quad E_{12} := F(\sqrt{m_1 m_2}), \quad E := E_1 \times E_2,$$

and let $\Gamma := \mathrm{Gal}(K/F) \cong C_2^2$ be the Galois group of K/F. We define the torus

$$G := R^{(1)}_{E/F}(\mathbb{G}_{\mathrm{m}})$$

and let

$$G' := R_{E/F}(\mathbb{G}_{\mathrm{m}})/\,\mathbb{G}_{\mathrm{m}}$$

be the dual torus of G. By definition, we have a short exact sequence

$$1 \to G \to R_{E/F}(\mathbb{G}_{\mathrm{m}}) \xrightarrow{N} \mathbb{G}_{\mathrm{m}} \to 1, \tag{3.1}$$

where N is the norm homomorphism.

The purpose of this section is the proof of Proposition 3.7, which expresses $\{BG\}$ and $\{G\}$ as rational functions in $\mathbb{L}$, with coefficients classes of étale algebras.

Let σ_1 and σ_2 be generators for Γ such that $E_1 = K^{\sigma_1}$ and $E_2 = K^{\sigma_2}$. Consider the Γ-action on $\mathbb{G}_{\mathrm{m}}^2$, where $\sigma_1(u, v) = (v^{-1}, u^{-1})$ and $\sigma_2(u, v) = (v, u)$, and set

$$T := {}^{\alpha}(\mathbb{G}_{\mathrm{m}}^2), \tag{3.2}$$

where $\alpha \in H^1(F, \Gamma)$ corresponds to the extension K/F.

Lemma 3.3. *We have*

$$\{T\} = \mathbb{L}^2 + (\{E_{12}\} - \{K\})\mathbb{L} + \{K\} - \{E_1\} - \{E_2\} + 1.$$

Proof. The embedding of $\mathbb{G}_{\mathrm{m}}$ in $\mathbb{P}^1$ as the complement of $Z := \{0, \infty\}$ gives an open embedding $\mathbb{G}_{\mathrm{m}}^2 \hookrightarrow (\mathbb{P}^1)^2$ such that the Γ-action on $\mathbb{G}_{\mathrm{m}}^2$ extends to $(\mathbb{P}^1)^2$. By definition

$${}^{\alpha}(\mathbb{P}^1)^2 = ((\mathbb{P}^1)^2 \times \operatorname{Spec} K)/\,\Gamma,$$

where $\Gamma = \langle \sigma_1, \sigma_2 \rangle$ acts diagonally. We first take the quotient by the subgroup $\langle \sigma_1\sigma_2 \rangle$. Since $\sigma_1\sigma_2(u, v) = (u^{-1}, v^{-1})$ and $E_{12} = K^{\sigma_1\sigma_2}$, by Lemma 2.6(b)

$${}^{\alpha}(\mathbb{P}^1)^2 = ((\mathbb{P}^1)^2 \times \operatorname{Spec} E_{12})/C_2,$$

where C_2 acts on $(\mathbb{P}^1)^2$ by switching the two factors. Here we are using the fact that every automorphism of $(\mathbb{P}^1)^2$ must respect the ruling (because it respects the intersection form), and so $\operatorname{Aut}((\mathbb{P}^1)^2) = (\operatorname{Aut}(\mathbb{P}^1))^2 \rtimes C_2$, where C_2 switches the two factors. By Lemma 2.1(b) we deduce that ${}^{\alpha}(\mathbb{P}^1)^2 \cong R_{E_{12}/F}(\mathbb{P}^1)$, so by Lemma 2.6(e)

$$\{{}^{\alpha}(\mathbb{P}^1)^2\} = \mathbb{L}^2 + \{E_{12}\}\mathbb{L} + 1. \tag{3.4}$$

We may partition $(\mathbb{P}^1)^2 \setminus \mathbb{G}_{\mathrm{m}}^2$ into two strata

$$Z_1 := Z \times Z, \quad Z_2 := (Z \times \mathbb{G}_{\mathrm{m}}) \amalg (\mathbb{G}_{\mathrm{m}} \times Z).$$

The Γ-action on Z_1 has two orbits, and Γ acts on Z_2 by transitively permuting the components as the Klein subgroup of S_4. By Lemma 2.1(a), ${}^{\alpha}Z_1 = \operatorname{Spec} E_1 \amalg \operatorname{Spec} E_2$ and ${}^{\alpha}Z_2 = \mathbb{G}_{\mathrm{m}} \times \operatorname{Spec} K$. By (3.4)

$$\begin{aligned}\{T\} &= \{{}^{\alpha}(\mathbb{P}^1)^2\} - \{{}^{\alpha}Z_1\} - \{{}^{\alpha}Z_2\} \\ &= \mathbb{L}^2 + \{E_{12}\}\mathbb{L} + 1 - \{E_1\} - \{E_2\} - \{K\}(\mathbb{L} - 1) \\ &= \mathbb{L}^2 + (\{E_{12}\} - \{K\})\mathbb{L} + \{K\} - \{E_1\} - \{E_2\} + 1. \qquad \square\end{aligned}$$

Proposition 3.5. *There is a short exact sequence of tori*

$$1 \to \mathbb{G}_{\mathrm{m}} \to G \to T \to 1,$$

where T is the torus of (3.2).

Proof. Let P, M and $\mathbb{Z}$ be the character lattices of $R_{E/F}(\mathbb{G}_{\mathrm{m}})$, G and $\mathbb{G}_{\mathrm{m}}$, respectively. We may view P as the Γ-lattice with a basis e_1, e_2, e_3, e_4, such that σ_1 acts by switching e_1 with e_2 and fixing e_3 and e_4, and σ_2 switches e_3 with e_4 and fixes e_1 and e_2. The sequence of Γ-lattices dual to (3.1) identifies M with the cokernel of the Γ-homomorphism $\mathbb{Z} \to P$ given by $1 \mapsto e_1 + e_2 + e_3 + e_4$; denote by $\overline{X}_i \in M$ the projection of e_i. Following Kunyavskiĭ [1987, Section 3, Proposition 1(b)], we consider an exact sequence of Γ-lattices

$$0 \to N \to M \xrightarrow{\pi} \mathbb{Z} \to 0. \tag{3.6}$$

The map π is defined by $\pi\left(\sum a_i \overline{X}_i\right) = a_1 + a_2 - a_3 - a_4$, and $N := \operatorname{Ker}\pi$. A basis for N is given by $v_1 := \overline{X}_1 + \overline{X}_3$ and $v_2 := \overline{X}_1 + \overline{X}_4$. With respect to the basis (v_1, v_2), the Γ-action on N is given by $\sigma_1(a, b) = (-b, -a)$ and $\sigma_2(a, b) = (b, a)$. It is now clear that N is the character lattice of the torus T of (3.2), hence the proof is complete. □

Proposition 3.7. (a) *BG is stably rational.*

(b) $\{BG\}\{G'\} = 1$ *in* $K_0(\mathrm{Stacks}_F)$.

Proof. Consider the sequence

$$1 \to \mathbb{G}_{\mathrm{m}} \to G' \to (R_{E_1/F}(\mathbb{G}_{\mathrm{m}})/\,\mathbb{G}_{\mathrm{m}}) \times (R_{E_2/F}(\mathbb{G}_{\mathrm{m}})/\,\mathbb{G}_{\mathrm{m}}) \to 1, \tag{3.8}$$

which exhibits G' as a $\mathbb{G}_{\mathrm{m}}$-torsor over a rational variety, by Lemma 2.6(c). We deduce that G' is rational, and now (a) and (b) follow from Lemma 2.3. □

Proposition 3.9. *We have*

$$\{G\} = (\mathbb{L} - 1)(\mathbb{L}^2 + (\{E_{12}\} - \{K\})\mathbb{L} + \{K\} - \{E_1\} - \{E_2\} + 1) \tag{3.10}$$

and

$$\{BG\}^{-1} = (\mathbb{L} - 1)(\mathbb{L} - \{E_1\} + 1)(\mathbb{L} - \{E_2\} + 1) \tag{3.11}$$

in $K_0(\mathrm{Stacks}_F)$.

Proof. By Proposition 3.5, G is a $\mathbb{G}_{\mathrm{m}}$-torsor over T. Since $\mathbb{G}_{\mathrm{m}}$ is special, $\{G\} = (\mathbb{L} - 1)\{T\}$. The class of T was determined in Lemma 3.3.

By Proposition 3.7(b), $\{BG\}^{-1} = \{G'\}$. Since $\mathbb{G}_{\mathrm{m}}$ is special, by (3.8),

$$\{G'\} = (\mathbb{L} - 1)\{R^{(1)}_{E_1/F}(\mathbb{G}_{\mathrm{m}})\}\{R^{(1)}_{E_2/F}(\mathbb{G}_{\mathrm{m}})\}.$$

Now (3.11) follows from Lemma 2.6(c). □

4. The refined Euler characteristic

Let F be a field of characteristic zero. Using the computations of the previous section, we will reduce Theorem 1.5(b) to the assertion that a certain polynomial in $\mathbb{L}$ with coefficients motivic classes of étale algebras is a nonzero element of $K_0(\mathrm{Var}_F)$. To prove the assertion, we will use a simplified version of the refined Euler characteristic, introduced by Ekedahl [2009b].

Fix a prime number p, and let $\mathcal{G}$ be a profinite group. The *representation ring* $a_p(\mathcal{G})$ of $\mathcal{G}$ is the Grothendieck ring of continuous $\mathcal{G}$-representations $[M]$ of finite dimension over $\mathbb{F}_p$, subject to the relations $[M \oplus N] = [M] + [N]$. Note that no relations for nonsplit short exact sequences are imposed. The product structure on $a_p(\mathcal{G})$ is given by tensor product of representations. The next observation is well-known when $\mathcal{G}$ is assumed to be finite; see [Benson 1991, Section 5.1].

Lemma 4.1. *As an abelian group,* $a_p(\mathcal{G})$ *is freely generated by the set of isomorphism classes of indecomposable representations.*

Proof. It is clear that $a_p(\mathcal{G})$ is generated by isomorphism classes of indecomposable representations. Assume that $\sum a_i[M_i] - \sum b_j[N_j] = 0$ in $a_p(\mathcal{G})$, for some positive integers a_i, b_j and some pairwise nonisomorphic indecomposable $\mathcal{G}$-representations M_i and N_j.

As a group, $a_p(\mathcal{G})$ is the quotient group F/I, where F is the free abelian group with one generator $\langle P\rangle$ for every isomorphism class of $\mathcal{G}$-representations P, and I is the subgroup generated by all elements of the form $\langle P \oplus Q\rangle - \langle P\rangle - \langle Q\rangle$. It follows that we may find a $\mathcal{G}$-representation X such that

$$(\oplus_i M_i^{\oplus a_i}) \oplus X \cong (\oplus_j N_j^{\oplus b_j}) \oplus X.$$

Let $\mathcal{G}_0$ be a finite quotient of $\mathcal{G}$ such that $\mathcal{G}$ acts on M_i, N_j and X through $\mathcal{G}_0$. Then $M \oplus X \cong N \oplus X$ as $\mathcal{G}_0$-representations, where $M := \oplus_i M_i^{\oplus a_i}$ and $N := \oplus_j N_j^{\oplus b_j}$. By the Krull–Schmidt Theorem applied to the group algebra $\mathbb{F}_p[\mathcal{G}_0]$, this implies $M \cong N$ as $\mathcal{G}_0$-modules, hence as $\mathcal{G}$-modules. This is impossible, because the indecomposable representations M_i and N_j are pairwise nonisomorphic. □

Proposition 4.2. *Let F be a field of characteristic zero, let* $\mathrm{Gal}(F)$ *be the absolute Galois group of F, and let $R_p := a_p(\mathrm{Gal}(F))$. There is a ring homomorphism*

$$\mu : K_0(\mathrm{Var}_F) \to R_p[t]$$

such that for every smooth complete variety X we have $\mu(X) = \sum_i [H^i(\overline{X}_{ét}, \mathbb{F}_p)]t^i$.

Proof. See the proof of [Ekedahl 2009b, Proposition 3.2(i)]. To show that μ is well-defined, one needs to assume that $\operatorname{char} F = 0$ in order to invoke Bittner's presentation of $K_0(\mathrm{Var}_F)$; see [Bittner 2004, Theorem 3.1]. □

5. Proof of Theorem 1.5

Theorem 1.5(a) was proved in Proposition 3.7(a), so we will focus on Theorem 1.5(b). We maintain the notation given at the beginning of Section 3.

Proof of Theorem 1.5(b). Assume by contradiction that $G = R^{(1)}_{E/F}(\mathbb{G}_{\mathrm{m}})$ satisfies (1.3). Then by Proposition 3.9 we have

$$(\mathbb{L}-1)(\mathbb{L}-\{E_1\}+1)(\mathbb{L}-\{E_2\}+1) = (\mathbb{L}-1)(\mathbb{L}^2+(\{E_{12}\}-\{K\})\mathbb{L}+\{K\}-\{E_1\}-\{E_2\}+1)$$

in $K_0(\mathrm{Stacks}_F)$. Since $\mathbb{L}-1$ is invertible in $K_0(\mathrm{Stacks}_F)$, we may divide by $\mathbb{L}-1$ on both sides. Subtracting $\mathbb{L}^2$ on the left and on the right, we arrive at

$$(2-\{E_1\}-\{E_2\})\mathbb{L}+(1-\{E_1\})(1-\{E_2\}) = (\{E_{12}\}-\{K\})\mathbb{L}+\{K\}-\{E_1\}-\{E_2\}+1,$$

that is

$$(\{K\}-\{E_1\}-\{E_2\}-\{E_{12}\}+2)\mathbb{L} = 0$$

in $K_0(\mathrm{Stacks}_F)$.

Recall that $K_0(\mathrm{Stacks}_F)$ is the localization of $K_0(\mathrm{Var}_F)$ at $\mathbb{L}$ and the cyclotomic polynomials in $\mathbb{L}$; see [Ekedahl 2009b, Theorem 1.2]. It follows that

$$(\{K\} - \{E_1\} - \{E_2\} - \{E_{12}\} + 2) f(\mathbb{L}) = 0 \tag{5.1}$$

in $K_0(\mathrm{Var}_F)$, where $f(x) \in \mathbb{Z}[x]$ is a monic polynomial of some degree n.

In order to obtain a contradiction, we now want to apply the homomorphism μ of Proposition 4.2, with respect to the prime $p = 2$. If L/F is an étale algebra of degree n, $\mu(\{L\})$ consists of the permutation representation of $\mathrm{Gal}(F)$ associated to L, concentrated in degree 0. Since we have chosen $p = 2$, $\mu(\{\mathbb{P}^1\})$ consists of one copy of the trivial representation in degree 0 and 2 (in the case $p > 2$ one would need a Tate twist in degree 2). Since $\mathbb{L} = \{\mathbb{P}^1\} - 1$, we deduce that $\mu(\mathbb{L}) = t^2$, and hence $\mu(f(\mathbb{L})) = f(t^2)$.

If X is a finite $\mathrm{Gal}(F)$-set, we denote by $\mathbb{F}_2[X]$ the permutation representation over $\mathbb{F}_2$ associated to X. Recall from Section 3 that we denote $\mathrm{Gal}(K/F)$ by $\Gamma = \langle \sigma_1, \sigma_2 \rangle$. Applying μ to (5.1) and looking at degree $2n$, we obtain

$$[\mathbb{F}_2[\Gamma]] - [\mathbb{F}_2[\Gamma/\langle\sigma_1\rangle]] - [\mathbb{F}_2[\Gamma/\langle\sigma_2\rangle]] - [\mathbb{F}_2[\Gamma/\langle\sigma_{12}\rangle]] + 2[\mathbb{F}_2] = 0$$

in R_2. This is a nontrivial relation of linear dependence in R_2 among classes of indecomposable representations. This is in contradiction with Lemma 4.1, hence $\{BG\} \neq \{G\}^{-1}$, as desired. □

Remark 5.2. By [Voskresensky 1998, Section 4.9, Example 7] every torus of rank 2 is rational, so by Proposition 3.5 the torus G is rational. By Lemma 2.3, BG' is stably rational and $\{BG'\} = \{G\}^{-1}$. By Proposition 3.7(b) we have $\{BG\} = \{G'\}^{-1}$, so $\{BG'\}\{G'\} = \{BG\}^{-1}\{G\}^{-1}$. Since $\{BG\}\{G\} \neq 1$, the conclusions of Theorem 1.5(a) and (b) hold for G' as well.

6. Proof of Theorem 1.6

We maintain the notation of Section 3.

Proof of Theorem 1.6. Let $\Gamma := \mathrm{Gal}(K/F)$, let M be the character lattice of G, so that $M/2M$ is the character module of A, and let P be the character lattice of $R_{E/F}(\mathbb{G}_\mathrm{m})$. As in the proof of Proposition 3.5, we view P as the lattice freely generated by e_1, e_2, e_3, e_4, such that σ_1 acts by switching e_1 with e_2, and σ_2 by switching e_3 with e_4. Using (3.1), we may construct a commutative diagram of Γ-modules

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Z} & \longrightarrow & P & \longrightarrow & M & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle\iota} & & \| & & \downarrow & & \\ 0 & \longrightarrow & N & \longrightarrow & P & \xrightarrow{\varphi} & M/2M & \longrightarrow & 0. \end{array} \tag{6.1}$$

with exact rows. Here $\mathbb{Z}$ denotes the trivial one-dimensional Γ-lattice, $\iota(1) := e_1 + e_2 + e_3 + e_4$, and N is the kernel of φ, that is,

$$N = \left\{ \sum_{i=1}^{4} a_i e_i : a_1 \equiv a_2 \equiv a_3 \equiv a_4 \pmod 2 \right\}.$$

Applying the snake lemma to (6.1), we obtain a short exact sequence

$$0 \to \mathbb{Z} \xrightarrow{\iota} N \to M \to 0.$$

Define $\pi : N \to \mathbb{Z}$ by sending $\sum a_i e_i$ to $(a_1 + a_2)/2$. Then π is a Γ-homomorphism and ι is a section of π. Therefore, we have an isomorphism $N \cong \mathbb{Z} \oplus M$.

Let S be an F-torus with character lattice N. Since $N \cong \mathbb{Z} \oplus M$, we have $S \cong \mathbb{G}_{\mathrm{m}} \times G$. Thus, the bottom row of (6.1) corresponds to the short exact sequence of group schemes

$$1 \to A \to R_{E/F}(\mathbb{G}_{\mathrm{m}}) \to \mathbb{G}_{\mathrm{m}} \times G \to 1.$$

By Lemma 2.2, we have $\{BA\} = \{\mathbb{G}_{\mathrm{m}}\}\{G\}/\{R_{E/F}(\mathbb{G}_{\mathrm{m}})\}$. Applying Lemma 2.2 to (3.1), we see that $\{BG\} = \{\mathbb{G}_{\mathrm{m}}\}/\{R_{E/F}(\mathbb{G}_{\mathrm{m}})\}$. Therefore, $\{BA\} = \{BG\}\{G\}$. By Theorem 1.6 we have $\{BG\} \neq \{G\}^{-1}$, hence $\{BA\} \neq 1$, as desired. □

Acknowledgments

I would like to thank my advisor Zinovy Reichstein for his guidance and for greatly improving the exposition, Mattia Talpo and Angelo Vistoli for helpful comments, and Boris Kunyavskiĭ for sending me a copy of his paper [Kunyavski 1987]. I am very grateful to the anonymous referee for finding a mistake in a previous version of the proof of Theorem 1.5, and for suggesting a fix.

References

[Behrend and Dhillon 2007] K. Behrend and A. Dhillon, "On the motivic class of the stack of bundles", *Adv. Math.* **212**:2 (2007), 617–644. MR Zbl

[Benson 1991] D. J. Benson, *Representations and cohomology, II: Cohomology of groups and modules*, Cambridge Stud. Adv. Math. **31**, Cambridge Univ. Press, 1991. MR Zbl

[Bergh 2016] D. Bergh, "Motivic classes of some classifying stacks", *J. Lond. Math. Soc.* (2) **93**:1 (2016), 219–243. MR Zbl

[Bittner 2004] F. Bittner, "The universal Euler characteristic for varieties of characteristic zero", *Compos. Math.* **140**:4 (2004), 1011–1032. MR Zbl

[Dhillon and Young 2016] A. Dhillon and M. B. Young, "The motive of the classifying stack of the orthogonal group", *Michigan Math. J.* **65**:1 (2016), 189–197. MR Zbl

[Ekedahl 2009a] T. Ekedahl, "A geometric invariant of a finite group", preprint, 2009. arXiv

[Ekedahl 2009b] T. Ekedahl, "The Grothendieck group of algebraic stacks", preprint, 2009. arXiv

[Favi and Florence 2008] G. Favi and M. Florence, "Tori and essential dimension", *J. Algebra* **319**:9 (2008), 3885–3900. MR Zbl

[Florence 2008] M. Florence, "On the essential dimension of cyclic p-groups", *Invent. Math.* **171**:1 (2008), 175–189. MR Zbl

[Florence and Reichstein 2018] M. Florence and Z. Reichstein, "The rationality problem for forms of $\overline{M}_{0,n}$", *Bull. Lond. Math. Soc.* **50**:1 (2018), 148–158. MR Zbl

[Florence et al. 2017] M. Florence, N. Hoffmann, and Z. Reichstein, "On the rationality problem for forms of moduli spaces of stable marked curves of positive genus", preprint, 2017. arXiv

[Gille and Szamuely 2006] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Stud. Adv. Math. **101**, Cambridge Univ. Press, 2006. MR Zbl

[Joyce 2007] D. Joyce, "Motivic invariants of Artin stacks and 'stack functions'", *Q. J. Math.* **58**:3 (2007), 345–392. MR Zbl

[Kunyavski 1987] B. È. Kunyavskiĭ, "Three-dimensional algebraic tori", pp. 90–111 in *Investigations in number theory*, edited by N. G. Chudakov, Saratov. Gos. Univ., 1987. In Russian; translated in *Selecta Math. Soviet.* **9**:1 (1990), 1–21. MR Zbl

[Martino 2016] I. Martino, "The Ekedahl invariants for finite groups", *J. Pure Appl. Algebra* **220**:4 (2016), 1294–1309. MR Zbl

[Martino 2017] I. Martino, "Introduction to the Ekedahl invariants", *Math. Scand.* **120**:2 (2017), 211–224. MR Zbl

[Pirisi and Talpo 2019] R. Pirisi and M. Talpo, "On the motivic class of the classifying stack of G_2 and the spin groups", *Int. Math. Res. Not.* **2019**:10 (2019), 3265–3298. MR Zbl

[Serre 1958] J.-P. Serre, "Espaces fibrés algébriques", exposé 1 in *Anneaux de Chow et applications*, Séminaire Claude Chevalley **3**, Sécretariat Math., Paris, 1958.

[Talpo and Vistoli 2017] M. Talpo and A. Vistoli, "The motivic class of the classifying stack of the special orthogonal group", *Bull. Lond. Math. Soc.* **49**:5 (2017), 818–823. MR Zbl

[Toën 2005] B. Toën, "Grothendieck rings of Artin n-stacks", preprint, 2005. arXiv

[Voskresensky 1998] V. E. Voskresenskiĭ, *Algebraic groups and their birational invariants*, Transl. Math. Monogr. **179**, Amer. Math. Soc., Providence, RI, 1998. MR Zbl

Communicated by Martin Olsson
Received 2018-08-07 Revised 2019-07-16 Accepted 2019-12-19

scavia@math.ubc.ca *University of British Columbia, Vancouver BC, Canada*

ALGEBRA AND NUMBER THEORY 14:4 (2020)
https://doi.org/10.2140/ant.2020.14.867

A representation theory approach to integral moments of L-functions over function fields

Will Sawin

We propose a new heuristic approach to integral moments of L-functions over function fields, which we demonstrate in the case of Dirichlet characters ramified at one place (the function field analogue of the moments of the Riemann zeta function, where we think of the character n^{it} as ramified at the infinite place). We represent the moment as a sum of traces of Frobenius on cohomology groups associated to irreducible representations. Conditional on a hypothesis on the vanishing of some of these cohomology groups, we calculate the moments of the L-function and they match the predictions of the Conry, Farmer, Keating, Rubinstein, and Snaith recipe (*Proc. Lond. Math. Soc.* (3) **91** (2005), 33–104).

In this case, the decomposition into irreducible representations seems to separate the main term and error term, which are mixed together in the long sums obtained from the approximate functional equation, even when it is dyadically decomposed. This makes our heuristic statement relatively simple, once the geometric background is set up. We hope that this will clarify the situation in more difficult cases like the L-functions of quadratic Dirichlet characters to squarefree modulus. There is also some hope for a geometric proof of this cohomological hypothesis, which would resolve the moment problem for these L-functions in the large degree limit over function fields.

1. Introduction

The heuristics of Conrey, Farmer, Keating, Rubinstein and Snaith [Conrey et al. 2005] give precise conjectures for the distribution of special values of L–functions in certain families. They were extended to function fields in [Andrade and Keating 2014]. Certain constants appearing in these predictions can be related to statistics of random matrices.

While these are conjectures in general, they are known for many families up to an error term of $O(1/\sqrt{q})$ in the function field setting (e.g., [Katz and Sarnak 1999; Katz 2013; 2015]). This error term hides everything but the random matrix term. However, the random matrix term appears in a particularly natural way. In the function field setting, the L-functions are equal to characteristic polynomials of the matrices giving the action of Frobenius elements on a certain Galois representation, and these matrices are random in a precise technical sense [Deligne 1980, Theorem 3.5.3].

We are not able today to prove the full conjecture of [Conrey et al. 2005] over function fields for any family of L-functions using the geometric approach initiated by [Katz and Sarnak 1999]. However, we propose a middle ground. Using the machinery of étale cohomology, and in particular the interpretation

MSC2010: primary 11M50; secondary 11M38, 11T55, 14F20, 20G05.
Keywords: L-functions, moments, function fields, polynomials, finite fields, integral moments.

of L-functions via representations of monodromy groups, we will describe a new heuristic which matches the predictions of [Conrey et al. 2005]. However, while the heuristics of [loc. cit.] require multiple manipulations, that do not make sense on their own, we will make a single assumption on vanishing of cohomology groups, which could well be true. This assumption also makes predictions for other problems, such as the variance of the divisor function in short intervals.

In this paper, we describe this heuristic, and verify its relationship to [loc. cit.], only for the "short interval" family of characters.

Definition 1.1. For n a natural number and $\mathbb{F}_q$ a finite field, consider a primitive even Dirichlet character $\psi : (\mathbb{F}_q[x]/x^{n+1})^\times \to \mathbb{C}^\times$. Here "primitive" means that the character is nontrivial on elements congruent to 1 mod x^n, and "even" means that it is trivial on $\mathbb{F}_q^\times$. Define a function χ on monic polynomials in $\mathbb{F}_q[T]$ by, for f monic of degree d,

$$\chi(f) = \psi(f(x^{-1})x^d).$$

It is easy to see that χ depends only on the $n+1$ leading terms of f. Let $S_{n,q}$ be the set of functions χ arising from primitive even Dirichlet characters ψ in this way. Because there are q^n even Dirichlet characters of which q^{n-1} are imprimitive, this set has cardinality $q^n - q^{n-1}$.

For $\chi \in S_{n,q}$, form the associated L-functions

$$L(s, \chi) = \sum_{\substack{f \in \mathbb{F}_q[T] \\ \text{monic}}} \chi(f)|f|^{-s}$$

where $|f| = q^{\deg f}$, with functional equation

$$L(s, \chi) = \epsilon_\chi q^{(n-1)(1/2-s)} L(1-s, \bar{\chi})$$

for a unique $\epsilon_\chi \in \mathbb{C}$.

Let $\text{Prim}_n = \mathbb{A}^n - \mathbb{A}^{n-1}$. Katz [2013, Sections 2 and 3] defined an explicit bijection between $\text{Prim}_n(\mathbb{F}_q)$ and $S_{n,q}$. We will reproduce the precise formula in Definition 2.1 below, but as the details are not relevant to the big picture, we will leave it as a black box here.

Definition 1.2. Let L_{univ} be the unique lisse sheaf of rank $n-1$ on Prim_n such that for a point $y \in \text{Prim}_n(\mathbb{F}_q)$ corresponding to a character χ under the correspondence of Definition 2.1, we have the identity

$$\det(1 - q^{-s}\,\text{Frob}_q, L_{\text{univ},y}) = L(s, \chi) \tag{1-1}$$

between the characteristic polynomial of Frobenius acting on the stalk of L_{univ} at y and the L-function of χ. Katz [2013, Lemma 4.1] proves the existence of this lisse sheaf by an explicit construction.

Hypothesis 1.3. *Let $n, r, \tilde{r}, w$ be natural numbers with $0 \le w \le n$.*

Let $\mathcal{F}$ be an irreducible lisse $\mathbb{Q}_\ell$-sheaf on $\text{Prim}_{n,\bar{\mathbb{F}}_q}$ that appears as a summand of

$$\det(L_{\text{univ}})^{-\tilde{r}} \otimes \bigotimes_{i=1}^{r+\tilde{r}} \wedge^{d_i} (L_{\text{univ}})$$

for some $0 \leq d_1, \ldots, d_{r+\tilde{r}} \leq n-1$, *but which does not appear as a geometric summand of* $L_{\mathrm{univ}}^{\otimes a} \otimes L_{\mathrm{univ}}^{\vee \otimes b}$ *for* $0 \leq a, b \leq n-1$.

We say that Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$ *is satisfied if, for all such* $\mathcal{F}$,

$$H_c^j(\mathrm{Prim}_{n,\bar{\mathbb{F}}_q}, \mathcal{F}) = 0$$

for all $j > n + w$.

Theorem 1.4. *Let* $n, r, \tilde{r}, w$ *be natural numbers with* $0 \leq w \leq n$ *and* $\mathbb{F}_q$ *a finite field. Assume that Hypothesis* $\mathrm{H}(n, r, \tilde{r}, w)$ *is satisfied. Assume also that* $n > 2\max(r, \tilde{r}) + 1$ *and if* $n = 4$ *or* 5 *that the characteristic of* $\mathbb{F}_q$ *is not* 2. *Let* $C_{r,\tilde{r}} = (2+\max(r, \tilde{r}))^{\max(r,\tilde{r})+1}$. *Let* $\alpha_1, \ldots, \alpha_{r+\tilde{r}}$ *be imaginary numbers. Let* ϵ_χ *be the* ϵ*-factor of* $L(s, \chi)$. *Then*

$$\frac{1}{q^n - q^{n-1}} \sum_{\chi \in S_{n,q}} \epsilon_\chi^{-\tilde{r}} \prod_{i=1}^{r+\tilde{r}} L\left(\tfrac{1}{2} - \alpha_i, \chi\right) \tag{1-2}$$

$$= \sum_{\substack{S \subseteq \{1,\ldots,r+\tilde{r}\} \\ |S| = \tilde{r}}} \prod_{i \notin S} q^{\alpha_i(n-1)} \sum_{\substack{f_1,\ldots,f_{r+\tilde{r}} \in \mathbb{F}_q[T] \\ \text{monic} \\ \prod_{i\in S} f_i / \prod_{i \notin S} f_i \in T^{\mathbb{Z}}}} \prod_{i \in S} |f_i|^{-1/2+\alpha_i} \prod_{i \notin S} |f_i|^{-1/2-\alpha_i} \tag{1-3}$$

$$+ O(q^{(w-n)/2} C_{r,\tilde{r}}^n n^{r+\tilde{r}}). \tag{1-4}$$

Remark 1.5. The parameters $n, r, \tilde{r}$ of Hypothesis 1.3 and Theorem 1.4 bear a clear relationship to the moment (1-2) — indeed, n determines the conductor of the characters while r and $\tilde{r}$ determine the powers we raise the L-function and its ϵ-factor to. The meaning of the parameter w is less clear, and so we explain it here.

In Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$, w determines the width of the region where we do not assume that the cohomology groups vanish. By Poincaré duality and Artin's affine theorem, $H_c^j(\mathrm{Prim}_n, \mathcal{F}) = 0$ for $j < n$, so under Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$, the only possible nonvanishing cohomology groups occur when j ranges from n to $n + w$. In particular, the larger w is, the weaker an assumption Hypothesis 1.3 is.

In Theorem 1.4, w does not affect the moment (1-2) nor the main term (1-3), but only the bound for the error term (1-4). In particular, the larger w is, the larger, and thus weaker, the bound (1-4).

Thus we can get by with a weaker geometric hypothesis, at the cost of a weaker analytic result. Depending on our purposes we can use the parameter w in two ways — either proving a cohomology vanishing statement, finding the least value of w for which it implies Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$, and deducing the corresponding bound, or determining a desired bound and finding the greatest value of w for which Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$ is sufficient to prove it.

Remark 1.6. It follows from Theorem 1.4 that the main term (1-3) is finite. In other words the sum of meromorphic functions is in fact a holomorphic function on some neighborhood of the locus where $\alpha_1, \ldots, \alpha_{r+\tilde{r}}$ are imaginary. We can also show this more directly.

The sum (1-3) is manifestly symmetric in the variables $\alpha_1, \dots, \alpha_{r+\tilde{r}}$, because each term is invariant under permuting the variables in S and the variables outside S, and summing over the possibilities for S makes it invariant under permuting all the variables. So if we multiply (1-3) by the Vandermonde determinant $\prod_{1\le i_1<i_2\le r+\tilde{r}}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})$, it becomes antisymmetric in the variables. When we multiply each individual term by the Vandermonde determinant, or even its factor $\prod_{i_1\in S, i_2\notin S}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})$, they become holomorphic (Lemma 3.4) — in fact we can express them by a convergent Euler product. So the sum times the Vandermonde is holomorphic. But any antisymmetric holomorphic function of q^{α_i} vanishes whenever $q^{\alpha_{i_1}}=q^{\alpha_{i_2}}$ for any i_1, i_2 and thus is divisible by the Vandermonde determinant, so (1-3) itself is holomorphic.

Lemma 3.7, below, clarifies some of the properties of the main term (1-3). It shows that (1-3) is a nonzero polynomial in n if $\alpha_1=\dots=\alpha_{r+\tilde{r}}=0$ and is a (nonzero) quasiperiodic function of n if $\alpha_1, \dots, \alpha_{r+\tilde{r}}$ are all distinct. In the polynomial case, the main term dominates the error term as long as $w<(1-\log_q C_{r,\tilde{r}}-\epsilon)n$ for any $\epsilon>0$, as then the error term decays exponentially with n. Similarly, in the quasiperiodic case, the main term will dominate the error term for most n as long as $w<(1-\log_q C_{r,\tilde{r}}-\epsilon)n$ for any $\epsilon>0$. More information about the main term, including the calculation of the leading term of this polynomial, is contained in Lemma 3.7.

Remark 1.7. (1) We explain why (1-3) is indeed the prediction of the CFKRS recipe [Conrey et al. 2005, Section 4] (or [Andrade and Keating 2014, Section 4.2]) for this family. This is obtained by the 5-step process (1) start with a product of shifted L-functions, (2) apply the "approximate" functional equation to each term (in the function field case, an exact formula, following from polynomiality and the usual functional equation), (3) average the sign of the functional equations, (4) replace each summand by its expected value when averaged over the family, (5) extend the sums by removing limits of summation.

In step (1) we start with (1-2). In step (2) we apply the "approximate" functional equation

$$L\left(\tfrac{1}{2}-\alpha_i, \chi\right)=\sum_{\substack{f\in\mathbb{F}_q[T]\\ \text{monic}\\ \deg f\le(n-1)/2}}\chi(f)|f|^{-1/2-\alpha_i}+\epsilon_\chi q^{-\alpha_i(n-1)}\sum_{\substack{f\in\mathbb{F}_q[T]\\ \text{monic}\\ \deg f<(n-1)/2}}\bar{\chi}(f)|f|^{-1/2+\alpha_i}$$

to obtain that (1-2) is

$$\sum_{S\subseteq\{1,\dots,r+\tilde{r}\}}\epsilon_\chi^{r-|S|}\prod_{i\notin S}q^{-\alpha_i(n-1)}\sum_{\substack{f_1,\dots,f_{r+\tilde{r}}\in\mathbb{F}_q[T]\\ \text{monic}\\ \deg f_i\le\frac{1}{2}(n-1-1_{i\notin S})}}\chi\left(\prod_{i\in S}f_i\right)\bar{\chi}\left(\prod_{i\notin S}f_i\right)\prod_{i\in S}|f_i|^{-1/2-\alpha_i}\prod_{i\notin S}|f_i|^{-1/2+\alpha_i}.$$

In step (3) we remove the terms where $r\ne|S|$, as the average of the root number $\epsilon_\chi^{r-|S|}$ cancels there. (We show it cancels as part of Lemma 2.10.)

In step (4) we observe that for n sufficiently large, the average over $\chi\in S_{n,q}$ of $\chi\left(\prod_{i\in S}f_i\right)\bar{\chi}\left(\prod_{i\notin S}f_i\right)$ vanishes unless $\prod_{i\in S}f_i/\prod_{i\notin S}f_i\in T^{\mathbb{Z}}$.

In step (5) we extend the sums by removing the degree condition, getting

$$\sum_{\substack{S\subseteq\{1,\dots,r+\tilde r\}\\ |S|=r}} \prod_{i\notin S} q^{-\alpha_i(n-1)} \sum_{\substack{f_1,\dots,f_{r+\tilde r}\in\mathbb{F}_q[T] \\ \text{monic} \\ \prod_{i\in S} f_i/\prod_{i\notin S} f_i \in T^{\mathbb{Z}}}} \prod_{i\in S} |f_i|^{-1/2-\alpha_i} \prod_{i\notin S} |f_i|^{-1/2+\alpha_i}.$$

(2) We can write the error term in Theorem 1.4 as

$$O((q^n)^{-1/2+w/n+((\max(r,\tilde r)+1)\log(\max(r,\tilde r)+2))/\log q+\epsilon}).$$

The error term predicted by [Conrey et al. 2005] is always the size of the family raised to the power $-\frac{1}{2}+\epsilon$. Our exponent approaches the predicted square-root cancellation as long as $w/n \to 0$ and $(\max(r,\tilde r)+1)\log(\max(r,\tilde r)+2)/\log q \to 0$.

In fact, we are able to verify some nontrivial cases of Hypothesis 1.3 using results from [Sawin 2018]. More precisely, we see in Lemma 5.3 that when $\mathbb{F}_q$ is a field of characteristic p, then Hypothesis $\mathrm{H}(n, r, 1, n+1-(p-2r)/(pr)\cdot n)$ is satisfied for any n, r. This gives the following unconditional estimate:

Corollary 1.8 (Corollary 5.4). *Let n, r be natural numbers and $\mathbb{F}_q$ a finite field of characteristic p. Assume also that $n > 2r+1$ and if $n = 4$ or 5 that the characteristic of $\mathbb{F}_q$ is not 2. Let $C_{r,1} = (2+r)^{r+1}$. Let $\alpha_1, \dots, \alpha_{r+1}$ be imaginary numbers. Let ϵ_χ be the ϵ-factor of $L(\chi)$. Then*

$$\frac{1}{(q^n - q^{n-1})} \sum_{\chi\in S_{n,q}} \epsilon_\chi^{-1} \prod_{i=1}^{r+1} L\bigl(\tfrac{1}{2}-\alpha_i, \chi\bigr)$$
$$= \sum_{j=1}^{r+1} q^{-\alpha_j(n-1)} \biggl(\frac{1}{1-q^{-1/2-\alpha_j}} \prod_{i\neq j} \frac{1-q^{-1+\alpha_i-\alpha_j}}{(1-q^{-1/2+\alpha_i})(1-q^{\alpha_i-\alpha_j})} \biggr) + O\Bigl(\sqrt{q}\,(q^{-(p-2r)/(2pr)} C_{r,1})^n n^{r+1}\Bigr).$$

As p goes to ∞ with fixed r, this bound converges to a power savings of $\frac{1}{2}r$.

The main term in the case $r = 2, \tilde r = 2$ is also given by an explicit, though complicated, rational function in q^{α_i} and q, see Lemma 3.8.

We present a summary of some of the key ideas in the proof of Theorem 1.4.

First note that one subtlety in Theorem 1.4 is that (1-3) is a sum of terms that have poles at the points we are most interested in studying, and that the poles only disappear when we sum all the terms. This makes it tricky to try to prove the theorem by splitting up the terms, as this could introduce infinities. As noted in Remark 1.6, we can remove the poles by multiplying by a suitable Vandermonde determinant, and this helps to prove that the sum is holomorphic. The first step in our proof is a purely algebraic description of what happens when we multiply (1-2) by the same Vandermonde determinant. We show that the coefficients of monomials in $q^{\alpha_1}, \dots, q^{\alpha_{r+\tilde r}}$ in this product will be averages over χ of Schur functions in the zeroes of $L(s, \chi)$ corresponding to irreducible representations of GL_{n-1} (Lemma 2.7). Before multiplying by the Vandermonde determinant, the coefficients were typically characters of highly reducible representations, so the Vandermonde determinant significantly simplifies (1-2) as well.

These coefficients will be crucial to proving Theorem 1.4. Both (1-2) and (1-3) can be expressed as Laurent series in the variables $q^{\alpha_1}, \dots, q^{\alpha_{r+\tilde{r}}}$, and they remain Laurent series when each is multiplied by the Vandermonde determinant. We will prove Theorem 1.4 by showing that the coefficient of each individual monomial $q^{-\sum_i \alpha_i d_i}$ in (1-2) times the Vandermonde is equal to the coefficient of the same monomial in (1-3) times the Vandermonde, up to a controlled error. We think of this strategy as being analogous to, in classical moment calculations, breaking a sum over integers into dyadic intervals and handling them separately, or in function field moments calculation breaking a sum over polynomials into many sums over polynomials of fixed degree.

Typically in moment estimates, we break up the sum into distinct ranges, and there will be some ranges in which we can show that the off-diagonal terms cancel using only orthogonality of characters. The tuples $(d_1, \dots, d_{r+\tilde{r}})$ for which orthogonality of characters is enough are described by Lemma 3.6. Using this fact, we match some of the terms from (1-2) with some of the terms to (1-3). We will then individually bound all the unmatched terms.

The unmatched terms in (1-2) turn out to be averages over χ of certain Schur functions in the zeroes of $L(s, \chi)$ — specifically, those corresponding to irreducible representations of GL_{n-1} that do not appear as a summand of $L_{\mathrm{univ}}^{\otimes a} \otimes L_{\mathrm{univ}}^{\vee \otimes b}$ for any $0 \le a, b \le n-1$. These are exactly the representations discussed in our Hypothesis 1.3. Using this hypothesis, combined with the Grothendieck–Lefschetz fixed point formula and some Betti number estimates (Lemma 2.13), we show these terms are small.

We bound the unmatched terms in (1-3) using estimates from Lemma 3.5. These estimates are proved by expressing the diagonal term times the Vandermonde determinant as an Euler product, controlling each term of the Euler product, deducing a region of holomorphicity for the Euler product and an upper bound near the boundary of that region, then using a contour integral to show the coefficients decay as we get further from the corner. This is in contrast to the situation before multiplying by the Vandermonde determinant, where the coefficients of many different monomials in the diagonal term are large.

Remark 1.9. We present some remarks on the hypothesis, with the first two from an analytic perspective and the remainder from a geometric perspective:

(1) To obtain predictions for moments, instead of Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$, we could make a purely analytic conjecture of square-root cancellation in the trace of the cohomology (equivalently, the sum of the Schur polynomial associated to this representation, evaluated at the roots of the L-function, over all primitive Dirichlet characters) for representations outside this special set.

Such a hypothesis is essentially equivalent to a uniform version of the conjecture of [Conrey et al. 2005] for shifted moments, as we can extract these individual coefficients by a Fourier series after multiplying by the Vandermonde determinant. However, Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$ would not follow directly from this unless the cohomology groups were proven to be pure.

If made uniform in $r, \tilde{r}$, such a hypothesis would imply conjectures for ratios and tuple correlations — presumably matching the predictions of [Conrey et al. 2008], and therefore [Conrey and Snaith 2007]. On the other hand, while Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$ can be stated uniformly in $r, \tilde{r}$, it would not imply a good estimate on the error term in the degree aspect unless stronger Betti number bounds than those in Section 2C were proven.

Despite these difficulties, we have stated Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$ in a geometric way to motivate it as a natural statement (we would not have come up with it if it weren't for geometry) and to suggest the potential of a geometric proof.

(2) We can view the averages of Schur polynomials of the roots of the L-function that appear in the analytic version of the hypothesis (equivalently, the functions $F(V)$ discussed below in Section 2) as being a distant analogue of the exponential sums considered in the circle method, as they are the averages of characters of irreducible representations of GL_{n-1} over an arithmetically natural finite set, while the exponential sums in the circle method are the averages of characters of irreducible representations of $\mathbb{Z}$ over an arithmetically natural finite set. In this case, the averages of characters of irreducible representations appearing in $\mathrm{std}^{\otimes a} \otimes \mathrm{std}^{\vee \otimes b}$ for $0 \leq a, b \leq n-1$ are the analogue of the major arcs, which we can calculate reasonably explicitly. The averages of characters of other irreducible representations are the analogue of the minor arcs, which we hope to prove cancel.

(3) As part of our proof, we will implicitly calculate the trace of Frobenius on the cohomology of sheaves $\mathcal{F}$ which do appear as a summand of $L_{\mathrm{univ}}^{\otimes a} \otimes L_{\mathrm{univ}}^{\vee \otimes b}$ for $0 \leq a, b \leq n-1$. So our hypothesis is a version of the usual heuristic that what we cannot calculate should cancel. Of course, such a heuristic may be overly optimistic. Instead, what is interesting here is that it is a very straightforward and geometrically natural heuristic.

(4) The calculations of the traces for the sheaves which do appear as a summand of $L_{\mathrm{univ}}^{\otimes a} \otimes L_{\mathrm{univ}}^{\vee \otimes b}$ for $0 \leq a, b \leq n-1$ are closely related to Katz's calculations [2015, Section 5]. (In the case $N = n < p$, the sheaf $\mathcal{F}$ defined in [loc. cit., Section 4] is the restriction of L_{univ} to a hyperplane section, and essentially the same calculations as in [loc. cit., Section 5] can be done in this setting.) So the failure of square-root cancellation he demonstrates does not cause a problem for us, as it occurs exactly in the cases where we do not assume square-root cancellation. In fact, we show that the non-square-root terms that he observes correspond exactly to the secondary terms predicted by [Conrey et al. 2005].

(5) Hypothesis $\mathrm{H}(n, r, \tilde{r}, n-1)$ is known for every $n, r, \tilde{r}$ by Poincaré duality in étale cohomology. This gives a bound in the q aspect whose error term is $O(q^{-1/2})$. This implies that the main term of Theorem 1.4 must match, to within $O(q^{-1/2})$, the main term obtained by applying Katz's equidistribution result [2013, Theorem 1.2] and performing a matrix integral. Our method in this case is simply a (more complicated) variant of the proof of Deligne's equidistribution theorem [1980, Theorem 3.5.3], which Katz uses in his proof, combined with the calculation of the matrix integral.

(6) It is possible that some very strong form of Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$ could be true. For instance, we could take the parameter w to be uniform in $n, r, \tilde{r}$. This would be equivalent to replacing the first condition on $\mathcal{F}$ by the condition that it appears as a summand of the tensor product of some tensor power of L_{univ} with some tensor power of its dual. Conceivably the uniform constant could be as low as $w = 2$.

However, it is likely to be easier to prove weaker special cases first, which is why we have stated it flexibly using multiple parameters.

Remark 1.10. We present some remarks on possible generalizations. We first discuss families that are harmonic in the sense of [Sarnak et al. 2016], and then geometric families:

(1) We expect that these results can be generalized to at least some families with orthogonal and symplectic symmetry type. The simplest cases for our method are probably the families of Dirichlet characters studied by Katz [2017], where both orthogonal and symplectic examples are given. One simply replaces the Vandermonde determinant with, for the r-th moment in the orthogonal case,

$$\prod_{1 \le i_1 < i_2 \le r} (q^{\alpha_{i_1}} - q^{\alpha_{i_2}})(q^{\alpha_{i_1}} q^{\alpha_{i_2}} - 1) \tag{1-5}$$

or, for the r-th moment in the symplectic case,

$$\prod_{1 \le i \le r} (q^{2\alpha_i} - 1) \prod_{1 \le i_1 < i_2 \le r} (q^{\alpha_{i_1}} - q^{\alpha_{i_2}})(q^{\alpha_{i_1}} q^{\alpha_{i_2}} - 1). \tag{1-6}$$

The hypothesis needed then has to do with the cohomology of sheaves generated from the universal sheaves constructed by Katz in that paper.

These formulas arise from the algebra of the orthogonal and symplectic group respectively, and thus which one to use should depend only on the symmetry type of the L-function. Specifically, they can be calculated by attempting to repeat the proof of Lemma 2.5 in the orthogonal or symplectic case. One starts with the orthogonal or symplectic Jacobi–Trudi identity, which relates the irreducible representations of the group to the determinant of a matrix whose entries are wedge powers of the standard representation [Fulton and Harris 1991, (24.25), Corollary 24.35, Corollary 24.45]. Using this, it is straightforward to express a multivariable power series whose coefficients are irreducible representations of the group as the determinant of a fixed Vandermonde-like matrix times a multivariable power series whose entries are tensor products of wedge powers of the standard representation. This matrix determinant can be evaluated by the product formula (1-5) or (1-6).

Alternately, one can use the identities [Bump and Gamburd 2006, Lemma 4 on page 249 and Lemma 5 on page 257] which express the product of L-functions as a sum of Schur functions of the zeroes associated to irreducible representations of the appropriate symplectic or orthogonal group times Schur functions of the variables q^{α_i} associated to irreducible representations of the "Howe dual" group Sp_{2r} or O_{2r} respectively. The appropriate replacement for the Vandermonde determinant is then the Weyl denominator for the characters of Sp_{2r} or O_{2r}, as appropriate.

(2) Similar results can be proven for moments of an L-function of a fixed Galois representation twisted by a varying Dirichlet character, again conditional on a cohomological hypothesis. However, the dependency on n in the error term may be worse or even ineffective, as Betti number bounds are more difficult in this case. If the Galois representation is an Artin representation splitting over the function field of a curve of bounded degree and genus, it should be possible to make the dependence on n an effective exponential.

(3) For other harmonic families of Dirichlet characters, such as those of squarefree modulus, stating properly an analogous hypothesis seems to require the use of higher-dimensional sheaf convolution Tannakian categories, which have not yet been connected to equidistribution. If that geometric setup is handled, there should not be any major new difficulties. The fourth absolute moment for Dirichlet characters of prime modulus was studied in [Tamam 2014].

(4) For families of automorphic forms on higher-rank groups, the $q \to \infty$ equidistribution theory is not yet available, which is a precondition for our method.

(5) New difficulties present themselves in the family of all quadratic Dirichlet characters with squarefree moduli of a given degree. This family has attracted the most attention in the function field setting, beginning with [Hoffstein and Rosen 1992] and [Andrade and Keating 2012] on the first moment. Recently, improved estimates for the first four moments were obtained in [Florea 2017a; 2017b; 2017c]. Improved estimates on the third moment were obtained in [Diaconu 2019], demonstrating the existence of a secondary term and thereby verifying a prediction from [Diaconu et al. 2003].

The difficulties in applying our method to this case start with the fact that there is no range of short sums where the off-diagonal terms cancel completely. Thus, there is no set of irreducible representations close to the trivial representation in highest weight space whose contributions can be exactly computed. Furthermore, the existence of a secondary term in the cubic case suggests that even for representations very far from the trivial representation in highest weight space, the contribution does not necessarily exhibit square-root cancellation and the term does not vanish above the middle degree. However, neither of these difficulties seems insurmountable, and it is possible that the representation-theoretical and cohomological approach can separate the main term from the secondary terms and shed light, if only conjecturally, on each.

(6) For general geometric families, the situation is likely similar to, but more complicated than, the situation for quadratic Dirichlet characters.

2. Representation theory and algebraic geometry

For any $d \geq 0$, define

$$\lambda_d(\chi) = q^{-d/2} \sum_{\substack{f \text{ monic} \\ \text{degree } d}} \chi(f),$$

so that $L(s,\chi) = \sum_{d=0}^{n-1} \lambda_d(\chi) q^{d(1/2-s)}$. Let $\epsilon_\chi = \lambda_{n-1}(\chi)$ be the ϵ-factor of $L(s,\chi)$, so that $\lambda_{n-1-d}(\chi) = \epsilon_\chi \overline{\lambda_d(\chi)}$. By the Riemann hypothesis or more directly from the explicit formula for ϵ_χ in terms of Gauss sums, we have $|\epsilon_\chi| = 1$ for all χ.

Definition 2.1. We define a map from points of $\mathrm{Prim}_n(\mathbb{F}_q)$ to primitive characters of $(\mathbb{F}_q[x]/x^{n+1})^\times$. In fact, recalling that Prim_n is $\mathbb{A}^n - \mathbb{A}^{n-1}$, we will define a map from $\mathbb{A}^n(\mathbb{F}_q)$ to characters of $(\mathbb{F}_q[x]/x^{n+1})^\times$. We defer to [Katz 2013, Sections 2 and 3] for the proof that this defines a bijection between even characters and points of $\mathbb{A}^n(\mathbb{F}_q)$, and that the primitive ones correspond to exactly the points that do not lie $\mathbb{A}^{n-1}$.

Recall that for a natural number l, the length l Witt vectors $W_l(\mathbb{F}_q)$ are a ring whose elements are l-tuples of elements in $\mathbb{F}_q$, with addition and multiplication defined by the Witt polynomials. For each $m \leq n$ prime to l let $l(m,n)$ be the least natural number such that $p^{l(m,n)}m > n$ and fix an additive character $\psi_m : W_{l(m,n)}(\mathbb{F}_p) = \mathbb{Z}/p^{l(m,n)} \to \mathbb{C}^\times$.

We have the Artin–Hasse exponential power series

$$AH(x) = e^{-\sum_{k=0}^{\infty} x^{p^k}/p^k}$$

whose coefficients are p-adic integers. Given an element in $(\mathbb{F}_q[x]/x^{n+1})^\times$, we can express it uniquely as

$$a_0 \prod_{1 \leq mp^e \leq n, m \text{ prime to } p} AH(a_{mp^e} x^{mp^e})^{1/m} \tag{2-1}$$

for $a_0 \in \mathbb{F}_q^\times$, $a_1, \dots, a_n \in \mathbb{F}_q$. This is because $AH(a_{mp^e} X^{mp^e})^{1/m} = 1 - a_{mp^e} x^{mp^e}/m + \cdots$ and so we can inductively choose each a_{mp^e} to fix the coefficient of the corresponding power of x.

For a tuple $b_1, \dots, b_n$ in $\mathbb{F}_q$ defining a point of $\mathbb{A}^n(\mathbb{F}_q)$, the associated character of $(\mathbb{F}_q[x]/x^{n+1})^\times$ is the one that sends (2-1) to

$$\sum_{m \leq n, m \text{ prime to } p, m \leq n} \psi_m\left(\operatorname{tr}^{W_{l(m,n)}(\mathbb{F}_p)}_{W_{l(m,n)}(\mathbb{F}_q)}((a_m, a_{pm}, \dots, a_{p^{(l(m,n)-1)}m}) \times (b_m, b_{pm}, \dots, b_{p^{(l(m,n)-1)}m}))\right)$$

where the multiplication of tuples denoted by $\times$ is taken in the ring of Witt vectors.

We recall that L_{univ} was defined, using Definition 2.1, in Definition 1.2.

Definition 2.2. Let m be the order of the geometric monodromy group of the determinant of L_{univ}.

Definition 2.3. Let μ be $(-1)^{m(n-1)}$ times the (unique) eigenvalue of Frob_q on the m-th power of the determinant of $L_{\text{univ}}\left(\frac{1}{2}\right)$. The Tate twist, which has the effect of multiplying all eigenvalues of Frobenius on L_{univ} by $q^{-1/2}$, normalizes these eigenvalues to have absolute value 1, so the eigenvalue on the m-th power of the determinant will also have absolute value 1, and thus μ will as well.

Let $R(\text{GL}_{n-1})$ be the representation ring of GL_{n-1} over $\mathbb{Z}$.

We fix throughout an embedding $\iota : \overline{\mathbb{Q}}_\ell \to \mathbb{C}$. Using it, we will abuse notation and identify elements of $\overline{\mathbb{Q}}_\ell$ with their images in $\mathbb{C}$ under ι.

Let F be the unique additive group homomorphism: $R(\text{GL}_{n-1}) \to \mathbb{C}$ whose value on a representation V is

$$F(V) = \sum_{j \in \mathbb{Z}} (-1)^j \operatorname{tr}\left(\text{Frob}_q, H^j_c\left(\text{Prim}_{n,\overline{\mathbb{F}}_q}, V\left(L_{\text{univ}}\left(\tfrac{1}{2}\right)\right)\right)\right).$$

2A. ***L-functions and irreducible representations.*** In this subsection, we relate the moments of L-functions that will be our main object of the study to the functions $F(V)$ for irreducible representations V. It culminates in Lemma 2.7, which expresses the moment of L-functions, times a Vandermonde determinant, as a sum of $F(V)$. To do this, we must first in Lemma 2.4 relate the L-function coefficients to $F(V)$, then in Lemma 2.5 prove an identity in the representation ring that lets us reduce to irreducible representations. The calculation of multiplicities in Lemma 2.6 is proved by the same methods and will be useful later in conjunction with our Betti number bounds.

Lemma 2.4. *For any* $r, \tilde{r}, d = (d_1, \dots, d_{r+\tilde{r}})$,

$$\sum_{\chi \in S_{n,q}} \epsilon_\chi^{-\tilde{r}} \prod_{i=1}^{r+\tilde{r}} \lambda_{d_i}(\chi) = (-1)^{\sum_{i=1}^{r+\tilde{r}} d_i} F\Bigl(\det{}^{-\tilde{r}} \otimes \bigotimes_{i=1}^{r+\tilde{r}} \bigwedge^{d_i}\Bigr).$$

Furthermore, $\epsilon_\chi^m = \mu$ *for all primitive* χ.

Proof. First we observe that, by the Grothendieck–Lefschetz fixed point formula [SGA 4½ 1977, Sommes trig. (1.1.1), page 169]

$$F(V) = \sum_{\chi \in \mathrm{Prim}_n(\mathbb{F}_q)} \mathrm{tr}\bigl(\mathrm{Frob}_{q,\chi}, V\bigl(L_{\mathrm{univ}}\bigl(\tfrac{1}{2}\bigr)\bigr)\bigr).$$

By construction, $\mathrm{Prim}_n(\mathbb{F}_q)$ is in bijection with $S_{n,q}$. It is therefore sufficient to prove that

$$\mathrm{tr}\bigl(\mathrm{Frob}_{q,\chi}, \wedge^d\bigl(L_{\mathrm{univ}}\bigl(\tfrac{1}{2}\bigr)\bigr)\bigr) = (-1)^d \lambda_d(\chi).$$

Because the trace of any matrix on $\wedge^d$ of the standard representation is $(-1)^d$ times the d-th coefficient of the characteristic polynomial of that matrix, this follows from the fact (1-1) that the characteristic polynomial of $\mathrm{Frob}_{q,\chi}$ acting on L_{univ} is $L(s, \chi)$.

A special case is that $\epsilon_\chi = \lambda_{n-1}(\chi) = (-1)^{n-1} \mathrm{tr}\bigl(\mathrm{Frob}_{q,\chi}, \det\bigl(L_{\mathrm{univ}}\bigl(\tfrac{1}{2}\bigr)\bigr)\bigr)$. Thus

$$\epsilon_\chi^m = (-1)^{m(n-1)} \mathrm{tr}\bigl(\mathrm{Frob}_{q,\chi}, \det{}^{\otimes m}\bigl(L_{\mathrm{univ}}\bigl(\tfrac{1}{2}\bigr)\bigr)\bigr) = \mu$$

by the definition of μ. ☐

For $0 \le d_1 \le \cdots \le d_{r+\tilde{r}} \le n-1$, let $V_{d_1,\dots,d_{r+\tilde{r}}}$ be the irreducible representation of GL_{n-1} associated to the conjugate partition of $d_{r+\tilde{r}}, \dots, 1$, in other words the representation whose highest weight character is $\lambda_1^{r+\tilde{r}} \cdots \lambda_{d_1}^{r+\tilde{r}} \lambda_{d_1+1}^{\tilde{r}+\tilde{r}-1} \cdots \lambda_{d_2}^{r+\tilde{r}-1} \cdots \lambda_{d_{r+\tilde{r}}}^1$. Let

$$V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde{r}}} = V_{d_1,\dots,d_{r+\tilde{r}}} \otimes \det{}^{-\tilde{r}},$$

so its highest weight character is $\lambda_1^r \cdots \lambda_{d_1}^r \lambda_{d_1+1}^{r-1} \cdots \lambda_{d_2}^{r-1} \cdots \lambda_{d_r} \lambda_{d_{r+1}+1}^{-1} \cdots \lambda_{n-1}^{-\tilde{r}}$.

Lemma 2.5. *In the ring* $R(\mathrm{GL}_{n-1})[q^{\alpha_1}, \dots, q^{\alpha_{r+\tilde{r}}}]$ *with formal variables* $q^{\alpha_1}, \dots, q^{\alpha_{r+\tilde{r}}}$,

$$\Bigl(\sum_{\sigma \in S_{r+\tilde{r}}} \mathrm{sgn}(\sigma) \prod_{i=1}^{r+\tilde{r}} q^{(\sigma(i)-1)\alpha_i}\Bigr) \sum_{0 \le d_1,\dots,d_{r+\tilde{r}} \le n-1} q^{\sum_{i=1}^{r+\tilde{r}} d_i\alpha_i} (-1)^{\sum_{i=1}^{r+\tilde{r}} d_i + \binom{r+\tilde{r}}{2}} \det{}^{-\tilde{r}} \otimes \bigotimes_{i=1}^{r+\tilde{r}} \wedge^{d_i} \tag{2-2}$$

$$= \sum_{\sigma \in S_{r+\tilde{r}}} \sum_{0 \le d_1 \le \cdots \le d_{r+\tilde{r}} \le n-1} \mathrm{sgn}(\sigma) q^{\sum_{i=1}^{r+\tilde{r}} (d_i+i-1)\alpha_{\sigma(i)}} (-1)^{\sum_{i=1}^{r+\tilde{r}} d_i} V_{d_1,\dots,d_r|d_{\tilde{r}+1},\dots,d_{r+\tilde{r}}}. \tag{2-3}$$

Proof. To check this, observe that (2-2) and (2-3) are antisymmetric in $\alpha_1, \dots, \alpha_{r+\tilde{r}}$. Hence it is sufficient to check that the coefficients of $q^{\sum_{i=1}^{r+\tilde{r}} (d_i+i-1)\alpha_i}$ in (2-2) and (2-3) agree for $d_1 \le d_2 \le \cdots \le d_{r+\tilde{r}}$. Only the trivial permutation contributes to the coefficient of $q^{\sum_{i=1}^{r+\tilde{r}} (d_i+i-1)\alpha_i}$ in (2-3), since the tuple

$(d_1, \dots, d_{r+\tilde r} + r + \tilde r - 1)$ is always in strictly increasing order, but applying any nontrivial permutation to that tuple will produce a tuple not in strictly increasing order. Thus the coefficient of $q^{\sum_{i=1}^{r+\tilde r}(d_i+i-1)\alpha_i}$ in (2-3) is $(-1)^{\sum_{i=1}^{r+\tilde r} d_i + \binom{r+\tilde r}{2}} V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}$.

In (2-2), every permutation σ contributes the amount

$$\operatorname{sgn}(\sigma)(-1)^{\sum_{i=1}^{r+\tilde r} d_i + \binom{r+\tilde r}{2}} \det{}^{-\tilde r} \otimes \bigotimes_{i=1}^{r+\tilde r} \wedge^{d_i+i-1-(\sigma(i)-1)}$$

to this coefficient, so it suffices to check that

$$V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}} = \left(\sum_{\sigma\in S_{r+\tilde r}} \operatorname{sgn}(\sigma) \det{}^{-\tilde r} \otimes \bigotimes_{i=1}^{r+\tilde r} \wedge^{d_i+i-\sigma(i)} \right).$$

Here we interpret wedge powers as vanishing if the power does not lie between 0 and $n-1$.

As $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}} = \det^{-\tilde r} \otimes V_{d_1,\dots,d_{r+\tilde r}}$, it suffices to check that

$$V_{d_1,\dots,d_{r+\tilde r}} = \sum_{\sigma\in S_{r+\tilde r}} \operatorname{sgn}(\sigma) \bigotimes_{i=1}^{r+\tilde r} \wedge^{d_i+i-\sigma(i)}. \tag{2-4}$$

Observe that the right side of (2-4) is the determinant of an $(r+\tilde r)\times(r+\tilde r)$ matrix whose i, j entry is $\wedge^{d_i+i-j}$. By the second Jacobi–Trudi identity for Schur functions [Fulton and Harris 1991, Formula A6], this determinant is equal to $V_{d_1,\dots,d_{r+\tilde r}}$ in the representation ring of GL_{n-1}. □

Lemma 2.6. *The multiplicity that $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}$ appears in $\bigoplus_{0\le e_1,\dots,e_{r+\tilde r}\le n-1} \det^{-s} \otimes \bigotimes_{i=1}^{r+\tilde r} \wedge^{e_i}$ is*

$$\lim_{\alpha_1,\dots,\alpha_r\to 1} \frac{\sum_{\sigma\in S_{r+\tilde r}} \operatorname{sgn}(\sigma) q^{(d_i+i-1)\alpha_{\sigma}(i)}}{\prod_{1\le i_1<i_2\le r+\tilde r} (q^{\alpha_{i_2}} - q^{\alpha_{i_1}})}.$$

Proof. This is obtained from progressively simplifying the formula of Lemma 2.5. First we make the substitution $q^{\alpha_i} \to -q^{\alpha_i}$, which removes some powers of -1. Then we apply the linear map from $R(\mathrm{GL}_{n-1})$ to $\mathbb{Z}$ that sends $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}$ to 1 and every other irreducible representation to 0. This gives

$$\left(\sum_{\sigma\in S_{r+\tilde r}} \operatorname{sgn}(\sigma) \prod_{i=1}^{r+\tilde r} q^{(\sigma(i)-1)\alpha_i} \right) \sum_{0\le d_1,\dots,d_{r+\tilde r}\le n-1} q^{\sum_{i=1}^{r+\tilde r} d_i\alpha_i} \operatorname{mult}\left(V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}, \det{}^{-\tilde r} \otimes \bigotimes_{i=1}^{r+\tilde r} \wedge^{d_i} \right) = \sum_{\sigma\in S_{r+\tilde r}} \operatorname{sgn}(\sigma) q^{\sum_{i=1}^{r+\tilde r}(d_i+i-1)\alpha_{\sigma(i)}}$$

(where $\operatorname{mult}(V, W)$ is the multiplicity of the irreducible representation V in the representation W).

We then divide both sides by $\sum_{\sigma\in S_{r+\tilde r}}\operatorname{sgn}(\sigma)\prod_{i=1}^{r+\tilde r} q^{(\sigma(i)-1)\alpha_i}=\prod_{1\le i_1<i_2\le r+\tilde r}(q^{\alpha_{i_2}}-q^{\alpha_{i_1}})$ and take the limit as $\alpha_1,\dots,\alpha_{r+s}$ go to 1, obtaining

$$\sum_{0\le d_1,\dots,d_{r+\tilde r}\le n-1}\operatorname{mult}\biggl(V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}},\det{}^{-\tilde r}\otimes\bigotimes_{i=1}^{r+\tilde r}\wedge^{d_i}\biggr)=\lim_{\alpha_1,\dots,\alpha_r\to 1}\frac{\sum_{\sigma\in S_{r+\tilde r}}\operatorname{sgn}(\sigma)q^{(d_i+i-1)\alpha_{\sigma}(i)}}{\prod_{1\le i_1<i_2\le r+\tilde r}(q^{\alpha_{i_2}}-q^{\alpha_{i_1}})}$$

and finally observe that the sum of the multiplicity of $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}$ in a sequence of representations is equal to the multiplicity of $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}$ in the direct sum of these representations. □

Lemma 2.7. *We have the identity*

$$\prod_{1\le i_1<i_2\le r+\tilde r}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})\sum_{\chi\in S_{n,q}}\epsilon_\chi^{-s}\prod_{i=1}^{r+\tilde r}L\bigl(\tfrac12-\alpha_i,\chi\bigr)$$
$$=\sum_{\sigma\in S_{r+\tilde r}}\sum_{0\le d_1\le\cdots\le d_{r+\tilde r}\le n-1}\operatorname{sgn}(\sigma)q^{\sum_{i=1}^{r+\tilde r}(d_i+i-1)\alpha_{\sigma(i)}}(-1)^{\sum_{i=1}^{r+\tilde r}d_i}F(V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}})$$

Proof. We have

$$\sum_{\chi\in S_{n,q}}\epsilon_\chi^{-\tilde r}\prod_{i=1}^{r+\tilde r}L\bigl(\tfrac12-\alpha_i,\chi\bigr)=\sum_{\chi\in S_{n,q}}\sum_{0\le d_1,\dots,d_{r+\tilde r}\le n-1}q^{\sum_{i=1}^{r+\tilde r}d_i\alpha_i}(-1)^{\sum_{i=1}^{r+\tilde r}d_i}\epsilon_\chi^{-\tilde r}\prod_{i=1}^{r+\tilde r}\lambda_{d_i}(\chi).$$

Applying Lemma 2.4, this is

$$\sum_{0\le d_1,\dots,d_{r+\tilde r}\le n-1}q^{\sum_{i=1}^{r+\tilde r}d_i\alpha_i}(-1)^{\sum_{i=1}^{r+\tilde r}d_i}F\biggl(\det{}^{-\tilde r}\otimes\bigotimes_{i=1}^{r+\tilde r}\wedge^{d_i}\biggr).$$

Now multiply by the Vandermonde factor

$$\prod_{1\le i_1<i_2\le r+\tilde r}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})=(-1)^{\binom{r+\tilde r}{2}}\sum_{\sigma\in S_{r+\tilde r}}\operatorname{sgn}(\sigma)\prod_{i=1}^{r+\tilde r}q^{\alpha_i(\sigma(i)-1)}$$

to obtain

$$\biggl(\sum_{\sigma\in S_{r+\tilde r}}\operatorname{sgn}(\sigma)\prod_{i=1}^{r+\tilde r}q^{\alpha_i(\sigma(i)-1)}\biggr)\sum_{0\le d_1,\dots,d_{r+\tilde r}\le n-1}q^{\sum_{i=1}^{r+\tilde r}d_i\alpha_i}(-1)^{\sum_{i=1}^{r+\tilde r}d_i+\binom{r+\tilde r}{2}}F\biggl(\det{}^{-\tilde r}\otimes\bigotimes_{i=1}^{r+\tilde r}\wedge^{d_i}\biggr)$$

which by Lemma 2.5 is

$$=\sum_{\sigma\in S_{r+\tilde r}}\sum_{0\le d_1\le\cdots\le d_{r+\tilde r}\le n-1}\operatorname{sgn}(\sigma)q^{\sum_{i=1}^{r+\tilde r}(d_i+i-1)\alpha_{\sigma(i)}}(-1)^{\sum_{i=1}^{r+\tilde r}d_i}F(V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}).$$ □

2B. *Auxiliary results.* In this subsection, we prove two lemmas that will be needed to concretely interpret the conditions on $\mathcal F$ in Hypothesis $\mathrm{H}(n,r,\tilde r,w)$. We describe which $d_1,\dots,d_{r+\tilde r}$ have $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}\bigl(L_{\mathrm{univ}}\bigl(\tfrac12\bigr)\bigr)$ appear as a summand of $L_{\mathrm{univ}}^{\otimes a}\otimes L_{\mathrm{univ}}^{\vee\otimes b}$, which turns out to depend on the constant m, and then we prove a lemma that gives us some control on m.

Lemma 2.8. *For $0 \le d_1 \le \cdots d_{r+\tilde r} \le n-1$ and $0 \le k \le r+\tilde r$, $V_{d_1,\dots,d_k|d_{k+1},\dots,d_{r+\tilde r}}$ appears as a summand $\mathrm{std}^{\otimes a} \otimes \mathrm{std}^{\vee\otimes b}$ for some $0 \le a, b \le n-1$ (possibly depending on $d_1, \dots, d_{r+\tilde r}$) if and only if $\sum_{i=1}^k d_i \le n-1$ and $\sum_{i=k+1}^{r+\tilde r}(n-1-d_i) \le n-1$*

Proof. For the only if direction, consider the element of GL_{n-1} depending on a real parameter $\lambda > 1$ whose eigenvalues are λ with multiplicity d_k and 1 with multiplicity $n-1-d_k$. Its eigenvalue on the highest weight vector of $V_{d_1,\dots,d_k|d_{k+1},\dots,d_{r+\tilde r}}$ is $\lambda^{\sum_{i=1}^k d_i}$. On the other hand, its eigenvalues on $\mathrm{std}^{\otimes a} \otimes \mathrm{std}^{\vee\otimes b}$ are at most λ^a because its eigenvalues on std are at most λ and its eigenvalues on $\mathrm{std}^\vee$ are at most 1. Similarly, the element whose eigenvalues are 1 with multiplicity $n-1-d_{k+1}$ and λ^{-1} with multiplicity d_k acts on the highest weight vector of $V_{d_1,\dots,d_k|d_{k+1},\dots,d_{r+\tilde r}}$ with eigenvalue $\lambda^{\sum_{i=k+1}^{r+\tilde r}(n-1-d_i)}$, but its eigenvalues on $\mathrm{std}^{\otimes a} \otimes \mathrm{std}^{\vee\otimes b}$ are at most λ^b.

For the if direction, we observe that the highest weight of

$$\left(\bigotimes_{i=1}^{k} \wedge^{d_i} \mathrm{std}\right) \otimes \left(\bigotimes_{i=k+1}^{r+\tilde r} \wedge^{n-1-d_i} \mathrm{std}^\vee\right) \tag{2-5}$$

equals the highest weight of $V_{d_1,\dots,d_k|d_{k+1},\dots,d_{r+\tilde r}}$, and that (2-5) is a summand of

$$\mathrm{std}^{\otimes \sum_{i=1}^k d_i} \otimes \mathrm{std}^{\vee\otimes \sum_{i=k+1}^{r+\tilde r}(n-1-d_i)}. \qquad \square$$

Lemma 2.9. *Assume $n \ge 3$, and, if $n = 3$, that the characteristic of $\mathbb{F}_q$ is not 2 or 5.*

For $0 \le d_1 \le \cdots d_{r+\tilde r} \le n-1$, the sheaf $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}\left(L_{\mathrm{univ}}\left(\tfrac12\right)\right)$ appears as a geometric summand of $L_{\mathrm{univ}}^{\otimes a} \otimes L_{\mathrm{univ}}^{\vee\otimes b}$ for $0 \le a, b \le n-1$ if and only if there is some k such that $\sum_{i=1}^k d_i \le n-1$, $\sum_{i=k+1}^{r+\tilde r}(n-1-d_i) \le n-1$, and $k \equiv r \mod m$

Proof. By [Katz 2013, Theorem 7.1], under these assumptions on n, the monodromy group of L_{univ} contains SL_{n-1}. Thus two irreducible representations of GL_{n-1} give isomorphic sheaves when composed with L_{univ} only if one is equal to the other twisted by an integer power of the determinant. Because m is the order of the geometric monodromy group of the determinant, in fact they are isomorphic if and only if the integer is a multiple of m. The claim then follows from Lemma 2.8. $\square$

Lemma 2.10. *The natural number m from Definition 2.2 is divisible by the largest power of p that is greater than or equal to $(n-1)/2$.*

Proof. Because $\mu = \epsilon_\chi^m$ for all χ and $|\epsilon_\chi| = 1$, we have $|\mu| = 1$.

Fix $\chi \in S_{n,q}$. There is a unique nontrivial character $\psi_\chi : \mathbb{F}_q \to \mathbb{C}^\times$ with $\chi(1+ax^n) = \psi_\chi(-a)$ for all $a \in \mathbb{F}_q$. We have

$$\epsilon_\chi = q^{-(n-1)/2} \sum_{a_1,\dots,a_{n-1}\in\mathbb{F}_q} \chi\left(1+\sum_i a_i x^i\right) = q^{-(n+1)/2} \sum_{a_1,\dots,a_n\in\mathbb{F}_q} \chi\left(1+\sum_i a_i x^i\right)\psi_\chi(a_n)$$

but for $\psi \ne \psi_\chi$,

$$q^{-(n+1)/2} \sum_{a_1,\dots,a_n\in\mathbb{F}_q} \chi\left(1+\sum_i a_i x^i\right)\psi(a_n) = 0.$$

Now fixing a nontrivial character ψ, there are q^{n-1} characters $\chi \in S_{n,q}$ with $\psi_\chi = \psi$, so

$$\sum_{\chi \in S_{n,q}} \left(q^{-(n+1)/2} \sum_{a_1,\dots,a_n \in \mathbb{F}_q} \chi\Big(1 + \sum_i a_i x^i\Big) \psi(a_n) \right)^m = \sum_{\substack{\chi \in S_{n,q} \\ \psi_\chi = \psi}} \epsilon_\chi^m = \sum_{\substack{\chi \in S_{n,q} \\ \psi_\chi = \psi}} \mu = q^{n-1}\mu. \tag{2-6}$$

On the other hand, we have

$$\begin{aligned}
&\sum_{\chi \in S_{n,q}} \left(q^{-(n+1)/2} \sum_{a_1,\dots,a_n \in \mathbb{F}_q} \chi\Big(1 + \sum_i a_i x^i\Big) \psi(a_n) \right)^m \\
&\quad = \sum_{\chi : (\mathbb{F}_q[T]/T^{n+1})^\times / \mathbb{F}_q^\times \to \mathbb{C}^\times} \left(q^{-(n+1)/2} \sum_{a_1,\dots,a_n \in \mathbb{F}_q} \chi\Big(1 + \sum_i a_i x^i\Big) \psi(a_n) \right)^m \\
&\quad = q^{-m((n+1)/2)} \sum_{a_{i,j} \in \mathbb{F}_q, i=1,\dots,n, j=1,\dots m} \; \sum_{\chi : (\mathbb{F}_q[T]/T^{n+1})^\times / \mathbb{F}_q^\times \to \mathbb{C}^\times} \chi\left(\prod_{j=1}^m \Big(1 + \sum_{i=1}^n a_{i,j} x^i\Big)\right) \psi\Big(\sum_{j=1}^m a_{n,j}\Big) \\
&\quad = q^{-m((n+1)/2)} \sum_{\substack{a_{i,j} \in \mathbb{F}_q, i=1,\dots,n, j=1,\dots m \\ \prod_{j=1}^m \left(1 + \sum_{i=1}^n a_{i,j} x^i\right) \equiv 1 \bmod x^{n+1}}} q^n \psi\Big(\sum_{j=1}^m a_{n,j}\Big)
\end{aligned} \tag{2-7}$$

where we extend the sum from primitive χ to all χ because $\sum_{a_n \in \mathbb{F}_q} \chi\big(1 + \sum_i a_i x^i\big)\psi(a_n) = 0$ for imprimitive χ and then expand the product and use orthogonality of characters.

Combining (2-6) with (2-7), we have

$$\sum_{\substack{a_{i,j} \in \mathbb{F}_q, i=1,\dots,n, j=1,\dots m \\ \prod_{j=1}^m \left(1 + \sum_{i=1}^n a_{i,j} x^i\right) \equiv 0 \bmod x^{n+1}}} \psi\Big(\sum_{j=1}^m a_{n,j}\Big) = q^{m(n+1)/2-1}\mu \tag{2-8}$$

where the power of q shows up as $q^{m(n+1)/2+(n-1)-n}$. If m is not divisible by a large power of p, we will derive a contradiction from $|\mu| = 1$ and the upper bound we will prove for the left side of (2-8).

The left side of (2-8) is a Kloosterman-type sum. By standard stationary phase analysis, inductively for i from 1 to $\lfloor \frac{n-1}{2} \rfloor$, the sum over $a_{n-i,j}$ vanishes unless $a_{i,j_1} = a_{i,j_2}$ for all j_1, j_2.

If n is odd, after restricting to this subset the number of terms remaining in (2-8) is $q^{(m-1)((n+1)/2)}$ times the number of $a_1, \dots, a_{(n-1)/2}$ satisfying

$$\left(1 + \sum_{i=1}^{(n-1)/2} a_i x^i\right)^m \equiv 1 \mod x^{(n-1)/2+1}.$$

The only way the total size is $q^{m((n+1)/2)-1}$ is if the number of such $a_1, \dots, a_{(n-1)/2}$ is $q^{(n-1)/2}$, which only happens if the largest power of p dividing m is at least $(n-1)/2$.

For n even, we observe that when $a_{1,j_1}, \dots, a_{n/2-1,j_2}$ for all $i \le n/2 - 1$ and all j_1, j_2, then the sum over $a_{n/2,1}, \dots, a_{n/2,m}$ is a quadratic Gauss sum in $m-1$ variables, which is nondegenerate unless $p \mid m$, in which case it has one-dimensional degeneracy locus. This means (2-8) is at most $q^{(m-1)(n+1)/2+1/2}$ times the number of $a_1, \dots, a_{n/2-1}$ satisfying

$$\left(1+\sum_{i=1}^{n/2-1} a_i x^i\right)^m \equiv 1 \mod x^{n/2}.$$

The only way this can be at least $q^{m((n+1)/2)-1}$ is if the number of such $a_1, \dots, a_{(n-1)/2}$ is at least $q^{(n-1)/2}$, which only happens if the largest power of p dividing m is at least $(n-1)/2$. □

2C. *Betti number bounds.* In this subsection, we bound the dimension of the cohomology groups $H^j_c\big(\mathrm{Prim}_{n,\overline{\mathbb{F}}_q}, V\big(L_{\mathrm{univ}}\big(\tfrac{1}{2}\big)\big)\big)$ that occur in the definition of $F(V)$. Combined with Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$, this will lead to bounds on $F(V)$. This will proceed by relating these cohomology groups to cohomology groups of varieties with constant coefficients, and then estimating those using Betti number bounds due to Katz. This relation is a geometric version of the standard arithmetic argument where, using the approximate functional equation and orthogonality of characters, a moment of L-functions of Dirichlet characters is reduced to the count (weighted by a smooth function) of tuples of natural numbers satisfying a congruence condition.

We now give a more geometric interpretation of the construction in Definition 2.1. Let the scheme Witt_n be $\prod_{m\geq 1 \text{ prime to } p, m\leq n} W_{l(m,n)}$ where $W_{l(m,n)} \cong \mathbb{A}^{l(m,n)}$ is the scheme parametrizing length $l(m,n)$ Witt vectors. This is a product of commutative unipotent group schemes and hence is a commutative unipotent group scheme itself, isomorphic to $\mathbb{A}^n$. Each $\mathbb{F}_q$-point corresponds to an even Dirichlet character $\mathbb{F}_q[x]/x^{n+1} \to \mathbb{C}^\times$ by Definition 2.1, and the multiplication in the group structure corresponds to multiplication of characters.

Katz constructs L_{univ} as $R^1\,\mathrm{pr}_{2,!}\,\mathcal{L}_{\mathrm{univ}}$ for a certain lisse rank one sheaf $\mathcal{L}_{\mathrm{univ}}$ on $\mathbb{A}^1 \times \mathrm{Prim}_n$, with pr_2 the projection onto Prim_n [Katz 2013, Section 4]. However, Katz's definition of $\mathcal{L}_{\mathrm{univ}}$ works equally well to construct a lisse rank one sheaf on $\mathbb{A}^1 \times \mathrm{Witt}_n$. We will also refer to this sheaf as $\mathcal{L}_{\mathrm{univ}}$ and the projection map $\mathbb{A}^1 \times \mathrm{Witt}_n \to \mathrm{Witt}_n$ as pr_2.

For natural numbers m_1, m_2, let Z_{n,m_1,m_2} be the subspace of $\mathbb{A}^{m_1} \times \mathbb{A}^{m_2}$ consisting of points $(a_1, \dots, a_{m_1}, b_1, \dots, b_{m_2})$ such that $\prod_{i=1}^{m_1}(1-a_i x) \equiv \prod_{i=1}^{m_2}(1-b_i x) \bmod x^{n+1}$.

Lemma 2.11. *For every $j \in \mathbb{Z}$, there is an $S_{m_1} \times S_{m_2}$-equivariant isomorphism*

$$H^j_c(\mathrm{Witt}_{n,\overline{\mathbb{F}}_q}, (R(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}})^{\otimes m_1} \otimes (R(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}})^{\otimes m_2}) = H^{j-2n}_c(Z_{n,m_1,m_2,\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell(-n)).$$

Proof. By applying the Künneth formula [SGA 4_3 1973, Exposé XVII, Théorèm 5.4.3],

$$\begin{aligned} &H^j_c(\mathrm{Witt}_{n,\overline{\mathbb{F}}_q}, (R(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}})^{\otimes m_1} \otimes (R(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}})^{\otimes m_2}) \\ &\qquad = H^j_c\left(\mathbb{A}^{m_1} \times \mathbb{A}^{m_2} \times \mathrm{Witt}_{n,\overline{\mathbb{F}}_q}, \bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^\vee_{\mathrm{univ}}(b_i, \omega)\right)\end{aligned}$$

where $(a_1, \dots, a_{m_1}, b_1, \dots, b_{m_2})$ are coordinates on $\mathbb{A}^{m_1} \times \mathbb{A}^{m_2}$ and ω is a coordinate on Witt_n.

Let pr_1 be the projection $\mathbb{A}^{m_1} \times \mathbb{A}^{m_2} \times \mathrm{Witt}_n \to \mathbb{A}^{m_1} \times \mathbb{A}^{m_2}$ and let $i : Z_{n,m_1,m_2} \to \mathbb{A}^{m_1} \times \mathbb{A}^{m_2}$ be the closed immersion. By applying the projection formula [SGA 4_3 1973, Exposé XVII, Proposition 5.2.9] to pr_1 on the left side and i on the right, it suffices to find an isomorphism

$$R(\mathrm{pr}_1)_!\biggl(\bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^{\vee}_{\mathrm{univ}}(b_i, \omega)\biggr) \cong i_!\mathbb{Q}_\ell[-2n](-n). \tag{2-9}$$

To do this, we will first check that the stalk of $R(\mathrm{pr}_1)_!\bigl(\bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^{\vee}_{\mathrm{univ}}(b_i, \omega)\bigr)$ vanishes outside the image of i. To do this, by proper base change [SGA 4_3 1973, Exposé XVII, Proposition 5.2.8], it suffices to check that the compactly supported cohomology of the fiber vanishes. It even suffices to check this for finite field-valued points, as the support is constructible. Let $(a_1, \dots, a_{m_1}, b_1, \dots, b_{m_2}) \in \mathbb{A}^{m_1} \times \mathbb{A}^{m_2}(\mathbb{F}_q)$ be a point over a possibly larger finite field extension $\mathbb{F}_q$, and let $\mathcal{L}'$ be the fiber of $\bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^{\vee}_{\mathrm{univ}}(b_i, \omega)$ over $(a_1, \dots, a_{m_1}, b_1, \dots, b_{m_2})$, a sheaf lisse of rank one on Witt_n. Over any finite field extension of $\mathbb{F}_q$, the trace function of $\mathcal{L}_{\mathrm{univ}}(a_i, \omega)$ is a Frob_q-invariant character of $\mathrm{Witt}_n(\mathbb{F}_q)$ evaluated at ω. Hence the trace function of $\mathcal{L}'$ is also a Frob_q-invariant character. Thus the pullback of the trace function of $\mathcal{L}'$ under the Lang isogeny $\mathrm{Witt}_n \to \mathrm{Witt}_n$, $g \mapsto \mathrm{Frob}_q(g)g^{-1}$ is trivial, so by Chebotarev the pullback of $\mathcal{L}'$ under the Lang isogeny is trivial. So $\mathcal{L}'$ is a summand of the pushforward of the constant sheaf by the Lang isogeny of Witt_n, and thus its cohomology is a summand of the cohomology of Witt_n, which is $\mathbb{Q}_\ell(-n)$ in degree $2n$ because $\mathrm{Witt}_n \cong \mathbb{A}^n$. Thus if $H^*_c(\mathrm{Witt}_n, \mathcal{L}')$ is nontrivial, it is equal to $\mathbb{Q}_\ell(-n)$, which implies that the sum of the trace function is q^n, so the character of $\mathrm{Witt}_n(\mathbb{F}_q)$ induced by $(a_1, \dots, a_{m_1}, b_1, \dots, b_{m_2})$ is trivial, which contradicts the claim that $(a_1, \dots, a_{m_1}, b_1, \dots, b_{m_2}) \notin Z_{n,m_1,m_2}$.

So in fact $R(\mathrm{pr}_1)_!\bigl(\bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^{\vee}_{\mathrm{univ}}(b_i, \omega)\bigr)$ is supported on the image of i.

Next we check that, restricting to the inverse image under pr_1 of the image of i, the trace function of $\bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^{\vee}_{\mathrm{univ}}(b_i, \omega)$ is the constant function 1. We use the fact that the trace function of $\mathcal{L}_{\mathrm{univ}}(a_i, \omega)$ is by construction the evaluation of the character corresponding to ω at $(1-a_i x)$. It follows from this and multiplicativity of χ that the trace function of $\bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^{\vee}_{\mathrm{univ}}(b_i, \omega)$ is the value of this character at $\prod_{i=1}^{m_1}(1-a_i x)/\prod_{i=1}^{m_2}(1-b_i x)$. By the definition on i, the image under pr_1 of the image of i is exactly the locus where $\prod_{i=1}^{m_1}(1-a_i x)/\prod_{i=1}^{m_2}(1-b_i x) \cong 1 \mod x^{m+1}$. Hence the value of the character on this element is 1, making the trace function 1.

Because its trace function is 1, $\bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^{\vee}_{\mathrm{univ}}(b_i, \omega) \cong \mathbb{Q}_\ell$, giving an isomorphism

$$i^* R(\mathrm{pr}_1)_!\biggl(\bigotimes_{i=1}^{m_1} \mathcal{L}_{\mathrm{univ}}(a_i, \omega) \otimes \bigotimes_{i=1}^{m_2} \mathcal{L}^{\vee}_{\mathrm{univ}}(b_i, \omega) \cong \mathbb{Q}_\ell\biggr) \cong i^* R(\mathrm{pr}_1)_!\mathbb{Q}_\ell \cong i^*\mathbb{Q}_\ell[-2n](n)$$

and thus by the support condition an isomorphism (2-9), as desired. □

Lemma 2.12. *For $0 \le d_1, \ldots, d_{r+\tilde r} \le n-1$, there is a long exact sequence of complexes of vector spaces*

$$H^*_c\biggl(\mathrm{Prim}_{n,\bar{\mathbb{F}}_q}, \bigotimes_{i=1}^{r} \wedge^{d_i}(L_{\mathrm{univ}})[-d_i] \otimes \bigotimes_{i=r+1}^{r+\tilde r} \wedge^{d_i}(L^\vee_{\mathrm{univ}})[-d_i]\biggr)$$
$$\to H^*_c(\mathrm{Witt}_{n,\bar{\mathbb{F}}_q}, (R(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}})^{\otimes \sum_{i=1}^r d_i} \otimes (R(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}})^{\otimes \sum_{i=r+1}^{r+\tilde r} d_i})^{S_{d_1}\times\cdots\times S_{d_{r+\tilde r}}}$$
$$\to H^*_c(\mathrm{Witt}_{n-1,\bar{\mathbb{F}}_q}, (R(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}})^{\otimes \sum_{i=1}^r d_i} \otimes (R(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}})^{\otimes \sum_{i=r+1}^{r+\tilde r} d_i})^{S_{d_1}\times\cdots\times S_{d_{r+\tilde r}}}$$

where $S_{d_1}\times\cdots\times S_{d_{r+\tilde r}} \subseteq S_{\sum_{i=1}^r d_i} \times S_{\sum_{i=r+1}^{r+\tilde r} d_i}$ in the obvious way.

Proof. In view of the excision long exact sequence [SGA 4_3 1973, Exposé XVII, (5.1.16.2)], it suffices to find an isomorphism (over Prim_n)

$$\bigotimes_{i=1}^{r} \wedge^{d_i}(L_{\mathrm{univ}})[-d_i] \otimes \bigotimes_{i=r+1}^{r+\tilde r} \wedge^{d_i}(L^\vee_{\mathrm{univ}})[-d_i]$$
$$\cong ((R(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}})^{\otimes \sum_{i=1}^r d_i} \otimes (R(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}})^{\otimes \sum_{i=r+1}^{r+\tilde r} d_i})^{S_{d_1}\times\cdots\times S_{d_{r+\tilde r}}}.$$

To do this, observe that by definition

$$L_{\mathrm{univ}} = R^1(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}}$$

and that

$$L^\vee_{\mathrm{univ}} = R^1(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}}$$

because they are lisse, irreducible, and have the same trace functions up to scaling. Note too that $\mathcal{L}_{\mathrm{univ}}$ and its dual have no higher and lower cohomology in the fibers of pr_1 over Prim_n, so that

$$L_{\mathrm{univ}}[-1] = R(\mathrm{pr}_1)_!\mathcal{L}_{\mathrm{univ}},\ L^\vee_{\mathrm{univ}}[-1] = R(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}}.$$

Because the tensor product of complexes is anticommutative in odd degrees, we have

$$(\wedge^d L_{\mathrm{univ}})[d] = \mathrm{Sym}^d(L_{\mathrm{univ}}[1]) = ((L_{\mathrm{univ}}[1])^{\otimes d})^{S_d} = ((R(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}})^{\otimes d})^{S_d}$$

and similarly for $\mathcal{L}^\vee_{\mathrm{univ}}$.

Tensoring these equalities for d_i from 1 to r, and the dual equalities for d_i from $r+1$ to $r+\tilde r$, we have the desired isomorphism. □

Lemma 2.13. *For $0 \le d_1, \ldots, d_{r+\tilde r} \le n-1$, we have the Betti number bound*

$$\sum_j \dim H^j_c\biggl(\mathrm{Prim}_{n,\bar{\mathbb{F}}_q}, \bigotimes_{i=1}^{r} \wedge^{d_i}(L) \otimes \bigotimes_{i=r+1}^{r+\tilde r} \wedge^{d_i}(L^\vee)\biggr) \le 4(2+\max(r,\tilde r))^{n+\sum_{i=1}^{r+\tilde r} d_i}. \tag{2-10}$$

Proof. We apply the exact sequence of Lemma 2.12 and then evaluate each term using Lemma 2.11. Because of this, it suffices to bound

$$\sum_j \dim(H_c^j(Z_{n,\sum_{i=1}^r d_i,\sum_{i=r+1}^{r+\tilde{r}} d_i,\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell))^{S_{d_1}\times\cdots\times S_{d_{r+\tilde{r}}}}$$

and

$$\sum_j \dim(H_c^j(Z_{n-1,\sum_{i=1}^r d_i,\sum_{i=r+1}^{r+\tilde{r}} d_i,\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell))^{S_{d_1}\times\cdots\times S_{d_{r+\tilde{r}}}}$$

separately. We have

$$(H_c^j(Z_{n,\sum_{i=1}^r d_i,\sum_{i=r+1}^{r+\tilde{r}} d_i,\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell))^{S_{d_1}\times\cdots\times S_{d_{r+\tilde{r}}}} = H_c^j(Z_{n,\sum_{i=1}^r d_i,\sum_{i=r+1}^{r+\tilde{r}} d_i,\overline{\mathbb{F}}_q}/(S_{d_1}\times\cdots\times S_{d_{r+\tilde{r}}}), \mathbb{Q}_\ell).$$

Because we can take the coordinates of $(\mathbb{A}^d)^{S_d}$ to be the coefficients of the polynomial $\prod_{i=1}^T (T-a_i)$ for $a_1,\dots,a_d$ the coordinates of $\mathbb{A}^d$, we can view $Z_{n,\sum_{i=1}^r d_i,\sum_{i=r+1}^{r+\tilde{r}} d_i,\overline{\mathbb{F}}_q}/(S_{d_1}\times\cdots\times S_{d_{r+\tilde{r}}})$ as the moduli space of tuples of monic polynomials $f_1,\dots,f_{r+\tilde{r}}$, with f_i of degree d_i, such that the leading $n+1$ coefficients of $\prod_{i=1}^r f_i$ and $\prod_{i=r+1}^{r+\tilde{r}} f_i$ agree. The equality of the leading coefficient is trivial, while the equality of the remaining n coefficients is a system of n polynomial equations of degrees $\max(r,\tilde{r})$. So this is the solution set of a system of n equations, of degree at most $\max(r,\tilde{r})$, in $\sum_{i=1}^{r+\tilde{r}} d_i$ variables. By [Katz 2001, Theorem 12],

$$\sum_j \dim_i H_c^j(Z_{n,\sum_{i=1}^r d_i,\sum_{i=r+1}^{r+\tilde{r}} d_i,\overline{\mathbb{F}}_q}/(S_{d_1}\times\cdots\times S_{d_{r+\tilde{r}}}), \mathbb{Q}_\ell) \le 3(2+\max(r,\tilde{r}))^{n+\sum_{i=1}^{r+\tilde{r}} d_i}.$$

For Z_{n-1}, the same argument gives a Betti number bound of

$$3(2+\max(r,\tilde{r}))^{n-1+\sum_{i=1}^{r+\tilde{r}} d_i} \le (2+\max(r,\tilde{r}))^{n+\sum_{i=1}^{r+\tilde{r}} d_i}$$

as $2+\max(r,\tilde{r}) \ge 3$.

Summing the bounds for Z_n and Z_{n-1}, we get (2-10). □

3. Analysis of the main term

For $S \subseteq \{1,\dots,r+\tilde{r}\}$, let

$$M_S(\alpha_1,\dots,\alpha_{r+\tilde{r}}) = \prod_{1\le i_1<i_2\le r+\tilde{r}} (q^{\alpha_{i_1}} - q^{\alpha_{i_2}}) \sum_{\substack{f_1,\dots,f_{r+\tilde{r}}\in\mathbb{F}_q[T] \\ \text{monic} \\ \prod_{i\in S} f_i/\prod_{i\notin S} f_i \in T^{\mathbb{Z}}}} \prod_{i\in S} q^{(-1/2+\alpha_i)\deg f_i} \prod_{i\notin S} q^{(-1/2-\alpha_i)\deg f_i}.$$

Note that this is independent of n.

The main goal of this section is to estimate the coefficients of M_S, viewed as a Laurent series in the q^{α_i}. In the first two lemmas we will establish some basic properties that will be useful later, in particular giving conditions for the coefficients of this series to be nonvanishing. In the next three lemmas, we will prove a bound on the coefficients using a contour integral argument. In Lemma 3.6, we will relate the coefficients of M_S to the coefficients of (1-3).

Lemma 3.1. *Let $d_1, \dots, d_{r+\tilde{r}}$ be integers. The coefficient of $\prod_i q^{\alpha_i d_i}$ in*

$$\sum_{\substack{f_1,\dots,f_{r+\tilde{r}} \in \mathbb{F}_q[T] \\ \text{monic} \\ \prod_{i\in S} f_i / \prod_{i \notin S} f_i \in T^{\mathbb{Z}}}} \prod_{i \in S} q^{(-1/2+\alpha_i)\deg f_i} \prod_{i \notin S} q^{(-1/2-\alpha_i)\deg f_i}$$

vanishes unless $d_i \geq 0$ for $i \in S$ and $d_i \leq 0$ for $i \notin S$. Furthermore, the coefficient of $\prod_i q^{\alpha_i d_i}$ in this expression is symmetric in the variables d_i for $i \in S$ and also in the d_i for $i \notin S$.

Proof. The vanishing is because, in each term of the sum, q^{α_i} appears only in nonnegative powers if $i \in S$ and in nonpositive powers in $i \in S$. The symmetry is because the definition is symmetric in the α_i for $i \in S$ and symmetric in the α_i for $i \notin S$, by permuting the corresponding f_is. □

Lemma 3.2. *M_S is antisymmetric in the α_i variables for $i \in S$, and also in the α_i for $i \notin S$. Expressed as a power series, the coefficient of $q^{\sum_i \alpha_i d_i}$ is nonzero only if there exists a permutation $\sigma \in S_{r+\tilde{r}}$ such that $\sigma(\{1,\dots,|S|\}) = S$, $d_{\sigma(i)} \geq i-1$ for all i from 1 to $|S|$, and $d_{\sigma(i)} \leq i-1$ for all i from $|S|+1$ to $r+\tilde{r}$.*

Proof. These follow from Lemma 3.1 once we adjust for the Vandermonde factor. The first claim follows because the Vandermonde is antisymmetric, and it is multiplied by a symmetric term, making the product antisymmetric. The support conditions follow from the support statement in Lemma 3.1 combined with the fact that the Vandermonde determinant is a sum of terms $q^{\sum_i \alpha_i d_i}$ where $d_1, \dots, d_{r+\tilde{r}}$, is a permutation of $0, \dots, r+\tilde{r}-1$. □

We next prove three lemmas that bound the coefficients of M_S. Lemma 3.3 will give a general means for converting bounds for the values of a power series into bounds for that power series, based on a contour integral. Lemma 3.4 will use this to bound the coefficients of a simplified version of M_S, where we include only some of the coefficients in the Vandermonde factor. Lemma 3.5 will apply this to bound M_S.

Lemma 3.3. *Let $F(\alpha_1, \dots, \alpha_{r+\tilde{r}})$ be a power series in q^{α_i} ($i \in S$) and $q^{-\alpha_i}$ ($i \notin S$) with complex coefficients which converges for any $(\alpha_1, \dots, \alpha_{r+\tilde{r}})$ such that $\operatorname{Re}\alpha_i$ is sufficiently small for i in S and sufficiently large for i not in S.*

Let $c_1, \dots, c_{r+\tilde{r}}$ be real numbers such that $F(\alpha_1, \dots, \alpha_{r+\tilde{r}})$ extends to a holomorphic function on the set of tuples $(\alpha_1, \dots, \alpha_{r+\tilde{r}})$ such that $\operatorname{Re}(\alpha_i) < c_i$ if $i \in S$ and $\operatorname{Re}(\alpha_i) > c_i$ if $i \notin S$. Suppose further that, for any $\epsilon > 0$, whenever $\operatorname{Re}(\alpha_i) \leq c_i - \epsilon$ if $i \in S$ and $\operatorname{Re}(\alpha_i) \geq c_i + \epsilon$ if $i \notin S$, this holomorphic function is $(1/(1-q^{\epsilon}))^{O(1)}$.

Then the coefficient of $q^{\sum_i d_i \alpha_i}$ in $F(\alpha_1, \dots, \alpha_{r+\tilde{r}})$ is

$$O\biggl(\biggl(1 + \sum_{i\in S} d_i - \sum_{i \notin S} d_i\biggr)^{O(1)} q^{-\sum_{i=1}^{r+\tilde{r}} c_i d_i}\biggr).$$

Proof. Let $\delta_i = -1$ if $i \in S$ and 1 if $i \notin S$. By the Cauchy integral formula, the coefficient of $q^{\sum_i d_i \alpha_i}$ in $F(\alpha_1, \dots, \alpha_{r+\tilde{r}})$ is

$$\frac{1}{(2\pi)^{r+\tilde{r}}} \int_{t_1,\dots,t_{r+\tilde{r}} \in [0,2\pi]} \frac{F(c_1 + \delta_i \epsilon + it_1, \dots, c_{r+\tilde{r}} + \delta_i \epsilon + it_{r+\tilde{r}})}{q^{(c_1+\delta_i\epsilon+it_1)d_1 + \cdots + (c_{r+\tilde{r}} + \delta_i \epsilon + it_{r+\tilde{r}})d_{r+\tilde{r}}}} \, dt_1 \cdots dt_{r+\tilde{r}} \tag{3-1}$$

where we have extended F as a holomorphic function from its region of absolute convergence. Then by our assumed bound, (3-1) is

$$\frac{1}{(2\pi)^{r+\tilde{r}}}\int_{t_1,\dots,t_{r+\tilde{r}}\in[0,2\pi]}\frac{O((1/(1-q^{\epsilon}))^{O(1)})}{q^{\sum_{i=1}^{r+\tilde{r}}c_id_i-\sum_{i\in S}\epsilon d_i+\sum_{i\notin S}\epsilon d_i}}\,dt_1\dots dt_{r+\tilde{r}}=\frac{O((1/(1-q^{\epsilon}))^{O(1)})}{q^{\sum_{i=1}^{r+\tilde{r}}c_id_i-\sum_{i\in S}\epsilon d_i+\sum_{i\notin S}\epsilon d_i}}.$$

Now let

$$\epsilon=\frac{1}{(1+\sum_{i\in S}d_i-\sum_{i\notin S}d_i)\log q}.$$

Then our bound for (3-1) specializes to

$$\frac{O\big(\big(1+\sum_{i\in S}d_i-\sum_{i\notin S}d_i\big)^{O(1)}\big)}{q^{\sum_{i=1}^{r+\tilde{r}}c_id_i-O(1)}}=O\Bigg(\bigg(1+\sum_{i\in S}d_i-\sum_{i\notin S}d_i\bigg)^{O(1)}q^{-\sum_{i=1}^{r+\tilde{r}}c_id_i}\Bigg).$$

□

To simplify M_S, observe that

$$\prod_{1\le i_1<i_2\le r+\tilde{r}}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})=\pm\bigg(\prod_{i_1\in S}\prod_{i_2\notin S}(q^{\alpha_{i_2}}-q^{\alpha_{i_1}})\bigg)\bigg(\prod_{\substack{i_1,i_2\in S\\ i_1<i_2}}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})\bigg)\bigg(\prod_{\substack{i_1,i_2\notin S\\ i_1<i_2}}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})\bigg) \tag{3-2}$$

and that we can further write

$$\prod_{i_1\in S}\prod_{i_2\notin S}(q^{\alpha_{i_2}}-q^{\alpha_{i_1}})=\bigg(\prod_{i\notin S}q^{r\alpha_i}\bigg)\bigg(\prod_{i_1\in S}\prod_{i_2\notin S}(1-q^{\alpha_{i_1}-\alpha_{i_2}})\bigg). \tag{3-3}$$

We will only use the factor $\prod_{i_1\in S}\prod_{i_2\notin S}(1-q^{\alpha_{i_1}-\alpha_{i_2}})$, which can be expressed nicely as an Euler product, in Lemma 3.4, and will add the rest of the formula in Lemma 3.5.

Lemma 3.4. *Let $d_1,\dots,d_{r+\tilde{r}}$ be integers with $\sum_{i\in S}d_i-\sum_{i\notin S}d_i\ge 0$. The coefficient of $\prod_i q^{\alpha_i d_i}$ in*

$$\bigg(\prod_{i_1\in S}\prod_{i_2\notin S}(1-q^{\alpha_{i_1}-\alpha_{i_2}})\bigg)\sum_{\substack{f_1,\dots,f_{r+\tilde{r}}\in\mathbb{F}_q[T]\\ \text{monic}\\ \prod_{i\in S}f_i/\prod_{i\notin S}f_i\in T^{\mathbb{Z}}}}\prod_{i\in S}q^{(-1/2+\alpha_i)\deg f_i}\prod_{i\notin S}q^{(-1/2-\alpha_i)\deg f_i} \tag{3-4}$$

is

$$O\Bigg(\bigg(1+\sum_{i\in S}d_i-\sum_{i\notin S}d_i\bigg)^{O(1)}\min\big(q^{(-\max_{i\in S}d_i+\min_{i\notin S}d_i)/2},q^{(-\sum_{i\in S}d_i)/2},q^{(\sum_{i\notin S}d_i)/2},q^{-|\sum_{i\in S}d_i+\sum_{i\notin S}d_i|/2}\big)\Bigg).$$

Note that, by Lemma 3.1, this coefficient vanishes unless $\sum_{i\in S}d_i-\sum_{i\notin S}d_i\ge 0$.

Proof. We view the upper bound as a conjunction of four upper bounds and prove each separately, by applying Lemma 3.3.

In each case, by Lemma 3.3, it suffices to prove that, for all $\epsilon>0$, the expression (3-4) is $(1/(1-q^{\epsilon}))^{O(1)}$ if $\alpha_1,\dots,\alpha_{r+\tilde{r}}$ satisfy one of the following:

(1) For some $i_1\in S$, we have $\operatorname{Re}\alpha_{i_1}\le\frac{1}{2}-\epsilon$, and $\operatorname{Re}\alpha_i\le-\epsilon$ for all other $i\in S$. For some $i_2\notin S$, we have $\operatorname{Re}\alpha_{i_2}\ge-\frac{1}{2}+\epsilon$, and $\operatorname{Re}\alpha_i\ge\epsilon$ for all other $i\notin S$.

(2) $\operatorname{Re}\alpha_i \le \frac{1}{2}-\epsilon$ for $i \in S$, $\operatorname{Re}\alpha_i \ge \epsilon$ for $i \notin S$.

(3) $\operatorname{Re}\alpha_i \le -\epsilon$ for $i \in S$, $\operatorname{Re}\alpha_i \ge \frac{1}{2}+\epsilon$ for $i \notin S$.

(4) $\operatorname{Re}\alpha_i \le \frac{1}{2}-\epsilon$ for all $i \in S$, $\operatorname{Re}\alpha_i \ge \frac{1}{2}+\epsilon$ for $i \notin S$.

(5) $\operatorname{Re}\alpha_i \le -\frac{1}{2}-\epsilon$ for all $i \in S$, $\operatorname{Re}\alpha_i \ge -\frac{1}{2}+\epsilon$ for $i \notin S$.

For all cases, we will use the Euler products

$$\sum_{\substack{f_1,\dots,f_{r+\tilde r}\in\mathbb{F}_q[T]\\ \text{monic}\\ \prod_{i\in S} f_i/\prod_{i\notin S} f_i\in T^{\mathbb{Z}}}} \prod_{i\in S} q^{(-1/2+\alpha_i)\deg f_i} \prod_{i\notin S} q^{(-1/2-\alpha_i)\deg f_i} \tag{3-5}$$

$$= \biggl(\prod_{i\in S}(1-q^{-1/2+\alpha_i})^{-1}\prod_{i\notin S}(1-q^{-1/2-\alpha_i})^{-1}\biggr) \prod_{\substack{\pi\in\mathbb{F}_q[T]\\ \text{monic}\\ \text{irreducible}\\ \pi\neq T}} \sum_{\substack{e_1,\dots,e_{r+\tilde r}\in\mathbb{N}\\ \sum_{i\in S} e_i=\sum_{i\notin S} e_i}} |\pi|^{-\sum_{i\in S} e_i+\sum_{i\in S}\alpha_i e_i-\sum_{i\notin S}\alpha_i e_i} \tag{3-6}$$

where the first term is the Euler factor at T, and

$$(1-q^{\alpha_{i_1}-\alpha_{i_2}}) = \prod_{\substack{\pi\in\mathbb{F}_q[T]\\ \text{monic}\\ \text{irreducible}}} (1-|\pi|^{\alpha_{i_1}-\alpha_{i_2}-1}). \tag{3-7}$$

Let us briefly discuss issues of convergence. The sum (3-5) is (absolutely) convergent whenever $\operatorname{Re}\alpha_i < -\frac{1}{2}$ for $i \in S$ and $\operatorname{Re}\alpha_i > \frac{1}{2}$ for $i \notin S$. (In fact, this holds even without the condition $\prod_{i\in S} f_i/\prod_{i\notin S} f_i \in T^{\mathbb{Z}}$.) Hence (3-6) converges to the value of the sum in the same region. The Euler product (3-7) is valid when $\operatorname{Re}\alpha_{i_1}-\operatorname{Re}\alpha_{i_2}<0$. So the holomorphic function (3-4) is equal to

$$\biggl(\prod_{i\in S}(1-q^{-1/2+\alpha_i})^{-1}\prod_{i\notin S}(1-q^{-1/2-\alpha_i})^{-1}\prod_{i_1\in S}\prod_{i_2\notin S}(1-q^{\alpha_{i_1}-\alpha_{i_2}-1})\biggr) \prod_{\substack{\pi\in\mathbb{F}_q[T]\\ \text{monic}\\ \text{irreducible}\\ \pi\neq T}} f_\pi(\alpha_1,\dots,\alpha_{r+\tilde r}) \tag{3-8}$$

as long as $\operatorname{Re}(\alpha_i)$ is sufficiently small for $i \in S$ and sufficiently large for $i \notin S$, where

$$f_\pi(\alpha_1,\dots,\alpha_{r+\tilde r}) = \prod_{i_1\in S}\prod_{i_2\notin S}(1-q^{\alpha_{i_1}-\alpha_{i_2}-1}) \sum_{\substack{e_1,\dots,e_{r+\tilde r}\in\mathbb{N}\\ \sum_{i\in S} e_i=\sum_{i\notin S} e_i}} |\pi|^{-\sum_{i\in S} e_i+\sum_{i\in S}\alpha_i e_i-\sum_{i\notin S}\alpha_i e_i}$$

We will show that (3-8) in fact converges in a larger region, which will suffice to apply Lemma 3.3.

Let us fix ϵ for the remainder of the proof.

In all five cases, the Euler factors at T is manifestly $O((1-q^{-\epsilon})^{-O(1)})$ so we focus on the other Euler factors, where it suffices to prove that for $\alpha_1,\dots,\alpha_{r+\tilde r}$ in these ranges

$$|f_\pi(\alpha_1,\dots,\alpha_{r+\tilde r})| \le (1-|\pi|^{-1-\epsilon})^{-O(1)}. \tag{3-9}$$

This is sufficient because

$$\prod_{\substack{\pi\in\mathbb{F}_q[T]\\ \text{monic}\\ \text{irreducible}\\ \pi\neq T}} (1-|\pi|^{-1-\epsilon})^{-O(1)} \le \prod_{\substack{\pi\in\mathbb{F}_q[T]\\ \text{monic}\\ \text{irreducible}}} (1-|\pi|^{-1-\epsilon})^{-O(1)} = (\zeta_{\mathbb{F}_q[T]}(1+\epsilon))^{O(1)} = (1-q^{-\epsilon})^{-O(1)}.$$

We have

$$f_\pi(\alpha_1,\dots,\alpha_{r+\tilde r}) = \sum_{\substack{e_1,\dots,e_{r+\tilde r}\in\mathbb{N}\\ \sum_{i\in S} e_i=\sum_{i\notin S} e_i}} \Bigg(|\pi|^{-\sum_{i\in S} e_i+\sum_{i\in S}\alpha_i e_i-\sum_{i\notin S}\alpha_i e_i} \sum_{\substack{J\subseteq S\times S^c\\ |J\cap p_S^{-1}(j)|\le e_j,\, j\in S\\ |J\cap p_{S^c}^{-1}(j)|\le e_j,\, j\notin S}} (-1)^{|J|}\Bigg) \tag{3-10}$$

where p_S and p_{S^c} are the projections onto S and S^c.

Observe first that when $e_1,\dots,e_{r+\tilde r}$ are not all zero,

$$\sum_{\substack{J\subseteq S\times S^c\\ |J\cap p_S^{-1}(j)|\le e_j,\, j\in S\\ |J\cap p_{S^c}^{-1}(j)|\le e_j,\, j\notin S}} (-1)^{|J|} \tag{3-11}$$

vanishes unless $\max_{i\in S} e_i < |\{i\notin S \mid e_i>0\}|$ or $\max_{i\notin S} e_i < |\{i\in S \mid e_i>0\}|$. Indeed, if neither of these is satisfied, letting $i_1\in S$ and $i_2\notin S$ be such that e_{i_1} and e_{i_2} attain their maximal values, we see that adding (i_1,i_2) to J or removing it from J preserves the conditions $|J\cap p_S^{-1}(j)|\le e_j$, $j\in S$, $|J\cap p_{S^c}^{-1}(j)|\le e_j$, $j\notin S$, so defines a sign-reversing involution of J, and thus the sum vanishes.

In particular, (3-11) vanishes for all but finitely many $e_1,\dots,e_{r+\tilde r}$. So in each of the five cases, to prove (3-9) it suffices to show that each term in (3-10), except the one where $e_1,\dots,e_{r+\tilde r}=0$, is $O(|\pi|^{-1-\epsilon})$. Furthermore, if $\sum_{i\in S} e_i=\sum_{i\notin S} e_i=1$, the sum over J has two terms which cancel each other, so we may assume $\sum_{i\in S} e_i\ge 2$.

Case (1) is the most difficult. However, using the inequality we have verified, it is straightforward. Suppose

$$\max_{i\in S} e_i < |\{i\notin S \mid e_i>0\}|,$$

so

$$\max_{i\in S} e_i \le |\{i\notin S \mid e_i>0\}|-1.$$

But

$$|\{i\notin S \mid e_i>0\}| + \max_{i\notin S} e_i - 1 \le \sum_{i\notin S} e_i,$$

so

$$\max_{i\in S} e_i + \max_{i\notin S} e_i \le \sum_{i\notin S} e_i$$

thus

$$\begin{aligned}-\sum_{i\in S} e_i+\sum_{i\in S}\mathrm{Re}(\alpha_i)e_i-\sum_{i\notin S}\mathrm{Re}(\alpha_i)e_i &\le -\sum_{i\in S} e_i+\frac{\max_{i\in S} e_i+\max_{i\notin S} e_i}{2}-2\epsilon\\ &\le -\sum_{i\in S} e_i+\frac{\sum_{i\notin S} e_i}{2}-2\epsilon\\ &\le -\frac{\sum_{i\in S} e_i}{2}-2\epsilon\le -1-2\epsilon\end{aligned}$$

so all terms are $O(|\pi|^{-1-2\epsilon})$.

For cases (2) and (3), we have

$$|\pi|^{-\sum_{i\in S} e_i+\sum_{i\in S}\mathrm{Re}(\alpha_i)e_i-\sum_{i\notin S}\mathrm{Re}(\alpha_i)e_i}\le |\pi|^{-(1/2+2\epsilon)\sum_{i\in S} e_i},$$

so the terms are $O(|\pi|^{-1-4\epsilon})$.

For case (4) and (5), we have

$$|\pi|^{-\sum_{i\in S} e_i+\sum_{i\in S}\mathrm{Re}(\alpha_i)e_i-\sum_{i\notin S}\mathrm{Re}(\alpha_i)e_i}\le |\pi|^{-(1+2\epsilon)\sum_{i\in S} e_i},$$

so all the terms are $O(|\pi|^{-2-4\epsilon})$. □

Lemma 3.5. *Let $d_1,\dots,d_{r+\tilde r}$ be integers satisfying the inequalities of Lemma 3.2. Then the coefficient of $\prod_i q^{\alpha_i d_i}$ in M_S is bounded by*

$$O\biggl(\Bigl(O(1)+\sum_{i\in S} d_i-\sum_{i\notin S} d_i\Bigr)^{O(1)}\biggr)$$
$$\times\min\bigl(q^{(-\max_{i\in S} d_i+\min_{i\notin S} d_i-1)/2},\ q^{(-\sum_{i\in S} d_i+\binom{|S|}{2})/2},\ q^{(\sum_{i\notin S} d_i+\binom{|S|}{2}-\binom{r+\tilde r}{2})/2},\ q^{-|\sum_{i\in S} d_i+\sum_{i\notin S} d_i-\binom{r+\tilde r}{2}|/2}\bigr).$$

Proof. By (3-2) and (3-3), M_S is equal to the expression (3-4) times

$$\pm\prod_{\substack{1\le i_1<i_2\le r+\tilde r\\ i_1,i_2\in S\text{ or } i_1,i_2\notin S}}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})\prod_{i\notin S} q^{|S|\alpha_i}.$$

This additional factor has bounded coefficients and is supported on those terms $q^{\sum_i\alpha_i d_i}$ where $\{d_i\mid i\in S\}=\{0,\dots,|S|-1\}$ and $\{d_i\mid i\notin S\}=\{|S|,\dots,r+\tilde r-1\}$.

Hence we can obtain bounds for M_S by subtracting from the exponents in Lemma 3.4 the minimal possible contribution of an element in the support of this additional factor to the exponent, which are as stated. In fact, in all cases but the first, we are minimizing a constant function. □

Lemma 3.6. *Assume $|S|-r$ is a multiple of m. The coefficients of $q^{\sum_i\alpha_i d_i}$ in the power series*

$$\prod_{1\le i_1<i_2\le r+\tilde r}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})\sum_{\chi\in S_{n,q}}\epsilon_\chi^{-\tilde r}\prod_{i=1}^{r+\tilde r}L\bigl(\tfrac12-\alpha_i,\chi\bigr)\quad\textit{and}\quad (q^n-q^{n-1})\mu^{(r-|S|)/m}\prod_{i\notin S}q^{\alpha_i(n-1)}M_S(\alpha_1,\dots,\alpha_{r+\tilde r})$$

agree as long as

$$\sum_{i\in S} d_i - \binom{|S|}{2},\ \sum_{i\notin S}(n-1-d_i) + \binom{r+\tilde{r}}{2} - \binom{|S|}{2} \le n-1.$$

Proof. We have

$$\epsilon_\chi^{-\tilde{r}} \prod_{i=1}^{r+\tilde{r}} L\bigl(\tfrac{1}{2}-\alpha_i, \chi\bigr) = \epsilon_\chi^{r+\tilde{r}-|S|-\tilde{r}} \prod_{i\in S} L\bigl(\tfrac{1}{2}-\alpha_i, \chi\bigr) \prod_{i\notin S} q^{(n-1)\alpha_i} L\bigl(\tfrac{1}{2}+\alpha_i, \bar{\chi}\bigr).$$

Because $r-|S|$ is divisible by m, $\epsilon_\chi^{r-|S|} = \mu^{(r-|S|)/m}$. Thus

$$\sum_{\chi\in S_{n,q}} \epsilon_\chi^{-\tilde{r}} \prod_{i=1}^{r+\tilde{r}} L\bigl(\tfrac{1}{2}-\alpha_i, \chi\bigr)$$
$$= \mu^{(r-|S|)/m} \prod_{i\notin S} q^{(n-1)\alpha_i} \sum_{\substack{f_1,\dots,f_{r+\tilde{r}}\in\mathbb{F}_q[T]\\ \text{monic}}} \sum_{\chi\in S_{n,q}} \chi\Bigl(\prod_{i\in S} f_i\Bigr) \bar{\chi}\Bigl(\prod_{i\notin S} f_i\Bigr) \prod_{i\in S} q^{(-1/2+\alpha_i)\deg f_i} \prod_{i\notin S} q^{(-1/2-\alpha_i)\deg f_i}.$$

Now

$$\sum_{\chi\in S_{n,q}} \chi\Bigl(\prod_{i\in S} f_i\Bigr) \bar{\chi}\Bigl(\prod_{i\notin S} f_i\Bigr)$$

is equal to the sum over all even characters of $(\mathbb{F}_q[x]/x^{n+1})^\times$ minus the sum over all even characters of $(\mathbb{F}_q[x]/x^n)^\times$. The characters modulo x^n depend on the leading n coefficients of the polynomial. Hence both sums vanish unless the leading n coefficients are equal. Thus if $\deg \prod_{i\in S} f_i$, $\deg \prod_{i\notin S} f_i \le n-1$, there are only n coefficients, and so both sums cancel unless $\prod_{i\in S} f_i / \prod_{i\notin S} f_i \in T^{\mathbb{Z}}$, in which case the sum over characters is $q^n - q^{n-1}$.

This occurs precisely in the coefficients $q^{\sum_i \alpha_i d_i}$ where $\sum_{i\in S} d_i, \sum_{i\notin S}(n-1-d_i) \le n-1$. Hence for d_i satisfying those inequalities, the coefficients of $q^{\sum_i \alpha_i d_i}$ in $\sum_{\chi\in S_{n,q}} \epsilon_\chi^{-\tilde{r}} \prod_{i=1}^{r+\tilde{r}} L\bigl(\tfrac{1}{2}-\alpha_i, \chi\bigr)$ and

$$(q^n - q^{n-1}) \mu^{(r-|S|)/m} \prod_{i\notin S} q^{(n-1)\alpha_i} \sum_{\substack{f_1,\dots,f_{r+\tilde{r}}\in\mathbb{F}_q[T]\\ \text{monic}\\ \prod_{i\in S} f_i / \prod_{i\notin S} f_i \in T^{\mathbb{Z}}}} \prod_{i\in S} q^{(-1/2+\alpha_i)\deg f_i} \prod_{i\notin S} q^{(-1/2-\alpha_i)\deg f_i}$$

are equal.

Multiplying the monomial $q^{\sum_i \alpha_i d_i}$ by the Vandermonde determinant produces a sum of monomials. In each monomial, $\sum_{i\in S} d_i$ is increased by at least $\binom{|S|}{2}$ and $\sum_{i\notin S}(n-1-d_i)$ is reduced by at most $\binom{r+\tilde{r}}{2} - \binom{|S|}{2}$. Hence the identity in the multiplied terms is satisfied as long as $\sum_{i\in S} d_i - \binom{|S|}{2} \le n-1$ and $\sum_{i\notin S}(n-1-d_i) + \binom{r+\tilde{r}}{2} - \binom{|S|}{2} \le n-1$. □

3A. ***Additional results.*** Here we prove some additional results that are not necessary to prove Theorem 1.4 but may be helpful to interpret it. We describe the general behavior of (1-3) and compute it in a special case.

Lemma 3.7. (1) *The holomorphic function* (1-3) *evaluated at* $\alpha_1 = \cdots = \alpha_{r+\tilde r} = 0$ (*where it is defined by analytic continuation*, *see Remark 1.6*) *is a polynomial in* n *of degree* $r\tilde r$ *whose leading term is*

$$a_{r,\tilde r} g_{r,\tilde r} n^{r\tilde r}/(r\tilde r)! \tag{3-12}$$

where $g_{r,\tilde r}$ *is the random matrix factor*

$$g_{r,\tilde r} = (r\tilde r)! \prod_{j=0}^{r-1} j!/(j+\tilde r)!$$

and $a_{r,\tilde r}$ *is the Euler product*

$$(1-q^{-1})^{r\tilde r}(1-q^{-1/2})^{r+\tilde r} \prod_{\substack{\pi\in\mathbb{F}_q[T]\\ \text{monic}\\ \text{irreducible}\\ \pi\neq T}} \biggl((1-|\pi|^{-1})^{r\tilde r} \sum_{e\in\mathbb{N}} \binom{e+r-1}{r-1}\binom{e+\tilde r-1}{\tilde r-1}|\pi|^{-e}\biggr). \tag{3-13}$$

(*In this Euler product*, $(1-q^{-1})^{r\tilde r}(1-q^{-1/2})^{r+\tilde r}$ *may be viewed as the Euler factor at* T.)

(2) *On the portion of the imaginary axis where* $\alpha_1, \dots, \alpha_{r+\tilde r}$ *are distinct modulo* $2\pi i/\log q$, *the individual terms*

$$\sum_{\substack{f_1,\dots,f_{r+\tilde r}\in\mathbb{F}_q[T]\\ \text{monic}\\ \prod_{i\in S} f_i/\prod_{i\notin S} f_i\in T^{\mathbb{Z}}}} \prod_{i\in S}|f_i|^{-1/2+\alpha_i} \prod_{i\notin S}|f_i|^{-1/2-\alpha_i}$$

of (1-3) *are holomorphic so the sum* (1-3) *is simply a linear combination of terms* $\prod_{i\notin S} q^{\alpha_i(n-1)}$, *i.e.*, *a quasiperiodic function of* n.

Proof. It follows from Lemma 3.4 that

$$\biggl(\prod_{i_1\in S}\prod_{i_2\notin S}(1-q^{\alpha_{i_1}-\alpha_{i_2}})\biggr) \sum_{\substack{f_1,\dots,f_{r+\tilde r}\in\mathbb{F}_q[T]\\ \text{monic}\\ \prod_{i\in S} f_i/\prod_{i\notin S} f_i\in T^{\mathbb{Z}}}} \prod_{i\in S} q^{(-1/2+\alpha_i)\deg f_i} \prod_{i\notin S} q^{(-1/2-\alpha_i)\deg f_i}, \tag{3-14}$$

when written as a power series in $q^{-\alpha_i}$, converges for α_i in a neighborhood of the imaginary axis. Thus

$$\sum_{\substack{f_1,\dots,f_{r+\tilde r}\in\mathbb{F}_q[T]\\ \text{monic}\\ \prod_{i\in S} f_i/\prod_{i\notin S} f_i\in T^{\mathbb{Z}}}} \prod_{i\in S} q^{(-1/2+\alpha_i)\deg f_i} \prod_{i\notin S} q^{(-1/2-\alpha_i)\deg f_i}$$

is equal to the quotient of a holomorphic function in this neighborhood by

$$\biggl(\prod_{i_1\in S}\prod_{i_2\notin S}(1-q^{\alpha_{i_1}-\alpha_{i_2}})\biggr).$$

We will use that property to prove both parts of this lemma.

To prove part (2), if $\alpha_1, \dots, \alpha_{r+\tilde r}$ are distinct modulo $2\pi i/\log q$ then the denominators $\prod_{i_1\in S}\prod_{i_2\notin S}(1-q^{\alpha_{i_1}-\alpha_{i_2}})$ are nonvanishing and so the terms in the sum over S in Theorem 1.4 are individually holomorphic in a neighborhood of $\alpha_1, \dots, \alpha_{r+\tilde r}$. This means summing them as meromorphic functions and analytically continuing is the same as summing them normally. Thus (1-3) is simply a linear combination of the functions $\prod_{i \notin S} q^{\alpha_i(n-1)}$ for distinct subsets S, i.e., it is a quasiperiodic function of n.

To prove part (1), it is simplest to view each α_i as a linear function of a distinct variable t and set $t = 0$. Because each term $(1 - q^{\alpha_{i_1} - \alpha_{i_2}})$ vanishes to order 1 at $t = 0$, the product $\prod_{i_1 \in S} \prod_{i_2 \notin S} (1 - q^{\alpha_{i_1} - \alpha_{i_2}})$ vanishes to order $|S|(r + \tilde{r} - |S|) = r\tilde{r}$ at $t = 0$. Thus

$$\sum_{\substack{f_1, \dots, f_{r+\tilde{r}} \in \mathbb{F}_q[T] \\ \text{monic} \\ \prod_{i \in S} f_i / \prod_{i \notin S} f_i \in T^{\mathbb{Z}}}} \prod_{i \in S} q^{(-1/2+\alpha_i) \deg f_i} \prod_{i \notin S} q^{(-1/2-\alpha_i) \deg f_i}$$

is a holomorphic function of t divided by $t^{r\tilde{r}}$. On the other hand, the coefficient of t^j in the Taylor series for $\prod_{i \notin S} q^{\alpha_i(n-1)}$ is a polynomial in n of degree $\leq j$. Multiplying by a holomorphic function in t preserves this property, as does summing, so all told the main term is a power series in t whose j-th coefficient is a polynomial in n of degree $\leq j$, all divided by $t^{r\tilde{r}}$. The value at $t = 0$ is the coefficient of $t^{r\tilde{r}}$ in the numerator, which is a polynomial in n of degree $\leq r\tilde{r}$.

Furthermore, we can calculate the leading coefficient of this polynomial in n. From the proof of polynomiality, we can see that the holomorphic function (3-14) enters into the calculation only by its value when $\alpha_1, \dots, \alpha_{r+\tilde{r}} = 0$. This value is independent of S, by permuting the variables. Call it $a_{r,\tilde{r}}$. Because only the value at $\alpha_1, \dots, \alpha_{r+\tilde{r}} = 0$ is significant, we can replace the term

$$\sum_{\substack{f_1, \dots, f_{r+\tilde{r}} \in \mathbb{F}_q[T] \\ \text{monic} \\ \prod_{i \in S} f_i / \prod_{i \notin S} f_i \in T^{\mathbb{Z}}}} \prod_{i \in S} q^{(-1/2+\alpha_i) \deg f_i} \prod_{i \notin S} q^{(-1/2-\alpha_i) \deg f_i}$$

with

$$\frac{a_{r,\tilde{r}}}{\prod_{i_1 \in S} \prod_{i_2 \notin S} (1 - q^{\alpha_{i_1} - \alpha_{i_2}})}$$

without affecting the leading coefficient of the polynomial in n. After replacing these terms in (1-3), we obtain

$$a_{r,\tilde{r}} \sum_{\substack{S \subseteq \{1, \dots, r+\tilde{r}\} \\ |S| = r}} \frac{\prod_{i \notin S} q^{\alpha_i(n-1)}}{\prod_{i_1 \in S} \prod_{i_2 \notin S} (1 - q^{\alpha_{i_1} - \alpha_{i_2}})}. \tag{3-15}$$

The sum in (3-15) is also the Weyl character formula for the representation of $\mathrm{GL}_{r+\tilde{r}}$ with highest weights 0 repeated r times and $n-1$ repeated s times, evaluated at the diagonal element with eigenvalues q^{α_i}. So when $\alpha_i = 0$, the sum in (3-15) is simply the dimension of this representation, which by the Weyl dimension formula is a polynomial in n with leading term $n^{r\tilde{r}} \prod_{j=0}^{r-1} j!/(j+\tilde{r})!$. Multiplying by $a_{r,\tilde{r}}$, the leading term in (3-15) is exactly (3-12), except that we have to prove the Euler product formula (3-13) for $a_{r,\tilde{r}}$.

To do this, we follow the method of Lemma 3.4 and use the Euler products (3-6) and (3-7).

When we combine these two Euler products, the factor for every $\pi \neq T$ is

$$\left(\prod_{i_1 \in S, i_2 \notin S} (1 - |\pi|^{-1+\alpha_{i_1}-\alpha_{i_2}})\right) \sum_{\substack{d_1,\dots,d_{r+\tilde{r}} \in \mathbb{N} \\ \sum_{i\in S} d_i = \sum_{i \notin S} d_i}} |\pi|^{-\sum_{i\in S} d_i + \sum_{i \in S} \alpha_i d_i - \sum_{i\notin S} \alpha_i d_i} = 1 + O(|\pi|^{-2})$$

so this Euler product converges. When we specialize $\alpha_1, \dots, \alpha_{r+\tilde{r}} = 0$, the Euler factor for $\pi \neq T$ becomes

$$(1 - |\pi|^{-1})^{r\tilde{r}} \sum_{\substack{d_1,\dots,d_{r+\tilde{r}} \in \mathbb{N} \\ \sum_{i\in S} d_i = \sum_{i \notin S} d_i}} |\pi|^{-\sum_{i\in S} d_i}.$$

For each natural number e, there are $\binom{e+r-1}{r-1}$ nonnegative integer solutions to $\sum_{i\in S} d_i = e$ and $\binom{e+\tilde{r}-1}{\tilde{r}-1}$ nonnegative integer solutions to $\sum_{i\notin S} d_i = e$ and so the coefficient of π^e in

$$\sum_{\substack{d_1,\dots,d_{r+\tilde{r}} \in \mathbb{N} \\ \sum_{i\in S} d_i = \sum_{i \notin S} d_i}} |\pi|^{-\sum_{i\in S} d_i}$$

is $\binom{e+r-1}{r-1}\binom{e+\tilde{r}-1}{\tilde{r}-1}$.

When we specialize $\alpha_1, \dots, \alpha_{r+\tilde{r}} = 0$, the Euler factor for $\pi = T$ becomes

$$(1 - q^{-1})^{r\tilde{r}}(1 - q^{-1/2})^{r+\tilde{r}}.$$

Combining these Euler factors, we get (3-13). □

Lemma 3.8. *When $r = \tilde{r} = 2$, (1-3) specializes to*

$$\sum_{\substack{S \subseteq \{1,\dots,4\} \\ |S|=2}} \prod_{i \notin S} q^{\alpha_i(n-1)} \frac{1 - q^{-1-\sum_{i\in S}\alpha_i + \sum_{i\notin S}\alpha_i}}{1 - q^{-2-\sum_{i\in S}\alpha_i + \sum_{i\notin S}\alpha_i}} \prod_{i\in S} \frac{1}{1 - q^{-1/2-\alpha_i}} \prod_{i \notin S} \frac{1}{1 - q^{-1/2+\alpha_i}} \prod_{i_1 \in S, i_2 \notin S} \frac{1 - q^{-1-\alpha_{i_1}+\alpha_{i_2}}}{1 - q^{-\alpha_{i_1}+\alpha_{i_2}}}. \tag{3-16}$$

Proof. Let us consider first the term where $S = \{1, 2\}$.

We have

$$\sum_{\substack{f_1,f_2,f_3,f_4 \in \mathbb{F}_q[T] \\ \text{monic} \\ f_1 f_2 / f_3 f_4 \in T^{\mathbb{Z}}}} |f_1|^{-\frac{1}{2}-\alpha_1} |f_2|^{-\frac{1}{2}-\alpha_2} |f_3|^{-\frac{1}{2}+\alpha_3} |f_4|^{-\frac{1}{2}+\alpha_4} =$$

$$\frac{1}{(1-q^{-\frac{1}{2}-\alpha_1})(1-q^{-\frac{1}{2}-\alpha_2})(1-q^{-\frac{1}{2}+\alpha_3})(1-q^{-\frac{1}{2}+\alpha_4})} \sum_{\substack{f_1,f_2,f_3,f_4 \in \mathbb{F}_q[T] \\ \text{monic} \\ \text{prime to } T \\ f_1 f_2 = f_3 f_4}} |f_1|^{-\frac{1}{2}-\alpha_1} |f_2|^{-\frac{1}{2}-\alpha_2} |f_3|^{-\frac{1}{2}+\alpha_3} |f_4|^{-\frac{1}{2}+\alpha_4}. \tag{3-17}$$

Now

$$\sum_{\substack{f_1,f_2,f_3,f_4\in\mathbb{F}_q[T]\\ \text{monic}\\ \text{prime to } T\\ f_1f_2=f_3f_4}} |f_1|^{-1/2-\alpha_1}|f_2|^{-1/2-\alpha_2}|f_3|^{-1/2+\alpha_3}|f_4|^{-1/2+\alpha_4} = \prod_{\substack{\pi\in\mathbb{F}_q[T]\\ \text{monic}\\ \text{irreducible}\\ \pi\neq T}} \sum_{\substack{e_1,e_2,e_3,e_4\in\mathbb{N}\\ e_1+e_2=e_3+e_4}} |\pi|^{-(e_1+e_2+e_3+e_4)/2-\alpha_1e_1-\alpha_2e_2+\alpha_3e_3+\alpha_4e_4} \tag{3-18}$$

Given $e_1, e_2, e_3, e_4 \in \mathbb{N}$ with $e_1+e_2 = e_3+e_4$, the number of ways of writing $e_1 = a+b$, $e_2 = c+d$, $e_3 = a+c$, $e_4 = b+d$ with $a, b, c, d \in \mathbb{N}$ is $\min(e_1, e_2, e_3, e_4)+1$ as we must have $b = e_1-a$, $c = e_3-a$, $d = e_2-e_3+a = e_4-e_1+a$ and the valid a are in the interval $[\max(e_3-e_2, 0), \min(e_1, e_3)]$ whose length is $\min(e_1, e_2, e_3, e_4)$.

Hence the number of solutions for e_1, e_2, e_3, e_4 minus the number of solutions for e_1-1, e_2-1, e_3-1, e_4-1 is exactly 1 if e_1, e_2, e_3, e_4 are nonnegative with $e_1+e_2 = e_3+e_4$ and zero otherwise. This gives

$$\sum_{\substack{e_1,e_2,e_3,e_4\in\mathbb{N}\\ e_1+e_2=e_3+e_4}} |\pi|^{-(e_1+e_2+e_3+e_4)/2-\alpha_1e_1-\alpha_2e_2+\alpha_3e_3+\alpha_4e_4} = \frac{1-|\pi|^{-2-\alpha_1-\alpha_2+\alpha_3+\alpha_4}}{(1-|\pi|^{-1-\alpha_1+\alpha_3})(1-|\pi|^{-1-\alpha_1+\alpha_4})(1-|\pi|^{-1-\alpha_2+\alpha_3})(1-|\pi|^{-1-\alpha_2+\alpha_4})}.$$

Hence using the zeta function of $\mathbb{F}_q[T]$, (3-18) is

$$\frac{1-q^{-1-\alpha_1-\alpha_2+\alpha_3+\alpha_4}}{1-q^{-2-\alpha_1-\alpha_2+\alpha_3+\alpha_4}} \prod_{i_1\in\{1,2\}, i_2\in\{3,4\}} \frac{1-q^{-1-\alpha_{i_1}+\alpha_{i_2}}}{1-q^{-\alpha_{i_1}+\alpha_{i_2}}}$$

and thus (3-17) is

$$\frac{1}{(1-q^{-1/2-\alpha_1})(1-q^{-1/2-\alpha_2})(1-q^{-1/2+\alpha_3})(1-q^{-1/2+\alpha_4})} \frac{1-q^{-1-\alpha_1-\alpha_2+\alpha_3+\alpha_4}}{1-q^{-2-\alpha_1-\alpha_2+\alpha_3+\alpha_4}} \times \prod_{i_1\in\{1,2\}, i_2\in\{3,4\}} \frac{1-q^{-1-\alpha_{i_1}+\alpha_{i_2}}}{1-q^{-\alpha_{i_1}+\alpha_{i_2}}}.$$

Now if $S \neq \{1, 2\}$, we get the same formula, except with the variables permuted by some fixed permutation $\sigma \in S_4$ sending $\{1, 2\}$ to S. Summing over the possible choices of S, we obtain (3-16). □

4. Conclusion

We prove here a slightly more general version of the main theorem.

Proposition 4.1. *Assume $n \geq 3$, if $n = 3$ that the characteristic of $\mathbb{F}_q$ is not 2 or 5, and if $n = 4$ or 5 that the characteristic of $\mathbb{F}_q$ is not 2.*

Assume Hypothesis $\mathrm{H}(n, r, \tilde{r}, w)$. *Let* $\alpha_1, \dots, \alpha_{r+\tilde{r}}$ *be imaginary. Let* $C_{r,\tilde{r}} = (\max(r, \tilde{r}) + 2)^{\max(r,\tilde{r})+1}$. *Then*

$$\prod_{1\le i_1<i_2\le r+\tilde{r}} (q^{\alpha_{i_1}} - q^{\alpha_{i_2}}) \sum_{\chi\in S_{n,q}} \epsilon_\chi^{-\tilde{r}} \prod_{i=1}^{r+\tilde{r}} L(1/2-\alpha_i, \chi) \tag{4-1}$$

$$= \sum_{\substack{S\subseteq\{1,\dots,r+\tilde{r}\}\\ m \mid r-|S|}} (q^n - q^{n-1})\mu^{(r-|S|)/m} \prod_{i\notin S} q^{\alpha_i(n-1)} M_S(\alpha_1, \dots, \alpha_{r+\tilde{r}}) \tag{4-2}$$

$$+O\biggl(\biggl(\prod_{1\le i_1<i_2\le r+\tilde{r}} |q^{\alpha_{i_1}} - q^{\alpha_{i_2}}|\biggr) n^{r+\tilde{r}} C_{r+\tilde{r}}^{n-1} q^{(n+w)/2}\biggr). \tag{4-3}$$

Proof. For $d_1, \dots, d_{r+\tilde{r}} \in \mathbb{Z}$, let $T(d_1, \dots, d_{r+\tilde{r}})$ be the coefficient of $\prod_{i=1}^{r+\tilde{r}} q^{\alpha_i d_i}$ in (4-1) and let $R(d_1, \dots, d_{r+\tilde{r}})$ be the coefficient of $\prod_{i=1}^{r+\tilde{r}} q^{\alpha_i d_i}$ in (4-2).

The two sides (4-1) and (4-2) are antisymmetric in the variables $\alpha_1, \dots, \alpha_{r+\tilde{r}}$. For (4-1) this is clear and for (4-2), this follows from Lemma 3.2. Thus we can write (4-1) as

$$\sum_{d_1<\cdots<d_{r+\tilde{r}}} \sum_{\sigma\in S_{r+\tilde{r}}} \operatorname{sgn}(\sigma) T(d_1, \dots, d_{r+\tilde{r}}) q^{\sum_{i=1}^{r+\tilde{r}} \alpha_i d_{\sigma(i)}}$$

and (4-2) as

$$\sum_{d_1<\cdots<d_{r+\tilde{r}}} \sum_{\sigma\in S_{r+\tilde{r}}} \operatorname{sgn}(\sigma) R(d_1, \dots, d_{r+\tilde{r}}) q^{\sum_{i=1}^{r+\tilde{r}} \alpha_i d_{\sigma(i)}}$$

so the difference of (4-1) and (4-2) is

$$\sum_{d_1<\cdots<d_{r+\tilde{r}}} (T(d_1, \dots, d_{r+\tilde{r}}) - R(d_1, \dots, d_{r+\tilde{r}})) \sum_{\sigma\in S_{r+\tilde{r}}} \operatorname{sgn}(\sigma) q^{\sum_{i=1}^{r+\tilde{r}} \alpha_i d_{\sigma(i)}}.$$

By the Weyl character formula,

$$\frac{\sum_{\sigma\in S_{r+\tilde{r}}} \operatorname{sgn}(\sigma) q^{\sum_{i=1}^{r+\tilde{r}} \alpha_i d_{\sigma(i)}}}{q^{\alpha_{i_1}} - q^{\alpha_{i_2}}}$$

is the trace of a diagonal matrix with entries $q^{\alpha_1}, \dots, q^{\alpha_{r+\tilde{r}}}$ on the irreducible representation of $\mathrm{GL}_{r+\tilde{r}}$ with highest weight vector $(d_{r+\tilde{r}} + 1 - r - \tilde{r}, \dots, d_2 - 1, d_1)$ [Fulton and Harris 1991, Theorem 24.2]. Because this diagonal matrix has eigenvalues of norm 1, the absolute value of its trace is at most the dimension of that highest weight representation, which is $\prod_{1\le i_1<i_2\le r+\tilde{r}} (d_{i_2} - d_{i_1})$ [loc. cit., Corollary 24.6]. This implies the estimate

$$\biggl|\sum_{\sigma\in S_{r+\tilde{r}}} \operatorname{sgn}(\sigma) q^{\sum_{i=1}^{r+\tilde{r}} \alpha_i d_{\sigma(i)}}\biggr| \le \prod_{1\le i_1<i_2\le r+\tilde{r}} |q^{\alpha_{i_1}} - q^{\alpha_{i_2}}| \frac{\prod_{1\le i_1<i_2\le r+\tilde{r}} |d_{i_1} - d_{i_2}|}{\prod_{i=1}^{r+\tilde{r}} (i-1)!} \tag{4-4}$$

Using (4-4), it suffices to prove that

$$\sum_{d_1<\cdots<d_{r+\tilde r}} |T(d_1,\dots,d_{r+\tilde r})-R(d_1,\dots,d_{r+\tilde r})| \frac{\prod_{1\le i_1<i_2\le r+\tilde r}|d_{i_1}-d_{i_2}|}{\prod_{i=1}^{r+\tilde r}(i-1)!} = O(n^{r+\tilde r}C_{r+\tilde r}^{n-1}q^{(n+w)/2}) \quad (4\text{-}5)$$

Let $R_S(d_1,\dots,d_{r+\tilde r})$ be the coefficient of $\prod_{i=1}^{r+\tilde r} q^{\alpha_i d_i}$ in

$$(q^n-q^{n-1})\mu^{(r-|S|)/m}\prod_{i\notin S} q^{\alpha_i(n-1)} M_S(\alpha_1,\dots,\alpha_{r+\tilde r})$$

so that

$$R(d_1,\dots,d_{r+\tilde r}) = \sum_{\substack{S\subseteq\{1,\dots,r+\tilde r\}\\ m\,|\,r-|S|}} R_S(d_1,\dots,d_{r+\tilde r}). \quad (4\text{-}6)$$

By Lemma 3.6,

$$T(d_1,\dots,d_{r+\tilde r}) = R_S(d_1,\dots,d_{r+\tilde r}) \quad (4\text{-}7)$$

as long as

$$0\le \sum_{i\in S} d_i - \binom{|S|}{2}, \quad \sum_{i\notin S}(n-1-d_i)+\binom{r+\tilde r}{2}-\binom{|S|}{2}\le n-1. \quad (4\text{-}8)$$

We can distinguish three types of tuples $d_1,\dots,d_{r+\tilde r}$. The first is where (4-8) is not satisfied for any S, the second where (4-8) is satisfied for a unique S, and the third where (4-8) is satisfied for more than one S. In the first case, by (4-6), we have

$$|T(d_1,\dots,d_{r+\tilde r})-R(d_1,\dots,d_{r+\tilde r}))| \le |T(d_1,\dots,d_{r+\tilde r})| + \sum_{\substack{S\subseteq\{1,\dots,r+\tilde r\}\\ m\,|\,r-|S|}} |R_S(d_1,\dots,d_{r+\tilde r})|.$$

In the second case, by (4-6) and (4-7), we have

$$|T(d_1,\dots,d_{r+\tilde r})-R(d_1,\dots,d_{r+\tilde r}))| \le \sum_{\substack{S\subseteq\{1,\dots,r+\tilde r\}\\ m\,|\,r-|S|\\ \text{(4-8) does not hold}}} |R_S(d_1,\dots,d_{r+\tilde r})|.$$

In the third case, by (4-6) and (4-7), we have

$$|T(d_1,\dots,d_{r+\tilde r})-R(d_1,\dots,d_{r+\tilde r}))| \le \sum_{\substack{S\subseteq\{1,\dots,r+\tilde r\}\\ m\,|\,r-|S|}} |R_S(d_1,\dots,d_{r+\tilde r})|$$

where we have added the canceled R_S term on the right side back in to simplify the expression, without affecting the validity of the inequality.

Combining all these, we have

$$\sum_{d_1<\cdots<d_{r+\tilde r}} |T(d_1,\dots,d_{r+\tilde r})-R(d_1,\dots,d_{r+\tilde r}))| \frac{\prod_{1\le i_1<i_2\le r+\tilde r}|d_{i_1}-d_{i_2}|}{\prod_{i=1}^{r+\tilde r}(i-1)!}$$

$$\le \sum_{\substack{d_1<\cdots<d_{r+\tilde r}\\ \text{(4-8) does not hold for any } S}} |T(d_1,\dots,d_{r+\tilde r})| \frac{\prod_{1\le i_1<i_2\le r+\tilde r}|d_{i_1}-d_{i_2}|}{\prod_{i=1}^{r+\tilde r}(i-1)!} \tag{4-9}$$

$$+\sum_{\substack{S\subseteq\{1,\dots,r+\tilde r\}\\ m\,|\,r-|S|}} (q^n-q^{n-1}) \sum_{\substack{d_1<\cdots<d_{r+\tilde r}\\ \text{(4-8) does not hold}}} |R_S(d_1,\dots,d_{r+\tilde r})| \frac{\prod_{1\le i_1<i_2\le r+\tilde r}|d_{i_1}-d_{i_2}|}{\prod_{i=1}^{r+\tilde r}(i-1)!} \tag{4-10}$$

$$+\sum_{\substack{S\subseteq\{1,\dots,r+\tilde r\}\\ m\,|\,r-|S|}} (q^n-q^{n-1}) \sum_{\substack{d_1<\cdots<d_{r+\tilde r}\\ \text{(4-8) holds for } S\\ \text{(4-8) holds for some } S'\neq S}} |R_S(d_1,\dots,d_{r+\tilde r})| \frac{\prod_{1\le i_1<i_2\le r+\tilde r}|d_{i_1}-d_{i_2}|}{\prod_{i=1}^{r+\tilde r}(i-1)!} \tag{4-11}$$

We next will prove in Lemmas 4.6, 4.5, and 4.4, bounds for (4-9), (4-10), and (4-11) respectively. This gives

$$\text{(4-9)}+\text{(4-10)}+\text{(4-11)} \le O(n^{O(1)}q^{n/2})+O(n^{O(1)}q^{n/2})+O(n^{r+\tilde r}C^{n-1}_{r+\tilde r}q^{(n+w)/2}) = O(n^{r+\tilde r}C^{n-1}_{r+\tilde r}q^{(n+w)/2})$$

since $n^{O(1)}=O(C^{n-1}_{r,\tilde r})$, which is the desired bound (4-5). □

Lemma 4.2. *For every $S\subseteq\{1,\dots,r+\tilde r\}$ with $m\mid r-|S|$, for each tuple $d_1<\cdots<d_{r+\tilde r}\in\mathbb{Z}$ that satisfies* (4-8), *if $R_S(d_1,\dots,d_{r+\tilde r})\neq 0$, then $0\le d_1<\cdots<d_{r+\tilde r}\le n+r+\tilde r-2$.*

Proof. By (4-7), the assumptions imply that $T(d_1,\dots,d_{r+\tilde r})$ is not zero. By definition $T(d_1,\dots,d_{r+\tilde r})$ is the coefficient of $q^{\sum_i d_i\alpha_i}$ in the product of $\prod_{1\le i_1<i_2\le r+\tilde r}(q^{\alpha_{i_1}}-q^{\alpha_{i_2}})$, which is a polynomial in the q^{α_i} of degree at most $r+\tilde r-1$ in each variable, with $\sum_{\chi\in S_{n,q}}\epsilon_\chi^{-s}\prod_{i=1}^{r+\tilde r}L\big(\tfrac12-\alpha_i,\chi\big)$, which is a polynomial in the q^{α_i} of degree at most $n-1$ in each variable. Hence this product is a polynomial of degree $\le n+r+\tilde r-2$ in each variable, and thus the coefficient $T(d_1,\dots,d_{r+\tilde r})$ can only be nonzero if $0\le d_1,\dots,d_{r+\tilde r}\le n+r+\tilde r-2$. □

Lemma 4.3. *For every pair $S, S'\subseteq\{1,\dots,r+\tilde r\}$ with $m\mid r-|S|$, $m\mid r-|S'|$, and $S\neq S'$, for each tuple $d_1<\cdots<d_{r+\tilde r}\in\mathbb{Z}$ that satisfies* (4-8), *we have $R_S(d_1,\dots,d_{r+\tilde r})=O(n^{O(1)}q^{n/2})$*

Proof. We may assume without loss of generality that $R_S(d_1,\dots,d_{r+\tilde r})\neq 0$. We will apply Lemma 3.5 to bound this term, noting by Lemma 4.2 that the factor $(O(1)+\sum_{i\in S}d_i-\sum_{i\notin S}d_i)^{O(1)}$ appearing in Lemma 3.5 is $n^{O(1)}$.

We can freely use the consequence of (4-7) that

$$R_S(d_1,\dots,d_{r+\tilde r})=T(d_1,\dots,d_{r+\tilde r})=R_{S'}(d_1,\dots,d_{r+\tilde r})$$

to pass between S and S'.

We will split into two cases depending on if $|S'|=|S|$ or not.

If $|S'| = |S|$, we choose i_1 in S but not S' and i_2 in S' but not S. We have

$$(d_{i_1} + (n-1-d_{i_2})) + ((n-1-d_{i_1}) + d_{i_2}) \geq 2(n-1),$$

so one is at least $n-1$. Applying the first part of Lemma 3.5 for S if the first one is smaller and S' if the second one is smaller, we see that

$$R_S(d_1, \dots, d_{r+\tilde{r}}) = O(n^{O(1)} q^{n-n/2}).$$

If $|S'| \neq |S|$, assume without loss of generality that $|S| < |S'|$. Then by Lemma 2.10, $|S'| - |S| \geq m \geq 3$. Hence

$$\left(\sum_{i\in S} d_i - \binom{r+\tilde{r}}{2} - (r+\tilde{r}-|S|)(n-1)\right) - \left(\sum_{i\in S'} d_i - \binom{r+\tilde{r}}{2} - (r+\tilde{r}-|S'|)(n-1)\right) \geq 3(n-1) > 2n-2$$

so one of these two terms must have absolute value at least n. Without loss of generality, it is the term associated to S. Then by the fourth part of Lemma 3.5, $R_S(d_1, \dots, d_{r+\tilde{r}}) = O(n^{O(1)} q^{n-n/2})$. □

Lemma 4.4. *The sum* (4-11) *is* $O(n^{O(1)} q^{n/2})$.

Proof. By Lemma 4.2, the number of nonzero terms in (4-11) is $n^{O(1)}$, and the factor $\prod_{1\leq i_1<i_2\leq r+\tilde{r}} |d_{i_1} - d_{i_2}|$ appearing in each term is $n^{O(1)}$. By Lemma 4.3, the factor $R_s(d_1, \dots, d_{r+\tilde{r}})$ appearing in each term is $O(n^{O(1)} q^{n/2})$. So the sum over all terms is $O(n^{O(1)} q^{n/2})$. □

Next we handle (4-10).

Lemma 4.5. *The sum* (4-10) *is* $O(n^{O(1)} q^{n/2})$.

Proof. Fix $d_1 < \cdots < d_{r+\tilde{r}}$ that contribute nontrivially to (4-10), i.e., (4-8) is violated and $R_S(d_1, \dots, d_{r+\tilde{r}})$ is nonzero. By Lemma 3.2, the nonvanishing implies that

$$0 \leq \sum_{i\in S} d_i - \binom{|S|}{2}, 0 \leq \sum_{i\notin S} (n-1-d_i) + \binom{r+\tilde{r}}{2} - \binom{|S|}{2},$$

so because (4-8) is violated, we must have either

$$\sum_{i\in S} d_i - \binom{|S|}{2} \geq n$$

or

$$\sum_{i\notin S} (n-1-d_i) + \binom{r+\tilde{r}}{2} - \binom{|S|}{2} \geq n.$$

Let $k = \max(\sum_{i\in S} d_i - \binom{|S|}{2}, \sum_{i\notin S}(n-1-d_i) + \binom{r+\tilde{r}}{2} - \binom{|S|}{2})$. Then there are $k^{O(1)}$ tuples $d_1, \dots, d_{r+\tilde{r}}$ leading to a given k, each of which has coefficients $O(k^{O(1)} q^{n-k/2})$ by the second and third parts of Lemma 3.5, and $\prod_{1\leq i_1<i_2\leq r+\tilde{r}} |d_{i_1} - d_{i_2}| = k^{O(1)}$, so in total (4-10) is $\sum_{k\geq n} k^{O(1)} q^{n-k/2} = O(n^{O(1)} q^{n/2})$. □

Lemma 4.6. *The sum* (4-9) *is* $O(n^{r+\tilde{r}} C^{n-1}_{r+\tilde{r}} q^{(n+w)/2})$.

Proof. By Lemma 2.7,

$$T(d_1, d_2+1, \dots, d_{r+\tilde{r}}+r+\tilde{r}-1) = \pm F(V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde{r}}})$$

With this renormalized set of coefficients, (4-9) is a sum over tuples $0 \le d_1 \le \cdots \le d_{r+\tilde{r}} \le n-1$ such that for no S of cardinality congruent to r mod m do we have $0 \le \sum_{i\in S} d_i, \sum_{i\notin S}(n-1-d_i) \le n-1$. Hence by Lemma 2.9, for these tuples $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde{r}}}(L_{\text{univ}})$ does not appear as a summand of $L_{\text{univ}}^{\otimes a} \otimes L_{\text{univ}}^{\vee\otimes b}$ for $0 \le a, b \le n-1$.

By definition, we have

$$F(V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde{r}}}) = \sum_{j\in\mathbb{Z}} (-1)^j \operatorname{tr}\bigl(\operatorname{Frob}_q, H_c^j\bigl(\operatorname{Prim}_{n,\overline{\mathbb{F}}_q}, V_{d_1,\dots,d_r|d_{r+1},\dots,r_{r+\tilde{r}}}\bigl(L_{\text{univ}}\bigl(\tfrac{1}{2}\bigr)\bigr)\bigr)\bigr).$$

Because $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde{r}}}(L_{\text{univ}})$ is irreducible and because it appears as a summand of $\det^{-s}(L_{\text{univ}}) \otimes \bigotimes_{i=1}^{r+\tilde{r}} \wedge^{d_i}(L_{\text{univ}})$ and does not appear as a summand of $L_{\text{univ}}^{\otimes a} \otimes L_{\text{univ}}^{\vee\otimes b}$ for $0 \le a, b \le n-1$, we may apply Hypothesis H$(n, r, \tilde{r}, w)$ to get $H_c^j(\operatorname{Prim}_{n,\overline{\mathbb{F}}_q}, V_{d_1,\dots,d_r|d_{r+1},\dots,r_{r+\tilde{r}}}(L_{\text{univ}}(\frac{1}{2}))) = 0$ for $j > n+w$.

Because the L-functions $L(s, \chi)$ satisfy the Riemann hypothesis, the eigenvalues of Frobenius on L_{univ} have size $\sqrt{q}$ (see Definition 1.2), and so L_{univ} is pure of weight 1. This implies that $L_{\text{univ}}(\frac{1}{2})$ is pure of weight 0 and thus $V_{d_1,\dots,d_r|d_{r+1},\dots,r_{r+\tilde{r}}}(L_{\text{univ}}(\frac{1}{2}))$ is also pure of weight 0. By Deligne's Riemann hypothesis [Deligne 1980, Theorem 1], all eigenvalues of Frob_q on $H_c^j(\operatorname{Prim}_{n,\overline{\mathbb{F}}_q}, V_{d_1,\dots,d_r|d_{r+1},\dots,r_{r+\tilde{r}}}(L_{\text{univ}}(\frac{1}{2})))$ are $\le q^{j/2} \le q^{(n+w)/2}$. Thus because the trace of an endomorphism of a vector space is at most the dimension times the size of the greatest eigenvalue,

$$|F(V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde{r}}})| \le q^{(n+w)/2} \sum_{j\in\mathbb{Z}} \dim H_c^j(\operatorname{Prim}_{n,\overline{\mathbb{F}}_q}, V_{d_1,\dots,d_r|d_{r+1},\dots,r_{r+\tilde{r}}}(L_{\text{univ}})).$$

Hence (4-9) is at most

$$\sum_{0\le d_1\le\cdots\le d_{r+\tilde{r}}\le n-1} \frac{\prod_{1\le i_1<i_2\le r+\tilde{r}} |d_{i_2}+i_2-d_{i_1}-i_1|}{\prod_{i=1}^{r+\tilde{r}}(i-1)!} q^{(n+w)/2} \sum_{j\in\mathbb{Z}} \dim H_c^j(\operatorname{Prim}_{n,\overline{\mathbb{F}}_q}, V_{d_1,\dots,d_r|d_{r+1},\dots,r_{r+\tilde{r}}}(L_{\text{univ}})).$$

By the Weyl character formula and Weyl dimension formula,

$$\frac{\prod_{1\le i_1<i_2\le r+\tilde{r}} |d_{i_2}+i_2-d_{i_1}-i_1|}{\prod_{i=1}^{r+\tilde{r}}(i-1)!} = \lim_{\alpha_1,\dots,\alpha_r\to 1} \frac{\sum_{\sigma\in S_{r+\tilde{r}}} \operatorname{sgn}(\sigma) q^{(d_i+i-1)\alpha_{\sigma(i)}}}{\prod_{1\le i_1<i_2\le r+\tilde{r}}(q^{\alpha_{i_2}}-q^{\alpha_{i_1}})}$$

which, by Lemma 2.6, is the multiplicity that $V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde{r}}}$ appears in

$$\bigoplus_{0\le e_1,\dots,e_{r+\tilde{r}}\le n-1} \det{}^{-\tilde{r}} \otimes \bigotimes_{i=1}^{r+\tilde{r}} \wedge^{e_i}.$$

Hence

$$\sum_{0\le d_1\le d_2\le\cdots\le d_{r+\tilde r}\le n-1}\frac{\prod_{1\le i_1<i_2\le r+\tilde r}|d_{i_2}+i_2-d_{i_1}-i_1|}{\prod_{i=1}^{r+\tilde r}(i-1)!}\sum_j \dim H^j_c(\mathrm{Prim}_{n,\overline{\mathbb F}_q}, V_{d_1,\dots,d_r|d_{r+1},\dots,d_{r+\tilde r}}(L_{\mathrm{univ}}))$$
$$\le \sum_{0\le e_1,\dots,e_{r+\tilde r}\le n-1}\dim H^j_c(\mathrm{Prim}_{n,\overline{\mathbb F}_q}, \det(L_{\mathrm{univ}})^{-\tilde r}\otimes\bigotimes_{i=1}^{r+\tilde r}\wedge^{e_i}(L_{\mathrm{univ}})).$$

The total number of terms here is $n^{r+\tilde r}$. For each term, we apply Lemma 2.13 and use the fact that

$$\det(L_{\mathrm{univ}})^{-\tilde r}\otimes\bigotimes_{i=1}^{r+\tilde r}\wedge^{e_i}(L_{\mathrm{univ}})=\bigotimes_{i=1}^{r}\wedge^{e_i}(L_{\mathrm{univ}})\otimes\bigotimes_{i=\tilde r+1}^{r+\tilde r}\wedge^{n-1-e_1}(L^\vee_{\mathrm{univ}}).$$

We can assume in each term that $e_1\le e_2\le\cdots\le e_{r+\tilde r}$, so that

$$\sum_{i=1}^{r}e_i+\sum_{i=r+1}^{r+\tilde r}(n-1-e_i)\le re_r+\tilde r(n-1-e_{r+1})\le re_r+\tilde r(n-1-e_r)\le(n-1)\max(r,\tilde r).$$

This gives exactly the bound $O(C^{n-1}_{r+\tilde r}q^{(n+w)/2})$ for each term and thus $O(n^{r+\tilde r}C^{n-1}_{r+\tilde r}q^{(n+w)/2})$ in total. □

Remark 4.7. It is unsurprising that, in this proof, dimensions of irreducible representations of $\mathrm{GL}_{r+\tilde r}$ appear as multiplicities of irreducible representations of GL_{n-1} in

$$\sum_{0\le e_1,\dots,e_{r+\tilde r}\le n-1}\det{}^{-\tilde r}\otimes\bigotimes_{i=1}^{r+\tilde r}\wedge^{e_i},$$

because $\otimes\bigotimes_{i=1}^{r+\tilde r}\wedge^{e_i}$ admits a natural action of $\mathrm{GL}_{n-1}\times\mathrm{GL}_{r+\tilde r}$ as it is the exterior algebra of the tensor product of the standard representations of GL_{n-1} and $\mathrm{GL}_{r+\tilde r}$, and this action is preserved after tensoring with the $-\tilde r$ power of the determinant of GL_{n-1}. This is part of the approach of Bump and Gamburd [2006] to moments of the characteristic polynomial of random matrices.

Corollary 4.8. *Assume $n\ge 3$, if $n=3$ that the characteristic of $\mathbb F_q$ is not 2 or 5, and if $n=4$ or 5 that the characteristic of $\mathbb F_q$ is not 2.*

Assume Hypothesis $\mathrm H(n,r,\tilde r,w)$. *Let $\alpha_1,\dots,\alpha_{r+\tilde r}$ be imaginary. Let $C_{r,\tilde r}=(\max(r,\tilde r)+2)^{\max(r,\tilde r)+1}$*

$$\frac{1}{q^n-q^{n-1}}\sum_{\chi\in S_{n,q}}\epsilon_\chi^{-\tilde r}\prod_{i=1}^{r+\tilde r}L\big(\tfrac12-\alpha_i,\chi\big)$$
$$=\sum_{\substack{S\subseteq\{1,\dots,r+\tilde r\}\\ m\,|\,r-|S|}}\mu^{(r-|S|)/m}\prod_{i\notin S}q^{\alpha_i(n-1)}\sum_{\substack{f_1,\dots,f_{r+\tilde r}\in\mathbb F_q[T]\\ \text{monic}\\ \prod_{i\in S}f_i/\prod_{i\notin S}f_i\in T^{\mathbb Z}}}\prod_{i\in S}|f_i|^{-1/2+\alpha_i}\prod_{i\notin S}|f_i|^{-1/2-\alpha_i}+O(q^{(w-n)/2}C^n_{r,\tilde r}).$$

If $n>2\max(r,\tilde r)+1$ then we need only the terms where $r=|S|$.

Proof. The first claim follows from Proposition 4.1 after dividing both sides by

$$(q^n - q^{n-1}) \prod_{1 \le i_1 < i_2 \le r+\tilde{r}} (q^{\alpha_{i_1}} - q^{\alpha_{i_2}}).$$

The second claim follows from Lemma 2.10, because then $m > \max(r, \tilde{r})$. □

In particular, the second claim is Theorem 1.4.

Remark 4.9. The terms (4-10) and (4-11) arise from the main term of Proposition 4.1, are totally explicit, and are of the same size as the square-root error term. Thus, they could be subtracted off, producing an equivalent statement. Should they be?

Sometimes, when there are two equivalent formulations of the main term, one is better because it conveys more information. For instance, when $\alpha_1, \dots, \alpha_{r+\tilde{r}} = 0$, our CFKRS-like main term agrees with its simpler leading term (3-12), up to an error with logarithmic savings. If we could prove an estimate with logarithmic savings, the full main term and its leading term would be equivalent. But the full main term would carry more information, because we expect it to remain true with a much smaller error term.

Because (4-10) and (4-11) have square root size, they are dominated by even the most optimistic error terms (except maybe when $r, \tilde{r}$ are very small). Thus, subtracting them off and leaving them in really do convey the same information, and so we have chosen to leave them in as that gives a simpler, and more familiar, formula.

5. Verification of the hypothesis in special cases

Lemma 5.1. *Let $\mathcal{F}$ be an irreducible lisse $\mathbb{Q}_\ell$-sheaf on $\mathrm{Prim}_{n,\bar{\mathbb{F}}_q}$ that appears as a summand of*

$$L_{\mathrm{univ}}^{\otimes a} \otimes L_{\mathrm{univ}}^{\vee \otimes b}$$

for some $a \ge n > b$. Then

$$H_c^j(\mathrm{Prim}_n, \mathcal{F}) = 0$$

for $j > n + b + \lfloor \frac{a}{p} \rfloor - \lfloor \frac{n}{p} \rfloor + 1$.

Proof. To do this, we will first examine the cohomology of the space $Z_{n,a,b}$ defined in Section 2C. We relate them to the spaces $X_{a,n,(c_1,\dots,c_n)}$ of [Sawin 2018], defined as the subspace of $\mathbb{A}^a$ with variables $x_1, \dots, x_a$ satisfying the system of n equations $\prod_{i=1}^a (1 - T a_i) = 1 + c_1 T + \cdots + c_n T^n \mod T^{n+1}$.

The key fact that we will derive from [Sawin 2018] is that the S_a action on the cohomology of $X_{a,n,(c_1,\dots,c_n)}$ is trivial in degrees greater than $a - n + \lfloor \frac{a}{p} \rfloor - \lfloor \frac{n}{p} \rfloor + 1$. To see this, first note that the cohomology in degree $> 2(a-n)$ vanishes for dimension reasons. Second, the cohomology in degree $< 2(a-n)$ is handled by the first part of [loc. cit., Proposition 2.5]. Finally the cohomology in degree is exactly $2(a-n)$ is handled by the second part of [loc. cit., Proposition 2.5], because our condition that the degree is greater than $a - n + \lfloor \frac{a}{p} \rfloor - \lfloor \frac{n}{p} \rfloor + 1$ is then equivalent to the condition $a - n > \lfloor \frac{a}{p} \rfloor - \lfloor \frac{n}{p} \rfloor + 1$ of [loc. cit., Proposition 2.5].

The space $Z_{n,a,b}$ admits a map to $\mathbb{A}^b$ by projection onto the last b coordinates, and the fibers of this map are the spaces $X_{a,n,(c_1,\dots,c_n)}$, where $c_1,\dots,c_m$ are the coefficients to the product of the last b linear factors. So by the proper base change theorem, the fact that the cohomological dimension of $\mathbb{A}^b$ is $2b$, and our key fact, the action of S_a on the cohomology of $Z_{n,a,b}$ is trivial in degrees greater than $a+2b-n\lfloor \frac{a}{p}\rfloor - \lfloor \frac{n}{p}\rfloor + 1$.

Hence by Lemma 2.11, the S_a action on

$$H_c^j(\mathrm{Witt}_{n,\overline{\mathbb{F}}_q}, (R(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}})^{\otimes a}\otimes (R(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}})^{\otimes b})$$

is trivial whenever $j > a+2b-n\lfloor \frac{a}{p}\rfloor - \lfloor \frac{n}{p}\rfloor + 1$. Applying this to $n-1$, we can see that the S_a action on

$$H_c^{j-1}(\mathrm{Witt}_{n-1,\overline{\mathbb{F}}_q}, (R(\mathrm{pr}_2)_!\mathcal{L}_{\mathrm{univ}})^{\otimes a}\otimes (R(\mathrm{pr}_2)_!\mathcal{L}^\vee_{\mathrm{univ}})^{\otimes b})$$

is trivial whenever

$$j-1 > a+2b+n-1+\left\lfloor \frac{a}{p}\right\rfloor - \left\lfloor \frac{n-1}{p}\right\rfloor + 1 \geq \left(a+2b+n+\left\lfloor \frac{a}{p}\right\rfloor - \left\lfloor \frac{n}{p}\right\rfloor + 1\right) - 1.$$

By the long exact sequence of Lemma 2.12, the action of S_a on

$$H_c^j\left(\mathrm{Prim}_{n,\overline{\mathbb{F}}_q}, \bigotimes_{i=1}^a L_{\mathrm{univ}}[-1]\otimes \bigotimes_{i=1}^b L^\vee_{\mathrm{univ}}[-1]\right)$$

is trivial for $j > a+2b+n+\lfloor \frac{a}{p}\rfloor - \lfloor \frac{n}{p}\rfloor + 1$. Hence by shifting, the action of S_a on

$$H_c^j\left(\mathrm{Prim}_{n,\overline{\mathbb{F}}_q}, \bigotimes_{i=1}^a L_{\mathrm{univ}}\otimes \bigotimes_{i=1}^b L^\vee_{\mathrm{univ}}\right)$$

factors through the sign character if $j > n+b+\lfloor \frac{a}{p}\rfloor - \lfloor \frac{n}{p}\rfloor + 1$. But the sign-equivariant part of $\bigotimes_{i=1}^a L_{\mathrm{univ}}$ is $\wedge^a L_{\mathrm{univ}}$, which vanishes, so the sign-equivariant part of the cohomology vanishes as well. Thus in fact

$$H_c^j\left(\mathrm{Prim}_{n,\overline{\mathbb{F}}_q}, \bigotimes_{i=1}^a L_{\mathrm{univ}}\otimes \bigotimes_{i=1}^b L^\vee_{\mathrm{univ}}\right) = 0$$

if $j > n+b+\lfloor \frac{a}{p}\rfloor - \lfloor \frac{n}{p}\rfloor + 1$. Hence the same is true for any summand of $\bigotimes_{i=1}^a L_{\mathrm{univ}}\otimes\bigotimes_{i=1}^b L^\vee_{\mathrm{univ}}$, such as $\mathcal{F}$. □

Lemma 5.2. *Hypothesis* $\mathrm{H}\big(n, r, 0, \lfloor \frac{(n-1)r}{p}\rfloor - \lfloor \frac{n}{p}\rfloor + 1\big)$ *is satisfied for any* n, r.

Proof. Any sheaf $\mathcal{F}$ that is a summand of $\bigotimes_{i=1}^r \wedge^{d_i}(L_{\mathrm{univ}})$ for some $0\le d_1,\dots,d_r\le n-1$ is a summand of $L_{\mathrm{univ}}^{\otimes(\sum_{i=1}^r d_r)}$. If $\sum_{i=1}^r d_i < n$ then the condition of Hypothesis 1.3 is not satisfied, so it is vacuously true. Otherwise, we apply Lemma 5.1 with $a=\sum_{i=1}^r d_i \le r(n-1)$ and $b=0$. □

However, combining this with Proposition 4.1 would simply recover [Sawin 2018, Theorems 1.2 and 1.3].

Lemma 5.3. *Hypothesis* $\mathrm{H}(n, r, 1, w = n+1-n(p-2r)/(pr))$ *is satisfied for any* n, r.

Proof. We may assume $p > 2r$ as if $p \leq 2r$ then the claim follows immediately from the fact that $H^j(\mathrm{Prim}_{n,\bar{\mathbb{F}}_q}, \mathcal{F}) = 0$ for $j > 2n$ because $\dim \mathrm{Prim}_n = n$.

Let $\mathcal{F}$ be a summand of $\det^{-1}(L_{\mathrm{univ}}) \otimes \bigotimes_{i=1}^{r+1} \wedge^{d_i}(L_{\mathrm{univ}})$ for some $0 \leq d_1, \dots, d_{r+1} \leq n-1$. Without loss of generality, $0 \leq d_1 \leq \cdots \leq d_r \leq d_{r+1} \leq n-1$. Then because $\det^{-1}(L_{\mathrm{univ}}) \otimes \wedge^{d_{r+1}}(L_{\mathrm{univ}}) = \wedge^{n-1-r}(L_{\mathrm{univ}}^\vee)$, $\mathcal{F}$ is a summand of $L_{\mathrm{univ}}^{\otimes\left(\sum_{i=1}^r d_i\right)} \otimes L_{\mathrm{univ}}^{\vee\otimes(n-1-d_{r+1})}$. If $\sum_{i=1}^r d_i < n$ then Hypothesis $\mathrm{H}(n, r, 1, n+1-n(p-2r)/(pr))$ is vacuously true. Otherwise, $d_{r+1} \geq d_r \geq n/r$ and we apply Lemma 5.1 to see that the cohomology groups vanish for

$$j > n + (n-1-d_{r+1}) + \left\lfloor \frac{\sum_{i=1}^r d_r}{p} \right\rfloor - \left\lfloor \frac{n}{p} \right\rfloor + 1,$$

and we have

$$\begin{aligned}(n-1-d_{r+1}) + \left\lfloor \frac{\sum_{i=1}^r d_r}{p} \right\rfloor - \left\lfloor \frac{n}{p} \right\rfloor + 1 &\leq n-1-d_{r+1} + \frac{\sum_{i=1}^r d_r}{p} - \frac{n}{p} + 2 \\ &\leq n-1-\frac{p-r}{p} d_{r+1} - \frac{n}{p} + 2 \\ &\leq n-1-\frac{p-r}{pr} n - \frac{n}{p} + 2 \\ &= n+1-\frac{p-2r}{pr} n\end{aligned}$$

as desired. □

We could apply the same technique with $r, \tilde{r} \geq 2$ but we would not obtain a nontrivial bound this way.

Corollary 5.4. *Let n, r be natural numbers and $\mathbb{F}_q$ a finite field of characteristic p. Assume also that $n > 2r+1$ and if $n = 4$ or 5 that the characteristic of $\mathbb{F}_q$ is not 2. Let $C_{r,1} = (2+r)^{r+1}$. Let $\alpha_1, \dots, \alpha_{r+1}$ be imaginary numbers. Let ϵ_χ be the ϵ-factor of $L(\chi)$. Then*

$$\begin{aligned}&\frac{1}{(q^n - q^{n-1})} \sum_{\chi \in S_{n,q}} \epsilon_\chi^{-1} \prod_{i=1}^{r+1} L\left(\tfrac{1}{2} - \alpha_i, \chi\right) \\ &= \sum_{j=1}^{r+1} q^{\alpha_j(n-1)} \left(\frac{1}{1-q^{-1/2-\alpha_j}} \prod_{i \neq j} \frac{1-q^{-1+\alpha_i-\alpha_j}}{(1-q^{-1/2+\alpha_i})(1-q^{\alpha_i-\alpha_j})} \right) + O\left(\sqrt{q}\left(q^{-(p-2r)/2pr} C_{r,1}\right)^n n^{r+1}\right). \quad (5\text{-}1)\end{aligned}$$

Proof. By plugging Lemma 5.3 into Theorem 1.4, we obtain an identical formula to (5-1), except that the main term is

$$\sum_{j=1}^{r+1} q^{\alpha_j(n-1)} \sum_{\substack{f_1, \dots, f_{r+1} \in \mathbb{F}_q[T] \\ \text{monic} \\ \left(\prod_{i \neq j} f_i\right)/f_j \in T^{\mathbb{Z}}}} \left(\prod_{i \neq j} |f_i|^{-1/2+\alpha_i} \right) |f_j|^{-1/2-\alpha_j}.$$

But by factoring out powers of T, and then noting that f_j is uniquely determined by the other variables, we obtain

$$\sum_{\substack{f_1,\dots,f_{r+1}\in\mathbb{F}_q[T] \\ \text{monic} \\ \left(\prod_{i\neq j} f_i\right)/f_j\in T^{\mathbb{Z}}}} \left(\prod_{i\neq j}|f_i|^{-1/2+\alpha_i}\right)|f_j|^{-1/2-\alpha_j}$$

$$= \left(\prod_{i\neq j}\frac{1}{1-q^{-1/2+\alpha_i}}\right)\frac{1}{1-q^{-\frac{1}{2}-\alpha_j}} \sum_{\substack{f_1,\dots,f_{r+1}\in\mathbb{F}_q[T] \\ \text{monic} \\ \left(\prod_{i\neq j} f_i\right)=f_j \\ f_1,\dots,f_{r+1} \text{ prime to } T}} \left(\prod_{i\neq j}|f_i|^{-1/2+\alpha_i}\right)|f_j|^{-1/2-\alpha_j}$$

$$= \left(\prod_{i\neq j}\frac{1}{1-q^{-1/2+\alpha_i}}\right)\frac{1}{1-q^{-1/2-\alpha_j}} \sum_{\substack{f_1,\dots,f_{j-1},f_{j+1},\dots f_{r+1}\in\mathbb{F}_q[T] \\ \text{monic} \\ f_1,\dots,f_{j-1},f_{j+1},\dots f_{r+1} \text{ prime to } T}} \left(\prod_{i\neq j}|f_i|^{-1+\alpha_i-\alpha_j}\right)$$

$$= \left(\prod_{i\neq j}\frac{1}{1-q^{-1/2+\alpha_i}}\right)\frac{1}{1-q^{-1/2-\alpha_j}} \prod_{i\neq j}\frac{1-q^{-1+\alpha_i-\alpha_j}}{1-q^{\alpha_i-\alpha_j}}.$$

Plugging this in, we get (5-1). □

Acknowledgments

While writing this paper, the author was supported by Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zürich Foundation, and while finishing it, served as a Clay Research Fellow. The author would like to thank Emmanuel Kowalski, Corentin Perret–Gentil, and the two anonymous referees for helpful comments on drafts of this paper.

References

[Andrade and Keating 2012] J. C. Andrade and J. P. Keating, "The mean value of $L(\frac{1}{2}, \chi)$ in the hyperelliptic ensemble", *J. Number Theory* **132**:12 (2012), 2793–2816. MR Zbl

[Andrade and Keating 2014] J. C. Andrade and J. P. Keating, "Conjectures for the integral moments and ratios of L-functions over function fields", *J. Number Theory* **142** (2014), 102–148. MR Zbl

[Bump and Gamburd 2006] D. Bump and A. Gamburd, "On the averages of characteristic polynomials from classical groups", *Comm. Math. Phys.* **265**:1 (2006), 227–274. MR Zbl

[Conrey and Snaith 2007] J. B. Conrey and N. C. Snaith, "Applications of the L-functions ratios conjectures", *Proc. Lond. Math. Soc.* (3) **94**:3 (2007), 594–646. MR Zbl

[Conrey et al. 2005] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith, "Integral moments of L-functions", *Proc. Lond. Math. Soc.* (3) **91**:1 (2005), 33–104. MR Zbl

[Conrey et al. 2008] B. Conrey, D. W. Farmer, and M. R. Zirnbauer, "Autocorrelation of ratios of L-functions", *Commun. Number Theory Phys.* **2**:3 (2008), 593–636. MR Zbl

[Deligne 1980] P. Deligne, "La conjecture de Weil, II", *Inst. Hautes Études Sci. Publ. Math.* **52** (1980), 137–252. MR Zbl

[Diaconu 2019] A. Diaconu, "On the third moment of $L(\frac{1}{2}, \chi_d)$, I: The rational function field case", *J. Number Theory* **198** (2019), 1–42. MR Zbl

[Diaconu et al. 2003] A. Diaconu, D. Goldfeld, and J. Hoffstein, "Multiple Dirichlet series and moments of zeta and L-functions", *Compositio Math.* **139**:3 (2003), 297–360. MR Zbl

[Florea 2017a] A. Florea, "The fourth moment of quadratic Dirichlet L-functions over function fields", *Geom. Funct. Anal.* **27**:3 (2017), 541–595. MR Zbl

[Florea 2017b] A. Florea, "The second and third moment of $L\left(\frac{1}{2}, \chi\right)$ in the hyperelliptic ensemble", *Forum Math.* **29**:4 (2017), 873–892. MR Zbl

[Florea 2017c] A. M. Florea, "Improving the error term in the mean value of $L\left(\frac{1}{2}, \chi\right)$ in the hyperelliptic ensemble", *Int. Math. Res. Not.* **2017**:20 (2017), 6119–6148. MR Zbl

[Fulton and Harris 1991] W. Fulton and J. Harris, *Representation theory: a first course*, Graduate Texts in Math. **129**, Springer, 1991. MR Zbl

[Hoffstein and Rosen 1992] J. Hoffstein and M. Rosen, "Average values of L-series in function fields", *J. Reine Angew. Math.* **426** (1992), 117–150. MR Zbl

[Katz 2001] N. M. Katz, "Sums of Betti numbers in arbitrary characteristic", *Finite Fields Appl.* **7**:1 (2001), 29–44. MR Zbl

[Katz 2013] N. M. Katz, "Witt vectors and a question of Keating and Rudnick", *Int. Math. Res. Not.* **2013**:16 (2013), 3613–3638. MR Zbl

[Katz 2015] N. M. Katz, "On a question of Rudnick: do we have square root cancellation for error terms in moment calculations?", *Philos. Trans. Roy. Soc. A* **373**:2040 (2015), art. id. 20140312. MR Zbl

[Katz 2017] N. M. Katz, "Witt vectors and a question of Rudnick and Waxman", *Int. Math. Res. Not.* **2017**:11 (2017), 3377–3412. MR Zbl

[Katz and Sarnak 1999] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, Amer. Math. Soc. Colloq. Publ. **45**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Sarnak et al. 2016] P. Sarnak, S. W. Shin, and N. Templier, "Families of L-functions and their symmetry", pp. 531–578 in *Families of automorphic forms and the trace formula* (Puerto Rico, 2014), edited by W. Müller et al., Springer, 2016. MR Zbl

[Sawin 2018] W. Sawin, "Square-root cancellation for sums of factorization functions over short intervals in function fields", preprint, 2018. arXiv

[SGA 4½ 1977] P. Deligne, *Cohomologie étale* (Séminaire de Géométrie Algébrique du Bois Marie), Lecture Notes in Math. **569**, Springer, 1977. MR Zbl

[SGA 4_3 1973] M. Artin, A. Grothendieck, and J. L. Verdier, *Théorie des topos et cohomologie étale des schémas, Tome 3: Exposés IX–XIX* (Séminaire de Géométrie Algébrique du Bois Marie 1963–1964), Lecture Notes in Math. **305**, Springer, 1973. MR Zbl

[Tamam 2014] N. Tamam, "The fourth moment of Dirichlet L-functions for the rational function field", *Int. J. Number Theory* **10**:1 (2014), 183–218. MR Zbl

Communicated by Andrew Granville
Received 2018-10-17 Revised 2019-12-12 Accepted 2020-02-06

sawin@math.columbia.edu *Department of Mathematics, Columbia University, New York, NY, United States*

ALGEBRA AND NUMBER THEORY 14:4 (2020)
https://doi.org/10.2140/ant.2020.14.907

Deformations of smooth complete toric varieties: obstructions and the cup product

Nathan Ilten and Charles Turo

Let X be a complete $\mathbb{Q}$-factorial toric variety. We explicitly describe the space $H^2(X, \mathcal{T}_X)$ and the cup product map $H^1(X, \mathcal{T}_X) \times H^1(X, \mathcal{T}_X) \to H^2(X, \mathcal{T}_X)$ in combinatorial terms. Using this, we give an example of a smooth projective toric threefold for which the cup product map does not vanish, showing that in general, smooth complete toric varieties may have obstructed deformations.

1. Introduction

Background and motivation. Let X be any variety over an algebraically closed field $\mathbb{K}$ of characteristic not equal to two or three. The deformation theory of X provides useful information on how X might fit into a moduli space. The abstract theory guarantees that in good situations (e.g., X complete or X an isolated singularity) X will possess a versal deformation, from which all deformations of X can be induced. However in practice, the versal deformation of a given variety may be very difficult to describe in its entirety. It is thus interesting to study classes of varieties for which one may more explicitly understand the deformation theory.

One special class of varieties whose deformation theory has been studied are *toric varieties*. Deformations of such varieties have applications ranging from mirror symmetry [Mavlyutov 2004; Coates et al. 2013] to Kähler–Einstein and extremal metrics [Rollin and Tipler 2014; Ilten and Süss 2017]. The deformation theory of *affine* toric varieties has been described extensively by Altmann. Combinatorial formulas exist for the tangent and obstruction spaces T^1_X and T^2_X as well as a combinatorial description of the cup product map [Altmann 1994; 1997a]; see also [Filip 2018]. A combinatorial recipe may be used to construct deformations of X over affine space [Altmann 1995], and in some cases (e.g., isolated Gorenstein singularities) there is an explicit combinatorial description of the entire versal deformation [Altmann 1997b].

In this paper, we will instead continue the program initiated by the first author in [Ilten 2011] of describing the deformation theory of *smooth, complete* toric varieties. Let X be a smooth complete

MSC2010: primary 14M25; secondary 14B12, 14D15.
Keywords: deformation theory, toric varieties, cup product.

toric variety corresponding to a fan Σ. Ilten [2011] gave a combinatorial description of the space $T^1_X = H^1(X, \mathcal{T}_X)$ of first order deformations:

$$H^1(X, \mathcal{T}_X) = \bigoplus_{\rho \in \Sigma^{(1)}} \bigoplus_{\substack{u \in M \\ \rho(u) = -1}} \widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K}), \tag{1}$$

where ρ ranges over all rays of the fan Σ, M is the character lattice of the torus of X, $\rho(u) \in \mathbb{Z}$ denotes the pairing between the primitive generator of ρ and u, and $\Gamma_{\rho,u}$ is a certain graph; see Section 3. Here, $\widetilde{H}$ denotes reduced cohomology.

Generalizing Altmann's construction in the affine case, Ilten and Vollmert [2012] gave a recipe for producing deformations of any toric variety X over affine spaces from combinatorial data; see also [Mavlyutov 2009; Petracci 2018]. In particular, when X is smooth and complete, each connected component of a graph $\Gamma_{\rho,u}$ appearing in (1) gives rise to a one-parameter deformation (over $\mathbb{A}^1$) lifting the corresponding first order deformation in T^1_X [Ilten and Vollmert 2012, Theorem 6.5]. In fact, for any character $u \in M$, one may use this construction to produce a deformation over $\mathbb{A}^m$ whose image in T^1_X spans the entire degree u piece. This is evidence that, despite in general having nonvanishing obstruction spaces, smooth complete toric varieties might have unobstructed deformations, similar to the situation of, e.g., Calabi–Yau varieties [Tian 1987; Todorov 1989]. However, we will see below that this is not the case.

Results. Throughout, X will be a complete $\mathbb{Q}$-factorial toric variety corresponding to a fan Σ with character lattice M. The description (1) of $H^1(X, \mathcal{T}_X)$ in the case X smooth also holds when X is only $\mathbb{Q}$-factorial; see Section 3. There is also a straightforward generalization of (1) for $H^2(X, \mathcal{T}_X)$:

Proposition 1.1 (Proposition 3.1). *The cohomology group $H^2(X, \mathcal{T}_X)$ may be decomposed as*

$$H^2(X, \mathcal{T}_X) = \bigoplus_{\rho \in \Sigma^{(1)}} \bigoplus_{\substack{u \in M \\ \rho(u) = -1}} H^1(K_{\rho,u}, \mathbb{K}), \tag{2}$$

where each $K_{\rho,u}$ is a simplicial complex determined from Σ; see Section 3.

Our main result is then to give a combinatorial description of the cup product map

$$H^1(X, \mathcal{T}_X) \times H^1(X, \mathcal{T}_X) \to H^2(X, \mathcal{T}_X)$$

using (1) and (2). When X is smooth, $H^2(X, \mathcal{T}_X)$ is the obstruction space T^2_X, and the cup product may be used to obtain the quadratic terms in the obstruction equations for the versal deformation of X. To describe the cup product, we will use Čech cohomology (with respect to a closed covering) to describe elements of the cohomology groups $\widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K})$ and $H^1(K_{\rho,u}, \mathbb{K})$. The closed covering we consider will be indexed by maximal cones $\sigma \in \Sigma$; the corresponding closed sets will be the intersections of σ with either $\Gamma_{\rho,u}$ or $K_{\rho,u}$.

Theorem 1.2 (Theorem 4.3). *Fix $\rho, \rho' \in \Sigma^{(1)}$ and $u, u' \in M$ with $\rho(u) = \rho'(u') = -1$.*

(1) *The image of*
$$\widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K}) \times \widetilde{H}^0(\Gamma_{\rho',u'}, \mathbb{K})$$
in $H^2(X, \mathcal{T}_X)$ under the cup product via (1) *is* 0 *unless $\rho(u') = 0$ or $\rho'(u) = 0$.*

(2) *Assume that $\rho(u') = 0$, and let $f = (f_\sigma)$ and $f' = (f'_\sigma)$ be Čech zero-cycles of $\Gamma_{u,\rho}$ and $\Gamma_{u',\rho'}$. Then the cup product of $(\bar{f}, \bar{f}')$ is contained in $H^1(K_{\rho,u+u'})$ via* (2) *and may be represented by the Čech one-cocycle $g = (g_{\sigma\tau})$, where*
$$g_{\sigma\tau} = \begin{cases} \frac{\rho'(u)}{2}(f_\sigma f'_\tau - f_\tau f'_\sigma) & \text{if } K_{\rho,u+u'} \cap \sigma \cap \tau \neq \varnothing, \\ 0 & \text{otherwise.} \end{cases}$$
A similar formula holds when $\rho'(u) = 0$.

While this theorem gives an explicit description of the cup product on the combinatorial level, it is perhaps not always immediately obvious when the one-cocycle $(g_{\sigma\tau})$ is nontrivial. To remedy this, we proceed as follows. Assume as in the second part of the theorem that $\rho(u') = 0$. Consider any simple cycle α in $K_{\rho,u+u'}$, and connected components Z and Z' of $\Gamma_{\rho,u}$ and $\Gamma_{\rho',u'}$. Then $H^1(\alpha, \mathbb{K}) = \mathbb{K}$ with canonical generator α_{fun}, and Z and Z' induce elements of $\widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K})$ and $\widetilde{H}^0(\Gamma_{\rho',u'}, \mathbb{K})$. The pullback of the cup product of these elements to $H^1(\alpha, \mathbb{K})$ is $(Z *_\alpha Z') \cdot \alpha_{\text{fun}}$, where $Z *_\alpha Z'$ is determined from the intersection behavior of Z and Z' along α; see Theorem 5.3 and the subsection before it for a precise statement.

This leads to a straightforward method to determine when the cup product vanishes. In particular, we may easily use this to construct examples of smooth toric threefolds where the cup product does not vanish:

Corollary 1.3 (Corollary 6.1). *There exists a smooth complete toric threefold with obstructed deformations.*

Murphy's law and future directions. Is this obstructedness result (Corollary 1.3) surprising? We would argue that although perhaps not surprising, it is far from obvious. On the one hand, Vakil [2006] has shown *Murphy's Law* for several classes of deformation problems, that is, that arbitrarily bad singularities of finite type over $\mathbb{Z}$ can occur in the versal deformations. For example, this is true for smooth projective n-folds ($n \geq 2$) with ample canonical class. Vakil writes that his results suggest that "unless there is some natural reason for the [deformation] space to be well-behaved, it will be arbitrarily badly behaved."

On the other hand, toric varieties are so special that there may well be a natural reason for the deformation space to be well-behaved. In fact, Murphy's Law is false for smooth toric varieties! This follows, e.g., from [Ilten and Vollmert 2012, Theorem 6.5], which implies in particular that the versal deformation space of a smooth complete toric variety cannot be a fat point.

This means that the deformation theory of smooth complete toric varieties may belong to the small class of deformation problems which are obstructed, yet one can still hope to completely describe in some explicit manner. The next natural question to address is:

Question 1.4. *Is the versal deformation of a smooth complete toric variety cut out by quadrics*?

In fact, if we knew that the versal deformation was cut out by quadrics, then our results here would completely determine those equations. At the moment, we have far too little evidence to posit an answer one way or the other.

The remainder of this paper is organized as follows. In Section 2, we recall basic facts of Čech cohomology and toric geometry. In Section 3, we prove Proposition 1.1, describing T_X^2 combinatorially. The main work of this paper is contained in Section 4, where we prove our combinatorial description of the cup product (Theorem 1.2). In Section 5 we show how the cup product pulls back to simple cycles α. Finally, in Section 6, we present an example of an obstructed smooth toric threefold, proving Corollary 1.3.

2. Preliminaries

Čech cohomology. We begin by recalling basics of Čech cohomology and fixing notation; see, e.g., [Bosch 2013, §7.6] for more details. Let X be a topological space and $\mathcal{U} = \{U_i\}_{i \in I}$ either an open or closed cover of X. For any sheaf $\mathcal{F}$ of abelian groups on X, the group $\check{C}^p_{\text{sing}}(\mathcal{U}, \mathcal{F})$ of singular p-th Čech cochains is

$$\check{C}^p_{\text{sing}}(\mathcal{U}, \mathcal{F}) = \prod_{(i_0,\ldots,i_p) \in I^{p+1}} \mathcal{F}(U_{i_0} \cap \cdots \cap U_{i_p}).$$

The differential $d^p : \check{C}^p_{\text{sing}}(\mathcal{U}, \mathcal{F}) \to \check{C}^p_{\text{sing}}(\mathcal{U}, \mathcal{F})$ is defined by $d^p(f) = g$, where

$$g_{i_0 \ldots i_{p+1}} = \sum_{k=0}^{p+1} (-1)^k (f_{i_0 \ldots \widehat{i_k} \ldots i_{p+1}})_{|U_{i_0} \cap \ldots \cap U_{i_{p+1}}}.$$

The p-th singular Čech cohomology group of $\mathcal{F}$ with respect to the cover $\mathcal{U}$ is the p-th cohomology $\check{H}^p_{\text{sing}}(\mathcal{U}, \mathcal{F})$ of the complex $(\check{C}^\bullet_{\text{sing}}, d^\bullet)$. Elements of the kernel of d^p are called singular Čech cocycles.

It is more common to work with either *alternating* or *ordered* Čech cohomology, since these have bounded length and involve fewer terms. We will opt to consistently work with alternating Čech cohomology: if we do not explicitly specify that we are talking about *singular* Čech cohomology, then we are referring to alternating Čech cohomology. This is defined as follows.

The group of (alternating) p-th Čech cochains $\check{C}^p(\mathcal{U}, \mathcal{F})$ is the subgroup of $\check{C}^p_{\text{sing}}(\mathcal{U}, \mathcal{F})$ consisting of elements f satisfying

$$f_{i_{\pi(0)} \ldots i_{\pi(p)}} = \text{sign}(\pi) f_{i_0 \ldots i_p}$$

for any permutation $\pi \in S_{p+1}$ of $0, \ldots, p$, and

$$f_{i_0 \ldots i_p} = 0$$

if any index i_j is repeated. After eliminating terms with doubled indices, the differential d^p on the singular Čech complex also gives a differential on the subcomplex $\check{C}^\bullet(\mathcal{U}, \mathcal{F})$. The p-th (alternating) Čech cohomology group of $\mathcal{F}$ with respect to $\mathcal{U}$ is the p-th cohomology $\check{H}^p(\mathcal{U}, \mathcal{F})$ of this subcomplex. Elements of $\check{C}^p(\mathcal{U}, \mathcal{F})$ in the kernel of d^p are called (alternating) Čech cocycles.

The inclusion of complexes $\check{C}^\bullet(\mathcal{U}, \mathcal{F}) \hookrightarrow \check{C}^\bullet_{\text{sing}}(\mathcal{U}, \mathcal{F})$ induces homomorphisms of cohomology groups $\check{H}^p(\mathcal{U}, \mathcal{F}) \to \check{H}^p_{\text{sing}}(\mathcal{U}, \mathcal{F})$. In fact, on the level of cohomology, these maps are isomorphisms; see [Bosch

2013, §7.6, Lemma 1]. For our purposes, we need a map $\check{C}^\bullet_{\text{sing}}(\mathcal{U},\mathcal{F}) \to \check{C}^\bullet(\mathcal{U},\mathcal{F})$ which on cohomology induces the inverse of this isomorphism:

Lemma 2.1. *Assume that $\mathcal{F}$ is a sheaf of $\mathbb{Q}$-modules. The maps*

$$\phi^p : \check{C}^p_{\text{sing}}(\mathcal{U},\mathcal{F}) \to \check{C}^p(\mathcal{U},\mathcal{F})$$

defined by

$$\phi^p(f)_{i_0\ldots i_p} = \frac{1}{(p+1)!}\sum_{\pi\in S_{p+1}} \text{sign}(\pi)\, f_{i_{\pi(0)}\ldots i_{\pi(p)}}$$

give a homomorphism of complexes. The induced map on cohomology is an isomorphism inverse to the map induced by the inclusion of $\check{C}^\bullet$ in $\check{C}^\bullet_{\text{sing}}$.

Proof. To show that $\phi^\bullet$ is a homomorphism of complexes, by linearity it suffices to consider images of elements $f \in \check{H}^p_{\text{sing}}(\mathcal{U},\mathcal{F})$ which are contained in a single summand. The equality $d^p(\phi(f)) = \phi^{p+1}(d^p(f))$ then follows from a direct computation.

It is straightforward to check that ϕ^p is a section to the inclusion of $\check{H}^p(\mathcal{U},\mathcal{F})$ in $\check{H}^p_{\text{sing}}(\mathcal{U},\mathcal{F})$. Since this inclusion induces an isomorphism on cohomology, it follows that ϕ^p does as well. □

Remark 2.2. If $\mathcal{F}$ is a sheaf of modules over a field $\mathbb{K}$ of characteristic $q \neq 0$, one may still define the map ϕ as in Lemma 2.1 for those p such that $p < q-1$. It follows that it will still induce an isomorphism of cohomology for $p < q-2$. In particular, since we are always assuming that our base field $\mathbb{K}$ doesn't have characteristic two or three, we will always obtain isomorphisms in cohomology for $p = 0, 1, 2$.

In the following, we will be using Čech cohomology in two situations. The first is when X is an algebraic variety, $\mathcal{F}$ is a coherent sheaf, and $\mathcal{U}$ is a particular open affine cover. In this case $\check{H}^p(\mathcal{U},\mathcal{F})$ is canonically isomorphic to the sheaf cohomology $H^p(X,\mathcal{F})$ [Hartshorne 1977, Theorem 4.5], so we will usually just write $H^p(X,\mathcal{F})$. The second situation is when X is a finite simplicial complex, $\mathcal{F}$ is the constant sheaf with coefficients in $\mathbb{K}$, and $\mathcal{U}$ is a particular cover by closed simplices, all of whose intersections are contractible. In this case, $\check{H}^p(\mathcal{U},\mathcal{F})$ is canonically isomorphic to the simplicial cohomology groups $H^p(X,\mathbb{K})$ [Godement 1958, §II.5.2], and we will again usually just write $H^p(X,\mathbb{K})$.

Cup products. Assume now that $\mathcal{F}$ is a sheaf of algebras on a topological space X with covering $\mathcal{U}$. The multiplication in $\mathcal{F}$ induces a *cup product* in cohomology

$$\check{H}^p(\mathcal{U},\mathcal{F}) \times \check{H}^q(\mathcal{U},\mathcal{F}) \to \check{H}^{p+q}(\mathcal{U},\mathcal{F}).$$

This is described for the *singular* Čech cohomology groups as follows; see, e.g., [Bosch 2013, §7.6 Exercise 6]. Given singular p- and q-cocycles $f = (f_{i_0\ldots i_p})$ and $f' = (f'_{i_0\ldots i_q})$, the cup product of the cohomology classes represented by f and f' is represented by the $p+q$ cocycle $g = (g_{i_0\ldots i_{p+q}})$ with

$$g_{i_0\ldots i_{p+q}} = f_{i_0\ldots i_p} * f'_{i_p\ldots i_{p+q}}, \tag{3}$$

where $*$ denotes the product on $\mathcal{F}$. This product gives $\bigoplus_p \check{H}^p_{\text{sing}}(\mathcal{U},\mathcal{F})$ the structure of a graded associative algebra.

For our purposes, we desire a description similar to (3) for the cup product between *alternating* Čech cohomology groups. This may be obtained by appropriately composing the maps between $\check{C}^\bullet(\mathcal{U},\mathcal{F})$ and $\check{C}^\bullet_{\text{sing}}(\mathcal{U},\mathcal{F})$ with the cup product on singular Čech cohomology.

We will do this explicitly for the case of interest to us, namely, when $\mathcal{F}=\mathcal{T}_X$ is the tangent sheaf on an algebraic variety X with product induced by the Lie bracket [,], and $p=q=1$:

Lemma 2.3. *Let $f=(f_{ij})$ and $f=(f'_{ij})$ be Čech one-cycles in $\check{C}^1(\mathcal{U},\mathcal{T}_X)$. Then the image of their cohomology classes under the cup product map*

$$\check{H}^1(\mathcal{U},\mathcal{T}_X)\times\check{H}^1(\mathcal{U},\mathcal{T}_X)\to\check{H}^2(\mathcal{U},\mathcal{T}_X)$$

is represented by the two-cycle $g=(g_{ijk})$ with

$$g_{ijk}=\tfrac{1}{6}\big([f_{ij},f'_{jk}]+[f_{ij},f'_{ik}]+[f_{ik},f'_{jk}]-[f_{ik},f'_{ij}]-[f_{jk},f'_{ik}]-[f_{jk},f'_{ij}]\big).$$

Proof. To compute the cup product, we first include f and f' in the group of singular Čech cochains $\check{C}^1_{\text{sing}}(\mathcal{U},\mathcal{T}_X)$, and then apply (3) to find a representative $\tilde{g}$ of the cup product as a singular two-cycle. We get

$$\tilde{g}_{ijk}=[f_{ij},f'_{jk}].$$

The claim now follows from Lemma 2.1 and a straightforward computation by setting $g=\phi^2(\tilde{g})$. □

Remark 2.4. Choosing a section to the inclusion of $\check{C}^\bullet(\mathcal{U},\mathcal{T}_X)\hookrightarrow\check{C}^\bullet_{\text{sing}}(\mathcal{U},\mathcal{T}_X)$ that is different from our preferred section ϕ would lead to a representation of the cup product on the cocycle level that is different from that of Lemma 2.3. Our choice of section ϕ was motivated by the symmetry of the expression for g_{ijk} in this lemma.

Toric varieties. We now fix notation and review some basic facts from toric geometry. See [Fulton 1993] or [Cox et al. 2011] for a more thorough introduction. Throughout the paper we will fix a lattice M which is the character lattice of the algebraic torus $T=\operatorname{Spec}\mathbb{K}[M]$. The lattice $N=\operatorname{Hom}(M,\mathbb{Z})$ is the lattice of one-parameter subgroups of T. We denote the $\mathbb{Q}$-vector spaces associated to M,N by $M_\mathbb{Q}$ and $N_\mathbb{Q}$.

Given a fan Σ in $N_\mathbb{Q}$, we associate a toric variety X_Σ; see [Cox et al. 2011, §3.1]. The variety X_Σ is covered by open affine varieties U_σ as σ ranges over maximal cones in the fan Σ, where

$$U_\sigma=\operatorname{Spec}\mathbb{K}[M\cap\sigma^\vee];\qquad \sigma^\vee=\{u\in M_\mathbb{Q}\mid v(u)\geq 0\ \text{ for all } v\in\sigma\}.$$

We denote the regular function on U_σ associated to $u\in M\cap\sigma^\vee$ by χ^u.

Important geometric properties of X_Σ can be translated into properties on Σ. For example, the variety X_Σ is complete if and only if the fan Σ is complete, that is, the union of all cones in Σ is all of $N_\mathbb{Q}$ [loc. cit., Theorem 3.4.6]. Likewise, the variety X_Σ is smooth if and only if Σ is smooth, that is, each maximal $\sigma\in\Sigma$ has rays whose primitive generators are a subset of a lattice basis of N [loc. cit., Theorem 3.1.19]. Slightly more generally, the variety X is $\mathbb{Q}$-factorial if and only if Σ is a simplicial fan, that is, each maximal $\sigma\in\Sigma$ has rays whose primitive generators are linearly independent [loc. cit., Proposition 4.2.7]. We will henceforth always assume that Σ is complete and simplicial. In other words, we will assume that X is $\mathbb{Q}$-factorial and complete.

We denote the rays of Σ by $\Sigma^{(1)}$; to any ray $\rho \in \Sigma^{(1)}$ and $u \in M$ we denote by $\rho(u)$ the evaluation of the primitive lattice generator of ρ at u. Prime torus invariant divisors of X_Σ are in bijection with rays in $\Sigma^{(1)}$ [loc. cit., §4.1]. We denote the divisor corresponding to ρ by D_ρ. Any torus invariant divisor D may be written uniquely as a sum

$$D = \sum_{\rho \in \Sigma^{(1)}} a_\rho D_\rho.$$

The sheaf $\mathcal{O}(D)$ has the following local description: the function χ^u is in $H^0(U_\sigma, \mathcal{O}(D))$ if and only if

$$\rho(u) + a_\rho \geq 0$$

for all $\rho \in \sigma \cap \Sigma^{(1)}$. In particular, fixing a ray ρ, $\chi^u \in H^0(U_\sigma, \mathcal{O}(D_\rho))$ if and only if for all $\epsilon \in \sigma \cap \Sigma^{(1)}$,

$$\epsilon(u) \geq \begin{cases} 0 & \text{if } \epsilon \neq \rho, \\ -1 & \text{if } \epsilon = \rho. \end{cases} \tag{4}$$

The Euler sequence. The fundamental tool for understanding the tangent bundle on a smooth toric variety is the *Euler sequence*. For $X = X_\Sigma$ complete and $\mathbb{Q}$-factorial, there is an exact sequence of sheaves

$$0 \to N^1 \otimes \mathcal{O}_X \to \bigoplus_{\rho \in \Sigma^{(1)}} \mathcal{O}_X(D_\rho) \xrightarrow[\eta]{} \mathcal{T}_X \to 0,$$

where N^1 is a finite-dimensional vector space; see [Cox et al. 2011, Theorem 8.1.6] (and dualize). This generalizes the standard Euler sequence on projective space. We will need an explicit description of the map η. Following through the construction in [loc. cit.] and dualizing, one obtains that

$$\eta(\chi^u) = \partial(\rho, u) \tag{5}$$

for χ^u a local section of $\mathcal{O}(D_\rho)$, where the derivation $\partial(\rho, u)$ is defined via

$$\partial(\rho, u)(\chi^v) = \rho(v)\chi^{u+v}$$

for any $v \in M$.

We will be interested in the cohomology groups of $\mathcal{T}_X$. The following was first observed by Jaczewski in the smooth case:

Lemma 2.5 [Jaczewski 1994]. *For $p \geq 1$, the map η induces isomorphisms*

$$\bigoplus_{\rho \in \Sigma^{(1)}} H^p(X, \mathcal{O}_X(D_\rho)) \to H^p(X, \mathcal{T}_X).$$

Proof. This follows directly from the Euler sequence, the long exact sequence of cohomology, and the vanishing of $H^p(X, \mathcal{O}_X)$ for $p \geq 1$; see [Cox et al. 2011, Theorem 9.2.3]. ☐

Cohomology of divisorial sheaves on toric varieties. In order to understand the cohomology of $\mathcal{T}_X$, Lemma 2.5 implies that it will be useful to have a combinatorial description of the cohomology groups of the sheaves $\mathcal{O}(D_\rho)$. Since T acts on $X = X_\Sigma$, it will also act on the spaces of sections of $\mathcal{T}_X$ and $\mathcal{O}(D)$ for any torus invariant divisor D. This induces an M-grading on the respective cohomology groups. We will follow [Cox et al. 2011, §9.1] to describe the graded pieces of these cohomology groups. We go into what might seem more detail than necessary since we will later need explicit descriptions of the maps between various isomorphic cohomology groups.

Let $D = \sum a_\rho D_\rho$ be any torus invariant divisor. Fixing some $u \in M$, we define the simplicial complex

$$V_{D,u} = \bigcup_{\sigma \in \Sigma} \operatorname{conv}\{n_\rho \mid \rho \in \sigma \cap \Sigma^{(1)} \text{ and } \rho(u) + a_\rho < 0\} \subset N_{\mathbb{Q}},$$

where n_ρ is the primitive generator of any ray ρ. For each $\sigma \in \Sigma$, there is a natural exact sequence

$$0 \to H^0(U_\sigma, \mathcal{O}_X(D))_u \to \mathbb{K} \to H^0(V_{D,u} \cap \sigma, \mathbb{K}) \to 0;$$

see [Cox et al. 2011, equation 9.1.10]. Here, the degree u piece of $H^0(U_\sigma, \mathcal{O}_X(D))$ is denoted by $H^0(U_\sigma, \mathcal{O}_X(D))_u$. Let I be the set of maximal cones in Σ; we consider the open cover $\mathcal{U} = \{U_\sigma\}_{\sigma \in I}$ of X_Σ. Likewise, we have a closed cover $\mathcal{V} = \{V_\sigma\}_{\sigma \in I}$ of $V_{D,u}$, where $V_\sigma = V_{D,u} \cap \sigma$. The above exact sequence thus leads to an exact sequence

$$0 \to \check{C}^p(\mathcal{U}, \mathcal{O}(D))_u \to \check{C}^p(\mathcal{W}, \mathbb{K}) \to \check{C}^p(\mathcal{V}, \mathbb{K}) \to 0, \tag{6}$$

where $\mathcal{W} = \{W_\sigma\}_{\sigma \in I}$ is the trivial closed cover of a single point x with each $W_\sigma = x$. This sequence is compatible with the Čech differentials, so we obtain an exact sequence of Čech complexes. Since $H^0(x, \mathbb{K}) = \mathbb{K}$ and $H^p(x, \mathbb{K}) = 0$ for $p > 0$, the long exact sequence of cohomology implies that the connecting homomorphisms

$$\check{H}^{p-1}(\mathcal{V}, \mathbb{K}) \to \check{H}^p(\mathcal{U}, \mathcal{O}(D))_u \tag{7}$$

are isomorphisms if $p \geq 2$, and for $p = 1$ we have the exact sequence

$$0 \to H^0(X, \mathcal{O}(D))_u \to \mathbb{K} \to H^0(V_{D,u}, \mathbb{K}) \to H^1(X, \mathcal{O}(D))_u \to 0.$$

This final exact sequence induces an isomorphism between the reduced cohomology $\widetilde{H}^0(V_{D,u}, \mathbb{K})$ and $H^1(X, \mathcal{O}(D))_u$; see [Cox et al. 2011, Theorem 9.1.3].

3. Tangent and obstruction spaces

As before, we consider a complete $\mathbb{Q}$-factorial toric variety $X = X_\Sigma$. For $\rho \in \Sigma^{(1)}$ and $u \in M$, we define

$$V_{\rho,u} := V_{D_\rho,u}$$

and note that the vertices of $V_{\rho,u}$ have a concrete description: for $\epsilon \in \Sigma^{(1)}$, $n_\epsilon \in V_{\rho,u}$ if and only if

(1) $\epsilon = \rho$ and $\epsilon(u) < -1$; or

(2) $\epsilon \neq \rho$ and $\epsilon(u) < 0$.

We define $\Gamma_{\rho,u}$ and $K_{\rho,u}$ to be the one- and two-skeleta of $V_{\rho,u}$, respectively. More generally, let $V_{\rho,u}^{(p)}$ denote the p-skeleton of $V_{\rho,u}$. Below we will come to see that we only need to consider the special case when $\rho(u) = -1$, in which case the description of the vertices of $V_{\rho,u}$ simplifies and n_ρ itself is never a vertex of $V_{\rho,u}$.

We briefly comment on the decomposition

$$H^1(X, \mathcal{T}_X) = \bigoplus_{\rho\in\Sigma^{(1)}} \bigoplus_{\substack{u\in M\\ \rho(u)=-1}} \widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K}).$$

This was shown in [Ilten 2011] in the smooth case; it was noted in [Mavlyutov 2009] that this also holds in the $\mathbb{Q}$-factorial case. The decomposition arises by combining Lemma 2.5 with the isomorphism between $\widetilde{H}^0(V_{D_\rho,u}, \mathbb{K})$ and $H^1(X, \mathcal{O}(D_\rho))$ described afterwards. One then notes that $\widetilde{H}^0(V_{D_\rho,u}, \mathbb{K}) = \widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K})$, and this is nonzero only if $\rho(u) = -1$.

A similar argument to the one above yields a description of $H^p(X, \mathcal{T}_X)$ for all $p \geq 1$:

Proposition 3.1. *For $p \geq 1$, the space $H^p(X, \mathcal{T}_X)$ may be decomposed as*

$$H^p(X, \mathcal{T}_X) = \bigoplus_{\substack{(\rho,u)\in\Sigma^{(1)}\times M\\ \rho(u)=-1}} \widetilde{H}^{p-1}(V_{\rho,u}, \mathbb{K}) = \bigoplus_{\substack{(\rho,u)\in\Sigma^{(1)}\times M\\ \rho(u)=-1}} \widetilde{H}^{p-1}(V_{\rho,u}^{(p)}, \mathbb{K}).$$

In particular, the space $H^2(X, \mathcal{T}_X)$ may be decomposed as

$$H^2(X, \mathcal{T}_X) = \bigoplus_{\rho\in\Sigma^{(1)}} \bigoplus_{\substack{u\in M\\ \rho(u)=-1}} H^1(K_{\rho,u}, \mathbb{K}).$$

Proof. By Lemma 2.5, we have an M-graded isomorphism

$$\bigoplus_{\rho\in\Sigma^{(1)}} H^p(X, \mathcal{O}_X(D_\rho)) \to H^p(X, \mathcal{T}_X).$$

Coupled with (7), we obtain

$$H^p(X, \mathcal{T}_X)_u \cong \bigoplus_{\rho\in\Sigma^{(1)}} \widetilde{H}^{p-1}(V_{\rho,u}, \mathbb{K}).$$

We now show that $\widetilde{H}^{p-1}(V_{\rho,u}, \mathbb{K}) = 0$ unless $\rho(u) = -1$. From the explicit description of $V_{\rho,u}$ above, we observe that if $\rho(u) \neq -1$, then $V_{\rho,\lambda\cdot u}$ is the same for any $\lambda \in \mathbb{Z}_{>0}$. In particular, if $\widetilde{H}^{p-1}(V_{\rho,u}, \mathbb{K}) \neq 0$, $H^p(X, \mathcal{O}(D_\rho))$ would be an infinite-dimensional $\mathbb{K}$-vector space, which is impossible since X is complete. We conclude that we must only consider those pairs (ρ, u) such that $\rho(u) = -1$.

Finally, the $(p-1)$-st reduced cohomology of $V_{\rho,u}$ is the same as that of its p-skeleton $V_{\rho,u}^{(p)}$. □

We will be interested in special zero-cocycles $f = (f_\sigma)$ representing elements of $\widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K})$ coming from a connected component Z of $\Gamma_{\rho,u}$. For such a connected component Z, we define $f(Z) = (f(Z)_\sigma)$ by

$$f(Z)_\sigma = \begin{cases} 1 & \text{if } \sigma \cap Z \neq \varnothing, \\ 0 & \text{if } \sigma \cap Z = \varnothing. \end{cases} \tag{8}$$

These will be useful cocycles for us, since the classes of $\{f(Z)\}$ form a basis for $H^0(\Gamma_{\rho,u},\mathbb{K})$ as Z ranges over all connected components of $\Gamma_{\rho,u}$. In particular, they provide a spanning set for $\widetilde{H}^0(\Gamma_{\rho,u},\mathbb{K})$. If we are instead considering a connected component Z' of $\Gamma_{\rho',u'}$, we will use the notation $f'(Z')$.

4. Combinatorial description of cup product

Mapping to $H^2(X,\mathcal{T}_X)$. Fix $\rho,\rho'\in\Sigma^{(1)}$ and $u,u'\in M$ satisfying $\rho(u)=\rho'(u')=-1$. We now describe the map

$$\widetilde{H}^0(\Gamma_{\rho,u},\mathbb{K})\times\widetilde{H}^0(\Gamma_{\rho',u'},\mathbb{K})\to H^2(X,\mathcal{T}_X)_{u+u'}$$

induced by the cup product in terms of Čech cocycles:

Lemma 4.1. *Let $f=(f_\sigma)$, $f'=(f'_\sigma)$ be Čech zero-cocycles of $\Gamma_{\rho,u}$ and $\Gamma_{\rho',u'}$. The image in $H^2(X,\mathcal{T}_X)$ of the corresponding reduced cohomology classes under the cup product is represented by the Čech two-cycle $\theta=(\theta_{\sigma\tau\gamma})$, where*

$$\theta_{\sigma\tau\gamma}=\tfrac{1}{2}\big(f_\sigma f'_\tau-f_\tau f'_\sigma+f_\gamma f'_\sigma-f_\sigma f'_\gamma+f_\tau f'_\gamma-f_\gamma f'_\tau\big)\cdot\big(\rho(u')\partial(\rho',u+u')-\rho'(u)\partial(\rho,u+u')\big).$$

Proof. We just need to trace through the inclusions of $\widetilde{H}^0(\Gamma_{\rho,u},\mathbb{K})$ and $\widetilde{H}^0(\Gamma_{\rho',u'},\mathbb{K})$ in $H^1(X,\mathcal{T}_X)$ and compose with the description of the cup product found in Lemma 2.3. First, mapping f to a cohomology class in $H^1(X,\mathcal{O}(D_\rho))_u$, we must use the first connecting homomorphism of (6). We do this by sending f to $a=(a_\sigma)\in\check{C}^0(\mathcal{W},\mathbb{K})$ with $a_\sigma=f_\sigma$ and applying the differential d^0 to obtain $b=d^0(a)\in\check{C}^1(\mathcal{W},\mathbb{K})$ with

$$b_{\sigma\tau}=a_\tau-a_\sigma=f_\tau-f_\sigma.$$

By construction, this is the image of the element $c\in\check{C}^1(\mathcal{U},\mathcal{O}(D_\rho))_u$ where

$$c_{\sigma\tau}=(f_\tau-f_\sigma)\cdot\chi^u.$$

Mapping further to $H^1(X,\mathcal{T}_X)_u$ using Lemma 2.5, we obtain the cocycle $g=(g_{\sigma\tau})$, where

$$g_{\sigma\tau}=(f_\tau-f_\sigma)\partial(\rho,u).$$

A similar computation holds for f'; we denote the corresponding one-cocycle in $\check{C}^1(\mathcal{U},\mathcal{T}_X)_{u'}$ by g'.

Before applying Lemma 2.3, we note the straightforward calculation

$$\begin{aligned}[\partial(\rho,u),\partial(\rho',u')]&=\partial(\rho,u)\circ\partial(\rho',u')-\partial(\rho',u')\circ\partial(\rho,u)\\&=\rho(u')\partial(\rho',u+u')-\rho'(u)\partial(\rho,u+u').\end{aligned}$$

Taking this into account while applying the lemma to g and g', we obtain the two-cocycle θ with

$$\begin{aligned}\theta_{\sigma\tau\gamma}=\tfrac{1}{6}\Big(&(f_\tau-f_\sigma)(f'_\gamma-f'_\tau)+(f_\tau-f_\sigma)(f'_\gamma-f'_\sigma)+(f_\gamma-f_\sigma)(f'_\gamma-f'_\tau)-(f_\gamma-f_\sigma)(f'_\tau-f'_\sigma)\\&-(f_\gamma-f_\tau)(f'_\gamma-f'_\sigma)-(f_\gamma-f_\tau)(f'_\tau-f'_\sigma)\Big)\cdot\Big(\rho(u')\partial(\rho',u+u')-\rho'(u)\partial(\rho,u+u')\Big).\end{aligned}$$

This simplifies to the expression in the claim. □

Lifting to $H^2(X, \mathcal{O}(D_\rho))$. We now show how to lift the cocycle θ of Lemma 4.1 to a cocycle representing an element of

$$H^2(X, \mathcal{O}(D_\rho))_{u+u'} \oplus H^2(X, \mathcal{O}(D_{\rho'}))_{u+u'}.$$

Lemma 4.2. *Assume that $\rho \neq \rho'$. With f, f', θ as in Lemma 4.1, define*

$$\kappa = (\kappa_{\sigma\tau\gamma}); \qquad \kappa' = (\kappa'_{\sigma\tau\gamma});$$

$$\kappa_{\sigma\tau\gamma} = \frac{\rho'(u)}{2}\big(f_\sigma f'_\tau - f_\tau f'_\sigma + f_\gamma f'_\sigma - f_\sigma f'_\gamma + f_\tau f'_\gamma - f_\gamma f'_\tau\big) \cdot \chi^{u+u'};$$

$$\kappa'_{\sigma\tau\gamma} = \frac{\rho(u')}{2}\big(f_\sigma f'_\tau - f_\tau f'_\sigma + f_\gamma f'_\sigma - f_\sigma f'_\gamma + f_\tau f'_\gamma - f_\gamma f'_\tau\big) \cdot \chi^{u+u'}.$$

Then κ and κ' are two-cocyles in $\check{C}^2(\mathcal{U}, \mathcal{O}(D_\rho)$ and $\check{C}^2(\mathcal{U}, \mathcal{O}(D_{\rho'})$, and the image of $\kappa' - \kappa$ under the map to $\check{C}^2(\mathcal{U}, \mathcal{T}_X)$ induced by η is exactly θ.

Proof. It follows from the explicit description of η in (5) that the image of $\kappa' - \kappa$ is indeed θ. So we just need to show that κ and κ' are two-cocycles. We will show below that for all $\sigma\tau\gamma$, $\kappa_{\sigma\tau\gamma}$ is an element of $H^0(U_{\sigma\cap\tau\cap\gamma}, \mathcal{O}(D_\rho))$. A similar statement will also hold for κ'. It then remains to show that κ and κ' are in the kernel of the differential d^2. But since for $\rho \neq \rho'$, $\partial(\rho, u+u')$ and $\partial(\rho', u+u')$ are linearly independent over $\mathbb{K}$, the images of κ and κ' under the map induced by η must lie in the kernel of the differential. It follows that κ and κ' must as well.

So we are left to show the claim that $\kappa_{\sigma\tau\gamma}$ is an element of $H^0(U_{\sigma\cap\tau\cap\gamma}, \mathcal{O}(D_\rho))$ for arbitrary choice of σ, τ, γ. By bilinearity of the cup product, it suffices to do this for the special cases when $f = f(Z)$ and $f' = f'(Z')$ as in (8) for connected components Z, Z' of $\Gamma_{\rho,u}$ and $\Gamma_{\rho',u'}$. In the following, we shall fix such components Z, Z'.

In order to show that $\kappa_{\sigma\tau\gamma}$ is a regular section as desired, we need to show that either $\rho'(u) = 0$, $\chi^{u+u'} \in H^0(U_{\sigma\cap\tau\cap\gamma}, \mathcal{O}(D_\rho))$, or

$$f_\sigma f'_\tau - f_\tau f'_\sigma + f_\gamma f'_\sigma - f_\sigma f'_\gamma + f_\tau f'_\gamma - f_\gamma f'_\tau = f_\sigma(f'_\tau - f'_\gamma) + f_\tau(f'_\gamma - f'_\sigma) + f_\gamma(f'_\sigma - f'_\tau) = 0. \quad (9)$$

Let us thus assume that neither $\rho'(u)$ nor the expression in (9) is zero. Thus, all of $f'_\sigma, f'_\tau, f'_\gamma$ can not be equal, and by symmetry the same is true for $f_\sigma, f_\tau, f_\gamma$. Using that $f = f(Z)$ and $f' = f'(Z')$ and the symmetry of the expression, we may assume without loss of generality that we are in one of two cases:

(1) $f_\sigma = 1$, $f_\tau = f_\gamma = 0$; or

(2) $f_\sigma = f_\tau = 1$, $f_\gamma = 0$.

In the first case, we must thus have $f'_\tau \neq f'_\gamma$; without loss of generality $f'_\tau = 1$ and $f'_\gamma = 0$. In other words, our connected component Z intersects σ but not τ and γ, whereas Z' intersects τ but not γ.

Consider any ray ϵ of $\sigma \cap \tau \cap \gamma$. By (4) we must show that $\epsilon(u+u') \geq 0$ if $\epsilon \neq \rho$, and $\epsilon(u+u') \geq -1$ if $\epsilon = \rho$. Suppose $\epsilon = \rho$; then $\epsilon(u+u') = \epsilon(u') - 1$. If $\epsilon(u') < 0$, then ϵ is a vertex of $\Gamma_{\rho',u'}$ (note we are assuming $\rho \neq \rho'$). Now, ϵ is in τ, and Z' intersects τ, so by convexity, $\epsilon \in Z'$. But ϵ is also in γ, contradicting $Z' \cap \gamma = \varnothing$. Hence, $\epsilon(u') \geq 0$, implying $\epsilon(u+u') \geq -1$ as required.

Suppose instead that $\epsilon = \rho'$; then $\epsilon(u+u') = \rho'(u+u') = \epsilon(u) - 1$. If $\epsilon(u) < 0$, then ϵ is a vertex of $\Gamma_{\rho,u}$. Since $\sigma \cap Z \neq \varnothing$, convexity again implies that $\epsilon \in Z$, but this contradicts $Z \cap \tau = \varnothing$. We have also assumed that $\rho'(u) \neq 0$, so we conclude that $\epsilon(u) \geq 1$, and $\epsilon(u+u') \geq 0$.

Finally, supposed that $\epsilon \neq \rho, \rho'$. Arguing similarly to above, we cannot have $\epsilon(u) < 0$, since then we would obtain $\epsilon \in Z$, contradicting $Z \cap \tau = \varnothing$. Likewise, we cannot have $\epsilon(u') < 0$. We thus conclude $\epsilon(u+u') \geq 0$. This concludes the argument for the first case.

In the second case, we notice that we cannot have $f'_\tau = f'_\sigma$, and may argue as in the first case after appropriately permuting the roles of σ, τ, and γ. □

Lifting to simplicial cohomology. We are now able to come to our main result:

Theorem 4.3. *Fix $\rho, \rho' \in \Sigma^{(1)}$ and $u, u' \in M$ with $\rho(u) = \rho'(u') = -1$.*

(1) *The image of*

$$\widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K}) \times \widetilde{H}^0(\Gamma_{\rho',u'}, \mathbb{K})$$

in $H^2(X, \mathcal{T}_X)$ under the cup product via (1) *is* 0 *unless $\rho(u') = 0$ or $\rho'(u) = 0$.*

(2) *Assume that $\rho(u') = 0$, and let $f = (f_\sigma)$ and $f' = (f'_\sigma)$ be Čech zero-cycles of $\Gamma_{\rho,u}$ and $\Gamma_{\rho',u'}$. Then the cup product of $(\bar{f}, \bar{f}')$ is contained in $H^1(K_{\rho,u+u'})$ via* (2) *and may be represented by the Čech one-cocycle $g = (g_{\sigma\tau})$, where*

$$g_{\sigma\tau} = \begin{cases} \dfrac{\rho'(u)}{2}(f_\sigma f'_\tau - f_\tau f'_\sigma) & \text{if } K_{\rho,u+u'} \cap \sigma \cap \tau \neq \varnothing, \\ 0 & \text{otherwise}. \end{cases}$$

A similar formula holds when $\rho'(u) = 0$.

Proof. Let f, f' be Čech zero-cocyles of $\Gamma_{\rho,u}$ and $\Gamma_{\rho',u'}$. We first show item (1). If $\rho = \rho'$, then Lemma 4.1 implies that the image of the cup product of the classes of f and f' in $H^2(X, \mathcal{T}_X)$ is zero. So we may henceforth assume that $\rho \neq \rho'$. By Lemma 4.2, the cup product of the classes of f and f' may be represented by a cocyle in

$$\bigoplus_{\rho \in \Sigma^{(1)}} H^2(X, \mathcal{O}(D_\rho))_{u+u'}$$

living entirely in the ρ and ρ' summands. But it follows from Proposition 3.1 that these vanish respectively unless $\rho(u+u') = -1$ or $\rho'(u+u') = -1$. Given $\rho(u) = \rho'(u') = -1$, this is equivalent to $\rho(u') = 0$ or $\rho'(u) = 0$. This completes the proof of the first item.

For item (2), assume now that $\rho(u') = 0$. By Lemma 4.2, it follows that the cup product of the classes of f and f' may be thought of as a class in $H^2(X, \mathcal{O}(D_\rho))_{u+u'}$, represented by the cocycle $-\kappa$ with

$$\kappa_{\sigma\tau\gamma} = \frac{\rho'(u)}{2}\big(f_\sigma f'_\tau - f_\tau f'_\sigma + f_\gamma f'_\sigma - f_\sigma f'_\gamma + f_\tau f'_\gamma - f_\gamma f'_\tau\big) \cdot \chi^{u+u'}.$$

We consider the element $\tilde{g} \in \check{C}^1(\mathcal{W}, \mathbb{K})$ defined by

$$\tilde{g}_{\sigma\tau} = \frac{\rho'(u)}{2}(f_\sigma f'_\tau - f_\tau f'_\sigma).$$

On the one hand, the image of $\tilde{g}$ in $\check{C}^1(\mathcal{V}, \mathbb{K})$ is exactly the one-cochain g from the statement of the theorem. On the other hand, we may compute that

$$d^1(\tilde{g})_{\sigma\tau\gamma} = \frac{\rho'(u)}{2}(f_\tau f'_\sigma - f_\sigma f'_\tau + f_\sigma f'_\gamma - f_\gamma f'_\sigma + f_\gamma f'_\tau - f_\tau f'_\gamma)$$

in $\check{C}^2(\mathcal{W}, \mathbb{K})$. This is exactly the image of $-\kappa$ under the inclusion

$$\check{C}^2(\mathcal{U}, \mathcal{O}(D_\rho))_{u+u'} \hookrightarrow \check{C}^2(\mathcal{W}, \mathbb{K}).$$

The exact sequence (6) and its compatibility with the Čech differentials implies that g is a cocycle in $\check{C}^1(\mathcal{V}, \mathbb{K})$, and that the image of its cohomology class under the connecting homomorphism is represented by $-\kappa$. This completes the proof of the second claim. □

Remark 4.4. With notation as in Theorem 4.3, one might wonder what happens to the cup product when *both* $\rho(u') = 0$ and $\rho'(u) = 0$. It follows directly from the second part of the theorem that the image of this part of the cup product vanishes.

5. Pulling back to cycles

Setup. Let α be a simple cycle in $K_{\rho,u+u'}$, that is, an oriented connected subgraph of the edges of $K_{\rho,u+u'}$ in which no edges are repeated and every vertex has degree 2. Such a cycle α gives rise to a one-cycle $[\alpha]$ in the simplicial homology $H_1(K_{\rho,u+u'}, \mathbb{K})$ by considering the sum

$$[\alpha] = \sum_{E\in\alpha} \pm E$$

with signs depending on the orientation of α and the chosen orientation of $K_{\rho,u+u'}$. This similarly determines a distinguished generator of $H_1(\alpha, \mathbb{K})$ which we will also denote by $[\alpha]$. We denote the element in $H^1(\alpha, \mathbb{K})$ dual to the class in $H_1(\alpha, \mathbb{K})$ of $[\alpha]$ by α_{fun}.

Definition 5.1. A Σ-*reduced cycle* is a simple cycle α in $K_{\rho,u+u'}$, with $[\alpha]$ not homologous to zero, such that no edges of α are contained in a common cone of Σ.

By the following lemma, it will suffice to consider only Σ-reduced cycles:

Lemma 5.2. *Any class in* $H_1(K_{\rho,u+u'}, \mathbb{K})$ *may be written as a sum of classes of* Σ*-reduced cycles.*

Proof. It suffices to show the lemma for classes represented by $[\beta]$, for some simple cycle β. Suppose that β is not Σ-reduced, with edges E_1 and E_2 contained in a cone σ. If $E_1 = [v_1, v_2]$ and $E_2 = [v_2, v_3]$, then we may replace these edges in β by $E = [v_1, v_3]$ and obtain an equivalent homology class. The resulting simple cycle β' has one fewer edge.

Similarly, if $E_1 \cap E_2 = \varnothing$, we may split β into simple cycles β', β'' such that $[\beta]$ is homologous to $[\beta'] + [\beta'']$ and both these cycles have fewer edges.

The claim now follows by infinite descent. □

Given a Σ-reduced cycle α, the closed cover $\mathcal{V}$ of $K_{\rho,u+u'}$ induces a closed cover $\mathcal{V}^\alpha = \{V_\sigma^\alpha\}$ of α:

$$V_\sigma^\alpha = \alpha \cap \sigma.$$

This closed cover is contractible, so $H^p(\mathcal{V}^\alpha, \mathbb{K}) = H^p(\alpha, \mathbb{K})$.

There is a natural map of complexes

$$\check{C}^\bullet(\mathcal{V}, \mathbb{K}) \to \check{C}^\bullet(\mathcal{V}^\alpha, \mathbb{K})$$

induced by restriction. The corresponding map of cohomology

$$\iota_\alpha^* : H^p(K_{\rho,u+u'}, \mathbb{K}) \to H^p(\alpha, \mathbb{K})$$

is the pullback morphism corresponding to the inclusion $\iota_\alpha : \alpha \hookrightarrow K_{\rho,u+u'}$. In particular, for any $\omega \in H^1(K_{\rho,u+u'}, \mathbb{K})$,

$$\iota_\alpha^*(\omega) = \langle \alpha, \omega \rangle \cdot \alpha_{\text{fun}}$$

where $\langle\ ,\ \rangle$ denotes the pairing between H_1 and H^1.

Computing the cup product. Fix a Σ-reduced cycle α, and let Z and Z' be connected components of $\Gamma_{\rho,u}$ and $\Gamma_{\rho',u'}$. As in Theorem 4.3, assume that $\rho(u') = 0$. We now show how to compute $\iota_\alpha^*(\omega)$ directly, where ω is the class of $H^1(K_{\rho,u+u'}, \mathbb{K})$ corresponding to the cup product of the classes of $f(Z)$ and $f'(Z')$.

Let $E(\alpha)$ denote the set of edges α. We will write

$$E(\alpha) = \{E_1, \ldots, E_k\}$$

with the edges ordered cyclically modulo k (E_{i+1} is the edge following E_i). For each edge E_i of α we fix a maximal cone σ_i with $\sigma_i \cap \alpha = E_i$. These cones exist since α is Σ-reduced.

To Z we associate a subset $\alpha(Z)$ of $E(\alpha)$:

$$\alpha(Z) = \{E_i \in E(\alpha) \mid Z \cap \sigma_i \neq \varnothing\}.$$

The set $\alpha(Z')$ is defined analogously. An index $1 \leq i \leq k$ is *relevant* if

$$\alpha(Z) \cap \{E_i, E_{i+1}\}, \quad \alpha(Z') \cap \{E_i, E_{i+1}\}$$

are not equal but both nonempty, and their union contains both E_i and E_{i+1}. Here, indices are taken modulo k.

We set

$$Z *_\alpha Z' := \frac{\rho'(u)}{2} \cdot \sum_{i \text{ relevant}} b_i, \qquad b_i = \begin{cases} 1 & \text{if } E_i \in \alpha(Z),\ E_{i+1} \in \alpha(Z'), \\ -1 & \text{if } E_i \in \alpha(Z'),\ E_{i+1} \in \alpha(Z). \end{cases}$$

See Figure 1 for an illustration of various values of b_i: for each relevant index i, the value of b_i is recorded in parenthesis next to the edge E_i. For example in the leftmost case, there are two relevant indices, each with $b_i = 1$.

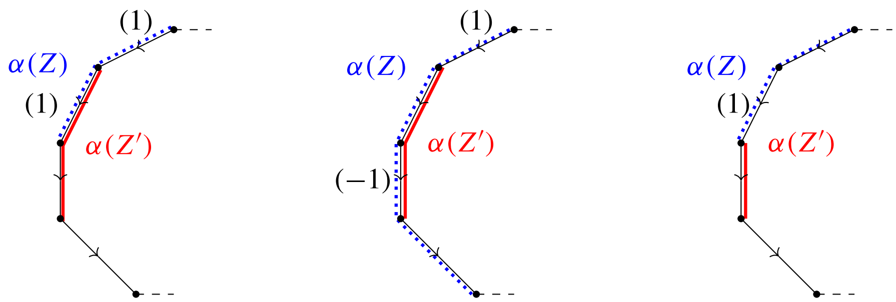

Figure 1. Values of b_i.

Theorem 5.3. *With the above notation,*

$$\iota_\alpha(\omega)^* = (Z *_\alpha Z') \cdot \alpha_{\text{fun}}.$$

In particular, the image of $\widetilde{H}^0(\Gamma_{\rho,u}, \mathbb{K}) \times \widetilde{H}^0(\Gamma_{\rho',u'}, \mathbb{K})$ *under the cup product is zero if and only if for all* Σ*-reduced cycles* α *and all connected components* Z, Z' *of* $\Gamma_{\rho,u}$ *and* $\Gamma_{\rho',u'}$, $Z *_\alpha Z' = 0$.

Proof. We know by Theorem 4.3 that ω is represented in $H^1(K_{\rho,u+u'}, \mathbb{K})$ by the cocyle $g = (g_{\sigma\tau})$, where

$$g_{\sigma\tau} = \begin{cases} \dfrac{\rho'(u)}{2}(f_\sigma f'_\tau - f_\tau f'_\sigma) & \text{if } K_{\rho,u+u'} \cap \sigma \cap \tau \neq \varnothing, \\ 0 & \text{otherwise.} \end{cases}$$

The image of this cocyle under the map

$$\check{C}^1(\mathcal{V}, \mathbb{K}) \to \check{C}^1(\mathcal{V}^\alpha, \mathbb{K})$$

is thus $(h_{\sigma\tau})$, where

$$h_{\sigma\tau} = \begin{cases} \dfrac{\rho'(u)}{2}(f_\sigma f'_\tau - f_\tau f'_\sigma) & \text{if } \alpha \cap \sigma \cap \tau \neq \varnothing, \\ 0 & \text{otherwise.} \end{cases}$$

There is an easier closed cover $\mathcal{W}^\alpha = (E_i)$ of α that we would like to use; it is indexed by $1 \le i \le k$. This is again a closed cover with all intersections contractible, so it also computes $H^1(\alpha, \mathbb{K})$. The assignment $i \mapsto \sigma_i$ lets us view $\mathcal{V}^\alpha$ as a refinement of $\mathcal{W}^\alpha$, and induces a map of Čech complexes

$$\check{C}^p(\mathcal{W}^\alpha, \mathbb{K}) \to \check{C}^p(V^\alpha, \mathbb{K});$$

see, e.g., [Hartshorne 1977, Exercise III.4.4]. This map has a natural section given by forgetting entries with indices not among the σ_i; both maps induce isomorphisms on cohomology.

Hence, we may represent $\iota_\alpha^*(\omega)$ as a Čech one-cocycle with respect to $\mathcal{W}^\alpha$ by (a_{ij}) with

$$a_{ij} = \begin{cases} \dfrac{\rho'(u)}{2}(f_{\sigma_i} f'_{\sigma_j} - f_{\sigma_j} f'_{\sigma_i}) & \text{if } i - j \equiv \pm 1 \mod k, \\ 0 & \text{otherwise.} \end{cases}$$

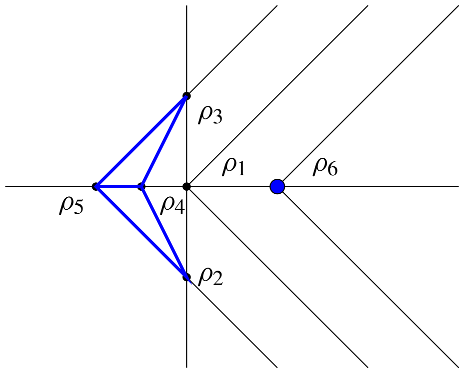

Figure 2. $\Sigma \cap [u=-1]$ for $u=(-1,0,0)$ and $\Gamma_{\rho_1,u}$.

On the other hand, a straightforward computation shows that for any Čech one-cocycle $c=(c_{ij})$ with respect to $\mathcal{W}^\alpha$, the class represented by c in $H^1(\alpha,\mathbb{K})$ is

$$\left(\sum_{i=1}^{k} c_{i(i+1)}\right)\cdot \alpha_{\mathrm{fun}},$$

where indices are taken modulo k.

We thus compute

$$\begin{aligned}\sum_{i=1}^{k} a_{i(i+1)} &= \sum_{i=1}^{k} \frac{\rho'(u)}{2}(f_{\sigma_i} f'_{\sigma_{i+1}} - f_{\sigma_{i+1}} f'_{\sigma_i})\\ &= \sum_{i\ \text{relevant}} \frac{\rho'(u)}{2}(f_{\sigma_i} f'_{\sigma_{i+1}} - f_{\sigma_{i+1}} f'_{\sigma_i}) = \sum_{i\ \text{relevant}} \frac{\rho'(u)}{2} b_i = Z *_\alpha Z'.\end{aligned}$$

□

6. An obstructed example

We now consider the following concrete example. Let $N=\mathbb{Z}^3$, and define rays

$$\begin{aligned}&\rho_1=\mathbb{Q}_{\geq 0}\cdot(1,0,0), && \rho_2=\mathbb{Q}_{\geq 0}\cdot(1,0,-1), && \rho_3=\mathbb{Q}_{\geq 0}\cdot(1,0,1),\\ &\rho_4=\mathbb{Q}_{\geq 0}\cdot(2,-1,0), && \rho_5=\mathbb{Q}_{\geq 0}\cdot(1,-1,0), && \rho_6=\mathbb{Q}_{\geq 0}\cdot(1,1,0),\\ &\rho_7=\mathbb{Q}_{\geq 0}\cdot(0,1,-1), && \rho_8=\mathbb{Q}_{\geq 0}\cdot(0,1,1), && \rho_9=\mathbb{Q}_{\geq 0}\cdot(-1,0,0).\end{aligned}$$

These ρ_i form the rays of a smooth complete fan Σ whose maximal cones are spanned by

$$\begin{array}{lll} \rho_1,\rho_2,\rho_4, & \rho_1,\rho_2,\rho_7, & \rho_1,\rho_3,\rho_4,\\ \rho_1,\rho_3,\rho_8, & \rho_1,\rho_6,\rho_7, & \rho_1,\rho_6,\rho_8,\\ \rho_2,\rho_4,\rho_5, & \rho_2,\rho_5,\rho_9, & \rho_2,\rho_7,\rho_9,\\ \rho_3,\rho_4,\rho_5, & \rho_3,\rho_5,\rho_9, & \rho_3,\rho_8,\rho_9,\\ & \rho_6,\rho_7,\rho_9, & \rho_6,\rho_8,\rho_9. \end{array}$$

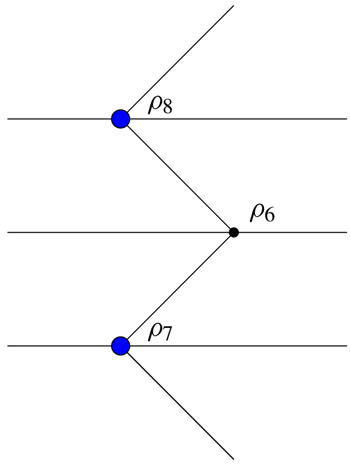

Figure 3. $\Sigma \cap [u' = -1]$ for $u' = (0, -1, 0)$ and $\Gamma_{\rho_6,u'}$.

We will see using Theorem 5.3 that X_Σ has nonvanishing cup-product, and hence obstructed deformations. This will show:

Corollary 6.1. *There exists a smooth complete toric threefold with obstructed deformations.*

The degrees where we will look for first-order deformations are $u = (-1, 0, 0)$ and $u' = (0, -1, 0)$. We picture the intersection of Σ with the hyperplane

$$\{v \in N_{\mathbb{Q}} \mid v(u) = -1\}$$

in Figure 2. The graph $\Gamma_{\rho_1,u}$ is also pictured in the figure in blue bold lines. It has two connected components: one component contains the generators of the rays ρ_2, ρ_3, ρ_4, ρ_5, and the other contains the generator of ρ_6. We will denote the first component by Z. Note that for any other choice of ray ρ, $\Gamma_{\rho,u}$ is connected. Hence, $H^1(X, \mathcal{T}_X)_u$ is one-dimensional.

In Figure 3 we picture the intersection of Σ with the hyperplane

$$\{v \in N_{\mathbb{Q}} \mid v(u') = -1\}$$

along with the graph $\Gamma_{\rho_6,u'}$. This graph has two connected components, consisting of the primitive generators of ρ_7 and ρ_8. We denote the first of these components by Z'. Again for any other choice of ray ρ, $\Gamma_{\rho,u'}$ is connected, so $H^1(X, \mathcal{T}_X)_{u'}$ is also one-dimensional.

We will now compute the cup product ω of the first-order deformations corresponding to Z and Z'. By Theorem 4.3, it is possible that this is nonzero, since $\rho_1(u') = 0$. This class will live in degree $u + u'$, so we picture the intersection of Σ with the hyperplane

$$\{v \in N_{\mathbb{Q}} \mid v(u + u') = -1\}$$

in Figure 4.

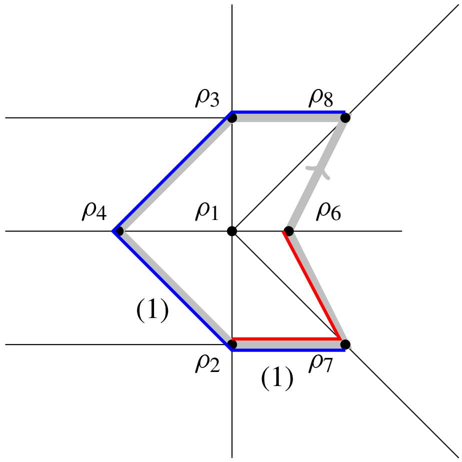

Figure 4. $\Sigma \cap [u+u'=-1]$, α, and $\alpha(Z)$, $\alpha(Z')$.

Let α be the Σ-reduced cycle pictured in gray oriented in counter-clockwise direction. That is, α has vertices (in order) at the primitive generators of

$$\rho_8, \rho_3, \rho_4, \rho_2, \rho_7, \rho_6, \rho_8.$$

This is indeed Σ-reduced. To apply Theorem 5.3, we must choose a maximal cone $\sigma_i \in \Sigma$ for each edge of α. In this instance there is a canonical choice: for the edge corresponding to rays ρ_i and ρ_j, we take the cone generated by ρ_i, ρ_j, ρ_1.

This gives rise to the sets $\alpha(Z)$ and $\alpha(Z')$, pictured in the figure in blue and red, respectively. The set $\alpha(Z)$ is a subgraph with vertices equal to the primitive generators of ρ_8, ρ_3, ρ_4, ρ_2, ρ_7, and the set $\alpha(Z')$ is a subgraph with vertices equal to the primitive generators of ρ_2, ρ_7, ρ_6.

The only *relevant* indices are for the edges corresponding to ρ_4, ρ_2 and ρ_2, ρ_7; each contributes a value of $b_i = 1$, so we obtain by Theorem 5.3 that

$$\iota_\alpha^*(\omega) = 1 \cdot \alpha_{\text{fun}}.$$

In particular, the cup product ω of the first-order deformations corresponding to Z and Z' is nonzero.

Acknowledgements

Both authors were partially supported by NSERC. Ilten was partially supported by the grant 346300 for IMPAN from the Simons Foundation and the matching 2015-2019 Polish MNiSW fund. We thank the anonymous referees for useful comments.

References

[Altmann 1994] K. Altmann, "Computation of the vector space T^1 for affine toric varieties", *J. Pure Appl. Algebra* **95**:3 (1994), 239–259. MR Zbl

[Altmann 1995] K. Altmann, "Minkowski sums and homogeneous deformations of toric varieties", *Tohoku Math. J.* (2) **47**:2 (1995), 151–184. MR Zbl

[Altmann 1997a] K. Altmann, "Infinitesimal deformations and obstructions for toric singularities", *J. Pure Appl. Algebra* **119**:3 (1997), 211–235. MR Zbl

[Altmann 1997b] K. Altmann, "The versal deformation of an isolated toric Gorenstein singularity", *Invent. Math.* **128**:3 (1997), 443–479. MR Zbl

[Bosch 2013] S. Bosch, *Algebraic geometry and commutative algebra*, Springer, 2013. MR Zbl

[Coates et al. 2013] T. Coates, A. Corti, S. Galkin, V. Golyshev, and A. Kasprzyk, "Mirror symmetry and Fano manifolds", pp. 285–300 in *European Congress of Math.* (Kraków, 2012), edited by R. Latała et al., Eur. Math. Soc., Zürich, 2013. MR Zbl

[Cox et al. 2011] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric varieties*, Graduate Studies in Math. **124**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[Filip 2018] M. Filip, "The Gerstenhaber product $\mathrm{HH}^2(A)\times\mathrm{HH}^2(A)\to\mathrm{HH}^3(A)$ of affine toric varieties", preprint, 2018. arXiv

[Fulton 1993] W. Fulton, *Introduction to toric varieties*, Ann. of Math. Studies **131**, Princeton Univ. Press, 1993. MR Zbl

[Godement 1958] R. Godement, *Topologie algébrique et théorie des faisceaux*, Actualités Sci. Indust. **1252**, Hermann, Paris, 1958. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Math. **52**, Springer, 1977. MR Zbl

[Ilten 2011] N. O. Ilten, "Deformations of smooth toric surfaces", *Manuscripta Math.* **134**:1-2 (2011), 123–137. MR Zbl

[Ilten and Süss 2017] N. Ilten and H. Süss, "K-stability for Fano manifolds with torus action of complexity 1", *Duke Math. J.* **166**:1 (2017), 177–204. MR Zbl

[Ilten and Vollmert 2012] N. O. Ilten and R. Vollmert, "Deformations of rational T-varieties", *J. Algebraic Geom.* **21**:3 (2012), 531–562. MR Zbl

[Jaczewski 1994] K. Jaczewski, "Generalized Euler sequence and toric varieties", pp. 227–247 in *Classification of algebraic varieties* (L'Aquila, 1992), edited by C. Ciliberto et al., Contemp. Math. **162**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

[Mavlyutov 2004] A. R. Mavlyutov, "Deformations of Calabi–Yau hypersurfaces arising from deformations of toric varieties", *Invent. Math.* **157**:3 (2004), 621–633. MR Zbl

[Mavlyutov 2009] A. Mavlyutov, "Deformations of toric varieties via Minkowski sum decompositions of polyhedral complexes", preprint, 2009. arXiv

[Petracci 2018] A. Petracci, "Homogeneous deformations of toric pairs", preprint, 2018. arXiv

[Rollin and Tipler 2014] Y. Rollin and C. Tipler, "Deformations of extremal toric manifolds", *J. Geom. Anal.* **24**:4 (2014), 1929–1958. MR Zbl

[Tian 1987] G. Tian, "Smoothness of the universal deformation space of compact Calabi–Yau manifolds and its Petersson–Weil metric", pp. 629–646 in *Mathematical aspects of string theory* (San Diego, CA, 1986), edited by S.-T. Yau, Adv. Ser. Math. Phys. **1**, World Sci., Singapore, 1987. MR Zbl

[Todorov 1989] A. N. Todorov, "The Weil–Petersson geometry of the moduli space of $\mathrm{SU}(n\geq 3)$ (Calabi–Yau) manifolds, I", *Comm. Math. Phys.* **126**:2 (1989), 325–346. MR Zbl

[Vakil 2006] R. Vakil, "Murphy's law in algebraic geometry: badly-behaved deformation spaces", *Invent. Math.* **164**:3 (2006), 569–590. MR Zbl

Communicated by Joseph Gubeladze
Received 2019-01-02 Revised 2019-11-25 Accepted 2020-02-06

nilten@sfu.ca Simon Fraser University, Burnaby BC, Canada

cturo@sfu.ca Simon Fraser University, Burnaby BC, Canada

ALGEBRA AND NUMBER THEORY 14:4 (2020)
dx.doi.org/10.2140/ant.2020.14.927

Mass equidistribution on the torus in the depth aspect

Yueke Hu

In this paper we prove the equidistribution of the restriction of the mass of automorphic newforms to a nonsplit torus in the depth aspect. This result is better than the current known results on the similar problem in the eigenvalue aspect. The method is relatively elementary and makes use of the known effective QUE result in the depth aspect.

1. introduction

1A. *Arithmetic QUE and QUER problems.* The QUE (quantum unique ergodicity) property in the arithmetic setting is a special case of the conjecture by Rudnick and Sarnak [1994] concerning the asymptotic behavior of the mass measure associated to a normalized holomorphic modular form or Maass form.

More specifically let f be a cuspidal Hecke eigenform with Laplace eigenvalue λ. For any fixed test function ϕ on the modular curve $\Gamma\backslash\mathbb{H}$, define the associated mass measure μ_f and the standard hyperbolic measure μ as follows:

$$\mu_f(\phi)=\int_{\Gamma\backslash\mathbb{H}}|f(z)|^2\phi(z)\,dz,\qquad \mu(\phi)=\int_{\Gamma\backslash\mathbb{H}}\phi(z)\,dz.$$

Then the arithmetic QUE property states that

$$\frac{\mu_f(\phi)}{\mu_f(1)}-\frac{\mu(\phi)}{\mu(1)}\to 0 \tag{1-1}$$

as $\lambda\to\infty$.

This result is now known by the work of Lindenstrauss [2006] and Soundararajan [2010]. Later on similar QUE results in the weight and level aspect were proven in a series of papers by different authors [Holowinsky and Soundararajan 2010; Marshall 2011; Nelson 2011; Nelson et al. 2014; Hu 2018].

It is natural to ask about the asymptotic behavior of the restriction of the mass measures to certain submanifolds $\mathcal{C}$ (especially geodesics). More specifically, for a test function ϕ on $\mathcal{C}$, a fixed Haar measure dt on $\mathcal{C}$, define the restricted measures as follows:

$$\mu_{\mathcal{C},f}(\phi)=\int_{\mathcal{C}}|f(t)|^2\phi(t)\,dt,\qquad \mu_{\mathcal{C}}(\phi)=\int_{\mathcal{C}}\phi(t)\,dt.$$

MSC2010: primary 58J51; secondary 11F67, 22E50, 37A45.
Keywords: Equidistribution in the depth aspect, restriction of the mass measure to torus.

As $\lambda \to \infty$, the property

$$\frac{\mu_{\mathcal{C},f}(\phi)}{\mu_f(1)} - \frac{\mu_{\mathcal{C}}(\phi)}{\mu(1)} \to 0 \tag{1-2}$$

is referred to as QUER in [Christianson et al. 2013] (here R stands for restriction). If only a density 1 subsequence of eigenfunctions satisfy (1-1) or (1-2), we shall refer to the corresponding property as QE or QER.

An obvious counterexample to QUER is the geodesic $\gamma_0 = \{iy\}$, as the restriction of odd Maass forms to it are always zero. The failure comes from the additional symmetry of the odd Maass forms with respect to this particular geodesic. The work in [Dyatlov and Zworski 2013; Toth and Zelditch 2013] implies that a *generic asymmetric* geodesic on the modular curve should satisfy the QER property. The work by Christianson, Toth and Zelditch [Christianson et al. 2013] reveals an intricate relation between QUE and QUER. It is however not clear how to directly apply these works to answer (1-2) for any given $\mathcal{C}$.

Ghosh, Reznikov and Sarnak [2013] showed that for long enough but fixed subsegments $\mathcal{S}$ of certain special geodesics and horocycles

$$1 \ll \int_{\mathcal{S}} |f(t)|^2 \, dt \ll \lambda^{\epsilon}. \tag{1-3}$$

When f is an Eisenstein series, Young [2018] proved QUER in the t-aspect when restricted to vertical geodesic segments.

For general geodesic segments of unit length, Marshall [2016] showed that

$$\int_{\mathcal{S}} |f(t)|^2 \, dt \ll \lambda^{3/14+\epsilon}. \tag{1-4}$$

It improves the exponent $\frac{1}{4}$ in the work of Burq, Gérard, and Tzvetkov [Burq et al. 2007], which holds for eigenfunctions on general compact Riemann surfaces.

The main purpose of this paper is to prove an analogue of QUER in the depth aspect when restricted to closed geodesics or CM points, together with an effective control over the rate of convergence. We shall directly formulate our result in the adelic language of automorphic forms and automorphic representations. See, for example, [Michel and Venkatesh 2006] on the relation between the classical language and adelic language for the torus.

Let $\mathbb{F}$ be a number field and $\mathbb{E}$ be a quadratic field extension over $\mathbb{F}$, with any fixed embedding into $\mathrm{GL}_2(\mathbb{F})$. Let v_0 be a fixed nonarchimedean place of the base field $\mathbb{F}$ and q be the cardinality of the residue field of $\mathbb{F}_{v_0}$. We assume throughout the paper that $2 \nmid q$. Let f be an automorphic cuspidal newform on GL_2 over $\mathbb{F}$, which is ramified at v_0, of finite conductor $N = q^c$, with trivial central character and bounded archimedean components (i.e., the associated local representations π_v at Archimedean places have bounded weight or eigenvalues, and the associated local components f_v come from K-types of bounded weight). From now on, we replace the domain of the integral for $\mu_f(\phi)$ and $\mu(\phi)$ by $[\mathrm{GL}_2] := \mathbb{A}_{\mathbb{F}}^{\times} \mathrm{GL}_2(\mathbb{F}) \backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{F}})$, and take $\mathcal{C} = [\mathbb{E}^{\times}] := \mathbb{A}_{\mathbb{F}}^{\times} \mathbb{E}^{\times} \backslash \mathbb{A}_{\mathbb{E}}^{\times}$.

Theorem 1.1. *For notations as above, the restricted mass measure $\mu_{[\mathbb{E}^\times],f}$ is weakly equidistributed as $c \to \infty$, in the sense that*

$$\frac{\mu_{[\mathbb{E}^\times],f}(\phi)}{\mu_f(1)} \to \frac{\mu_{[\mathbb{E}^\times]}(\phi)}{\mu(1)}$$

for any test function $\phi \in C_c^\infty(\mathcal{C})$. Furthermore if Ω is a fixed Hecke character on $[\mathbb{E}^\times]$, we have the following estimate for the rate of convergence:

$$\left|\frac{\mu_{[\mathbb{E}^\times],f}(\Omega)}{\mu_f(1)} - \frac{\mu_{[\mathbb{E}^\times]}(\Omega)}{\mu(1)}\right| \ll_{\mathbb{E},C(\Omega),q,\epsilon} q^{(\theta-1/4+\epsilon)c}. \tag{1-5}$$

Here θ is a bound towards the Ramanujan conjecture, and $\theta < \frac{7}{64}$ by [Blomer and Brumley 2011].

Note that as $\mathcal{C}$ is compact, any test function $\phi \in C_c^\infty(\mathcal{C})$ can be written as a linear combination of Hecke characters. Thus to prove the theorem it suffice to check (1-5). The approach we shall take is relatively simple. We shall do a spectral decomposition essentially for $|f|^2$. But instead of a long spectral sum, the additional invariance of Ω allows us to do a short sum. The integral against Ω for the residual spectrum gives the main term. The contribution from the cuspidal and continuous spectra can be controlled following the strategy in [Nelson et al. 2014; Hu 2017], by using the convexity bound of the triple product/Rankin–Selberg L-function, together with a power saving upper bound for the local integrals (which we shall slightly generalize in Section 3).

Remark 1.2. (1) When $\mathbb{E}$ is a real quadratic extension over $\mathbb{Q}$, our main result corresponds to QUER property for closed geodesics in the depth aspect. When $\mathbb{E}$ is imaginary, it corresponds to QUER for CM points.

(2) We do not specify the embedding of $\mathbb{E}$ in the theorem as long as it is fixed. Our main result has additional flexibility in the sense that for any fixed $g \in \mathrm{GL}_2(\mathbb{F})$, the theorem still hold when we take $\mathcal{C}' = [\mathbb{E}^\times]g$. This is because $|f|^2(tg) = |f|^2(g^{-1}tg)$, so the QUER problem for $\mathcal{C}'$ is effectively equivalent to the QUER problem for $\mathcal{C}'' = g^{-1}[\mathbb{E}^\times]g$, which is already solved by the theorem for the conjugated embedding.

(3) The dependence of the implied constant on q can be worked out explicitly. It comes from the bound of the local period integrals in Section 3. Actually from the proof, one can see that the implied constant can be easily controlled by q. Thus the same strategy can be applied to show QUER on $[\mathbb{E}^\times]$ when $q \to \infty$ and c is large enough.

(4) Furthermore, we expect similar strategy to work for $\mathcal{C}$ being a split torus or a unipotent subgroup, as long as its embedding into GL_2 is not upper triangular. The reason for this expectation is that Lemma 4.2, the main ingredient to shorten the spectral sum, only requires an element whose lower left entry is nontrivial.

There will be however an additional issue. In these cases, we need to take the test function to be a compactly supported smooth function on $\mathcal{C}$, which can be represented as an integral over continuous spectra. Applying the same strategy as for the nonsplit torus case, one will run into integrals of the continuous spectrum on GL_2 against the continuous spectrum on $\mathcal{C}$, which is not absolutely convergent. So one need proper regularization (I believe the regularization in [Michel and Venkatesh 2010, Section 4.3] should suffice) for this strategy to work. We shall leave the details to interested readers.

Remark 1.3. The idea of the current approach comes from helpful discussions with Paul Nelson. The author originally used the spectral decomposition for $|f|^2$ directly, and made use of the vanishing result for Waldspurger's period integral to get a short sum. The current approach is simpler and allows for slightly more general situations.

2. Notations

Let $\mathbb{F}$ be a number field and $\mathbb{F}_v$ be the corresponding local field of $\mathbb{F}$ at a place v. Let O_v be the ring of integers of $\mathbb{F}_v$ and ϖ_v be a local uniformizer. Let $q = |\varpi_v|_v^{-1}$.

For an additive character ψ over a local field $\mathbb{F}_v$, its level $c(\psi)$ is the least integer such that ψ is trivial on $\varpi_v^{c(\psi)} O_v$. Without loss of generality we shall fix ψ to be unramified (or level 0). For a multiplicative character χ over $O_v^\times$, its level $c(\chi)$ is the least integer such that χ is trivial on $1+\varpi_v^{c(\chi)} O_v$. When χ is trivial on $O_v^\times$, we say that it is unramified or $c(\chi) = 0$.

Let $\mathbb{E}$ be a quadratic field extension over $\mathbb{F}$. Let $\mathbb{E}_v$ be the completion of $\mathbb{E}$ with respect to v. When $\mathbb{E}_v$ is a field extension over $\mathbb{F}_v$, let $O_{\mathbb{E}_v}$ be its ring of integers, and $\varpi_{\mathbb{E}_v}$ be a local uniformizer of $\mathbb{E}_v$. Define $U_{\mathbb{E}_v}(j) = 1+\varpi_{\mathbb{E}_v}^j O_{\mathbb{E}_v}$, and $U_{\mathbb{E}_v}(0) = O_{\mathbb{E}_v}^\times$ by convention. If $\mathbb{E}_v$ is split over $\mathbb{F}_v$, fix an isomorphism $\iota_v : \mathbb{E}_v \to \mathbb{F}_v \times \mathbb{F}_v$. For $\Omega_v = (\Omega_{1,v} \otimes \Omega_{2,v})\circ\iota_v$ a character of $\mathbb{E}_v^\times$, let $c(\Omega_v) = \max\{c(\Omega_{i,v})\}$. Define $U_{\mathbb{E}_v}(j) = \iota_v^{-1}((U_{\mathbb{F}_v} j) \times U_{\mathbb{F}_v}(j))$.

For the group GL_2, let Z be its center, B be its standard upper triangular Borel subgroup and N be the associated unipotent subgroup. Over a nonarchimedean place v, let $K_v = \mathrm{GL}_2(O_v)$ be the maximal compact open subgroup of $\mathrm{GL}_2(\mathbb{F}_v)$. For $n \geq 1$, define the following compact open subgroup of K_v:

$$K_0(\varpi_v^n) = \left\{ g \in K_v \mid g \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod \varpi_v^n \right\}.$$

Globally for a fixed finite place v_0 and $N = q^c$ as in the introduction, denote $K_0(N) = \prod_{v \nmid v_0 \infty} K_v \times K_0(\varpi_{v_0}^c)$.

For a representation π_v of $\mathrm{GL}_2(\mathbb{F}_v)$ with trivial central character, define $c(\pi_v)$ to be the smallest integer such that the subspace of π_v invariant by $K_0(\varpi_v^{c(\pi_v)})$ is nontrivial. A nontrivial element invariant by $K_0(\varpi_v^{c(\pi_v)})$ is called a newform. It is unique up to a scalar according to [Casselman 1973].

For $\mu_{1,v}, \mu_{2,v}$ two characters of $\mathbb{F}_v^\times$, let $\pi(\mu_{1,v}, \mu_{2,v}, s)$ denote the parabolically induced representation which contains smooth functions φ on $\mathrm{GL}_2(\mathbb{F}_v)$ satisfying

$$\varphi\left(\begin{pmatrix} a_1 & n \\ 0 & a_2 \end{pmatrix} g\right) = \mu_{1,v}(a_1)\mu_{2,v}(a_2)\left|\frac{a_1}{a_2}\right|_{\mathbb{F}_v}^s \varphi(g). \tag{2-1}$$

When $s = \frac{1}{2}$, we simply write $\pi(\mu_{1,v}, \mu_{2,v}) = \pi\left(\mu_{1,v}, \mu_{2,v}, \frac{1}{2}\right)$.

When π_v is unitary, let $\langle \cdot , \cdot \rangle$ be the unitary pairing. For any $\varphi \in \pi_v$, let W_φ be the associated Whittaker function and $\Phi_\varphi(g) = \langle \pi_v(g)\varphi, \varphi \rangle$ be the associated matrix coefficient.

A unitary irreducible representation π_v satisfies the bound θ towards the Ramanujan conjecture, if either π_v is tempered, or $\pi_v \simeq \pi(\mu_v, \mu_v, s)$ is a complementary series representation with μ_v unitary and $\left|s-\frac{1}{2}\right| < \theta$.

3. Upper bounds for the local Rankin–Selberg integral and the triple product integral

Everything in this section is local and we shall omit the subscript v. Let π_i be representations of GL_2 with trivial central characters, with $c(\pi_2) = c(\pi_3) = c$, $c(\pi_1) = c_1$. Let $\varphi_i^0 \in \pi_i$ for $i = 1, 2, 3$ be L^2-normalized newforms.

Consider first the case where χ is a character of $\mathbb{F}^\times$, and $\varphi_1 = \varphi_{1,s} \in \pi_1 = \pi(\chi, \chi^{-1}, s)$ satisfies (2-1). In this case denote by

$$I^{\mathrm{RS}}(\varphi_1, \varphi_2, \varphi_3) = \int_{Z(\mathbb{F})N\backslash \mathrm{GL}_2(\mathbb{F})} W_{\varphi_2}(g) W^-_{\varphi_3}(g)\varphi_1(g)\, dg \tag{3-1}$$

the local integral for the Rankin–Selberg integral. Here W_φ is the Whittaker function associated to φ with respect to the fixed additive character ψ, while W^-_φ is for $\psi^-(x) = \psi(-x)$.

For general $\varphi_1 \in \pi_1$, denote by

$$I^{\mathrm{T}}(\varphi_1, \varphi_2, \varphi_3) = \int_{\mathbb{F}^\times\backslash \mathrm{GL}_2(\mathbb{F})} \prod_{i=1}^{3} \Phi_{\varphi_i}(g)\, dg \tag{3-2}$$

the local integral for the triple product formula.

In this section we shall prove the following upper bounds for I^{T} and I^{RS} when $\varphi_1 = \pi_1\left(\left(\begin{smallmatrix}\varpi^{-n} & 0\\ 0 & 1\end{smallmatrix}\right)\right)\varphi_1^0$, $\varphi_i = \varphi_i^0$ for $i = 2, 3$. They will be used later on to control the contributions from the cuspidal spectra and the continuous spectra.

Proposition 3.1. *Suppose that π_i satisfies the bound θ towards the Ramanujan conjecture, for $i = 1, 2, 3$. Suppose that $c_1 = c(\pi_1)$ is fixed and $c > c_1$. When π_1 is a principal series representation or a special representation, we have*

$$\left|I^{\mathrm{RS}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0, \varphi_2^0, \varphi_3^0\right)\right| = \left|I^{\mathrm{RS}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-c+n+c_1} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0, \varphi_2^0, \varphi_3^0\right)\right| \ll_{c_1,q,\epsilon} \frac{1}{q^{(1/2-\theta-\epsilon)\max\{n, c-c_1-n\}}}. \tag{3-3}$$

For general π_1, we have

$$\left|I^{\mathrm{T}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0, \varphi_2^0, \varphi_3^0\right)\right| = \left|I^{\mathrm{T}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-c+n+c_1} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0, \varphi_2^0, \varphi_3^0\right)\right| \ll_{c_1,q,\epsilon} \frac{1}{q^{(1-2\theta-\epsilon)\max\{n, c-c_1-n\}}}. \tag{3-4}$$

The rest of this section is dedicated to the proof of this result.

Remark 3.2. This result generalizes the similar upper bounds used in [Hu 2018]. The computations there cover the range $c_1+2n<c$, are based on case-by-case check, and can be vague in some situations. Here we employ similar ideas, but cover the whole range for n while giving slightly more uniform and explicit treatments.

Remark 3.3. This result is of independent interest and may potentially be useful for proving the subconvexity bounds for L-functions in the hybrid range.

3A. *Preparations.*

3A1. *Double coset decomposition.* From [Hu 2017; 2018] we have the following variant of the Iwasawa decomposition:

Lemma 3.4. *For every positive integer c,*

$$\mathrm{GL}_2(\mathbb{F}) = \coprod_{0\le i\le c} B\begin{pmatrix}1 & 0\\ \varpi^i & 1\end{pmatrix} K_0(\varpi^c).$$

Here B is the Borel subgroup of GL_2. Furthermore if f is a $ZK_0(\varpi^c)$-invariant function, then

$$\int_{\mathbb{F}^\times\backslash \mathrm{GL}_2(\mathbb{F})} f(g)\,dg = \sum_{0\le i\le c} A_i \int_{\mathbb{F}^\times\backslash B(\mathbb{F})} f\left(b\begin{pmatrix}1 & 0\\ \varpi^i & 1\end{pmatrix}\right) db. \tag{3-5}$$

Here we normalize the Haar measure on $\mathrm{GL}_2(\mathbb{F})$ such that $K=\mathrm{GL}_2(O)$ has volume 1, db is the left Haar measure on $\mathbb{F}^\times\backslash B(\mathbb{F})$ such that $Z(O)\backslash B(O)$ has volume 1, and $A_i \asymp q^{-i}$ are fixed constants.

3A2. *Decay of matrix coefficients.* We first recall the following result on the decay of matrix coefficients from [Venkatesh 2010, Lemma 9.1]. Let π be a representation of GL_2 satisfying the bound θ towards the Ramanujan conjecture. Define

$$\sigma_n = \begin{pmatrix}\varpi^n & 0\\ 0 & 1\end{pmatrix}. \tag{3-6}$$

Lemma 3.5. *For w_1, w_2 any two K-finite L^2-normalized elements of π, and any integer $n\ge 0$,*

$$\langle \pi(\sigma_n)w_1, w_2\rangle \ll_{\epsilon,q} \dim(Kw_1)^{1/2}\dim(Kw_2)^{1/2} q^{(\theta-1/2+\epsilon)n}. \tag{3-7}$$

In this paper we shall need the following (weaker) variant of the above result.

Corollary 3.6. *Let φ_0 be the newform of π and Φ_{φ_0} be the associated matrix coefficient. Then*

$$\sup_{g\in ZK\sigma_n K} \Phi_{\varphi_0}(g) \ll_{\epsilon,q,c(\pi)} q^{(\theta-1/2+\epsilon)n}. \tag{3-8}$$

For notational simplicity, from now on we will just write

$$\sup_{g\in K\sigma_n K} \Phi_{\varphi_0}(g) \ll_{c(\pi)} q^{(\theta-1/2+\epsilon)n}. \tag{3-9}$$

For applications, we need the following lemma:

Lemma 3.7. *For $g \in \mathrm{GL}_2(\mathbb{F})$, $g = (g_{ij})$, define $v_{\min}(g) = \min\{v(g_{ij})\}$. Then*

$$g \in ZK\sigma_{v(\det(g))-2v_{\min}(g)}K. \tag{3-10}$$

Proof. By applying $\omega = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ on the left and/or right if necessary, we can assume without loss of generality that $v(g_{22}) = v_{\min}(g)$. Then we have

$$\begin{pmatrix} 1 & -g_{22}^{-1}g_{12} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -g_{22}^{-1}g_{21} & 1 \end{pmatrix} = \begin{pmatrix} g'_{11} & 0 \\ 0 & g_{22} \end{pmatrix}, \tag{3-11}$$

where $\begin{pmatrix} 1 & -g_{22}^{-1}g_{12} \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -g_{22}^{-1}g_{21} & 1 \end{pmatrix} \in K$. By considering the determinant on both sides, we get that $v(g'_{11}) = v(\det(g)) - v(g_{22}) \geq v(g_{22})$. The claim then follows easily. □

3A3. *The Whittaker function and matrix coefficients.*

Lemma 3.8. *Let $m \in \mathbb{F}$ with $v(m) = -j < 0$, and μ be a character of $O^\times$ with $c(\mu) = k > 0$. Then*

$$\left|\int_{v(x)=0} \psi(mx)\mu^{-1}(x)\, d^*x\right| = \begin{cases} \sqrt{q/((q-1)^2q^{k-1})} & \text{if } j = k, \\ 0, & \text{otherwise.} \end{cases} \tag{3-12}$$

This follows directly from, for example, [Corbett and Saha 2018, Lemma 2.3].

Definition 3.9. Define

$$\mathbf{1}_{\chi,n}(x) = \begin{cases} \chi(u) & \text{if } x = u\varpi^n \text{ for } u \in O^\times, \\ 0 & \text{otherwise.} \end{cases} \tag{3-13}$$

We will say that a function $f(x)$ consists of level i components (with coefficients) of L^2-norm h, if we can write

$$f(x) = \sum_{c(\chi)=i} \sum_{n\in\mathbb{Z}} a_{\chi,n}\mathbf{1}_{\chi,n}(x), \tag{3-14}$$

where each χ is a character of $O^\times$, and $h = \left(\sum_{c(\chi)=i}\sum_n |a_{\chi,n}|^2\right)^{1/2}$.

The following result is from [Hu 2018, Proposition 2.12].

Proposition 3.10. *Let π be a supercuspidal representation with $c(\pi) = c$, or a parabolically induced representation $\pi(\mu_1, \mu_2)$ where $c(\mu_1) = c(\mu_2) = k = c/2$. Let W be the L^2-normalized Whittaker function for a newform of π, and define*

$$W^{(i)}(a) = W\left(\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \varpi^i & 1 \end{pmatrix}\right).$$

Then:

1. $W^{(c)}(a) = \mathbf{1}_{1,0}(a)$.
2. *For $i = c-1 > 1$, $W^{(c-1)}(a)$ is supported only on $O^\times$, where it consists of level 1 components with L^2-norm $\sqrt{q(q-2)/(q-1)^2}$, and also a level 0 component with coefficient $-1/(q-1)$.*
3. *In general for $0 \leq i < c-1$, $i \neq c/2$, $W^{(i)}(a)$ is supported only on $\{a \in \mathbb{F} : v(a) = \min\{0, 2i - c\}\}$, where it consists of level $c - i$ components with L^2-norm 1.*

(4) *When $i = k > 1$, $W^{(c/2)}$ is supported on O, where it consists of level $c/2$ components with L^2-norm 1. When $i = k = 1$, $W^{(1)}(a)$ consists of a level 0 component on $O^\times$ with coefficient $-1/(q-1)$, and level 1 components on O with L^2-norm $\sqrt{q(q-2)/(q-1)^2}$.*

We need however to know more about $W^{(c/2)}(a)$.

Lemma 3.11. *When π is a (twist-)minimal supercuspidal representation with $c(\pi) = 2k$, then $W^{(k)}(a)$ is supported on $O^\times$. When $\pi = \pi(\mu_1, \mu_2)$ with $c(\mu_1) = c(\mu_2) = c(\mu_2/\mu_1) = k$, we have*

$$\sup_{v(a)=j>0} |W^{(k)}(a)| \ll_q \frac{1}{q^{j/2}}. \tag{3-15}$$

Proof. For the minimal supercuspidal representation case, the proof is essentially the same as for [Hu 2017, Corollary 2.18], where unramified central character is assumed.

Now suppose that $\pi = \pi(\mu_1, \mu_2)$ with $c(\mu_i) = k$. By [Hu 2017, Lemma 2.12], we can write

$$W^{(k)}(a) = C_0^{-1} \int_{\substack{v(u) \le -k, \\ u \notin \varpi^{-k}(-1+\varpi O_{\mathbb{F}})}} \mu_1^{-1}(1+u\varpi^k)\mu_2(-au)\psi(-au) \left| \frac{\varpi^k}{au(1+u\varpi^k)} \right|^{1/2} q^{-v(a)}\, du, \tag{3-16}$$

where $C_0 = \int_{u \in O_{\mathbb{F}}^\times} \mu_2(\varpi^{-k}u)\psi(\varpi^{-k}u)\, du$.

By Lemma 3.8, $|C_0| \asymp \frac{1}{q^{k/2}}$. We claim that

$$\left| \int_{\substack{v(u) \le -k, \\ u \notin \varpi^{-k}(-1+\varpi O_{\mathbb{F}})}} \mu_1^{-1}(1+u\varpi^k)\mu_2(-au)\psi(-au) \left| \frac{\varpi^k}{au(1+u\varpi^k)} \right|^{1/2} q^{-v(a)}\, du \right| \ll_q \frac{1}{q^{(k+v(a))/2}}. \tag{3-17}$$

Then

$$|W^{(k)}(a)| \ll_q \frac{1}{q^{v(a)/2}}. \tag{3-18}$$

To prove the claim, we shall use the p-adic stationary phase analysis.

Note that when $v(x) \ge k/2$, $\mu_i(1+x)$ becomes an additive character in x. Thus there exists $\alpha_i \in \mathbb{F}$ such that $v(\alpha_i) = -k$ and

$$\mu_i(1+x) = \psi(\alpha_i x) \tag{3-19}$$

when $v(x) \ge k/2$. The condition $c(\mu_2/\mu_1) = k$ implies that $\alpha_1 \not\equiv \alpha_2 \mod \varpi^{-k+1} O_{\mathbb{F}}$.

Recall that $v(a) = j > 0$. Now we write $u = u_0(1+\Delta u)$ for u_0 modulo $U_{\mathbb{F}}(\lceil k/2 \rceil) = 1 + \varpi^{\lceil k/2 \rceil} O_{\mathbb{F}}$ multiplicatively, $v(u_0) \le -k$, $u_0 \notin \varpi^{-k}(-1+\varpi O_{\mathbb{F}})$, $\Delta u \in \varpi^{\lceil k/2 \rceil} O_{\mathbb{F}}$. Then by (3-19) and the observation above on the nonzero contribution, the integral on the left-hand side of (3-17) can be rewritten as

$$q^{-j/2} \sum_{u_0} \int_{\Delta u} \mu_1^{-1}(1+u_0\varpi^k)\mu_2(-au_0)\psi(-au_0)\psi\left(-\alpha_1 \frac{u_0 \Delta u \varpi^k}{1+u_0\varpi^k} + \alpha_2 \Delta u - au_0 \Delta u\right) d(\Delta u). \tag{3-20}$$

For the integral in Δu to be nonvanishing, we need that

$$-\alpha_1 \frac{u_0 \varpi^k}{1+u_0\varpi^k} + \alpha_2 - a u_0 \equiv 0 \bmod \varpi^{-\lceil k/2 \rceil}. \tag{3-21}$$

Since $\alpha_1 \not\equiv \alpha_2$ and $v(a) > 0$, the above equation has solutions only when $v(u_0) = -k$ or $-k-j$, and in each of these cases one can solve a unique solution of u_0 modulo $U_{\mathbb{F}}(\lfloor k/2 \rfloor)$ (or at most q solutions modulo $U_{\mathbb{F}}(\lceil k/2 \rceil)$). When nonvanishing, the integral in Δu gives an additional factor of absolute value $q^{-\lceil k/2 \rceil}$, concluding the proof of (3-17). □

Remark 3.12. From the proof, the implied constant in Lemma 3.11 can be controlled by $\sqrt{q}$, though one can expect further square-root cancellation from the q solutions of u_0 when $\lceil k/2 \rceil > \lfloor k/2 \rfloor$.

From now on let π be unitary, and $\Phi(g)$ be the matrix coefficient associated to a newform φ, normalized so that $\Phi(1) = 1$. It is right $K_0(\varpi^c)$-invariant. By Lemma 3.4, to understand $\Phi(g)$, it suffices to understand $\Phi\left(\left(\begin{smallmatrix} a & m \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ \varpi^i & 1 \end{smallmatrix}\right)\right)$ for $0 \le i \le c$. So we define

$$\Phi^{(i)}(a,m) = \Phi\left(\begin{pmatrix} a & m \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \varpi^i & 1 \end{pmatrix}\right).$$

Remark 3.13. Note that when $v(a)$ and $v(m)$ are fixed, $\Phi^{(i)}(a,m)$ only depends on m/a, as Φ is actually bi-$K_0(\varpi^c)$-invariant. So we can think of it as a one-parameter function and talk about its level.

By [Hu 2018, Proposition 3.1], we have the following result on the matrix coefficient of the newform.

Proposition 3.14. *Let π be as in Proposition 3.10 and Φ be the normalized matrix coefficient of the newform in π:*

(i) *For $c-1 \le i \le c$, $\Phi^{(i)}(a,m)$ is supported on $\{(a,m) : a \in O^\times, v(m) \ge -1\}$. On the support, we have*

$$\Phi^{(i)}(a,m) = \begin{cases} 1 & \text{if } v(m) \ge 0 \text{ and } i = c, \\ -1/(q-1) & \text{if } v(m) = -1 \text{ and } i = c, \\ -1/(q-1) & \text{if } v(m) \ge 0 \text{ and } i = c-1. \end{cases} \tag{3-22}$$

When $v(a) = 0$, $v(m) = -1$ and $i = c-1 > 1$, $\Phi^{(i)}(a,m)$ consists of level 1 components with L^2-norm $q\sqrt{q-2}/(q-1)^2$, and also a level 0 component with coefficient $1/(q-1)^2$.

(ii) *For $0 \le i < c-1$, $i \ne c/2$, $\Phi^{(i)}(a,m)$ is supported on $\{(a,m) : v(a) = \min\{0, 2i-c\}, v(m) = i-c\}$, where it consists of level $c-i$ components with L^2-norm $\sqrt{q/((q-1)^2 q^{c-i-1})}$.*

(iii) *When $c = 2k$ is even and $i = c/2 = k > 1$, $\Phi^{(i)}(a,m)$ is supported on $\{(a,m) : v(a) \ge 0, v(m) = -k\}$, where it consists of level k components with L^2-norm $\sqrt{q/((q-1)^2 q^{c/2-1})}$.*

When $i = k = 1$, $\Phi^{(i)}(a,m)$ is supported on $\{(a,m) : v(a) \ge 0, v(m) \ge -1\}$. When $v(m) \ge 0$, its value is as in (i). When $v(m) = -1$, it consists of a level 0 component at $v(a) = 0$ with coefficient $1/(q-1)^2$, and level 1 components at $v(a) \ge 0$ with L^2-norm $q\sqrt{q-2}/(q-1)^2$.

Again we need more knowledge about $\Phi^{(c/2)}(a,m)$.

Lemma 3.15. *When π is a minimal supercuspidal representation with $c(\pi)=2k$, then $\Phi^{(k)}(a,m)$ is vanishing when $v(a)>0$. When $\pi=\pi(\mu_1,\mu_2)$ with $c(\mu_1)=c(\mu_2)=c(\mu_2/\mu_1)=k$, we have for $j>0$, $v(m)=-k$*

$$\int_{v(a)=j}|\Phi^{(k)}(a,m)|^2\,d^*a \ll_q \frac{1}{q^{k+j}}. \tag{3-23}$$

Proof. In general the unitary pairing in π can be computed from the Whittaker functions as follows:

$$\langle\varphi_1,\varphi_2\rangle=\int_{\mathbb{F}^\times} W_{\varphi_1}\left(\begin{pmatrix}\alpha & 0\\ 0 & 1\end{pmatrix}\right)\overline{W_{\varphi_2}\left(\begin{pmatrix}\alpha & 0\\ 0 & 1\end{pmatrix}\right)}\,d^*\alpha. \tag{3-24}$$

Here $\varphi_i\in\pi$, W_{φ_i} are associated Whittaker functions. First of all by Proposition 3.10 (1) and the fact that

$$W\left(\begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix}g\right)=\psi(x)W(g)$$

for any Whittaker function, we have

$$\Phi^{(i)}(a,m)=\int_{v(\alpha)=0}\psi(m\alpha)W^{(i)}(a\alpha)\,d^*\alpha. \tag{3-25}$$

The claim for the supercuspidal representation case follows directly from this and Lemma 3.11. Suppose from now on that $\pi=\pi(\mu_1,\mu_2)$. For any character χ on $O^\times$, we extend it to be a character of $\mathbb{F}^\times$ by requiring $\chi(\varpi)=1$. By Proposition 3.14(iii) and the Parseval–Plancherel identity,

$$\int_{v(a)=j}|\Phi^{(k)}(a,m)|^2\,d^*a=\sum_{c(\chi)=k}\left|\int_{v(a)=j}\Phi^{(i)}(a,m)\chi(a)\,d^*a\right|^2. \tag{3-26}$$

Note that

$$\left|\int_{v(a)=j}\Phi^{(i)}(a,m)\chi(a)\,d^*a\right|=\left|\int_{v(\alpha)=0}\psi(m\alpha)\chi^{-1}(\alpha)\,d^*\alpha\int_{v(a)=j}W^{(i)}(a)\chi(a)\,d^*a\right|. \tag{3-27}$$

By Lemma 3.8, $\left|\int_{v(\alpha)=0}\psi(m\alpha)\chi^{-1}(\alpha)\,d^*\alpha\right|\asymp 1/q^{k/2}$ is independent of χ as long as $c(\chi)=k$, and $W^{(k)}(a)$ also consists only of level k components. So we have

$$\int_{v(a)=j}|\Phi^{(k)}(a,m)|^2\,d^*a\asymp\frac{1}{q^k}\int_{v(a)=j}|W^{(i)}(a)|^2\,d^*a\ll\frac{1}{q^{k+j}}. \tag{3-28}$$

Here the last inequality follows from Lemma 3.11. □

3A4. *The relation between the local Rankin–Selberg integral and the local triple product integral.* By [Hsieh 2017, Proposition 5.1] we have the following:

Lemma 3.16. *Suppose that π_1 is a parabolically induced representation, and π_i satisfies the bound $\theta<\frac{1}{6}$ towards the Ramanujan conjecture. Let $\tilde{\pi}_i$ be the contragredient representation of π_i, $\varphi_i\in\pi_i$ and $\tilde{\varphi}_i\in\tilde{\pi}_i$.*

Let $(\cdot\,,\cdot)$ *be the natural* GL_2*-invariant pairing between* π_i *and* $\tilde{\pi}_i$*. Then*

$$\int_{Z\backslash\,\mathrm{GL}_2(\mathbb{F})}\prod_i(\pi_i(g)\varphi_i,\tilde{\varphi}_i)\,dg$$
$$=\zeta_{\mathbb{F}}(1)\int_{Z(\mathbb{F})N\backslash\,\mathrm{GL}_2(\mathbb{F})}W_{\varphi_2}(g)W_{\varphi_3}(Jg)\varphi_1(g)\,dg\int_{Z(\mathbb{F})N\backslash\,\mathrm{GL}_2(\mathbb{F})}W_{\tilde{\varphi}_2}(Jg)W_{\tilde{\varphi}_3}(g)\tilde{\varphi}_1(g)\,dg. \quad (3\text{-}29)$$

Here $J=\left(\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right)$ *and* $W_\varphi(Jg)$ *is the Whittaker function associated to* φ *with respect to* ψ^-*.*

In our setting, this lemma implies the following:

Corollary 3.17. *For* $i=1,2,3$, *suppose that the central character of* π_i *is trivial, and* φ_i *is a newform or a single translate of a newform,* L^2*-normalized. Suppose that* π_1 *is a parabolically induced representation. Then*

$$|I^T(\varphi_1,\varphi_2,\varphi_3)|=\zeta_{\mathbb{F}}(1)|I^{RS}(\varphi_1,\varphi_2,\varphi_3)|^2. \quad (3\text{-}30)$$

Proof. When π_i has the trivial central character, we have $\tilde{\pi}_i\simeq\pi_i$. We choose $\tilde{\varphi}_i\in\tilde{\pi}_i$ by requiring the same invariance for $\tilde{\varphi}_i$ as for φ_i. (Note that a newform φ^0 can be identified as being $K_0(\varpi^c)$-invariant, and $\pi(g)\varphi_0$ can be identified as being $gK_0(\varpi^c)g^{-1}$-invariant.) Then up to a constant of absolute value 1, the left-hand side of (3-29) can be identified with $I^T(\varphi_1,\varphi_2,\varphi_3)$. On the other hand, we have $\int_{Z(\mathbb{F})N\backslash\,\mathrm{GL}_2}W_{\varphi_2}(g)W_{\varphi_3}(Jg)\varphi_1(g)\,dg=I^{RS}(\varphi_1,\varphi_2,\varphi_3)$, and

$$\begin{aligned}\left|\int_{Z(\mathbb{F})N\backslash\,\mathrm{GL}_2(\mathbb{F})}W_{\tilde{\varphi}_2}(Jg)W_{\tilde{\varphi}_3}(g)\tilde{\varphi}_1(g)\,dg\right|&=\left|\int_{Z(\mathbb{F})N\backslash\,\mathrm{GL}_2(\mathbb{F})}W_{\tilde{\varphi}_2}(g)W_{\tilde{\varphi}_3}(Jg)\tilde{\varphi}_1(Jg)\,dg\right|\\&=\left|\int_{Z(\mathbb{F})N\backslash\,\mathrm{GL}_2(\mathbb{F})}W_{\tilde{\varphi}_2}(g)W_{\tilde{\varphi}_3}(Jg)\tilde{\varphi}_1(g)\,dg\right|\\&=|I^{RS}(\varphi_1,\varphi_2,\varphi_3)|.\end{aligned} \quad (3\text{-}31)$$

This completes the proof. □

Remark 3.18. Lemma 3.16 was originally found in [Michel and Venkatesh 2010, Lemma 3.4.2], which additionally requires π_2 or π_3 to be tempered. It was subsequently extended in [Nelson 2019; Nelson et al. 2014; Hsieh 2017] for various settings to nontempered cases. It was mainly used to reduce the computation of the local triple product integral to that of the local Rankin–Selberg integral. In this paper, we will use the same approach when π_1 is a complementary series representation. On the other hand when π_1 is tempered, we will instead use the lemma to reduce the computation of the local Rankin–Selberg integral to that of the local triple product integral.

3B. ***Proof of Proposition 3.1.*** We first show the symmetry between n and $c-c_1-n$ by using the Atkin–Lehner operator. Let $a_i\in\mathbb{C}^\times$ be the Atkin–Lehner eigenvalues of φ_i^0 for $i=1,2,3$, satisfying $|a_i|=1$. More specifically for $\omega_c=\left(\begin{smallmatrix}0&1\\-\varpi^c&0\end{smallmatrix}\right)$ which stabilizes the congruence subgroup $K_0(\varpi^c)$, we have by the uniqueness of the newform,

$$\pi_i(\omega_c)\varphi_i^0=a_i\varphi_i^0 \quad (3\text{-}32)$$

for $i=2,3$. On the other hand for $c_1=c(\pi_1)$ and $\omega_{c_1}=\left(\begin{smallmatrix}0 & 1\\ -\varpi^{c_1} & 0\end{smallmatrix}\right)$, we also have

$$\pi_1(\omega_{c_1})\varphi_1^0=a_1\varphi_1^0. \tag{3-33}$$

Thus

$$\begin{aligned}\left|I^{\mathrm{RS}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0,\varphi_2^0,\varphi_3^0\right)\right| &= \left|I^{\mathrm{RS}}\left(\pi_1\left(\omega_c\begin{pmatrix}\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0,\pi_2(\omega_c)\varphi_2^0,\pi_3(\omega_c)\varphi_3^0\right)\right|\\ &= \left|I^{\mathrm{RS}}\left(\pi_1\left(\varpi^{c-n-c_1}\begin{pmatrix}\varpi^{-c+n+c_1} & 0\\ 0 & 1\end{pmatrix}\omega_{c_1}\right)\varphi_1^0,\varphi_2^0,\varphi_3^0\right)\right|\\ &= \left|I^{\mathrm{RS}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-c+n+c_1} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0,\varphi_2^0,\varphi_3^0\right)\right|. \end{aligned}\tag{3-34}$$

The same equality is true for the absolute value of the triple product integral.

3B1. *Bounding the Rankin–Selberg integral.* We first consider the case when π_1 is a principal series representation satisfying the bound θ towards the Ramanujan conjecture. By the discussion above we shall assume from now on that

$$n\geq c-c_1-n. \tag{3-35}$$

Note that $\pi_1\left(\left(\begin{smallmatrix}\varpi^{-n} & 0\\ 0 & 1\end{smallmatrix}\right)\right)\varphi_1^0$ is an old form of level c_1+n. Let $c'=\max\{c,n+c_1\}$. By the definition of the Rankin–Selberg integral and Lemma 3.4,

$$\begin{aligned}&\left|I^{\mathrm{RS}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0,\varphi_2^0,\varphi_3^0\right)\right|\\ &\qquad\leq\sum_{i=0}^{c'}A_i\int_{a\in\mathbb{F}^\times}|W_{\varphi_2^0}W^-_{\varphi_3^0}|\left(\begin{pmatrix}a & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \varpi^i & 1\end{pmatrix}\right)|\varphi_1^0|\left(\begin{pmatrix}a\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \varpi^{i-n} & 1\end{pmatrix}\right)|a|^{-1}\,d^*a.\end{aligned}\tag{3-36}$$

According to Proposition 3.10, $|W_{\varphi_2^0}W^-_{\varphi_3^0}|\left(\left(\begin{smallmatrix}a & 0\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & 0\\ \varpi^i & 1\end{smallmatrix}\right)\right)$ is in general supported on $\{a:v(a)=\min\{0,2i-c\}\}$. Then by the Cauchy–Schwarz inequality and the bounds for the L^2-norms for the individual Whittaker functions in Proposition 3.10, we have

$$\int_{v(a)=\min\{0,2i-c\}}|W_{\varphi_2^0}W^-_{\varphi_3^0}|\left(\begin{pmatrix}a & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \varpi^i & 1\end{pmatrix}\right)|a|^{-1}\,d^*a\ll_q q^{\min\{0,2i-c\}}. \tag{3-37}$$

This is still true when π_2,π_3 are principal series representations, $i=c/2$, and $v(a)=j>0$, as the decrease in $W_{\varphi_i^0}\left(\left(\begin{smallmatrix}a & 0\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & 0\\ \varpi^i & 1\end{smallmatrix}\right)\right)$ by Lemma 3.11 cancels the increase of $|a|^{-1}$ in the integral.

To bound the contribution from φ_1^0, we first assume that $c_1=0$, φ_1^0 is spherical and $\varphi_1^0(1)=1$. By our assumption on n, we get that $n\geq c/2$. We perform the standard Iwasawa decomposition for $\left(\begin{smallmatrix}a\varpi^{-n} & 0\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & 0\\ \varpi^{i-n} & 1\end{smallmatrix}\right)$ as follows: When $i\geq n$, it is already in the standard form. When $i<n$, we have

$$\begin{pmatrix}a\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \varpi^{i-n} & 1\end{pmatrix}=\begin{pmatrix}a\varpi^{-i} & a\varpi^{-n}\\ 0 & \varpi^{i-n}\end{pmatrix}\begin{pmatrix}0 & -1\\ 1 & \varpi^{n-i}\end{pmatrix}. \tag{3-38}$$

We first treat the cases $i \neq c/2$ or $v(a) \leq 0$. Then by the definition of φ_1^0 and that $v(a) = \min\{0, 2i - c\}$,

$$|\varphi_1^0|\left(\begin{pmatrix} a\varpi^{-n} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \varpi^{i-n} & 1 \end{pmatrix}\right) \ll \begin{cases} q^{(1/2+\theta)n} & \text{when } i \geq n \geq c/2, \\ q^{(1/2+\theta)\max\{2i-n,c-n\}} & \text{when } i < n,\ \max\{2i-n, c-n\} \geq 0, \\ q^{(1/2-\theta)\max\{2i-n,c-n\}} & \text{when } i < n,\ \max\{2i-n, c-n\} \leq 0. \end{cases} \tag{3-39}$$

Then we have

$$\begin{aligned} &\left| I^{\mathrm{RS}}\left(\pi_1\left(\begin{pmatrix} \varpi^{-n} & 0 \\ 0 & 1 \end{pmatrix}\right)\varphi_1^0, \varphi_2^0, \varphi_3^0\right)\right| \\ &\qquad \ll \sum_{i=0}^{c'} A_i q^{\min\{0,2i-c\}} \sup_{v(\alpha)=\min\{0,2i-c\}} |\varphi_1^0|\left(\begin{pmatrix} a\varpi^{-n} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \varpi^{i-n} & 1 \end{pmatrix}\right) \\ &\qquad \ll_{q,\epsilon} \frac{1}{q^{(1/2-\theta-\epsilon)n}}. \end{aligned} \tag{3-40}$$

The last inequality follows from the fact that the main contribution comes from $i = n$ if $\theta > 0$, and $c/2 \leq i \leq n$ when $\theta = 0$. Now we consider the possible contribution from the pieces where $i = c/2$ and $v(a) = j > 0$. In these cases (3-37) and (3-38) are still true and

$$|\varphi_1^0|\left(\begin{pmatrix} a\varpi^{-n} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \varpi^{c/2-n} & 1 \end{pmatrix}\right) \ll \begin{cases} q^{(1/2+\theta)(c-n-j)} & \text{if } n + j \leq c, \\ q^{(1/2-\theta)(c-n-j)} & \text{if } n + j \geq c. \end{cases} \tag{3-41}$$

Thus the contribution from these pieces are controlled by the piece $j = 0$.

Now suppose that π_1 is a principal series representation induced from two unitary ramified characters of equal levels, and in particular $c_1 > 0$ is even. The newform φ_1^0 in this case is supported on $B\left(\begin{smallmatrix} 1 & 0 \\ \varpi^{c_1/2} & 1 \end{smallmatrix}\right)K_0(\varpi^{c_1})$. For simplicity we normalize it such that $\varphi_1^0\left(\left(\begin{smallmatrix} 1 & 0 \\ \varpi^{c_1/2} & 1 \end{smallmatrix}\right)\right) = 1$. To L^2-normalize it there will be an additional factor involving c_1.

Then only the term $i - n = c_1/2$ remains in the sum in (3-36). Note that $i = n + c_1/2 \geq c/2$ for this term so $v(a) = 0$. Note that $\left(\begin{smallmatrix} a\varpi^{-n} & 0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ \varpi^{i-n} & 1 \end{smallmatrix}\right)$ is already written in the shape of $B\left(\begin{smallmatrix} 1 & 0 \\ \varpi^{c_1/2} & 1 \end{smallmatrix}\right)K_0(\varpi^{c_1})$ so

$$|\varphi_1^0|\left(\begin{pmatrix} a\varpi^{-n} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \varpi^{i-n} & 1 \end{pmatrix}\right) \ll q^{(1/2+\theta)n}.$$

Thus

$$\left| I^{\mathrm{RS}}\left(\pi_1\left(\begin{pmatrix} \varpi^{-n} & 0 \\ 0 & 1 \end{pmatrix}\right)\varphi_1^0, \varphi_2^0, \varphi_3^0\right)\right| \ll \frac{1}{q^i} q^{(1/2+\theta)n} \ll_{c_1} \frac{1}{q^{(1/2-\theta-\epsilon)n}}. \tag{3-42}$$

When $i = c/2$ and π_2, π_3 are principal series representation, the previous argument on the contribution from $v(a) = j > 0$ still applies here.

Now if π_1 is a special representation, then in particular it is tempered. Instead of bounding the Rankin–Selberg integral directly in this case, we use Corollary 3.17 to reduce the problem to bounding the triple product integral, which is to be done immediately below.

3B2. *Bounding the triple product integral.* We first consider the case when π_1 is tempered. Then its matrix coefficient satisfies (3-8) with $\theta = 0$.

Again by symmetry we can assume that $n \geq (c-c_1)/2$. Then for $c' = \max\{c, n+c_1\}$,

$$I^{\mathrm{T}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0, \varphi_2^0, \varphi_3^0\right) = \sum_{i=0}^{c'} A_i \int_{a,m} \Phi_{\varphi_2^0}\Phi_{\varphi_3^0}\left(\begin{pmatrix}a & m\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \varpi^i & 1\end{pmatrix}\right)\Phi_{\varphi_1^0}\left(\begin{pmatrix}a & m\varpi^n\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \varpi^{i-n} & 1\end{pmatrix}\right)|a|^{-1}\, d^*a\, dm. \tag{3-43}$$

Like before, we shall use the individual sup norm bound for $\Phi_{\varphi_1^0}$ on each piece, and use the Cauchy–Schwarz inequality and the bounds for the L^2-norms of $\Phi_{\varphi_2^0}$, $\Phi_{\varphi_3^0}$ to bound the integrals. For simplicity, let

$$J_i(S) = \int_{m\in S}\int_{v(a)=\min\{0,2i-c\}} |\Phi_{\varphi_2^0}\Phi_{\varphi_3^0}|\left(\begin{pmatrix}a & m\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \varpi^i & 1\end{pmatrix}\right)|a|^{-1}\, d^*a\, dm. \tag{3-44}$$

According to Proposition 3.14 and the Cauchy–Schwarz inequality, we have

$$J_i(S) \ll \begin{cases} q^{\min\{0,2i-c\}} & \text{if } 0 \leq i < c-1,\ S = \{v(m) = i-c\},\\ 1 & \text{if } i = c-1,\ S = \{v(m) = -1\},\\ q^{-2} & \text{if } i = c-1,\ S = \{v(m) \geq 0\},\\ q^{-1} & \text{if } i \geq c,\ S = \{v(m) = -1\},\\ 1 & \text{if } i \geq c,\ S = \{v(m) \geq 0\}.\end{cases} \tag{3-45}$$

To control $|\Phi_{\varphi_1^0}|$, consider the case $n \geq c/2$ first. Let $g = \left(\begin{smallmatrix}a & m\varpi^n\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & 0\\ \varpi^{i-n} & 1\end{smallmatrix}\right)$. We shall use Lemma 3.7 to identify the double-K-coset for g. Consider the cases $i \neq c/2$ or $v(a) \leq 0$ first:

(1) When $i > n \geq c/2$, it is clear that $g \in K$.

(2) When $c/2 \leq i \leq n$, $v(a) = 0$ and $v(m\varpi^n) \geq 0$. So $v_{\min}(g) = i-n$, $v(\det(g)) = 0$ and $g \in ZK\sigma_{2(n-i)}K$, where $v_{\min}(g)$ is as in Lemma 3.7 and σ_n is as in (3-6).

(3) When $i \leq c/2$, $v(a) = 2i-c$. So $v_{\min}(g) \leq i-n$, $v(\det(g)) = 2i-c$ and $g \in ZK\sigma_j K$ for some $j \geq 2n-c$.

Then by Corollary 3.6 we have the bound

$$\left|\Phi_{\varphi_1^0}\left(\begin{pmatrix}a & m\varpi^n\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \varpi^{i-n} & 1\end{pmatrix}\right)\right| \ll_{c_1} \begin{cases} 1 & \text{if } i > n,\\ q^{-2\delta_0(n-i)} & \text{if } c/2 \leq i \leq n,\\ q^{-2\delta_0(n-c/2)} & \text{if } i \leq c/2.\end{cases} \tag{3-46}$$

Here $\delta_0 = \frac{1}{2} - \epsilon$ as π_1 is tempered.

By applying (3-45) and (3-46) to (3-43), one can see that the main contribution comes from $c/2 \leq i \leq n$, and

$$\left|I^{\mathrm{T}}\left(\pi_1\left(\begin{pmatrix}\varpi^{-n} & 0\\ 0 & 1\end{pmatrix}\right)\varphi_1^0, \varphi_2^0, \varphi_3^0\right)\right| \ll_{c_1} \frac{1}{q^{(1-\epsilon)n}}. \tag{3-47}$$

Note that when $i = c/2$ and π_2, π_3 are principal series representations, their matrix coefficients can be nonvanishing when $v(a) > 0$. For fixed $v(a) = j \geq 0$, Lemma 3.15 implies that

$$\int_{v(m)=i-c} \int_{v(a)=j>0} |\Phi_{\varphi_2^0} \Phi_{\varphi_3^0}| \left(\begin{pmatrix} a & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \varpi^i & 1 \end{pmatrix} \right) |a|^{-1} \, d^*a \, dm \ll_q 1, \tag{3-48}$$

similar to (3-45). On the other hand, $v_{\min}(g) = i - n$, $v(\det(g)) = j$, so by Lemma 3.7 $g \in ZK\sigma_{2n-2i+j}K$. Then by Corollary 3.6, we have

$$\left| \Phi_{\varphi_1^0} \left(\begin{pmatrix} a & m\varpi^n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \varpi^{i-n} & 1 \end{pmatrix} \right) \right| \ll_{c_1} \frac{1}{q^{(1/2-\epsilon)(j+2n-2i)}}. \tag{3-49}$$

So the main contribution is still from $j = 0$.

Consider the case $(c - c_1)/2 \leq n < c/2$ now. The arguments are similar so we shall skip some details here. By the Cartan decomposition and Corollary 3.6,

$$\left| \Phi_{\varphi_1^0} \left(\begin{pmatrix} a & m\varpi^n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \varpi^{i-n} & 1 \end{pmatrix} \right) \right| \ll_{c_1} \begin{cases} 1 & \text{if } i > c - n, \\ q^{-2\delta_0(c-n-i)} & \text{if } c/2 \leq i \leq c - n, \\ q^{-2\delta_0(c/2-n)} & \text{if } i \leq c/2. \end{cases} \tag{3-50}$$

The main contribution comes from $c/2 \leq i \leq c - n$ and

$$\left| I^{\mathrm{T}} \left(\pi_1 \left(\begin{pmatrix} \varpi^{-n} & 0 \\ 0 & 1 \end{pmatrix} \right) \varphi_1^0, \varphi_2^0, \varphi_3^0 \right) \right| \ll_{c_1} \frac{1}{q^{(1-\epsilon)(c-n)}} \ll_{c_1} \frac{1}{q^{(1-\epsilon)n}}. \tag{3-51}$$

The last inequality follows from the condition $(c - c_1)/2 \leq n < c/2$. One can similarly argue as above for the case where $i = c/2$ and π_2, π_3 are principal series representations. We shall skip it here.

Now if π_1 is not tempered, it is automatically a principal series representation. In this case we use Corollary 3.17, and apply the upper bound for the Rankin–Selberg integral discussed in Section 3B1.

4. Proof of the main result

From now on we work in the global setting. We first need some more preparations.

4A. ***The global period integrals.*** Let $\chi = \otimes \chi_v$ be a Hecke character of $\mathbb{A}_{\mathbb{F}}^{\times}$. Let $\varphi_{1,s} \in \pi(\chi, \chi^{-1}, s)$, $\varphi_{1,s} = \otimes \varphi_{1,v,s}$ such that each local component $\varphi_{1,v,s} \in \pi(\chi_v, \chi_v^{-1}, s)$ satisfies (2-1). Denote by

$$E_{\varphi_{1,s}} = \sum_{\gamma \in B(\mathbb{F}) \backslash \mathrm{GL}_2(\mathbb{F})} \varphi_{1,s}(\gamma g) \tag{4-1}$$

According to, for example, [Bump 1997, Proposition 3.8.2],

$$\int_{[\mathrm{GL}_2]} \varphi_2(g)\varphi_3(g) E_{\varphi_{1,s}}(g) \, dg = \frac{L(\pi_2 \times \pi_3 \times \chi, s)}{L(2s, \chi^2)} \prod_{v \in S} I_v^{RS,0}(\varphi_{1,s,v}, \varphi_{2,v}, \varphi_{3,v}) \tag{4-2}$$

Here S contains Archimedean places as well as finite places where any of the local test vectors is not spherical. $I_v^{RS}(\varphi_{1,s,v}, \varphi_{2,v}, \varphi_{3,v})$ is as in (3-1),

$$I_v^{RS,0}(\varphi_{1,s,v}, \varphi_{2,v}, \varphi_{3,v}) = \frac{L_v(2s, \chi^2)}{L_v(\pi_2 \times \pi_3 \times \chi, s)} I_v^{RS}(\varphi_{1,s,v}, \varphi_{2,v}, \varphi_{3,v}) \tag{4-3}$$

at finite places, and $I_v^{RS,0}(\varphi_{1,s,v}, \varphi_{2,v}, \varphi_{3,v}) = I_v^{RS}(\varphi_{1,s,v}, \varphi_{2,v}, \varphi_{3,v})$ at archimedean places.

For general π_1, by Ichino's work [2008], we have

$$\left|\int_{[\mathrm{GL}_2]} \varphi_1(g)\varphi_2(g)\varphi_3(g)\,dg\right|^2 = \frac{\zeta_{\mathbb{F}}^2(2)L\big(\pi_1\times\pi_2\times\pi_3,\frac{1}{2}\big)}{8L(\pi_1\times\pi_2\times\pi_3,Ad,1)}\prod_{v\in S} I_v^{T,0}(\varphi_{1,v},\varphi_{2,v},\varphi_{3,v}). \tag{4-4}$$

Here $I_v^T(\varphi_{1,v},\varphi_{2,v},\varphi_{3,v})$ is as in (3-2),

$$I_v^{T,0}(\varphi_{1,v},\varphi_{2,v},\varphi_{3,v}) = \frac{L_v(\Pi_v,Ad,1)}{\zeta_v^2(2)L_v(\Pi_v,1/2)} I_v^T(\varphi_{1,v},\varphi_{2,v},\varphi_{3,v}) \tag{4-5}$$

at finite places and $I_v^{T,0}(\varphi_{1,v},\varphi_{2,v},\varphi_{3,v}) = I_v^T(\varphi_{1,v},\varphi_{2,v},\varphi_{3,v})$ at archimedean places.

4B. ***The spectral decomposition.*** Let F be a rapidly decreasing automorphic form on $\mathrm{GL}_2(\mathbb{A}_{\mathbb{F}})$ that is invariant by $K_0(N)$. For a unitary cuspidal automorphic representation π, let $\mathcal{B}(\pi,N)$ be an orthonormal basis for its subspace of $K_0(N)$-invariant elements. Similarly for $\pi_{\chi,s}=\pi(\chi,\chi^{-1},s)$ where χ is unitary and $s=\frac{1}{2}+it$, let $\mathcal{B}(\pi_{\chi,s},N)$ be an orthonormal basis for the subspace of $K_0(N)$-invariant elements under the unitary pairing

$$\langle\varphi_{1,s},\varphi_{2,s}\rangle = \prod_v \int_{K_v} \varphi_{1,s,v}(k)\overline{\varphi_{2,s,v}(k)}\,dk. \tag{4-6}$$

Note that all local components of this pairing are $\mathrm{GL}_2(\mathbb{F}_v)$-invariant according to [Bump 1997, Section 2.6]. Let $E_{\varphi_{\chi,s}}$ be the Eisenstein series associated to $\varphi_{\chi,s}\in\mathcal{B}(\pi_{\chi,s},N)$ as in (4-1)

Then we have the following variant of spectral decomposition for F (see, for example, [Michel and Venkatesh 2010, Section 2.2] for a more general version).

Lemma 4.1.

$$F = \frac{\langle F,1\rangle}{\langle 1,1\rangle}1 + \sum_{\pi,C(\pi)\,|\,N}\ \sum_{\varphi\in\mathcal{B}(\pi,N)} \langle F,\varphi\rangle\varphi + \sum_{\chi,C(\chi)^2\,|\,N}\ \sum_{\varphi_{\chi,s}\in\mathcal{B}(\pi_{\chi,1/2+it},N)} \int_t \langle F,E_{\varphi_{\chi,1/2+it}}\rangle E_{\varphi_{\chi,1/2+it}}\,dt \tag{4-7}$$

4C. ***The proof.*** Let $\Omega=\otimes\Omega_v$ be a Hecke character of $\mathbb{A}_{\mathbb{E}}^\times$. Recall that if $\mathbb{E}_v$ is split over $\mathbb{F}_v$ with isomorphism $\iota_v:\mathbb{E}_v\to\mathbb{F}_v\times\mathbb{F}_v$, we can identify Ω_v with $(\Omega_{1,v}\otimes\Omega_{2,v})\circ\iota_v$, and define $c(\Omega_v)=\max\{c(\Omega_{i,v})\}$, $U_{\mathbb{E}_v}(j)=\iota_v^{-1}(U_{\mathbb{F}_v}(j)\times U_{\mathbb{F}_v}(j))$. In a slightly more general setting we first show that the embedding of $\mathbb{E}_v$ does not affect the result as long as it is fixed.

Lemma 4.2. *Suppose that $c(\Omega_v)=j$. If $\mathbb{E}_v$ embedded in $M_2(\mathbb{F}_v)$ is not upper triangular, then there exists $i>0$ dependent only on j and the embedding of $\mathbb{E}_v$ such that $K_0(\varpi_v^i)\subset U_{\mathbb{E},v}(j)B(O_v)\subset U_{\mathbb{E},v}(j)K_0(\varpi_v^n)$ for any $n>0$.*

Proof. By assumption there exists $t=\left(\begin{smallmatrix} * & * \\ c & d\end{smallmatrix}\right)\in\mathbb{E}_v$ for $c\neq 0$. Let l be large enough such that for any $a\in O_v^\times$ and $b\in O_v$ the followings are satisfied:

$$a+b\varpi_v^l t = \begin{pmatrix} * & * \\ b\varpi_v^l c & a+b\varpi_v^l d\end{pmatrix} \in O_v^\times U_{\mathbb{E},v}(j)\cap K_v,\ v(\varpi_v^l c),\ v(\varpi_v^l d)>0.$$

The integer l depends only on j and the embedding of $\mathbb{E}_v$. Then for $i=v(\varpi_v^l c)$ and any $\left(\begin{smallmatrix} k_1 & k_2 \\ k_3 & k_4\end{smallmatrix}\right)\in K_0(\varpi_v^i)$, the equation

$$b\varpi_v^l ck_1 + (a+b\varpi_v^l d)k_3 = 0 \tag{4-8}$$

has a unique solution $b \in O_v$ for any fixed $a \in O_v^\times$. This implies that $(a+b\varpi_v^l t)\left(\begin{smallmatrix} k_1 & k_2 \\ k_3 & k_4 \end{smallmatrix}\right) \in B \cap K_v = B(O_v)$, and thus $K_0(\varpi_v^i) \subset U_{\mathbb{E},v}(j)B(O_v)$. Note that $O_v^\times \subset B(O_v)$. □

Corollary 4.3. *For f and $\mathbb{E}$ as in Theorem 1.1, there exists i dependent only on Ω and the embedding of $\mathbb{E}$ such that*

$$\int_{[\mathbb{E}^\times]} |f|^2(e)\Omega(e)\,de = \int_{[\mathbb{E}^\times]} \left(\frac{1}{\mathrm{Vol}(K_0(\varpi_v^i))} \int_{k \in K_0(\varpi_v^i)} \rho(k)|f|^2(e)\,dk \right) \Omega(e)\,de. \tag{4-9}$$

Here $k \in K_0(\varpi_v^i)$ is considered as an element in $\mathrm{GL}_2(\mathbb{A}_\mathbb{F})$ with all the other local components being 1, and $\rho(k)$ is the right regular representation of $\mathrm{GL}_2(\mathbb{A}_\mathbb{F})$ on the space of automorphic forms.

Proof. First of all, the right-hand side of (4-9) is equal to

$$\frac{1}{\mathrm{Vol}(K_0(\varpi_v^i))} \int_{k \in K_0(\varpi_v^i)} \int_{[\mathbb{E}^\times]} |f|^2(ek)\Omega(e)\,de\,dk. \tag{4-10}$$

Recall that $\mathbb{E}$ is a quadratic field extension, so $\mathbb{E}_v$ embedded in $\mathrm{GL}_2(\mathbb{F}_v)$ is indeed not upper triangular. For any $k \in K_0(\varpi_v^i)$, write $k = tk_0$ for $k_0 \in K_0(\varpi_v^c)$ and $t \in U_{\mathbb{E},v}(j)$ by Lemma 4.2. Then

$$\int_{[\mathbb{E}^\times]} |f|^2(ek)\Omega(e)\,de = \int_{[\mathbb{E}^\times]} |f|^2(etk_0)\Omega(e)\,de = \int_{[\mathbb{E}^\times]} |f|^2(e)\Omega(et^{-1})\,de = \int_{[\mathbb{E}^\times]} |f|^2(e)\Omega(e)\,de. \tag{4-11}$$

Here we have used the $K_0(\varpi_v^c)$-invariance for f and $U_{\mathbb{E},v}(j)$-invariance for Ω. □

Now let $F(g) = 1/\mathrm{Vol}(K_0(\varpi_v^i)) \int_{k \in K_0(\varpi_v^i)} |f|^2(gk)\,dk$. By construction F is $K_0(\varpi_v^i)$-invariant. Thus when spectrally decomposing F, the length of the spectral sum is fixed for fixed Ω and $\mathbb{E}$.

We can now prove Theorem 1.1. Let Ω be a fixed Hecke character of $\mathbb{A}_\mathbb{E}^\times$ that is invariant by $U_{\mathbb{E},v}(j)$. We assume without loss of generality that $\mu_f(1) = \langle |f|^2, 1 \rangle = 1$. Let $N = q^i$ be fixed as $c \to \infty$. Applying the spectral decomposition, we get that

$$\begin{aligned}
\int_{[\mathbb{E}^\times]} &|f|^2(e)\Omega(e)\,de \\
&= \int_{[\mathbb{E}^\times]} F(e)\Omega(e)\,de \\
&= \frac{\langle F, 1 \rangle}{\langle 1, 1 \rangle} \int_{[\mathbb{E}^\times]} \Omega(e)\,de + \sum_{\sigma, C(\sigma) \mid N} \sum_{\varphi \in \mathcal{B}(\sigma, N)} \langle F, \varphi \rangle \int_{[\mathbb{E}^\times]} \varphi(e)\Omega(e)\,de \\
&\qquad + \sum_{\chi, C(\chi)^2 \mid N} \sum_{\varphi_{\chi,s} \in \mathcal{B}(\pi_{\chi,1/2+it}, N)} \int_t \langle F, E_{\varphi_{\chi,1/2+it}} \rangle \int_{[\mathbb{E}^\times]} E_{\varphi_{\chi,1/2+it}}(e)\Omega(e)\,de\,dt \\
&= \frac{\langle |f|^2, 1 \rangle}{\langle 1, 1 \rangle} \int_{[\mathbb{E}^\times]} \Omega(e)\,de + \sum_{\sigma, C(\sigma) \mid N} \sum_{\varphi \in \mathcal{B}(\sigma, N)} \langle |f|^2, \varphi \rangle \int_{[\mathbb{E}^\times]} \varphi(e)\Omega(e)\,de \\
&\qquad + \sum_{\chi, C(\chi)^2 \mid N} \sum_{\varphi_{\chi,s} \in \mathcal{B}(\pi_{\chi,1/2+it}, N)} \int_t \langle |f|^2, E_{\varphi_{\chi,1/2+it}} \rangle \int_{[\mathbb{E}^\times]} E_{\varphi_{\chi,1/2+it}}(e)\Omega(e)\,de\,dt.
\end{aligned} \tag{4-12}$$

In the last equality, we have used the fact that for $\varphi' = \varphi$, 1, or $E_{\varphi_{\chi,1/2+it}}$,

$$\begin{aligned}\langle F, \varphi'\rangle &= \frac{1}{\mathrm{Vol}(K_0(\varpi_v^i))}\int_{k\in K_0(\varpi_v^i)} \langle \rho(k)|f|^2, \varphi'\rangle\, dk \\ &= \frac{1}{\mathrm{Vol}(K_0(\varpi_v^i))}\int_{k\in K_0(\varpi_v^i)} \langle |f|^2, \rho(k^{-1})\varphi'\rangle\, dk \\ &= \langle |f|^2, \varphi'\rangle. \end{aligned}\tag{4-13}$$

Note that φ and $E_{\varphi_{\chi,1/2+it}}$ must have trivial central characters. The main term is the constant term

$$\frac{\langle |f|^2, 1\rangle}{\langle 1, 1\rangle}\int_{[\mathbb{E}^\times]} \Omega(e)\, de = \frac{1}{\mu(1)}\int_{[\mathbb{E}^\times]} \Omega(e)\, de$$

by normalization, which is exactly what we want. So we need to prove a power saving in the depth aspect for both the cuspidal spectrum and the continuous spectrum.

We control $\langle |f|^2, \varphi\rangle$ and $\langle |f|^2, E_{\varphi_{\chi,1/2+it}}\rangle$ by following the strategy in [Nelson et al. 2014] and [Hu 2018]. In particular suppose that $f \in \pi$ and $\varphi \in \sigma$. By the convexity bound,

$$L\big(\sigma\times\pi\times\pi, \tfrac{1}{2}\big) \ll_{C(\sigma)} q^{(1/2+\epsilon)c(\pi)}. \tag{4-14}$$

The other L-functions in (4-4) can be bounded by $q^{\epsilon c(\pi)}$.

Then we use Proposition 3.1 to control the local triple product integrals. Note that in general the diagonal translates of newforms do not provide an orthonormal basis. But since N is fixed, one can write an orthonormal basis in terms of linear combinations of diagonal translates of newforms, with all coefficients fixed. (Interested readers can see [Blomer and Milićević 2015, Lemma 9] for the explicit coefficients, which is not necessary for us. Note that their paper works with $\mathbb{Q}$, but that particular lemma is purely local and can be easily extended to general situations.) Thus by applying Proposition 3.1 for each individual term, we get that for $\varphi \in \mathcal{B}(\sigma, N)$,

$$I_v^T(\varphi_v, f_v, f_v) \ll_{C(\sigma),q,\epsilon} \frac{1}{q^{(1-2\theta-\epsilon)c(\pi)}}. \tag{4-15}$$

Thus by the triple product formula (4-4), and the fact that $C(\sigma)\mid N$ depends only on $C(\Omega)$ and the embedding of $\mathbb{E}$, we have

$$\begin{aligned}|\langle |f|^2, \varphi\rangle|^2 &\ll_{\Omega,\mathbb{E},q,\epsilon} q^{(1/2+\epsilon)c(\pi)}\frac{1}{q^{(1-2\theta-\epsilon)c(\pi)}} = q^{(2\theta-1/2+\epsilon)c(\pi)}, \\ |\langle |f|^2, \varphi\rangle| &\ll_{\Omega,\mathbb{E},q,\epsilon} q^{(\theta-1/4+\epsilon)c(\pi)}. \end{aligned}\tag{4-16}$$

One can obtain a similar upper bound for $\langle |f|^2, E_{\varphi_{\chi,1/2+it}}\rangle$ using the Rankin–Selberg integral formula (4-2), the convexity bound for the L-function and the bound for its local integral in Proposition 3.1.

Note here that as the Archimedean parameters of φ or t in $E_{\varphi_{\chi,1/2+it}}$ go to infinity, the corresponding period integral $\langle |f|^2, \varphi\rangle$ or $\langle |f|^2, E_{\varphi_{\chi,1/2+it}}\rangle$ becomes rapidly decreasing. This is because the local integrals in (4-2) and (4-4) at archimedean places give additional Gamma factors that are rapidly decreasing. Effectively we can consider the Archimedean parameters of φ and t to be bounded.

Since $\mathbb{E}$, Ω are fixed and φ, $E_{\varphi_{\chi,1/2+it}}$ have bounded levels and archimedean components, we claim that

$$\int_{[\mathbb{E}^\times]} \varphi(e)\Omega(e)\,de,\ \int_{[\mathbb{E}^\times]} E_{\varphi_{\chi,1/2+it}}(e)\Omega(e)\,de \ll_{\Omega,\mathbb{E}} 1. \tag{4-17}$$

One can use the period integral formula in [Waldspurger 1985] to see this if necessary, but basically everything in the integral is either fixed or bounded.

Now combining (4-16), (4-17) with (4-12), we get the control over the error terms as claimed in Theorem 1.1.

Acknowledgements

The author thanks Paul Nelson for helpful discussions from which the new approach was discovered. The author would also like to thank Simon Marshall and Philippe Michel for helpful conversations, and the Max Planck Institute of Mathematics for support as this work was partially done during his visit there.

References

[Blomer and Brumley 2011] V. Blomer and F. Brumley, "On the Ramanujan conjecture over number fields", *Ann. of Math.* (2) **174**:1 (2011), 581–605. MR Zbl

[Blomer and Milićević 2015] V. Blomer and D. Milićević, "The second moment of twisted modular L-functions", *Geom. Funct. Anal.* **25**:2 (2015), 453–516. MR Zbl

[Bump 1997] D. Bump, *Automorphic forms and representations*, Cambridge Studies in Advanced Mathematics **55**, Cambridge University Press, 1997. MR Zbl

[Burq et al. 2007] N. Burq, P. Gérard, and N. Tzvetkov, "Restrictions of the Laplace–Beltrami eigenfunctions to submanifolds", *Duke Math. J.* **138**:3 (2007), 445–486. MR Zbl

[Casselman 1973] W. Casselman, "On some results of Atkin and Lehner", *Math. Ann.* **201** (1973), 301–314. MR Zbl

[Christianson et al. 2013] H. Christianson, J. A. Toth, and S. Zelditch, "Quantum ergodic restriction for Cauchy data: interior QUE and restricted QUE", *Math. Res. Lett.* **20**:3 (2013), 465–475. MR Zbl

[Corbett and Saha 2018] A. Corbett and A. Saha, "On the order of vanishing of newforms at cusps", *Math. Res. Lett.* **25**:6 (2018), 1771–1804. MR Zbl

[Dyatlov and Zworski 2013] S. Dyatlov and M. Zworski, "Quantum ergodicity for restrictions to hypersurfaces", *Nonlinearity* **26**:1 (2013), 35–52. MR Zbl

[Ghosh et al. 2013] A. Ghosh, A. Reznikov, and P. Sarnak, "Nodal domains of Maass forms I", *Geom. Funct. Anal.* **23**:5 (2013), 1515–1568. MR Zbl

[Holowinsky and Soundararajan 2010] R. Holowinsky and K. Soundararajan, "Mass equidistribution for Hecke eigenforms", *Ann. of Math.* (2) **172**:2 (2010), 1517–1528. MR Zbl

[Hsieh 2017] M.-L. Hsieh, "Hida families and p-adic triple product L-functions", preprint, 2017. arXiv

[Hu 2017] Y. Hu, "Triple product formula and the subconvexity bound of triple product L-function in level aspect", *Amer. J. Math.* **139**:1 (2017), 215–259. MR Zbl

[Hu 2018] Y. Hu, "Triple product formula and mass equidistribution on modular curves of level N", *Int. Math. Res. Not.* **2018**:9 (2018), 2899–2943. MR Zbl

[Ichino 2008] A. Ichino, "Trilinear forms and the central values of triple product L-functions", *Duke Math. J.* **145**:2 (2008), 281–307. MR Zbl

[Lindenstrauss 2006] E. Lindenstrauss, "Invariant measures and arithmetic quantum unique ergodicity", *Ann. of Math.* (2) **163**:1 (2006), 165–219. MR Zbl

[Marshall 2011] S. Marshall, "Mass equidistribution for automorphic forms of cohomological type on GL_2", *J. Amer. Math. Soc.* **24**:4 (2011), 1051–1103. MR Zbl

[Marshall 2016] S. Marshall, "Geodesic restrictions of arithmetic eigenfunctions", *Duke Math. J.* **165**:3 (2016), 463–508. MR Zbl

[Michel and Venkatesh 2006] P. Michel and A. Venkatesh, "Equidistribution, L-functions and ergodic theory: on some problems of Yu. Linnik", pp. 421–457 in *International Congress of Mathematicians*, vol. II, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Michel and Venkatesh 2010] P. Michel and A. Venkatesh, "The subconvexity problem for GL_2", *Publ. Math. Inst. Hautes Études Sci.* **111** (2010), 171–271. MR Zbl

[Nelson 2011] P. D. Nelson, "Equidistribution of cusp forms in the level aspect", *Duke Math. J.* **160**:3 (2011), 467–501. MR Zbl

[Nelson 2019] P. D. Nelson, "Subconvex equidistribution of cusp forms: reduction to Eisenstein observables", *Duke Math. J.* **168**:9 (2019), 1665–1722. MR Zbl

[Nelson et al. 2014] P. D. Nelson, A. Pitale, and A. Saha, "Bounds for Rankin–Selberg integrals and quantum unique ergodicity for powerful levels", *J. Amer. Math. Soc.* **27**:1 (2014), 147–191. MR Zbl

[Rudnick and Sarnak 1994] Z. Rudnick and P. Sarnak, "The behaviour of eigenstates of arithmetic hyperbolic manifolds", *Comm. Math. Phys.* **161**:1 (1994), 195–213. MR Zbl

[Soundararajan 2010] K. Soundararajan, "Quantum unique ergodicity for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$", *Ann. of Math.* (2) **172**:2 (2010), 1529–1538. MR Zbl

[Toth and Zelditch 2013] J. A. Toth and S. Zelditch, "Quantum ergodic restriction theorems: manifolds without boundary", *Geom. Funct. Anal.* **23**:2 (2013), 715–775. MR Zbl

[Venkatesh 2010] A. Venkatesh, "Sparse equidistribution problems, period bounds and subconvexity", *Ann. of Math.* (2) **172**:2 (2010), 989–1094. MR Zbl

[Waldspurger 1985] J.-L. Waldspurger, "Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie", *Compositio Math.* **54**:2 (1985), 173–242. MR Zbl

[Young 2018] M. P. Young, "Equidistribution of Eisenstein series on geodesic segments", *Adv. Math.* **340** (2018), 1166–1218. MR Zbl

Communicated by Peter Sarnak
Received 2019-01-04 Revised 2019-11-13 Accepted 2019-12-13

huyueke2012@gmail.com *Yau Mathematical Sciences Center, Tsinghua University, Beijing, China*

ALGEBRA AND NUMBER THEORY 14:4 (2020)
dx.doi.org/10.2140/ant.2020.14.947

The basepoint-freeness threshold and syzygies of abelian varieties

Federico Caucci

We show how a natural constant introduced by Jiang and Pareschi for a polarized abelian variety encodes information about the syzygies of the section ring of the polarization. As a particular case this gives a quick and characteristic-free proof of Lazarsfeld's conjecture on syzygies of abelian varieties, originally proved by Pareschi in characteristic zero.

1. Introduction

Throughout this paper we will work with abelian varieties over an algebraically closed field $\mathbb{K}$. Jiang and Pareschi [2017] introduced and studied the (generic) cohomological ranks $h^i(A, \mathcal{F}\langle x\underline{l}\rangle)$ of a (bounded complex of) $\mathbb{Q}$-twisted coherent sheaf on a polarized abelian variety $(A, \underline{l})$. This defines *cohomological rank functions* of $\mathcal{F}$ with respect to the polarization $\underline{l}$

$$h^i_{\mathcal{F},\underline{l}} : \mathbb{Q} \to \mathbb{Q}_{\geq 0},$$ [1]

as follows

$$h^i_{\mathcal{F},\underline{l}}(x) := h^i(A, \mathcal{F}\langle x\underline{l}\rangle).$$

In [loc. cit.] it is observed that these functions are already very interesting in the case $\mathcal{F} = \mathcal{I}_p$, where $\mathcal{I}_p$ is the ideal sheaf of a closed point $p \in A$. Indeed the *basepoint-freeness threshold*

$$\epsilon_1(\underline{l}) := \mathrm{Inf}\{x \in \mathbb{Q} \mid h^1_{\mathcal{I}_p,\underline{l}}(x) = 0\},$$ [2]

has the following properties:

(a) $\epsilon_1(\underline{l}) \leq 1$ *and* $\epsilon_1(\underline{l}) < 1$ *if and only if the polarization* $\underline{l}$ *is basepoint-free*, i.e., any line bundle L representing $\underline{l}$ has no base points.

(b) $\epsilon_1(\underline{l}) < \frac{1}{2}$ *if and only if* $\underline{l}$ *is projectively normal,* meaning that L is projectively normal for all line bundles L representing the class $\underline{l}$ [Jiang and Pareschi 2017, Corollary E].

In this paper we go further on item (b), proving that $\epsilon_1(\underline{l})$ indeed encodes information about the syzygies of the section algebra of L. In recent years syzygies of abelian varieties has received considerable attention. On the one hand Pareschi [2000] (see also [Pareschi and Popa 2004]), building partially on previous works of Kempf [1989; 1991], proved, in characteristic zero, Lazarsfeld's conjecture on syzygies of

MSC2010: primary 14K05; secondary 14C20, 14F17, 14Q20.
Keywords: syzygies, abelian varieties, Fourier–Mukai transform.

[1] In [loc. cit.] such functions are extended to (continuous) real functions, but in this paper we don't need this.

[2] In [loc. cit.] this is denoted by $\beta(\underline{l})$.

abelian varieties endowed with a polarization which is a multiple of a given one. This was is in turn a generalization of classical results of Koizumi and Mumford. On the other hand, more recently Küronya and Lozovanu [2019], Ito [2018] and Lozovanu [2018], building on previous work of Hwang and To [2011] and Lazarsfeld, Pareschi and Popa [Lazarsfeld et al. 2011], used completely different methods — involving local positivity and Nadel vanishing theorem — to prove (over $\mathbb{C}$) effective statements for the syzygies of abelian varieties of dimension 2 and 3 endowed with *any* polarization, in particular with a primitive polarization.

In this paper we show a general result, Theorem 1.1 below, partially generalizing (b) to higher syzygies. This provides at the same time a surprisingly quick proof of Lazarsfeld's conjecture, extending it to abelian varieties defined over a ground field of arbitrary characteristic, and a proof of the criterion of [Lazarsfeld et al. 2011] relating local positivity and syzygies.

Turning to details, we first recall some terminology about syzygies of projective varieties. Let X be a projective variety and let L be an ample line bundle on X. For an integer $p \geq 0$, the line bundle L is said to *satisfy the property* (N_p) if the first p steps of the minimal graded free resolution of the section algebra $R_L = \bigoplus_m H^0(X, L^m)$ over the polynomial ring $S_L = \operatorname{Sym} H^0(X, L)$ are linear (we refer to Section 4 for the precise definition). Thus (N_0) means that R_L is generated in degree 0 as an S_L-module, i.e., that L is projectively normal (*normally generated* in Mumford's terminology [1970]); (N_1) means that in addition the homogeneous ideal $I_{X/\mathbb{P}}$ of X in $\mathbb{P} = \mathbb{P}(H^0(X, L)^\vee)$ is generated by quadrics (*normally presented* in [Mumford 1970]); (N_2) means that the relations among these quadrics are generated by linear ones (this is the first nonclassical condition) and so on. These notions were introduced by Green [1984] and the present terminology was introduced in [Green and Lazarsfeld 1986]. Our main result is the following:

Theorem 1.1. *Let $(A, \underline{l})$ be a polarized abelian variety defined over an algebraically closed field $\mathbb{K}$ and let p be a nonnegative integer. If*

$$\epsilon_1(\underline{l}) < \frac{1}{p+2},$$

then the property (N_p) holds for $\underline{l}$, i.e., it holds for any line bundle L representing $\underline{l}$.

Corollary 1.2. *Let $m \in \mathbb{N}$. If*

$$\epsilon_1(\underline{l}) < \frac{m}{p+2},$$

then the polarization $m\underline{l}$ satisfies the property (N_p).

Proof. By definition (see Section 2) we have $h^1_{\mathcal{I}_p, m\underline{l}}(x) = h^1_{\mathcal{I}_p, \underline{l}}(mx)$, therefore $\epsilon_1(m\underline{l}) = \epsilon_1(\underline{l})/m$. Now Theorem 1.1 applies to $m\underline{l}$, because $\epsilon_1(m\underline{l}) < 1/(p+2)$. ☐

A classical result of Koizumi [1976] states that if L is an ample line bundle on a complex abelian variety and $m \geq 3$, then L^m is projectively normal (see [Sekiguchi 1976; 1977; Sasaki 1977] for a proof of the analog result in positive characteristic, based on Mumford's ideas). Moreover, a well-known theorem of Mumford and Kempf says that, when $m \geq 4$, the homogeneous ideal of A in the embedding given by L^m is generated by quadrics [Mumford 1970; Kempf 1991, Theorem 6.13], i.e., L^m is normally presented. Based on these classical facts and motivated by a result of Green [1984] on higher syzygies for curves, Lazarsfeld conjectured that, for an ample line bundle L on an abelian variety, L^m satisfies the property (N_p) if $m \geq p+3$ [Lazarsfeld 1989, Conjecture 1.5.1]. This was proved by Pareschi [2000] in characteristic zero. Pareschi and Popa [2004] also proved a stronger version of it.

We have that Corollary 1.2 gives a very quick — and characteristic-free — proof of Lazarsfeld's conjecture. Indeed, by (a) above,

$$\epsilon_1(\underline{l}) \le 1 < \frac{p+3}{p+2}.$$

Moreover it also implies that the polarization $m\underline{l}$ satisfies the property (N_p), as soon as $m \ge p+2$ and $\underline{l}$ is basepoint-free (see [Pareschi and Popa 2004] for a more precise result). Indeed, if $\underline{l}$ is basepoint-free, then

$$\epsilon_1(\underline{l}) < 1 = \frac{p+2}{p+2},$$

thanks again to (a) above.

More in general, defining

$$t(\underline{l}) := \max\{t \in \mathbb{N} \mid \epsilon_1(\underline{l}) \le \tfrac{1}{t}\},$$

we have

Theorem 1.3. *Let p and t be nonnegative integers with $p+1 \ge t$. Let $\underline{l}$ be a basepoint-free polarization on A such that $t(\underline{l}) \ge t$. Then the property (N_p) holds for $m\underline{l}$, as soon as $m \ge p+3-t$.*

However, one of the main feature of Theorem 1.1 is the chance to be applied to *primitive* polarizations, i.e., those that cannot be written as a multiple of another one. This is one of the reasons why it would be quite interesting the compute, or at least bound from above, the invariant $\epsilon_1(\underline{l})$ of polarized abelian varieties $(A, \underline{l})$. In this perspective, as already mentioned, an interesting issue arises in connection with a criterion of Lazarsfeld, Pareschi and Popa [2011], where they prove

> *if there exists an effective $\mathbb{Q}$-divisor F such that its multiplier ideal $\mathcal{J}(A, F)$ is the ideal sheaf of the identity point of the abelian variety A and $\frac{1}{p+2}\underline{l} - F$ is ample, then $\underline{l}$ satisfies the property N_p* (see [Küronya and Lozovanu 2019; Ito 2018; Lozovanu 2018]).

Therefore one is lead to consider the threshold

$$r(\underline{l}) := \mathrm{Inf}\{r \in \mathbb{Q} \mid \exists \text{ an effective } \mathbb{Q}\text{-divisor } F \text{ on } A \text{ such that } r\underline{l} - F \text{ is ample and } \mathcal{J}(A, F) = \mathcal{I}_0\}.^3$$

The relation with the basepoint-freeness threshold is in the following proposition, based on Nadel's vanishing.

Proposition 1.4. *Assume $\mathbb{K} = \mathbb{C}$. Then $\epsilon_1(\underline{l}) \le r(\underline{l})$.*

This, combined with Theorem 1.1, provides a different and simpler proof of the criterion of [Lazarsfeld et al. 2011].

Finally, we note that in the papers [Küronya and Lozovanu 2019; Ito 2018] for dimension 2 and [Lozovanu 2018] for dimension 3, the authors, in the spirit of Green's and Green and Lazarsfeld's conjectures on curves, show explicit geometric conditions ensuring the property (N_p) by means of upper bounds on the threshold $r(\underline{l})$ (or related invariants) and applying the criterion of [Lazarsfeld et al. 2011].

---

[3] Note that this set is nonempty, i.e., $r(\underline{l}) < +\infty$. Proof: Let k be a sufficiently large positive integer such that the Seshadri constant of $M = L^k$ is strictly bigger than $2 \dim A$. Such a k exists because of the homogeneity of the Seshadri constant. Then, by Lemma 1.2 of [Lazarsfeld et al. 2011], there exists an effective $\mathbb{Q}$-divisor F on A such that $\mathcal{J}(A, F) = \mathcal{I}_0$ and $F \equiv_{\mathrm{num}} \frac{1}{2}(1-c)M$, for some $0 < c \ll 1$. If we now take $r > \frac{1}{2}(1-c)k$, we have that $r\underline{l} - F$ is ample.

This suggests to look for similar estimates directly for the basepoint-freeness threshold $\epsilon_1(\underline{l})$. Namely one could ask if $\epsilon_1(\underline{l})$ is less or equal to

$$\mathrm{Inf}\{r \in \mathbb{Q}^+ \mid (D_r^{\dim Z} \cdot Z) > (\dim Z)^{\dim Z} \text{ for any abelian subvariety } \{0\} \neq Z \subseteq A\},$$

where $D_r := rL$ (see in particular [Ito 2018, Question 4.2]). This is true for complex abelian surfaces, thanks to the Proposition 1.4 and [Ito 2018].

The paper is organized as follows: In Section 2 we recall the definition and some basic properties of cohomological rank functions, and show that, despite the fact that in [Jiang and Pareschi 2017] the authors assume that the characteristic of the ground field is zero, the basic theory of cohomological rank functions works over an algebraically closed ground field of arbitrary characteristic as well. Finally, in this section we prove Proposition 1.4.

In Section 3 we prove the basic properties of the threshold $\epsilon_1(\underline{l})$ needed in the proof of the main results.

In Section 4 we show a criterion, due to Kempf [1989], reducing the property (N_p) of syzygies to the surjectivity of certain multiplication maps of global sections, inductively defined. This is easily proved and well-known in characteristic zero (see e.g., [Ein and Lazarsfeld 1993, proof of Corollary 2.2] or [Pareschi 2000, Lemma 4.1(a)]). Kempf's approach is more complicated, but has the advantage of working in arbitrary characteristic. Since Kempf's argument is somewhat obscure, we provide full details. We hope that this will be useful for extending to arbitrary characteristic some of known results concerning syzygies of projective varieties in characteristic zero.

In Section 5 we prove the Theorems 1.1 and 1.3.

Notation. Let A be an abelian variety over an algebraically closed field, and let $\dim A = g$. For $b \in \mathbb{Z}$,

$$\mu_b : A \to A, \quad x \mapsto bx$$

denotes the multiplication-by-b isogeny of degree b^{2g}. A polarization $\underline{l}$ on A is the class of an ample line bundle L in $\mathrm{NS}(A) = \mathrm{Pic} A/\mathrm{Pic}^0 A$. For a polarization $\underline{l}$ on A, the corresponding isogeny is denoted

$$\varphi_{\underline{l}} : A \to \hat{A},$$

where $\hat{A} = \mathrm{Pic}^0 A$ is the dual abelian variety. Recall that $\deg(\varphi_{\underline{l}}) = \chi(\underline{l})^2 = (h^0(\underline{l}))^2$. We denote by $\mathcal{P}$ the normalized Poincaré line bundle on $A \times \hat{A}$, and by $R\Phi_{\mathcal{P}} : \mathrm{D}^b(A) \to \mathrm{D}^b(\hat{A})$ the Fourier–Mukai–Poincaré equivalence [Mukai 1981]. Here $\mathrm{D}^b(A)$ denotes the bounded derived category of coherent sheaves on A. For $\alpha \in \hat{A}$, the corresponding line bundle on A is denoted by $P_\alpha = \mathcal{P}|_{A\times\{\alpha\}}$. Given a complex $\mathcal{F} \in \mathrm{D}^b(A)$, we denote by $\mathcal{F}^\vee = R\mathcal{H}om(\mathcal{F}, \mathcal{O}_A)$ its derived dual, and by $h^i_{\mathrm{gen}}(A, \mathcal{F})$ the dimension of the hypercohomology $H^i(A, \mathcal{F} \otimes P_\alpha)$, for α general in $\hat{A}$.

2. Cohomological rank functions on abelian varieties

Given $\mathcal{F} \in \mathrm{D}^b(A)$, $i \in \mathbb{Z}$ and a polarization $\underline{l}$ on A, Jiang and Pareschi [2017] considered cohomological rank functions

$$h^i_{\mathcal{F},\underline{l}} : \mathbb{Q} \to \mathbb{Q}_{\geq 0}$$

defined as follows:

$$h^i_{\mathcal{F},\underline{l}}(x) = h^i_{\mathcal{F}}(x\underline{l}) := \frac{1}{b^{2g}} h^i_{\mathrm{gen}}(A, \mu_b^*(\mathcal{F}) \otimes L^{ab}),$$

where $x = a/b \in \mathbb{Q}$ and $b > 0$. Since $\mu_b^*(\underline{l}) = b^2\underline{l}$, the pullback via μ_b of the rational class $(a/b)\underline{l}$ is $ab\underline{l}$. Moreover $\deg(\mu_b) = b^{2g}$, therefore, as explained in Remark 2.2 of [loc. cit.], one may think of $h^i_{\mathcal{F},\underline{l}}(x)$ as the (generic) cohomological rank $h^i(A, \mathcal{F}\langle x\underline{l}\rangle)$ of the $\mathbb{Q}$-*twisted complex* $\mathcal{F}\langle x\underline{l}\rangle$, which is defined similarly to [Lazarsfeld 2004, Section 6.2A]. Namely, $\mathcal{F}\langle x\underline{l}\rangle$ is the equivalence class of the pair $(\mathcal{F}, x\underline{l})$, where the equivalence is by definition

$$(\mathcal{F}\otimes L^m, x\underline{l}) \sim (\mathcal{F}, (m+x)\underline{l}),$$

for *any* line bundle L representing $\underline{l}$ and $m \in \mathbb{Z}$. Note that an "untwisted" object $\mathcal{F}$ may be naturally seen as the $\mathbb{Q}$-twisted object $\mathcal{F}\langle 0\underline{l}\rangle$. Moreover we have that $\mathcal{F}\otimes P_\alpha\langle x\underline{l}\rangle = \mathcal{F}\langle x\underline{l}\rangle$, for any $\alpha \in \mathrm{Pic}^0(A)$.

In [Jiang and Pareschi 2017] the authors introduced such notion assuming that the characteristic of the ground field $\mathbb{K}$ is zero. However the above definition makes sense in any characteristic. The main point consists in showing that it does not depend on the representation $x = a/b$. To this purpose we need to verify that the quantity $h^i_{\mathrm{gen}}(A, \mathcal{F})$ is multiplicative with respect to any isogeny μ_m:

$$h^i_{\mathrm{gen}}(A, \mu_m^*\mathcal{F}) = m^{2g} h^i_{\mathrm{gen}}(A, \mathcal{F}). \tag{2-1}$$

This is checked in [loc. cit.] under the assumption that $\mathrm{char}(\mathbb{K}) = 0$. However the same thing can be checked removing such assumption as follows. By cohomology and base change, $h^i_{\mathrm{gen}}(A, \mu_m^*\mathcal{F})$ is the generic rank of the Fourier–Mukai–Poincaré transform $R^i\Phi_{\mathcal{P}}(\mu_m^*\mathcal{F})$. Moreover $R^i\Phi_{\mathcal{P}}(\mu_m^*\mathcal{F}) = \hat{\mu}_{m*}R^i\Phi_{\mathcal{P}}(\mathcal{F})$ [Mukai 1981, (3.4)], where $\hat{\mu}_m : \hat{A} \to \hat{A}$ is the dual isogeny of μ_m, i.e., it is the multiplication-by-m isogeny of $\hat{A}$. Since the morphism $\hat{\mu}_m$ is in any case flat, the generic rank of $\hat{\mu}_{m*}R^i\Phi_{\mathcal{P}}(\mathcal{F})$ is that of $R^i\Phi_{\mathcal{P}}(\mathcal{F})$ multiplied by the degree of $\hat{\mu}_m$. Therefore we get (2-1). Granting this, $h^i_{\mathcal{F}}(x\underline{l})$ is well-defined: if we take another representation of x, say $x = (am)/(bm)$, then

$$\begin{aligned} h^i_{\mathcal{F}}(x\underline{l}) &= \frac{1}{(bm)^{2g}} h^i_{\mathrm{gen}}(A, \mu_{bm}^*(\mathcal{F})\otimes L^{abm^2}) \\ &= \frac{1}{(bm)^{2g}} h^i_{\mathrm{gen}}(A, \mu_m^*(\mu_b^*(\mathcal{F})\otimes L^{ab})) \\ &= \frac{1}{b^{2g}} h^i_{\mathrm{gen}}(A, \mu_b^*(\mathcal{F})\otimes L^{ab}). \end{aligned}$$

Remark 2.1. Although we won't need this in this paper, we remark that from the above discussion it follows that the basic properties satisfied by the cohomological rank functions described in Section 2 of [Jiang and Pareschi 2017] — especially the fundamental transformation formula with respect to the Fourier–Mukai–Poincaré transform Proposition 2.3 of [loc. cit.] and its consequences — work in any characteristic.

Using the cohomological rank functions it is possible to introduce several invariants attached to a polarized abelian variety $(A, \underline{l})$. Let us recall that, given a line bundle L that represents the class $\underline{l}$, the *kernel bundle* M_L associated to L is by definition the kernel of the evaluation map $H^0(A, L)\otimes\mathcal{O}_A \to L$. If L is basepoint-free, then M_L sits in the exact sequence

$$0 \to M_L \to H^0(A, L)\otimes\mathcal{O}_A \to L \to 0.$$

Definition 2.2. Let $(A, \underline{l})$ be a polarized abelian variety. Then we consider

$$\epsilon_1(\underline{l}) := \mathrm{Inf}\{x \in \mathbb{Q} \mid h^1_{\mathcal{I}_p}(x\underline{l}) = 0\},$$

where $\mathcal{I}_p$ is the ideal sheaf of a closed point $p \in A$ and, if $\underline{l}$ is basepoint-free

$$\kappa_1(\underline{l}) := \mathrm{Inf}\{x \in \mathbb{Q} \mid h^1_{M_L}(x\underline{l}) = 0\},$$

where M_L is the kernel bundle associated to a line bundle L representing $\underline{l}$.

Remark 2.3. The above invariants are well-defined, i.e., $\epsilon_1(\underline{l})$ does not depend on the point p, and $\kappa_1(\underline{l})$ is independent from the representing line bundle L. We point out that — although there no examples so far — $\epsilon_1(\underline{l})$ and $\kappa_1(\underline{l})$ could be irrational. However, as will be clear later on, this does not create any trouble.

The relation between the above two constants was established by Jiang and Pareschi.

Theorem 2.4 [Jiang and Pareschi 2017, Theorem D]. *Let $\underline{l}$ be a basepoint-free polarization. Then*

$$\kappa_1(\underline{l}) = \frac{\epsilon_1(\underline{l})}{1-\epsilon_1(\underline{l})}.$$

Remark 2.5. From this result, in [loc. cit.] it is derived that $\kappa_1(\underline{l}) < 1$, i.e., $\underline{l}$ is projectively normal, if and only if $\epsilon_1(\underline{l}) < \frac{1}{2}$ (see in particular [loc. cit., Corollary 8.2(b)]). Our Theorem 1.1 is an extension of the "if" implication to higher syzygies.

Proof of Proposition 1.4. Only in this subsection we make the assumption that the ground field $\mathbb{K}$ is $\mathbb{C}$.

Let $r \in \mathbb{Q}$ such that there exists an effective $\mathbb{Q}$-divisor F on A with

$$rL - F \text{ ample}, \tag{2-2}$$

$$\mathcal{I}_0 = \mathcal{J}(A, F). \tag{2-3}$$

To prove the proposition we need to prove that

$$h^1_{\mathcal{I}_0}(r\underline{l}) = 0. \tag{2-4}$$

Writing $r = a/b$ with $b > 0$, this means that

$$h^1_{\mathrm{gen}}(L^{ab} \otimes \mu_b^* \mathcal{I}_0) = 0.$$

But, by (2-3), the left-hand side is $h^1_{\mathrm{gen}}(L^{ab} \otimes \mu_b^* \mathcal{J}(A, F)) = h^1_{\mathrm{gen}}(L^{ab} \otimes \mathcal{J}(A, \mu_b^* F))$, where we used that forming multiplier ideals commutes with pulling back under étale morphism (see [Lazarsfeld 2004, Example 9.5.44]). Since $\mu_b^* F \equiv_{\mathrm{num}} b^2 F$, it follows from (2-2) that $L^{ab} - \mu_b^* F$ is ample. Therefore (2-4) follows from Nadel's vanishing.

3. Generic vanishing of $\mathbb{Q}$-twisted sheaves on abelian varieties

Following Section 5 of [Jiang and Pareschi 2017], one can extend the usual notions of *generic vanishing* to the $\mathbb{Q}$-twisted setting.

Definition/Theorem 3.1 [Jiang and Pareschi 2017, Theorem 5.1]. (1) *A $\mathbb{Q}$-twisted* sheaf $\mathcal{F}\langle x\underline{l}\rangle$, *with $x = a/b$, is said to be GV if*

$$\operatorname{codim}_{\hat{A}} \operatorname{Supp}(R^i\Phi_{\mathcal{P}}((\mu_b^*\mathcal{F})\otimes L^{ab})) \geq i, \quad \textit{for all } i > 0.$$

Equivalently the transform[4] $R\Phi_{\mathcal{P}^\vee}((\mu_b^*\mathcal{F}^\vee)\otimes L^{-ab})$ *is a sheaf concentrated in cohomological degree g, i.e.,*

$$R\Phi_{\mathcal{P}^\vee}((\mu_b^*\mathcal{F}^\vee)\otimes L^{-ab}) = R^g\Phi_{\mathcal{P}^\vee}((\mu_b^*\mathcal{F}^\vee)\otimes L^{-ab})[-g].$$

(2) *It is said to be $IT(0)$ if the transform*

$$R\Phi_{\mathcal{P}}((\mu_b^*\mathcal{F})\otimes L^{ab}) = R^0\Phi_{\mathcal{P}}((\mu_b^*\mathcal{F})\otimes L^{ab})$$

is concentrated in cohomological degree 0.

Remark 3.2. (a) The above definitions do not depend on the representation $x = a/b$. For example for any i, $R^i\Phi_{\mathcal{P}}(\mu_m^*\mathcal{F}) = \hat{\mu}_{m*}R^i\Phi_{\mathcal{P}}(\mathcal{F})$ [Mukai 1981, (3.4)] where $\hat{\mu}_m$ is the dual isogeny of μ_m, therefore by cohomology and base change we see that $\operatorname{Supp}(R^i\Phi_{\mathcal{P}}(\mu_m^*\mathcal{F}))$ corresponds to the image of $\operatorname{Supp}(R^i\Phi_{\mathcal{P}}(\mathcal{F}))$ via the isogeny $\hat{\mu}_m$.

(b) They neither depend on the line bundle L representing the class $\underline{l}$. Indeed, thanks to the exchange of translations and tensor product by elements of $\operatorname{Pic}^0 A$ [Mukai 1981, (3.1)], if L_0 is another line bundle algebraically equivalent to L, then $R^i\Phi_{\mathcal{P}}((\mu_b^*\mathcal{F})\otimes L_0^{ab})$ is a translate of $R^i\Phi_{\mathcal{P}}((\mu_b^*\mathcal{F})\otimes L^{ab})$.

By cohomology and base change one has that

$$\operatorname{Supp}(R^i\Phi_{\mathcal{P}}((\mu_b^*\mathcal{F})\otimes L^{ab})) \subseteq \{\alpha \in \hat{A} \mid H^i(A, (\mu_b^*\mathcal{F})\otimes L^{ab}\otimes P_\alpha) \neq 0\} =: V^i((\mu_b^*\mathcal{F})\otimes L^{ab}) \tag{3-1}$$

and, if $V^{i+1}((\mu_b^*\mathcal{F})\otimes L^{ab}) = \varnothing$, then equality holds. Moreover we have that the $\mathbb{Q}$-twisted sheaf $\mathcal{F}\langle x\underline{l}\rangle$ is *GV* if and only if

$$\operatorname{codim}_{\hat{A}} V^i((\mu_b^*\mathcal{F})\otimes L^{ab}) \geq i,$$

for all $i > 0$ and for any representation $x = a/b$ [Pareschi and Popa 2011a, Lemma 3.6]. By cohomology and base change again, $\mathcal{F}\langle x\underline{l}\rangle$ is $IT(0)$ if and only if

$$V^i((\mu_b^*\mathcal{F})\otimes L^{ab}) = \varnothing$$

for all $i > 0$ and for any representation $x = a/b$. In particular we see that an $IT(0)$ $\mathbb{Q}$-twisted sheaf is *GV*.

These generic vanishing concepts are strongly related to the invariants introduced in Definition 2.2, as explained in Section 8 of [Jiang and Pareschi 2017]. Namely, we have:

Lemma 3.3 [Jiang and Pareschi 2017, page 25]. *Given two polarizations $\underline{l}$ and $\underline{n}$ — with $\underline{n}$ basepoint-free — and a rational number x, the fact that $\epsilon_1(\underline{l}) < x$ (resp. $\kappa_1(\underline{n}) < x$) is equivalent to the fact that the $\mathbb{Q}$-twisted sheaf $\mathcal{I}_p\langle x\underline{l}\rangle$ (resp. $M_N\langle x\underline{n}\rangle$) is $IT(0)$.*

For the reader's convenience we explicitly write down the case of $\epsilon_1(\underline{l})$: assume that $\epsilon_1(\underline{l}) < x \in \mathbb{Q}$ and fix a sufficiently small $\eta > 0$ such that $x_0 := \epsilon_1(\underline{l}) + \eta \in \mathbb{Q}$ and $x_0 < x$. By (3-1), $\mathcal{I}_p\langle x_0\underline{l}\rangle$ is *GV*, therefore Hacon's criterion (see [Jiang and Pareschi 2017, Theorem 5.2(a)]) implies that $\mathcal{I}_p\langle (x_0+(x-x_0))\underline{l}\rangle = \mathcal{I}_p\langle x\underline{l}\rangle$ is $IT(0)$. Conversely suppose that $\mathcal{I}_p\langle x\underline{l}\rangle$ is $IT(0)$, then $\mathcal{I}_p\langle (x-y)\underline{l}\rangle$ is still $IT(0)$, for a sufficiently small

[4] Recall that $R\Phi_{\mathcal{P}^\vee} = (-1_{\hat{A}})^* R\Phi_{\mathcal{P}}$.

$y \in \mathbb{Q}^+$ [Jiang and Pareschi 2017, Theorem 5.2(c)]. Then $\epsilon_1(\underline{l}) < x - y < x$. For $\kappa_1(\underline{n})$, the argument is similar.

The following is a $\mathbb{Q}$-twisted analog of a well-known property of "preservation of vanishing"[Pareschi and Popa 2011b, Proposition 3.1].

Proposition 3.4. *Assume that $\mathcal{F}$ and $\mathcal{G}$ are coherent sheaves, and that one of them is locally free. If $\mathcal{F}\langle x\underline{l}\rangle$ is $IT(0)$ and $\mathcal{G}\langle y\underline{l}\rangle$ is GV, then $\mathcal{F}\langle x\underline{l}\rangle \otimes \mathcal{G}\langle y\underline{l}\rangle := (\mathcal{F}\otimes\mathcal{G})\langle (x+y)\underline{l}\rangle$ is $IT(0)$.*

Proof. Let $x = a/b$ and $y = c/d$, with $b, d > 0$. So $x + y = (ad + bc)/(bd)$. We want to prove that $\mu_{bd}^*(\mathcal{F}\otimes\mathcal{G})\otimes L^{(ad+bc)bd}$ is an $IT(0)$ sheaf. By hypothesis $\mathcal{F}\langle x\underline{l}\rangle$ is $IT(0)$, hence

$$\mu_d^*((\mu_b^*\mathcal{F})\otimes L^{ab}) = (\mu_{bd}^*\mathcal{F})\otimes L^{abd^2}$$

is an $IT(0)$ sheaf, because $R\Phi_{\mathcal{P}}(\mu_d^*((\mu_b^*\mathcal{F})\otimes L^{ab})) = \hat{\mu}_{d*}R\Phi_{\mathcal{P}}((\mu_b^*\mathcal{F})\otimes L^{ab}) = \hat{\mu}_{d*}R^0\Phi_{\mathcal{P}}((\mu_b^*\mathcal{F})\otimes L^{ab})$ [Mukai 1981, (3.4)] is concentrated in degree 0, where $\hat{\mu}_d : \hat{A} \to \hat{A}$ is the dual isogeny of μ_d. Likewise, if $\mathcal{G}\langle y\underline{l}\rangle$ is GV, by using the equivalence in Definition/Theorem 3.1(1), we have that

$$\mu_b^*((\mu_d^*\mathcal{G})\otimes L^{cd}) = (\mu_{bd}^*\mathcal{G})\otimes L^{b^2cd}$$

is a GV sheaf. Since

$$\mu_{bd}^*(\mathcal{F}\otimes\mathcal{G})\otimes L^{(ad+bc)bd} = ((\mu_{bd}^*\mathcal{F})\otimes L^{abd^2})\otimes((\mu_{bd}^*\mathcal{G})\otimes L^{b^2cd}),$$

we conclude by applying the "preservation of vanishing" for (untwisted) coherent sheaves [Pareschi and Popa 2011b, Proposition 3.1]. □

For our purposes, the central result of this section is the following:

Proposition 3.5. *Let p be a nonnegative integer. If*

$$\epsilon_1(\underline{l}) < \frac{1}{p+2},$$

then $M_L^{\otimes(p+1)}\otimes L^h$ is $IT(0)$ for all $h \geq 1$.

Proof. Let L be a line bundle representing $\underline{l}$, and let M_L be the kernel of the evaluation morphism $H^0(A, L)\otimes\mathcal{O}_A \to L$. The assumption on $\epsilon_1(\underline{l})$ implies, in particular, that $\underline{l}$ is basepoint-free and, using Theorem 2.4, we get

$$\kappa_1(\underline{l}) = \frac{\epsilon_1(\underline{l})}{1-\epsilon_1(\underline{l})} = -1 + \frac{1}{1-\epsilon_1(\underline{l})} < -1 + \frac{p+2}{p+1} = \frac{1}{p+1}.$$

By Lemma 3.3, this is equivalent to say that $M_L\langle 1/(p+1)\underline{l}\rangle$ is an $IT(0)$ $\mathbb{Q}$-twisted sheaf. Fix now an integer $h \geq 1$ and write $M_L^{\otimes(p+1)}\otimes L^h = M_L^{\otimes(p+1)}\otimes L\otimes L^{h-1}$ as the $\mathbb{Q}$-twisted sheaf

$$M_L^{\otimes(p+1)}\Big\langle\Big(\frac{p+1}{p+1} + h - 1\Big)\underline{l}\Big\rangle = \Big(M_L\Big\langle\frac{1}{p+1}\underline{l}\Big\rangle\Big)^{\otimes(p+1)}\otimes\mathcal{O}_A\langle(h-1)\underline{l}\rangle.$$

Since L^{h-1} is ample — hence $IT(0)$ — if $h > 1$, or it is trivial if $h = 1$, and $M_L\langle 1/(p+1)\underline{l}\rangle$ is $IT(0)$, we have that $M_L^{\otimes(p+1)}\otimes L^h$ is $IT(0)$ thanks to the "preservation of vanishing" (Proposition 3.4). □

4. Syzygies and the property (N_p)

We recall the definition and geometric meaning of the property (N_p) in more detail. Let X be a projective variety, defined over an algebraically closed field $\mathbb{K}$. If L gives an embedding

$$\phi_{|L|} : X \hookrightarrow \mathbb{P} = \mathbb{P}(H^0(X, L)^\vee),$$

then L is said to *satisfy the property* (N_p) if the first p steps of the minimal graded free resolution $E_\bullet(L)$ of the algebra $R_L := \bigoplus_m H^0(X, L^m)$ over the polynomial ring $S_L := \mathrm{Sym} H^0(X, L)$ are linear, i.e., of the form

$$\begin{array}{ccccccccccc} S_L(-p-1)^{\oplus i_p} & \longrightarrow & S_L(-p)^{\oplus i_{p-1}} & \longrightarrow \cdots \longrightarrow & S_L(-2)^{\oplus i_1} & \longrightarrow & S_L & \longrightarrow & R_L & \longrightarrow & 0 \\ \| & & \| & & \| & & \| & & & & \\ E_p(L) & & E_{p-1}(L) & & E_1(L) & & E_0(L) & & & & \end{array}$$

Thus (N_0) means that L is projectively normal (and in this case a resolution of the homogeneous ideal $I_{X/\mathbb{P}}$ of X in $\mathbb{P}$ is given by $\cdots \to E_1(L) \to I_{X/\mathbb{P}} \to 0$); (N_1) means that $I_{X/\mathbb{P}}$ is generated by quadrics; (N_2) means that the relations among these quadrics are generated by linear ones and so on.

Writing $\mathbb{K} = S_L/S_{L+}$ as the quotient of the polynomial ring S_L by the irrelevant maximal ideal $S_{L+} := \bigoplus_{m\geq 1} \mathrm{Sym}^m H^0(X, L)$, it is well-known that $\dim_{\mathbb{K}}(\mathrm{Tor}_i^{S_L}(R_L, \mathbb{K})_j)$ computes the cardinality of any minimal set of homogeneous generators of $E_i(L)$ of degree j, therefore

$$E_i(L) = \bigoplus_j \mathrm{Tor}_i^{S_L}(R_L, \mathbb{K})_j \otimes_{\mathbb{K}} S_L(-j)$$

and L satisfies the property (N_p) if and only if

$$\mathrm{Tor}_p^{S_L}(R_L, \mathbb{K})_j = 0 \quad \text{for all } j \geq p+2.^5 \tag{4-1}$$

A well established condition ensuring the property (N_p) for L in *characteristic zero* is the vanishing

$$H^1(X, M_L^{\otimes(p+1)} \otimes L^h) = 0 \quad \text{for all } h \geq 1. \tag{4-2}$$

Indeed, tensoring the Koszul resolution of $\mathbb{K}$ by R_L and taking graded pieces, we see that the property (N_p) for L is equivalent to the exactness in the middle of the Koszul complex

$$\Lambda^{p+1} H^0(X, L) \otimes H^0(X, L^h) \to \Lambda^p H^0(X, L) \otimes H^0(X, L^{h+1}) \to \Lambda^{p-1} H^0(X, L) \otimes H^0(X, L^{h+2})$$

for all $h \geq 1$ (see [Lazarsfeld 1989, pages 510–511] for details). This can be expressed in terms of the kernel bundle of L. Namely, taking wedge products of the exact sequence

$$0 \to M_L \to H^0(X, L) \otimes \mathcal{O}_X \to L \to 0,$$

we get

$$0 \to \Lambda^{p+1} M_L \to \Lambda^{p+1} H^0(X, L) \otimes \mathcal{O}_X \to \Lambda^p M_L \otimes L \to 0.$$

[5] $\mathrm{Tor}_0^{S_L}(R_L, \mathbb{K})_1$ is always trivial, because we are dealing with the complete linear series $|L|$ and the corresponding embedding is linearly normal. Moreover, the vanishing $\mathrm{Tor}_p^{S_L}(R_L, \mathbb{K})_j = 0$ for all $j \geq p+2$, forces $\mathrm{Tor}_i^{S_L}(R_L, \mathbb{K})_j = 0$ for all $0 \leq i \leq p$ and $j \geq i+2$ (see the proof of Proposition 1.3.3 in [Lazarsfeld 1989] for details).

Tensoring it by L^h and taking global section, we see that the exactness of the Koszul complex above is equivalent to the surjectivity of the map

$$\Lambda^{p+1} H^0(X, L) \otimes H^0(X, L^h) \to H^0(X, \Lambda^p M_L \otimes L^{h+1}),$$

that in turn follows from the vanishing

$$H^1(X, \Lambda^{p+1} M_L \otimes L^h) = 0 \quad \text{for all } h \geq 1. \tag{4-3}$$

Now, if $\text{char}(\mathbb{K}) = 0$, $\Lambda^{p+1} M_L$ is a *direct summand* of $M_L^{\otimes(p+1)}$ and in particular (4-2) implies (4-3); otherwise said L satisfies the property (N_p). If $\text{char}(\mathbb{K}) > 0$, the exterior power $\Lambda^{p+1} M_L$ may no longer be a direct summand of the tensor power $M_L^{\otimes(p+1)}$, hence the above discussion does not apply. Nevertheless in this section, following an approach essentially due to G. Kempf, we prove that (4-2) implies the property (N_p) for L, even in *positive characteristic*.

Proposition 4.1. *Let X be a projective variety defined over an algebraically closed field $\mathbb{K}$. Let L be an ample and globally generated line bundle on X, and let p be a nonnegative integer. If*

$$H^1(X, M_L^{\otimes(p+1)} \otimes L^h) = 0 \quad \textit{for all } h \geq 1,$$

then the property (N_p) holds for L.

Let us start by recalling two definitions and an algebraic lemma of Kempf [1989] (see also [Rubei 2000, Section 2]).

Definition 4.2. For any L_i (not necessarily ample) line bundles on X, let $K(L_1) = H^0(X, L_1)$ and, for $n > 1$, define $K(L_1, \dots, L_n)$ inductively by the exact sequence:

$$0 \to K(L_1, \dots, L_n) \to K(L_1, L_3, \dots, L_n) \otimes K(L_2) \to K(L_1 \otimes L_2, L_3, \dots, L_n).$$

In particular, $K(L_1, L_2)$ is the kernel of the multiplication map of global sections

$$H^0(X, L_1) \otimes H^0(X, L_2) \to H^0(X, L_1 \otimes L_2).$$

Definition 4.3. Let S be a polynomial ring over $\mathbb{K}$ and let R be a finitely generated graded S-module:

(1) Define $T^0(R) := R$, $T^1(R) := \text{Ker}[R(-1) \otimes_{\mathbb{K}} S_1 \to R]$ and inductively

$$T^i(R) := T^{i-1}(T^1(R)).$$

(2) Define

$$d^i(R) := \min\{d \in \mathbb{Z} \mid T^i(R) \text{ is generated over } S \text{ by elements of degree} \leq d\}.$$

Lemma 4.4 [Kempf 1989, Lemma 16]. *Let $S = \mathbb{K}[x_0, \dots, x_r]$ be a polynomial ring, graded in the standard way, over $\mathbb{K} = S/(x_0, \dots, x_r)$. Let R be a finitely generated graded S-module. If $j > p - i + d^i(R)$ for all $0 \leq i \leq p$, then*

$$\text{Tor}_p^S(R, \mathbb{K})_j = 0.$$

Due to some obscurities in Kempf's argument and for the sake of self-containedness, we prefer to give a proof of the above lemma, which closely follows that of Kempf.

Proof of Lemma 4.4. Consider the exact sequence

$$0 \to T^1(R) \to R(-1)\otimes_{\mathbb{K}} S_1 \xrightarrow{\alpha} R.$$

The image R' of α is a graded submodule of R. The quotient module $Q = R/R'$ is of finite length, hence its Castelnuovo–Mumford regularity $\operatorname{reg}(Q) = \max\{d \mid Q_d \neq 0\}$ (see [Eisenbud 2005, Corollary 4.4]). Moreover Q is zero in degrees $> d^0(R)$, therefore

$$\operatorname{Tor}^S_p(Q, \mathbb{K}) \text{ is zero in degrees } > p + d^0(R). \tag{4-4}$$

Indeed, if $\operatorname{Tor}^S_p(Q, \mathbb{K})_j \neq 0$ for a $j > p + d^0(R)$, then $\operatorname{reg}(Q) \leq d^0(R) < j - p$. But, by definition, $\operatorname{reg}(Q) = \operatorname{Sup}\{k - i \mid \dim_{\mathbb{K}}(\operatorname{Tor}^S_i(Q, \mathbb{K})_k) \neq 0\}$ and so we get a contradiction. Now (4-4) implies that the map

$$\operatorname{Tor}^S_p(R', \mathbb{K}) \to \operatorname{Tor}^S_p(R, \mathbb{K})$$

is surjective in degrees $> p + d^0(R)$. Therefore, in order to prove the statement, it is enough to prove that $\operatorname{Tor}^S_p(R', \mathbb{K})_j = 0$, if $j > p + d^0(R)$. From the long exact sequence associated to

$$0 \to T^1(R) \to R(-1)\otimes_{\mathbb{K}} S_1 \xrightarrow{\alpha} R' \to 0,$$

we get

$$\operatorname{Tor}^S_p(R(-1)\otimes_{\mathbb{K}} S_1, \mathbb{K}) \xrightarrow{\alpha_*} \operatorname{Tor}^S_p(R', \mathbb{K}) \xrightarrow{\delta} \operatorname{Tor}^S_{p-1}(T^1(R), \mathbb{K}).$$

Note that α_* is the multiplication by S_1 in the first variable. Since α_* is also the multiplication by S_1 in the second variable, it is the zero map. Therefore δ gives an inclusion

$$\operatorname{Tor}^S_p(R', \mathbb{K}) \subseteq \operatorname{Tor}^S_{p-1}(T^1(R), \mathbb{K})$$

and we may repeat this procedure p times, obtaining

$$\operatorname{Tor}^S_{-1}(T^{p+1}(R), \mathbb{K}) = 0. \qquad \square$$

If now L is an ample line bundle on X, $S = S_L$ and $R = R_L$, the link between the previous definitions is given by

$$T^i(R_L) = \bigoplus_{m \geq i} K(L^{m-i}, \underbrace{L, \ldots, L}_{i}). \tag{4-5}$$

Proof. If $i = 0$, then $T^0(R_L) = R_L$ and $K(L^m) = H^0(X, L^m)$. So (4-5) is true. By definition

$$T^i(R_L) = T^{i-1}(T^1(R_L)) = T^{i-1}(\operatorname{Ker}[R_L(-1)\otimes_{\mathbb{K}} H^0(X, L) \to R_L]),$$

and

$$0 \to \bigoplus_{m \geq i} K(L^{m-i}, \underbrace{L, \ldots, L}_{i}) \to \bigoplus_{m \geq i} K(L^{m-i}, \underbrace{L, \ldots, L}_{i-1}) \otimes H^0(X, L) \to \bigoplus_{m \geq i} K(L^{m-i+1}, \underbrace{L, \ldots, L}_{i-1}).$$

Therefore (4-5) holds, by induction on i. $\square$

The next lemma allows to reduce the property (N_p) for L to the vanishing (4-2), in a way that avoids the exterior power of M_L.

Lemma 4.5. (1) *For all $n \geq 0$ and $h \geq 1$, one has* $H^0(X, M_L^{\otimes n} \otimes L^h) = K(L^h, \underbrace{L, \ldots, L}_{n})$, *if L is basepoint-free.*

(2) *Let $i \geq 0$ and $h \geq 1$. If L is basepoint-free and $H^1(X, M_L^{\otimes(i+1)} \otimes L^h) = 0$, then the multiplication map*

$$K(L^h, \underbrace{L, \ldots, L}_{i}) \otimes H^0(X, L) \to K(L^{h+1}, \underbrace{L, \ldots, L}_{i})$$

is surjective.

(3) [Rubei 2000, page 2578] *If the multiplication maps*

$$K(L^h, \underbrace{L, \ldots, L}_{i}) \otimes H^0(X, L) \to K(L^{h+1}, \underbrace{L, \ldots, L}_{i})$$

are surjective for all $h \geq 1$, then $d^i(R_L) = i + 1$.

Proof. (1) If $n = 0$, then by definition $H^0(X, L^h) = K(L^h)$ for all $h \geq 1$. Suppose $n \geq 1$. The kernel bundle M_L sits in the short exact sequence

$$0 \to M_L \to H^0(X, L) \otimes \mathcal{O}_X \to L \to 0. \tag{4-6}$$

Tensoring it by $M_L^{\otimes(n-1)} \otimes L^h$, one obtains

$$0 \to M_L^{\otimes n} \otimes L^h \to H^0(X, L) \otimes M_L^{\otimes(n-1)} \otimes L^h \to M_L^{\otimes(n-1)} \otimes L^{h+1} \to 0. \tag{4-7}$$

Taking global sections of (4-7) and using the inductive hypothesis, we obtain

$$0 \to H^0(X, M_L^{\otimes n} \otimes L^h) \to H^0(X, L) \otimes K(L^h, \underbrace{L, \ldots, L}_{n-1}) \to K(L^{h+1}, \underbrace{L, \ldots, L}_{n-1}).$$

Therefore, by definition, $H^0(X, M_L^{\otimes n} \otimes L^h) = K(L^h, \underbrace{L, \ldots, L}_{n})$.

(2) Tensoring (4-6) by $M_L^{\otimes i} \otimes L^h$, we have

$$0 \to M_L^{\otimes(i+1)} \otimes L^h \to H^0(X, L) \otimes M_L^{\otimes i} \otimes L^h \to M_L^{\otimes i} \otimes L^{h+1} \to 0. \tag{4-8}$$

From the long exact sequence in cohomology associated to (4-8), and thanks to the point (1), one has

$$H^0(X, L) \otimes K(L^h, \underbrace{L, \ldots, L}_{i}) \xrightarrow{\alpha} K(L^{h+1}, \underbrace{L, \ldots, L}_{i}) \to H^1(X, M_L^{\otimes(i+1)} \otimes L^h) = 0.$$

Therefore the multiplication map α is surjective.

(3) By (4-5) and the hypothesis we have that $T^i(R_L)$ is generated over S_L by

$$K(\underbrace{L, \ldots, L}_{i+1}).$$

This means that it is generated by the piece of degree m with $m - i = 1$, i.e., $m = i + 1$. Therefore $d^i(R_L) = i + 1$. ☐

5. Proof of the Theorems 1.1 and 1.3

Proof of Theorem 1.1. Let L be a representative of the class $\underline{l}$. For all $0 \leq i \leq p$, we have

$$\epsilon_1(\underline{l}) < \frac{1}{p+2} \leq \frac{1}{i+2}.$$

Therefore L is basepoint-free and, thanks to the Proposition 3.5, we know that $M_L^{\otimes(i+1)} \otimes L^h$ is $IT(0)$, for all $h \geq 1$. This implies, in particular, that $H^1(A, M_L^{\otimes(i+1)} \otimes L^h) = 0$ for all $h \geq 1$. Hence, by Lemma 4.5(2) and (3), we obtain

$$d^i(R_L) = i+1.$$

Now, if $j > p - i + d^i(R_L) = p+1$, Kempf's Lemma 4.4 implies that

$$\mathrm{Tor}_p^{S_L}(R_L, \mathbb{K})_j = 0.$$

As explained in (4-1), this is equivalent to the property (N_p) for L. □

Proof of Theorem 1.3. Note that we have already proved the $t = 0$ case — even without the basepoint-freeness assumption — and the $t = 1$ case (Corollary 1.2). Hence we may assume $t > 1$. By Theorem 1.1, it suffices to show that $\epsilon_1(m\underline{l}) < 1/(p+2)$. We have

$$\epsilon_1(m\underline{l}) = \frac{\epsilon_1(\underline{l})}{m} \leq \frac{\epsilon_1(\underline{l})}{p+3-t} \leq \frac{1}{t(p+3-t)},$$

where the last inequality follows by definition. Let us impose now the inequality

$$\frac{1}{t(p+3-t)} < \frac{1}{p+2},$$

or equivalently

$$t^2 - (p+3)t + p + 2 < 0.$$

This is satisfied if and only if $1 < t < p+2$ and, by hypothesis, we have $1 < t \leq p+1$. □

Acknowledgments

This work is part of my PhD thesis. I would like to thank my advisor, Giuseppe Pareschi, for his invaluable guidance.

References

[Ein and Lazarsfeld 1993] L. Ein and R. Lazarsfeld, "Syzygies and Koszul cohomology of smooth projective varieties of arbitrary dimension", *Invent. Math.* **111**:1 (1993), 51–67. MR Zbl

[Eisenbud 2005] D. Eisenbud, *The geometry of syzygies*, Graduate Texts in Math. **229**, Springer, 2005. MR Zbl

[Green 1984] M. L. Green, "Koszul cohomology and the geometry of projective varieties", *J. Differential Geom.* **19**:1 (1984), 125–171. MR Zbl

[Green and Lazarsfeld 1986] M. Green and R. Lazarsfeld, "On the projective normality of complete linear series on an algebraic curve", *Invent. Math.* **83**:1 (1986), 73–90. MR Zbl

[Hwang and To 2011] J.-M. Hwang and W.-K. To, "Buser–Sarnak invariant and projective normality of abelian varieties", pp. 157–170 in *Complex and differential geometry*, edited by W. Ebeling et al., Springer Proc. Math. **8**, Springer, 2011. MR Zbl

[Ito 2018] A. Ito, "A remark on higher syzygies on abelian surfaces", *Comm. Algebra* **46**:12 (2018), 5342–5347. MR Zbl

[Jiang and Pareschi 2017] Z. Jiang and G. Pareschi, "Cohomological rank functions on abelian varieties", 2017. To appear in *Ann. Sci. École Norm. Sup.* arXiv

[Kempf 1989] G. R. Kempf, "Projective coordinate rings of abelian varieties", pp. 225–235 in *Algebraic analysis, geometry, and number theory* (Baltimore, 1988), edited by J.-I. Igusa, Johns Hopkins Univ. Press, Baltimore, MD, 1989. MR Zbl

[Kempf 1991] G. R. Kempf, *Complex abelian varieties and theta functions*, Springer, 1991. MR Zbl

[Koizumi 1976] S. Koizumi, "Theta relations and projective normality of abelian varieties", *Amer. J. Math.* **98**:4 (1976), 865–889. MR Zbl

[Küronya and Lozovanu 2019] A. Küronya and V. Lozovanu, "A Reider-type theorem for higher syzygies on abelian surfaces", *Algebr. Geom.* **6**:5 (2019), 548–570. MR Zbl

[Lazarsfeld 1989] R. Lazarsfeld, "A sampling of vector bundle techniques in the study of linear series", pp. 500–559 in *Lectures on Riemann surfaces* (Trieste, Italy, 1987), edited by M. Cornalba et al., World Sci., Teaneck, NJ, 1989. MR Zbl

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, II: Positivity for vector bundles, and multiplier ideals*, Ergebnisse der Mathematik (3) **49**, Springer, 2004. MR Zbl

[Lazarsfeld et al. 2011] R. Lazarsfeld, G. Pareschi, and M. Popa, "Local positivity, multiplier ideals, and syzygies of abelian varieties", *Algebra Number Theory* **5**:2 (2011), 185–196. MR Zbl

[Lozovanu 2018] V. Lozovanu, "Singular divisors and syzygies of polarized abelian threefolds", preprint, 2018. arXiv

[Mukai 1981] S. Mukai, "Duality between $D(X)$ and $D(\hat{X})$ with its application to Picard sheaves", *Nagoya Math. J.* **81** (1981), 153–175. MR Zbl

[Mumford 1970] D. Mumford, "Varieties defined by quadratic equations", pp. 29–100 in *Questions on algebraic varieties* (Varenna, Italy, 1969), edited by E. Marchionna, CIME Summer Schools **51**, Edizioni Cremonese, Rome, 1970. MR Zbl

[Pareschi 2000] G. Pareschi, "Syzygies of abelian varieties", *J. Amer. Math. Soc.* **13**:3 (2000), 651–664. MR Zbl

[Pareschi and Popa 2004] G. Pareschi and M. Popa, "Regularity on abelian varieties, II: Basic results on linear series and defining equations", *J. Algebraic Geom.* **13**:1 (2004), 167–193. MR Zbl

[Pareschi and Popa 2011a] G. Pareschi and M. Popa, "GV-sheaves, Fourier–Mukai transform, and generic vanishing", *Amer. J. Math.* **133**:1 (2011), 235–271. MR Zbl

[Pareschi and Popa 2011b] G. Pareschi and M. Popa, "Regularity on abelian varieties, III: Relationship with generic vanishing and applications", pp. 141–167 in *Grassmannians, moduli spaces and vector bundles*, edited by D. A. Ellwood and E. Previato, Clay Math. Proc. **14**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[Rubei 2000] E. Rubei, "On syzygies of abelian varieties", *Trans. Amer. Math. Soc.* **352**:6 (2000), 2569–2579. MR Zbl

[Sasaki 1977] R. Sasaki, "Theta relations and their applications in abstract geometry", *Sci. Rep. Tokyo Kyoiku Daigaku Sect. A* **13**:366-382 (1977), 290–312. MR Zbl

[Sekiguchi 1976] T. Sekiguchi, "On projective normality of abelian varieties", *J. Math. Soc. Japan* **28**:2 (1976), 307–322. MR Zbl

[Sekiguchi 1977] T. Sekiguchi, "On projective normality of abelian varieties, II", *J. Math. Soc. Japan* **29**:4 (1977), 709–727. MR Zbl

Communicated by Gavril Farkas
Received 2019-02-26 Accepted 2020-02-06

caucci@mat.uniroma1.it *Sapienza Università di Roma, Italy*

ALGEBRA AND NUMBER THEORY 14:4 (2020)
dx.doi.org/10.2140/ant.2020.14.961

On the Ekedahl–Oort stratification of Shimura curves

Benjamin Howard

We study the Hodge–Tate period domain associated to a quaternionic Shimura curve at a prime of bad reduction, and give an explicit description of its Ekedahl–Oort stratification.

1. Introduction

Fix a prime p, and let C be the completion of an algebraic closure of $\mathbb{Q}_p$. Denote by $\mathcal{O} \subset C$ its ring of integers, and by $k = \mathcal{O}/\mathfrak{m}$ its residue field.

1.1. *Stratifications of p-adic periods domains.* Let G be a p-divisible group over $\mathcal{O}$. It has a p-adic Tate module

$$T_p(G) = \mathrm{Hom}(\mathbb{Q}_p/\mathbb{Z}_p, G)$$

and a module of invariant differential forms $\Omega(G)$. These are free of finite rank over $\mathbb{Z}_p$ and $\mathcal{O}$, respectively. Using the canonical trivialization $\Omega(\mu_{p^\infty}) \cong \mathcal{O}$, we define the *Hodge–Tate morphism*

$$T_p(G) \cong \mathrm{Hom}(G^\vee, \mu_{p^\infty}) \xrightarrow{\mathrm{HT}} \mathrm{Hom}(\Omega(\mu_{p^\infty}), \Omega(G^\vee)) \cong \Omega(G^\vee), \tag{1.1.1}$$

where $G^\vee$ is the p-divisible group dual to G.

Theorem A [Scholze and Weinstein 2013]. *There is an equivalence between the category of p-divisible groups over $\mathcal{O}$ and the category of pairs (T, W) in which*

- *T is a free $\mathbb{Z}_p$-module of finite rank,*
- *$W \subset T \otimes_{\mathbb{Z}_p} C$ is a C-subspace.*

The equivalence sends G to its p-adic Tate module $T = T_p(G)$, endowed with its Hodge–Tate filtration

$$W = \ker\big(T_p(G) \otimes_{\mathbb{Z}_p} C \xrightarrow{\mathrm{HT}} \Omega(G^\vee) \otimes_{\mathcal{O}} C\big).$$

Fix a free $\mathbb{Z}_p$-module T of finite rank, and consider the $\mathbb{Q}_p$-scheme

$$X = \mathrm{Gr}_d(T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)$$

parametrizing subspaces of $T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ of some fixed dimension $d \le \mathrm{rank}(T)$. By the theorem of Scholze–Weinstein, every point $W \in X(C)$ determines a p-divisible group G over $\mathcal{O}$, whose reduction to the residue field we denote by G_k. Let $G_k[p]$ be the group scheme of p-torsion points in G_k.

This research was supported in part by NSF grants DMS1501583 and DMS1801905.
MSC2010: primary 11G18; secondary 14G35.
Keywords: Shimura varieties, Ekedahl–Oort stratification.

If we declare two points $W, W' \in X(C)$ to be equivalent when the corresponding reductions G_k and G'_k are isogenous, the resulting partition is the *Newton stratification* of $X(C)$. Alternatively, if we declare $W, W' \in X(C)$ to be equivalent when the p-torsion group schemes $G_k[p]$ and $G'_k[p]$ are isomorphic, the resulting partition is the *Ekedahl–Oort stratification* of $X(C)$.

There are similar partitions when X is replaced by a more sophisticated flag variety, called the *Hodge–Tate period domain*, associated to a Shimura datum of Hodge type and a prime p. This period domain and its Newton stratification were studied by Caraiani and Scholze [2017], who proved that each Newton stratum in $X(C)$ can be realized as the C-points of a locally closed subset of the associated adic space. For the Ekedahl–Oort stratification of $X(C)$ there is nothing in the existing literature, and it is not known if it has any structure other than set-theoretic partition.

In the case of modular curves, the Hodge–Tate period domain is the projective line $\mathbb{P}^1$ over $\mathbb{Q}_p$. In this case the Newton stratification and the Ekedahl–Oort stratification agree, and there are two strata: the *ordinary locus* $\mathbb{P}^1(\mathbb{Q}_p)$, and the *supersingular locus* $\mathbb{P}^1(C) \smallsetminus \mathbb{P}^1(\mathbb{Q}_p)$.

For the compact Shimura curve determined by an indefinite quaternion algebra over $\mathbb{Q}$, and a prime p at which the quaternion algebra is ramified, the Hodge–Tate period domain X is a twisted form of $\mathbb{P}^1$. All points of $X(C)$ give rise to supersingular p-divisible groups over k, and the Newton stratification consists of a single stratum, $X(C)$ itself. In contrast, the Ekedahl–Oort stratification is nontrivial, and the goal of this paper is to make it explicit.

Although the methods used here are fairly direct, it is not clear how far they can be extended. The case of Hilbert modular surfaces may already require new ideas.

For background on the classical Ekedahl–Oort stratification of reductions of Shimura varieties (as opposed to their Hodge–Tate period domains), we refer the reader to [Oort 2001; Moonen 2004; Viehmann and Wedhorn 2013; Zhang 2018].

1.2. ***The Shimura curve period domain.*** Let $\mathbb{Q}_{p^2} \subset C$ be the unique unramified quadratic extension of $\mathbb{Q}_p$, and let $\mathbb{Z}_{p^2} \subset \mathcal{O}$ be its ring of integers. Denote by $x \mapsto \bar{x}$ the nontrivial automorphism of $\mathbb{Q}_{p^2}$. Define a noncommutative $\mathbb{Z}_p$-algebra of rank 4 by

$$\Delta = \mathbb{Z}_{p^2}[\Pi],$$

where Π is subject to the relations $\Pi^2 = p$ and $\Pi \cdot x = \bar{x} \cdot \Pi$ for all $x \in \mathbb{Z}_{p^2}$. In other words, Δ is the unique maximal order in the unique quaternion division algebra over $\mathbb{Q}_p$.

Let T be a free Δ-module of rank one, and let X be the smooth projective variety over $\mathbb{Q}_p$ with functor of points

$$X(S) = \{\mathcal{O}_S\text{-module local direct summands } W \subset T \otimes_{\mathbb{Z}_p} \mathcal{O}_S \text{ of rank 2 that are stable under } \Delta\} \quad (1.2.1)$$

for any $\mathbb{Q}_p$-scheme S. This is the Hodge–Tate period domain associated to a quaternionic Shimura curve.

As we explain in Section 4.1, our period domain becomes isomorphic to the projective line after base change to $\mathbb{Q}_{p^2}$, and any choice of Δ-module generator $\lambda \in T$ determines a bijection

$$X(C) \cong C \cup \{\infty\}. \quad (1.2.2)$$

After fixing such a choice, we normalize the valuation $\mathrm{ord} : C \to \mathbb{R} \cup \{\infty\}$ by $\mathrm{ord}(p) = 1$, extend it to $C \cup \{\infty\}$ by $\mathrm{ord}(\infty) = -\infty$, and use (1.2.2) to view ord as a function

$$\mathrm{ord} : X(C) \to \mathbb{R} \cup \{-\infty, \infty\}.$$

The theorem of Scholze–Weinstein provides a canonical bijection

$$X(C) \cong \left\{ \begin{array}{c} \text{isomorphism classes of } p\text{-divisible groups } G \text{ over } \mathcal{O} \text{ of height 4 and dimension 2,} \\ \text{endowed with an action of } \Delta \text{ and a } \Delta\text{-linear isomorphism } T_p(G) \cong T \end{array} \right\}.$$

By forgetting the level structure $T_p(G) \cong T$, reducing to the residue field, and then taking p-torsion subgroups, we obtain a function

$$X(C) \to \{\text{isomorphism classes of finite group schemes over } k\text{, endowed with an action of } \Delta/p\Delta\}$$

sending $G \mapsto G_k[p]$, whose fibers are the *Ekedahl–Oort strata* of $X(C)$.

Hypothesis. For the rest of this introduction, we assume $p > 2$. Theorems B and C below are presumably true without this hypothesis, but we are unable to provide a proof. See the remarks following Theorem 2.2.2.

It is convenient to organize the strata into two types: those on which the p-torsion group scheme $G_k[p]$ is superspecial (in the sense of Section 3.2), and those on which it is not. The two theorems that follow show that there are three superspecial strata, and two infinite families of nonsuperspecial strata. These results are proved in Section 4.2, where the reader will also find an explicit recipe for computing the Dieudonné module of the p-torsion group scheme $G_k[p]$ attached to a point of $X(C)$.

Theorem B. *The conditions*

$$\frac{1}{p+1} < \mathrm{ord}(\tau) < \frac{p}{p+1}$$

on $\tau \in X(C)$ define an Ekedahl–Oort stratum, as do each one of

$$\mathrm{ord}(\tau) < \frac{1}{p+1}, \qquad \frac{p}{p+1} < \mathrm{ord}(\tau).$$

The union of these three strata is the locus of points with superspecial reduction. In particular, the isomorphism class of the finite group scheme $G_k[p]$ is the same all on three strata, but the isomorphism class of $G_k[p]$ with its Δ-action is not.

Now consider the locus of points

$$\left\{\tau \in C : \mathrm{ord}(\tau) = \frac{1}{p+1}\right\} \cup \left\{\tau \in C : \mathrm{ord}(\tau) = \frac{p}{p+1}\right\} \subset X(C) \tag{1.2.3}$$

at which the corresponding p-divisible group does not have superspecial reduction. The isomorphism class of the p-torsion group scheme $G_k[p]$ is constant on (1.2.3), but the isomorphism class of $G_k[p]$ with its Δ-action varies. In fact, the Δ-action varies so much that (1.2.3) decomposes as an infinite disjoint union of Ekedahl–Oort strata.

Theorem C. *The fibers of the composition*

$$\left\{\tau \in C : \mathrm{ord}(\tau) = \frac{1}{p+1}\right\} \xrightarrow{\tau \mapsto p/\tau^{p+1}} \mathcal{O}^\times \to k^\times$$

are Ekedahl–Oort strata, as are the fibers of the composition

$$\left\{\tau \in C : \mathrm{ord}(\tau) = \frac{p}{p+1}\right\} \xrightarrow{\tau \mapsto \tau^{p+1}/p^p} \mathcal{O}^\times \to k^\times.$$

Both unlabeled arrows are reduction to the residue field.

Remark 1.2.1. The infinitude of Ekedahl–Oort strata is a pathology arising from the nonsmooth reduction of compact Shimura curves. Similar pathologies for the reductions of Hilbert modular varieties at ramified primes are described in the appendix to [Andreatta and Goren 2003].

1.3. ***Notation and conventions.*** Throughout the paper p is a fixed prime. We allow $p = 2$ unless otherwise stated. Let $k = \mathcal{O}/\mathfrak{m}$ as above, and denote by $\sigma : k \to k$ the absolute Frobenius $\sigma(x) = x^p$.

The rings $\mathbb{Z}_{p^2} \subset \mathcal{O}$ and $\Delta = \mathbb{Z}_{p^2}[\Pi]$ have the same meaning as above. We label the embeddings

$$j_0, j_1 : \mathbb{Z}_{p^2} \to \mathcal{O} \tag{1.3.1}$$

in such a way that j_0 is the inclusion and $j_1(x) = j_0(\bar{x})$ is its conjugate.

2. Integral p-adic Hodge theory

In this section we recall the integral p-adic Hodge theory of an arbitrary p-divisible group G over $\mathcal{O}$. The quaternion order Δ plays no role whatsoever.

Following [Fargues 2015; Lau 2018; Scholze and Weinstein 2020], we will attach to G a Breuil–Kisin–Fargues module, and explain how to extract from it invariants of G such as its Hodge–Tate morphism $T_p(G) \to \Omega(G^\vee)$, and the Dieudonné module of its reduction to k.

2.1. ***A ring of periods.*** Let $C^\flat$ be the tilt of C, with ring of integers $\mathcal{O}^\flat$. Thus

$$\mathcal{O}^\flat = \varprojlim_{x \mapsto x^p} \mathcal{O}/(p)$$

is a local domain of characteristic p, fraction field $C^\flat$, and residue field $k = \mathcal{O}^\flat/\mathfrak{m}^\flat$. An element $x \in \mathcal{O}^\flat$ is given by a sequence $(x_0, x_1, x_2, \ldots)$ of elements $x_\ell \in \mathcal{O}/(p)$ satisfying $x_{\ell+1}^p = x_\ell$. After choosing arbitrary lifts $x_\ell \in \mathcal{O}$, set

$$x^\sharp = \lim_{\ell \to \infty} x_\ell^{p^\ell}.$$

The construction $x \mapsto x^\sharp$ defines a multiplicative function $\mathcal{O}^\flat \to \mathcal{O}$, and we define $\mathrm{ord} : \mathcal{O}^\flat \to \mathbb{R} \cup \{\infty\}$ by $\mathrm{ord}(x) = \mathrm{ord}(x^\sharp)$.

Denote by $\sigma : \mathcal{O}^\flat \to \mathcal{O}^\flat$ the absolute Frobenius $x \mapsto x^p$, and in the same way the induced automorphism of the local domain

$$A_{\mathrm{inf}} = W(\mathcal{O}^\flat).$$

There is a canonical homomorphism of $\mathbb{Z}_p$-algebras

$$\Theta : A_{\mathrm{inf}} \to \mathcal{O}$$

satisfying $\Theta([x]) = x^\sharp$, where $[\,\cdot\,] : \mathcal{O}^\flat \to A_{\mathrm{inf}}$ is the Teichmüller lift.

The kernel of Θ is a principal ideal. To construct a generator, first fix a $\mathbb{Z}_p$-module generator

$$\zeta = (\zeta_p, \zeta_{p^2}, \zeta_{p^3}, \dots) \in T_p(\mu_{p^\infty})$$

and define $\epsilon = (1, \zeta_p, \zeta_{p^2}, \dots) \in \mathcal{O}^\flat$. The element

$$\xi = [1] + [\epsilon^{1/p}] + [\epsilon^{2/p}] + \cdots + [\epsilon^{(p-1)/p}] \in A_{\mathrm{inf}}$$

generates $\ker(\Theta)$. If we denote by

$$\varpi = 1 + \epsilon^{1/p} + \epsilon^{2/p} + \cdots + \epsilon^{(p-1)/p} \in \mathcal{O}^\flat$$

its image under the reduction map $A_{\mathrm{inf}} \to A_{\mathrm{inf}}/(p) = \mathcal{O}^\flat$, then $\mathrm{ord}(\varpi) = 1$, and there are canonical isomorphisms

$$\mathcal{O}/(p) \cong A_{\mathrm{inf}}/(\xi, p) \cong \mathcal{O}^\flat/(\varpi).$$

The following lemma will be needed in the proof of Proposition 3.4.4.

Lemma 2.1.1. *The reduction map $\mathcal{O}^\times \to k^\times$ sends $\varpi^\sharp/p \mapsto -1$.*

Proof. By definition, $\varpi^\sharp = \lim_{\ell\to\infty} x_\ell^{p^\ell}$, where

$$x_\ell = 1 + \zeta_{p^{\ell+1}} + \zeta_{p^{\ell+1}}^2 + \cdots + \zeta_{p^{\ell+1}}^{p-1} = \frac{\zeta_{p^\ell} - 1}{\zeta_{p^{\ell+1}} - 1} \in \mathcal{O}.$$

The binomial theorem implies that

$$\zeta_{p^\ell} = (\zeta_{p^{\ell+1}} - 1 + 1)^p = (\zeta_{p^{\ell+1}} - 1)^p + sp(\zeta_{p^{\ell+1}} - 1) + 1$$

for some $s \in \mathcal{O}$. From this we deduce first that

$$x_\ell \equiv (\zeta_{p^{\ell+1}} - 1)^{p-1} \pmod{p\mathcal{O}},$$

and then that

$$x_\ell^{p^\ell} \equiv (\zeta_{p^{\ell+1}} - 1)^{(p-1)p^\ell} \pmod{p^{\ell-1}\mathcal{O}}. \tag{2.1.1}$$

For $1 \le i \le p-1$ set

$$u_i = \frac{1-\zeta_p^i}{1-\zeta_p} = 1 + \zeta_p + \cdots + \zeta_p^{i-1} \in \mathcal{O}^\times,$$

and note that Wilson's theorem implies $u_1 \cdots u_{p-1} \equiv -1 \pmod{\mathfrak{m}}$. Taking $X = 1$ in the factorization

$$X^{p-1} + \cdots + X + 1 = (X - \zeta_p)\cdots(X - \zeta_p^{p-1})$$

shows that $p = (1-\zeta_p)^{p-1} u_1 \cdots u_{p-1}$, and hence

$$\frac{(1-\zeta_p)^{p-1}}{p} \equiv -1 \pmod{\mathfrak{m}}.$$

Combining this with (2.1.1) shows that

$$\frac{x_\ell^{p^\ell}}{p} \equiv \frac{(\zeta_{p^{\ell+1}}-1)^{(p-1)p^\ell}}{p} \equiv -\left(\frac{(\zeta_{p^{\ell+1}}-1)^{p^\ell}}{1-\zeta_p}\right)^{p-1} \pmod{\mathfrak{m}}.$$

As $\mathbb{Q}_p(\zeta_{p^{\ell+1}})$ is totally ramified over $\mathbb{Q}_p$, the reduction of

$$\frac{(\zeta_{p^{\ell+1}}-1)^{p^\ell}}{1-\zeta_p} \in \mathcal{O}^\times$$

lies in the subgroup $\mathbb{F}_p^\times \subset k^\times$. It follows that $x_\ell^{p^\ell}/p \equiv -1 \pmod{\mathfrak{m}}$, and hence $\varpi^\sharp/p \equiv -1 \pmod{\mathfrak{m}}$. □

2.2. *Breuil–Kisin–Fargues modules.* There is an equivalence of categories between p-divisible groups over $\mathcal{O}$ and Breuil–Kisin–Fargues modules, whose definition we now recall.

Definition 2.2.1. A *Breuil–Kisin–Fargues module* is a triple (M, ϕ, ψ), in which M is a free module of finite rank over A_{inf}, and

$$\phi, \psi : M \to M$$

are homomorphisms of additive groups satisfying

$$\phi(am) = \sigma(a)\phi(m), \quad \psi(\sigma(a)m) = a\psi(m),$$

for all $a \in A_{\mathrm{inf}}$ and $m \in M$, as well as $\phi \circ \psi = \xi$.

Suppose (M, ϕ, ψ) is a Breuil–Kisin–Fargues module. Denote by

$$\sigma^* M = A_{\mathrm{inf}} \otimes_{\sigma, A_{\mathrm{inf}}} M$$

the Frobenius twist of M, and by N the image of the A_{inf}-linear map

$$M \xrightarrow{x \mapsto 1 \otimes \psi(x)} \sigma^* M.$$

It is easy to see that $\xi\sigma^* M \subset N \subset \sigma^* M$. We construct various realizations of M as follows:

- The *de Rham realization*

  $$M_{\mathrm{dR}} = \sigma^* M / \xi \sigma^* M,$$

  sits in the short exact sequence

  $$0 \to N/\xi\sigma^* M \to M_{\mathrm{dR}} \to \sigma^* M/N \to 0.$$

  of free $\mathcal{O}$-modules. Indeed, the freeness of M_{dR} is clear, the freeness of $\sigma^* M/N$ follows from the proof of [Lau 2018, Lemma 9.5], and the freeness of $N/\xi\sigma^* M$ is a consequence of this.
- The *étale realization* is the torsion-free $\mathbb{Z}_p$-module

  $$M_{\mathrm{et}} = M^{\psi=1}.$$

Its *Hodge–Tate filtration*

$$F_{\mathrm{HT}}(M) \subset M_{\mathrm{et}} \otimes_{\mathbb{Z}_p} C$$

is the kernel of the C-linear extension of

$$M_{\mathrm{et}} \xrightarrow{x \mapsto 1 \otimes \psi(x)} N/\xi\sigma^* M.$$

- The *crystalline realization*

$$M_{\mathrm{crys}} = W(k) \otimes_{\sigma, A_{\mathrm{inf}}} M$$

is a free $W(k)$-module, endowed with operators

$$F(a \otimes m) = \sigma(a) \otimes \phi(m), \qquad V(a \otimes m) = \sigma^{-1}(a) \otimes \psi(m).$$

These give M_{crys} the structure of a Dieudonné module.

The following theorem is no doubt known to the experts, but for lack of a reference we will explain in the next subsection how to deduce it from the results of [Lau 2018].

Theorem 2.2.2 (Fargues, Scholze and Weinstein, Lau). *Assume that $p > 2$. The category of p-divisible groups over $\mathcal{O}$ is equivalent to the category of Breuil–Kisin–Fargues modules. Moreover, the Breuil–Kisin–Fargues module (M, ϕ, ψ) associated to a p-divisible group G enjoys the following properties:*

(1) *There are isomorphisms of $\mathcal{O}$-modules*

$$\Omega(G^\vee) \cong N/\xi\sigma^* M, \qquad \mathrm{Lie}(G) \cong \sigma^* M/N. \tag{2.2.1}$$

(2) *If G_k denotes the reduction of G to the residue field $k = \mathcal{O}/\mathfrak{m}$, the covariant Dieudonné module of G_k is isomorphic to M_{crys}.*

(3) *There is an isomorphism $T_p(G) \cong M_{\mathrm{et}}$ making the diagram*

$$\begin{array}{ccc} T_p(G) & = & M_{\mathrm{et}} \\ {\scriptstyle \mathrm{HT}}\downarrow & & \downarrow \\ \Omega(G^\vee) & = & N/\xi\sigma^* M \end{array} \tag{2.2.2}$$

commute, where the vertical arrow on the right is the restriction to $M_{\mathrm{et}} \subset M$ of the $\mathcal{O}$-linear map

$$M \xrightarrow{x \mapsto 1 \otimes \psi(x)} N/\xi\sigma^* M.$$

All of these isomorphisms are functorial.

Some comments on this theorem are warranted, particularly regarding the restriction to $p > 2$. A functor[1] from Breuil–Kisin–Fargues modules to p-divisible groups over $\mathcal{O}$, but not a proof that it is an equivalence of categories, first appeared in the work Fargues [2015, §4.8.1]. His construction of the

[1] Fargues only considers formal p-divisible groups, and imposes a corresponding restriction on Breuil–Kisin–Fargues modules.

functor makes essential use of the theory of *windows* introduced by Zink [2001] and extended by Lau [2010; 2018], and assumes that $p > 2$.

A proof of the equivalence of categories is found in [Scholze and Weinstein 2020, Theorem 14.1.1],[2] where the result is attributed to Fargues. The construction of the functor in [Scholze and Weinstein 2020] is very different from the construction of [Fargues 2015], and does not use of the theory of windows. Instead, what is proved in [Scholze and Weinstein 2020] is that the category of Breuil–Kisin–Fargues modules is equivalent to the category of pairs (T, W) appearing in Theorem A, and hence is equivalent to the category of p-divisible groups. This proof comes with no restriction on p.

The identification of M_{crys} with the Dieudonné module of G_k is [Scholze and Weinstein 2020, Corollary 14.4.4], and the isomorphism $T_p(G) \cong M_{\mathrm{et}}$ can be deduced by carefully tracing through the construction of the equivalence. Unfortunately, the isomorphisms of $\mathcal{O}$-modules (2.2.1) seem difficult to deduce from the description of the equivalence found in [Scholze and Weinstein 2020].

Because of this, our equivalence of categories will be the one appearing in [Lau 2018], which follows Fargues. What Lau proves is that, when $p > 2$, the categories of Breuil–Kisin–Fargues modules and p-divisible groups over $\mathcal{O}$ are both equivalent to the category of windows. The various properties of the equivalence listed in Theorem 2.2.2 can be read off from the constructions of the two functors into the category of windows, which are quite simple and direct (of course, the proof that they are equivalences is not).

The invocation of Theorem 2.2.2 in the calculations of Section 3 is the only reason why the assumption $p > 2$ is imposed in the introduction. Our approach in the sequel will be to allow arbitrary p, but to take the conclusions of Theorem 2.2.2 as hypotheses.

2.3. *Proof of Theorem 2.2.2.* As we have already indicated, Theorem 2.2.2 is proved by relating the categories of Breuil–Kisin–Fargues modules and p-divisible groups to the category of windows introduced by Zink [2001] and extended by Lau [2010; 2018].

Our windows will be modules over the ring A_{crys}, which is defined as the p-adic completion of the subring

$$A^0_{\mathrm{crys}} = A_{\mathrm{inf}}[\xi^n/n! : n = 1, 2, 3, \ldots] \subset A_{\mathrm{inf}}[1/p].$$

It is an integral domain endowed with a ring homomorphism

$$\Theta_{\mathrm{crys}} : A_{\mathrm{crys}} \to \mathcal{O} \tag{2.3.1}$$

extending $\Theta : A_{\mathrm{inf}} \to \mathcal{O}$, and divided powers on the kernel $I = \ker(\Theta_{\mathrm{crys}})$.

The subring $A^0_{\mathrm{crys}} \subset A_{\mathrm{inf}}[1/p]$ is stable under σ, and there is a unique continuous extension to an injective ring homomorphism $\sigma : A_{\mathrm{crys}} \to A_{\mathrm{crys}}$ reducing to the usual p-power Frobenius on $A_{\mathrm{crys}}/pA_{\mathrm{crys}}$. Moreover, [Scholze and Weinstein 2013, Lemma 4.1.8] and the comments following [Lau 2018, (9.1)] show that

$$\sigma(I) \subset pA_{\mathrm{crys}} \quad \text{and} \quad \frac{\sigma(\xi)}{p} \in A^{\times}_{\mathrm{crys}}.$$

[2] Our conventions for Breuil–Kisin–Fargues modules and the equivalence of categories differ from those of [Scholze and Weinstein 2020]. The discrepancy amounts to a Tate twist.

The following definition of a window is taken from [Lau 2018, §2], where it would be called a *window over the frame*

$$\underline{A}_{\mathrm{crys}} = (A_{\mathrm{crys}}, I, \mathcal{O} = A_{\mathrm{crys}}/I, \sigma, \sigma_1),$$

with $\sigma_1 : I \to A_{\mathrm{crys}}$ defined by $\sigma_1(x) = \sigma(x)/p$.

Definition 2.3.1. A *window* is a quadruple (P, Q, Φ, Φ_1) consisting of a projective A_{crys}-module P of finite rank, a submodule $Q \subset P$, and σ-semilinear maps

$$\Phi : P \to P, \qquad \Phi_1 : Q \to P$$

satisfying the following properties:

- there exist A_{crys}-submodules $L, T \subset P$ such that

$$Q = L \oplus IT, \qquad P = L \oplus T,$$

- $a \otimes x \mapsto a\Phi_1(x)$ defines a surjection $\sigma^* Q \to P$ of A_{crys}-modules,
- $\Phi(ax) = p\Phi_1(ax)$ for all $a \in I$ and $x \in P$.

Remark 2.3.2. Taking $a = \xi$ in the final condition yields

$$\Phi(x) = \frac{p}{\sigma(\xi)} \cdot \Phi_1(\xi x)$$

for all $x \in P$. This implies $\Phi(x) = p\Phi_1(x)$ for all $x \in Q$, and shows that each one of Φ and Φ_1 determines the other.

Remark 2.3.3. Note that $IP \subset Q$, and that Q/IP and P/Q are projective (hence free) over $A_{\mathrm{crys}}/I \cong \mathcal{O}$.

Suppose G is a p-divisible group over $\mathcal{O}$. Let P be its crystalline Dieudonné module, evaluated at the divided power thickening (2.3.1). This is a projective A_{crys}-module of rank equal to the height of G, equipped with a σ-semilinear operator $\Phi : P \to P$ and a short exact sequence

$$0 \to \Omega(G^\vee) \to P/IP \to \mathrm{Lie}(G) \to 0$$

of free $\mathcal{O}$-modules. Define $Q \subset P$ as the kernel of $P \to \mathrm{Lie}(G)$. One can show that $\Phi(Q) \subset pP$, allowing us to define $\Phi_1 : Q \to P$ by

$$\Phi_1(x) = \frac{1}{p} \cdot \Phi(x).$$

The following is a special case of [Lau 2018, Proposition 9.7].

Theorem 2.3.4 (Lau). *The construction $G \mapsto (P, Q, \Phi, \Phi_1)$ just given defines a functor from the category of p-divisible groups over $\mathcal{O}$ to the category of windows. It is an equivalence of categories if $p > 2$.*

Now suppose we start with a Breuil–Kisin–Fargues module (M, ϕ, ψ). Set

$$P = A_{\mathrm{crys}} \otimes_{\sigma, A_{\mathrm{inf}}} M, \tag{2.3.2}$$

and define $\Phi : P \to P$ by $\Phi(a \otimes m) = \sigma(a) \otimes \phi(m)$ for all $a \in A_{\mathrm{crys}}$ and $m \in M$. The submodule $Q \subset P$, defined as the kernel of the composition

$$\begin{array}{ccc} A_{\mathrm{crys}} \otimes_{\sigma, A_{\mathrm{inf}}} M & \longrightarrow & A_{\mathrm{crys}}/I A_{\mathrm{crys}} \otimes_{\sigma, A_{\mathrm{inf}}} M \\ & & \downarrow \cong \\ & & A_{\mathrm{inf}}/\xi A_{\mathrm{inf}} \otimes_{\sigma, A_{\mathrm{inf}}} M \\ & & \downarrow \cong \\ & & \sigma^* M/\xi \sigma^* M \longrightarrow \sigma^* M/N \end{array}$$

is alternately characterized the A_{crys}-submodule generated by all elements of the form $1 \otimes \psi(m)$ and $a \otimes m$ with $m \in M$ and $a \in I$. There is a unique σ-semilinear map $\Phi_1 : Q \to P$ whose effect on these generators is

$$\Phi_1(1 \otimes \psi(m)) = \frac{\sigma(\xi)}{p} \otimes m, \qquad \Phi_1(a \otimes m) = \frac{\sigma(a)}{p} \otimes \phi(m).$$

The following is a special case of [Lau 2018, Theorem 1.5].

Theorem 2.3.5 (Lau). *The construction $(M, \phi, \psi) \mapsto (P, Q, \Phi, \Phi_1)$ just given defines a functor from the category of Breuil–Kisin–Fargues modules to the category of windows. It is an equivalence of categories if $p > 2$.*

Given a window (P, Q, Φ, Φ_1), define its *étale realization*

$$P_{\mathrm{et}} = \{x \in Q : \Phi_1(x) = x\}$$

as in [Lau 2019, §3]. This is a torsion-free $\mathbb{Z}_p$-module equipped with a *Hodge–Tate filtration*

$$F_{\mathrm{HT}}(P_{\mathrm{et}}) \subset P_{\mathrm{et}} \otimes_{\mathbb{Z}_p} C,$$

defined as the kernel of the C-linear extension of $P_{\mathrm{et}} \to Q/IP$.

Denote by HTpair the category of pairs (T, W) in which T is a torsion-free $\mathbb{Z}_p$-module, and $W \subset T \otimes_{\mathbb{Z}_p} C$ is a subspace. Using the obvious notation for the categories of Breuil–Kisin–Fargues modules, p-divisible groups over $\mathcal{O}$, and windows, we now have functors

$$\begin{array}{ccccc} \text{BKF-Mod} & \xrightarrow{a} & \text{Win} & \xleftarrow{b} & p\text{-DivGrp} \\ & {\scriptstyle d} \searrow & \downarrow {\scriptstyle c} & \swarrow {\scriptstyle e} & \\ & & \text{HTpair} & & \end{array} \tag{2.3.3}$$

Here a is given by Theorem 2.3.5, b is given by Theorem 2.3.4, c sends a window to its étale realization, d does the same for Breuil–Kisin–Fargues modules, and e sends a p-divisible group over $\mathcal{O}$ to its p-adic Tate module endowed with its Hodge filtration.

Remark 2.3.6. It is not obvious from the definitions that (2.3.3) commutes. When $p > 2$ the commutativity is a byproduct of the following proof.

Proof of Theorem 2.2.2. Assume that $p > 2$. In particular the functors of Theorems 2.3.4 and 2.3.5 are equivalences of categories, and their composition gives the desired equivalence of categories between p-divisible groups over $\mathcal{O}$ and Breuil–Kisin–Fargues modules.

Suppose G is a p-divisible group over $\mathcal{O}$, and let (P, Q, Φ, Φ_1) and (M, ϕ, ψ) be its corresponding window and Breuil–Kisin–Fargues module. The isomorphisms

$$\Omega(G^\vee) \cong Q/IP \cong N/\xi\sigma^* M \quad \text{and} \quad \mathrm{Lie}(G) \cong P/Q \cong \sigma^* M/N$$

are clear from the constructions of the functors of Theorems 2.3.4 and 2.3.5.

The quotient map $\mathcal{O} \to k$ induces a ring homomorphism $A_{\mathrm{inf}} \to W(k)$ sending $\xi \mapsto p$. It follows that there is a unique continuous extension to $A_{\mathrm{crys}} \to W(k)$ and, by (2.3.2), canonical isomorphisms

$$W(k) \otimes_{A_{\mathrm{crys}}} P \cong W(k) \otimes_{\sigma, A_{\mathrm{inf}}} M \cong M_{\mathrm{crys}}. \tag{2.3.4}$$

The functor of Theorem 2.3.4 is constructed in such a way that the leftmost $W(k)$-module in (2.3.4) is identified with the value of the Dieudonné crystal of G_k at the divided power thickening $W(k) \to k$, which is just the usual covariant Dieudonné module of G_k.

The window of the constant p-divisible group $\mathbb{Q}_p/\mathbb{Z}_p$ over $\mathcal{O}$ consists of

$$P^0 = A_{\mathrm{crys}} \quad \text{and} \quad Q^0 = A_{\mathrm{crys}}$$

endowed with the operators $\Phi : P^0 \to P^0$ and $\Phi_1 : Q^0 \to P^0$ defined by

$$\Phi(x) = p\sigma(x) \quad \text{and} \quad \Phi_1(x) = \sigma(x).$$

In particular there is a canonical isomorphism $Q^0/IP^0 \cong \mathcal{O}$.

The Breuil–Kisin–Fargues module of $\mathbb{Q}_p/\mathbb{Z}_p$ consists of

$$M^0 = A_{\mathrm{inf}}$$

endowed with the operators

$$\phi(x) = \xi\sigma(x) \quad \text{and} \quad \psi(x) = \sigma^{-1}(x).$$

The distinguished submodule $N^0 \subset \sigma^* M^0$ defined in Section 2.2 is all of $\sigma^* M^0 = \sigma^* A_{\mathrm{inf}}$, so is free of rank one generated by $1 \otimes 1$. Hence there is a canonical isomorphism $N^0/\xi\sigma^* M^0 \cong \mathcal{O}$.

From the equivalence of categories of Theorem 2.3.4 we obtain the commutative diagram

$$\begin{array}{ccc}
T_p(G) & \xrightarrow{\mathrm{HT}} & \Omega(G^\vee) \\
\| & & \| \\
\mathrm{Hom}_{\mathrm{p-DivGrp}}(\mathbb{Q}_p/\mathbb{Z}_p, G) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\Omega(\mu_{p^\infty}), \Omega(G^\vee)) \\
\| & & \| \\
\mathrm{Hom}_{\mathrm{Win}}(P^0, P) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(Q^0/IP^0, Q/IP) \\
\| & & \| \\
P_{\mathrm{et}} & \longrightarrow & Q/IP
\end{array} \tag{2.3.5}$$

Similarly, from the equivalence of categories of Theorem 2.3.5 we obtain the commutative diagram

$$\begin{array}{ccc} M_{\mathrm{et}} & \longrightarrow & N/\xi\sigma^* M \\ \| & & \| \\ \mathrm{Hom}_{\mathrm{BKF}}(M^0, M) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(N^0/\xi\sigma^* M^0, N/\xi\sigma^* M) \\ \| & & \| \\ \mathrm{Hom}_{\mathrm{Win}}(P^0, P) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(Q^0/IP^0, Q/IP) \\ \| & & \| \\ P_{\mathrm{et}} & \longrightarrow & Q/IP \end{array} \tag{2.3.6}$$

Combining these gives (2.2.2), completing the proof of Theorem 2.2.2.

As a final comment, we note that the diagrams (2.3.5) and (2.3.6) show that P_{et} and M_{et} are finitely generated $\mathbb{Z}_p$-modules, and that (2.3.3) commutes. If we denote by FinHTpair $\subset$ HTpair the full subcategory of pairs (T, W) with T of finite rank over $\mathbb{Z}_p$, we obtain a commutative diagram

$$\begin{array}{ccccc} \text{BKF-Mod} & \xrightarrow{a} & \text{Win} & \xleftarrow{b} & p\text{-DivGrp} \\ & \searrow^{d} & \downarrow{\scriptstyle c} & \swarrow_{e} & \\ & & \text{FinHTpair} & & \end{array}$$

in which the arrows a, b, and e are equivalences of categories (the last one by Theorem A). Hence all arrows are equivalences of categories. □

3. Bounding the Hodge–Tate periods

Let G be a p-divisible group of height four and dimension two over $\mathcal{O}$, endowed with an action $\Delta \to \mathrm{End}(G)$.

Throughout Section 3 we do not require $p > 2$. Instead we allow p to be arbitrary, but assume the conclusion of Theorem 2.2.2.

3.1. ***Hodge–Tate periods.*** The embeddings (1.3.1) determine a decomposition

$$\Omega(G^\vee) = \Omega_0(G^\vee) \oplus \Omega_1(G^\vee), \tag{3.1.1}$$

in which each summand is free of rank one over $\mathcal{O}$, and $\mathbb{Z}_{p^2} \subset \Delta$ acts on them through j_0 and j_1, respectively. The operator Π maps each summand injectively into the other. Applying $\otimes_{\mathcal{O}} k$ to (3.1.1) yields a decomposition

$$\Omega(G_k^\vee) = \Omega_0(G_k^\vee) \oplus \Omega_1(G_k^\vee)$$

into one-dimensional k-vector spaces.

Composing the Hodge–Tate morphism (1.1.1) with the two projections yields two *partial Hodge–Tate morphisms*

$$T_p(G) \xrightarrow{\mathrm{HT}_0} \Omega_0(G^\vee), \qquad T_p(G) \xrightarrow{\mathrm{HT}_1} \Omega_1(G^\vee).$$

By fixing isomorphisms

$$\Omega_0(G^\vee) \cong \mathcal{O}, \qquad \Omega_1(G^\vee) \cong \mathcal{O}, \tag{3.1.2}$$

we view these as $\mathcal{O}$-valued $\mathbb{Z}_p$-linear functionals on $T_p(G)$.

Lemma 3.1.1. *The Δ-module $T_p(G)$ is free of rank* 1.

Proof. As $\Delta \otimes \mathbb{Q}_p$ is a division ring, its module $T_p(G) \otimes \mathbb{Q}_p$ is necessarily free. Comparing $\mathbb{Q}_p$-dimensions shows that it is free of rank one, and hence $T_p(G)$ is isomorphic to some (left) Δ-submodule of $\Delta \otimes \mathbb{Q}_p$. As Δ admits a discrete valuation [Vignéras 1980, Lemme II.1.4] with uniformizer Π, every such submodule is principal and generated by a power of Π. ☐

Fix a Δ-module generator $\lambda \in T_p(G)$, and define

$$\tau_0 = \frac{\mathrm{HT}_0(\Pi\lambda)}{\mathrm{HT}_0(\lambda)}, \qquad \tau_1 = \frac{\mathrm{HT}_1(\Pi\lambda)}{\mathrm{HT}_1(\lambda)}.$$

These are the *Hodge–Tate periods* of G. In each fraction the numerator or denominator may vanish, but not simultaneously. Thus the Hodge–Tate periods lie in $\mathbb{P}^1(C) = C \cup \{\infty\}$. They do not depend on the choice of (3.1.2), but do depend on the choice of generator λ.

Proposition 3.1.2. *The Hodge–Tate periods satisfy $\tau_0 \cdot \tau_1 = p$.*

Proof. The action of Π on $\Omega_0(G^\vee) \oplus \Omega_1(G^\vee)$ is given by

$$(\omega_0, \omega_1) \mapsto (s_0\omega_1, s_1\omega_0)$$

for some $s_0, s_1 \in \mathcal{O}$ satisfying $s_0 s_1 = p$. From the Δ-linearity of the Hodge–Tate morphism we deduce first

$$\mathrm{HT}_0(\Pi\lambda) = s_0 \cdot \mathrm{HT}_1(\lambda), \qquad \mathrm{HT}_1(\Pi\lambda) = s_1 \cdot \mathrm{HT}_0(\lambda),$$

and then

$$\tau_0 \cdot \tau_1 = \frac{\mathrm{HT}_0(\Pi\lambda)}{\mathrm{HT}_0(\lambda)} \cdot \frac{\mathrm{HT}_1(\Pi\lambda)}{\mathrm{HT}_1(\lambda)} = s_0 \cdot s_1 = p.$$ ☐

3.2. *Reduction to the residue field.* Let G_k be the reduction of G to the residue field $k = \mathcal{O}/\mathfrak{m}$, and let (D, F, V) be its covariant Dieudonné module.

Definition 3.2.1. Let H be the p-divisible group of a supersingular elliptic curve over k. In other words, H is the unique connected p-divisible group of height two and dimension one. The reduction G_k is said to be

(1) *supersingular* if it is isogenous to $H \times H$,

(2) *superspecial* if it is isomorphic to $H \times H$.

Remark 3.2.2. Our notions of supersingular and superspecial depend only on the p-divisible group G_k, and not on its Δ-action. This differs from the meaning of superspecial in some literature on Shimura curves, e.g., [Kudla and Rapoport 2000].

The following proposition, which implies that the notion of superspecial depends only on the p-torsion subgroup scheme $G_k[p] \subset G_k$, is well-known. For lack of a reference we provide the proof.

Proposition 3.2.3. *The reduction G_k is supersingular, and the following are equivalent*:

(1) *G_k is superspecial.*

(2) *There is an isomorphism of group schemes $G_k[p] \cong H[p] \times H[p]$.*

(3) *$V^2 D \subset pD$.*

(4) *$FD = VD$.*

Proof. The supersingularity of G_k follows from the Dieudonné–Manin classification of isocrystals: one can list all isogeny classes of p-divisible groups over k of height four and dimension two, and the supersingular isogeny class is the only one whose endomorphism algebra contains a quaternion division algebra.

The implication (1) $\Longrightarrow$ (2) is trivial. For the implication (2) $\Longrightarrow$ (3) it suffices to check that V^2 kills the Dieudonné module of $H[p]$, which we leave to the reader.

Next we prove (3) $\Longrightarrow$ (4). If $D' \subset D$ is any $W(k)$-lattice stable under both F and V, then its corresponding p-divisible group G'_k is isogenous to G_k. In particular it has dimension 2, and hence

$$D'/VD' \cong \mathrm{Lie}(G'_k)$$

is a 2-dimensional k vector space. Applying this with $D' = D$ and $D' = VD$ shows that D/V^2D has length 4 as a $W(k)$-module. On the other hand, D/pD also has length 4, proving the first implication in

$$V^2 D \subset pD \Longrightarrow V^2 D = pD \Longrightarrow VD = FD.$$

Finally, we prove (4) $\Longrightarrow$ (1). Let α_p be the finite flat group scheme whose Dieudonné module is the $W(k)$-module k, endowed with the operators $F = 0$ and $V = 0$. If $FD = VD$ then, using the self-duality of α_p, we see that

$$\begin{aligned} \mathrm{Hom}(\alpha_p, G_k^\vee) &\cong \mathrm{Hom}(G_k[p], \alpha_p) \cong \mathrm{Hom}_k(D/(FD+VD), k) \\ &\cong \mathrm{Hom}_k(D/VD, k) \cong \mathrm{Hom}_k(\mathrm{Lie}(G), k) \end{aligned}$$

is a 2-dimensional k-vector space. It follows from [Oort 1975, Theorem 2] that $G_k^\vee$ is superspecial, and hence so is G_k. □

Let (M, ϕ, ψ) be the Breuil–Kisin–Fargues module of G. The quotient

$$M^\flat = M/pM$$

is a free module over $\mathcal{O}^\flat \cong A_{\mathrm{inf}}/(p)$, endowed with operators $\phi, \psi : M^\flat \to M^\flat$ satisfying $\phi \circ \psi = \varpi$.

Denote by $N^\flat = N/pN$ the image of

$$M^\flat \xrightarrow{m\mapsto 1\otimes\psi(m)} \sigma^* M^\flat.$$

Each of our embeddings $j_0, j_1 : \mathbb{Z}_{p^2} \to \mathcal{O}$ determines a map

$$\mathbb{Z}_{p^2} \to \mathcal{O}/p\mathcal{O} \cong \mathcal{O}^\flat/\varpi\mathcal{O}^\flat,$$

and these two maps lift uniquely to $j_0, j_1 : \mathbb{Z}_{p^2} \to \mathcal{O}^\flat$. The action of Δ on G determines an action on $M^\flat$, which induces a decomposition

$$M^\flat = M_0^\flat \oplus M_1^\flat$$

analogous to (3.1.1). It follows from the next proposition that each factor is free of rank two over $\mathcal{O}^\flat$.

Proposition 3.2.4.

(1) *D is free of rank one over $\Delta \otimes_{\mathbb{Z}_p} W(k)$.*

(2) *M is free of rank one over $\Delta \otimes_{\mathbb{Z}_p} A_{\mathrm{inf}}$.*

Proof. Reduce (1.3.1) to ring homomorphisms $j_0, j_1 : \mathbb{Z}_{p^2} \to k$, and denote again by $j_0, j_1 : \mathbb{Z}_{p^2} \to W(k)$ the unique lifts. There is a decomposition of $W(k)$-modules

$$D = D_0 \oplus D_1$$

in such a way that $\mathbb{Z}_{p^2} \subset \Delta$ acts on the two summands via j_0 and j_1, respectively. As in [Kudla and Rapoport 2000, §1], these summands are free of rank 2 over $W(k)$, and satisfy

$$pD_0 \subsetneq VD_1 \subsetneq D_0, \qquad pD_1 \subsetneq VD_0 \subsetneq D_1.$$

Moreover, either $\Pi D_0 = VD_0$ or $\Pi D_1 = VD_1$ (or both).

Without loss of generality, we may assume that $\Pi D_0 = VD_0$, and hence

$$pD_1 \subsetneq \Pi D_0 \subsetneq D_1.$$

Applying Π to these inclusions shows that

$$pD_0 \subsetneq \Pi D_1 \subsetneq D_0.$$

If we choose any $f_0 \in D_0$ and $f_1 \in D_1$ with nonzero images in $D_0/\Pi D_1$ and $D_1/\Pi D_0$, respectively, then $f_0, f_1, \Pi f_0, \Pi f_1 \in D$ reduce to a k-basis of D/pD. Using Nakayama's lemma it is easy to see that D is generated by $f_0 + f_1$ as a $\Delta \otimes W(k)$-module, and the first claim of the proposition follows.

Theorem 2.2.2 gives us an isomorphism

$$D/pD \cong \sigma^*(M/\mathfrak{m}M)$$

of $\Delta\otimes_{\mathbb{Z}_p} k$-modules, and from what was said above we deduce that $M/\mathfrak{m}M$ is free of rank one over $\Delta\otimes_{\mathbb{Z}_p} k$. The second claim of the proposition follows easily from this and Nakayama's lemma. □

3.3. ***The case*** $\mathbf{\Pi\Omega(G_k^\vee) = 0}$**.** We assume throughout Section 3.3 that

$$\Pi\Omega(G_k^\vee) = 0.$$

We will analyze the structure of $M^\flat$, with its operators ϕ and ψ, and use this to bound the Hodge–Tate periods of G. The first step is to choose a convenient basis.

Lemma 3.3.1. *There are $\mathcal{O}^\flat$-bases $e_0, f_0 \in M_0^\flat$ and $e_1, f_1 \in M_1^\flat$ such that the operator $\Pi \in \Delta$ satisfies*

$$\Pi e_0 = 0, \qquad \Pi e_1 = 0, \quad \Pi f_0 = e_1, \quad \Pi f_1 = e_0, \tag{3.3.1}$$

and such that ψ satisfies

$$\psi(e_0) = t_0 e_1, \quad \psi(e_1) = t_1 e_0, \quad \psi(f_0) = e_1 + t_1 f_1, \quad \psi(f_1) = e_0 + t_0 f_0$$

for scalars $t_0, t_1 \in \mathcal{O}^\flat$ satisfying $\mathrm{ord}(t_0) > 0$, $\mathrm{ord}(t_1) > 0$, *and*

$$\mathrm{ord}(t_0) + \mathrm{ord}(t_1) = 1/p.$$

Proof. As $M^\flat$ is free of rank one over $\Delta \otimes_{\mathbb{Z}_p} \mathcal{O}^\flat$, we may choose a basis such that (3.3.1) holds, and the relation $\psi \circ \Pi = \Pi \circ \psi$ then implies

$$\psi(e_0) = t_0 e_1, \quad \psi(e_1) = t_1 e_0, \quad \psi(f_0) = u_1 e_1 + t_1 f_1, \quad \psi(f_1) = u_0 e_0 + t_0 f_0$$

for some $u_0, u_1, t_0, t_1 \in \mathcal{O}^\flat$. The submodule $N^\flat \subset \sigma^* M^\flat$ is generated by

$$\begin{aligned} 1 \otimes \psi(e_0) &= t_0^p \otimes e_1, & 1 \otimes \psi(e_1) &= t_1^p \otimes e_0, \\ 1 \otimes \psi(f_0) &= u_1^p \otimes e_1 + t_1^p \otimes f_1, & 1 \otimes \psi(f_1) &= u_0^p \otimes e_0 + t_0^p \otimes f_0. \end{aligned}$$

Recall that $\mathfrak{m}^\flat \subset \mathcal{O}^\flat$ is the maximal ideal. The first isomorphism in (2.2.1) identifies $\Omega(G_k^\vee)$ with the image of $N^\flat$ in $(\sigma^* M^\flat)/\mathfrak{m}^\flat(\sigma^* M^\flat)$, and by hypothesis this k-vector space is annihilated by Π. It is easy to see from this that $\mathrm{ord}(t_0)$ and $\mathrm{ord}(t_1)$ are positive.

Using Theorem 2.2.2, we see that

$$\sigma^* M^\flat / N^\flat \cong (\sigma^* M / N) \otimes_{\mathcal{O}} \mathcal{O}/(p) \cong \mathrm{Lie}(G) \otimes_{\mathcal{O}} \mathcal{O}/(p)$$

is free of rank two over $\mathcal{O}/(p) \cong \mathcal{O}^\flat/(\varpi)$. On the other hand, $\sigma^* M^\flat / N^\flat$ is isomorphic to the cokernel of the matrix

$$\begin{pmatrix} t_1^p & & u_0^p & \\ & t_0^p & & u_1^p \\ & & t_0^p & \\ & & & t_1^p \end{pmatrix} \in M_4(\mathcal{O}^\flat),$$

whose reduction to $M_4(k)$ must therefore have rank 2. This implies that u_0 and u_1 are units, and using elementary row and column operations one sees that

$$\sigma^* M^\flat / N^\flat \cong \mathcal{O}^\flat/(t_0 t_1)^p \oplus \mathcal{O}^\flat/(t_0 t_1)^p.$$

Hence $(t_0t_1)^p = (\varpi)$. Finally, having already seen that u_0 and u_1 are units, an easy calculation shows that our basis elements may be rescaled in order to make $u_0 = 1$ and $u_1 = 1$. □

Fix a basis as in Lemma 3.3.1. Theorem 2.2.2 identifies

$$T_p(G)/pT(G) = M^{\psi=1}/pM^{\psi=1} \subset (M^\flat)^{\psi=1},$$

and the image of our fixed generator $\lambda \in T_p(G)$ has the form

$$a_0e_0 + a_1e_1 + b_0 f_0 + b_1 f_1 \in M^\flat$$

for some coefficients $a_0, a_1, b_0, b_1 \in \mathcal{O}^\flat$ satisfying

$$a_0^p = a_1t_1^p + b_1, \quad a_1^p = a_0t_0^p + b_0, \quad b_0^p = b_1t_0^p, \quad b_1^p = b_0t_1^p. \tag{3.3.2}$$

The first isomorphism of (2.2.1) identifies

$$\Omega(G^\vee)/p\Omega(G^\vee) = N/(pN + \xi\sigma^*M) = N^\flat/\varpi\sigma^*M^\flat$$

with the direct summand of $\sigma^*M^\flat/\varpi\sigma^*M^\flat$ generated by the reductions of

$$1 \otimes \psi(f_0) = 1 \otimes e_1 + t_1^p \otimes f_1 \in \sigma^*M^\flat, \quad 1 \otimes \psi(f_1) = 1 \otimes e_0 + t_0^p \otimes f_0 \in \sigma^*M^\flat.$$

If we use this basis to identify

$$\Omega(G^\vee)/p\Omega(G^\vee) = N^\flat/\varpi\sigma^*M^\flat \cong \mathcal{O}^\flat/(\varpi) \oplus \mathcal{O}^\flat/(\varpi)$$

then, again using Theorem 2.2.2, the partial Hodge–Tate morphisms

$$T_p(G)/pT_p(G) \xrightarrow{\mathrm{HT}_0} \Omega_0(G^\vee)/p\Omega_0(G^\vee) \cong \mathcal{O}^\flat/(\varpi)$$
$$T_p(G)/pT_p(G) \xrightarrow{\mathrm{HT}_1} \Omega_1(G^\vee)/p\Omega_1(G^\vee) \cong \mathcal{O}^\flat/(\varpi)$$

are given by

$$\begin{aligned} \mathrm{HT}_0(\lambda) &= a_1^p, \qquad \mathrm{HT}_0(\Pi\lambda) = b_0^p, \\ \mathrm{HT}_1(\lambda) &= a_0^p, \qquad \mathrm{HT}_1(\Pi\lambda) = b_1^p. \end{aligned} \tag{3.3.3}$$

Lemma 3.3.2. *For $i \in \{0, 1\}$, we have*

$$\mathrm{ord}(b_i) = \frac{1}{p^2-1} + \frac{p \cdot \mathrm{ord}(t_i)}{p+1}.$$

Proof. As $\Pi\lambda \in T_p(G)$ has nonzero image in

$$T_p(G)/pT_p(G) \subset M^\flat,$$

we must have $b_0e_1 + b_1e_0 \neq 0$. Therefore one of b_0 and b_1 is nonzero. The relations (3.3.2) then imply first that both b_0 and b_1 are nonzero, and then that

$$b_i^{p^2-1} = (t_0t_1)^p \cdot t_i^{p(p-1)}.$$

The claim follows by applying ord to both sides of this equality. □

Lemma 3.3.3. *If we assume that*

$$\frac{1}{p^2(p-1)} < \mathrm{ord}(t_1),$$

then

$$\mathrm{ord}(a_0) = \frac{1}{p(p^2-1)} + \frac{\mathrm{ord}(t_1)}{p+1}, \qquad \mathrm{ord}(a_1) = \frac{1}{p^2-1} - \frac{\mathrm{ord}(t_1)}{p+1}.$$

Of course there is a similar result if t_1 is replaced by t_0.

Proof. Recall the equality $a_0^p = a_1 t_1^p + b_1$ from (3.3.2). The only way this can hold is if (at least) one of the three relations

- $p \cdot \mathrm{ord}(a_0) = \mathrm{ord}(b_1) \le \mathrm{ord}(t_1^p a_1)$
- $p \cdot \mathrm{ord}(a_0) = \mathrm{ord}(t_1^p a_1) \le \mathrm{ord}(b_1)$
- $\mathrm{ord}(b_1) = \mathrm{ord}(t_1^p a_1) \le p \cdot \mathrm{ord}(a_0)$

is satisfied. The second and third relations cannot be satisfied, as each implies

$$0 \le \mathrm{ord}(a_1) \le \mathrm{ord}(b_1) - p \cdot \mathrm{ord}(t_1) = \frac{1}{p^2-1} - \frac{p^2 \cdot \mathrm{ord}(t_1)}{p+1} < 0.$$

Hence the first relation holds, and Lemma 3.3.2 shows that

$$p \cdot \mathrm{ord}(a_0) = \mathrm{ord}(b_1) = \frac{1}{p^2-1} + \frac{p \cdot \mathrm{ord}(t_1)}{p+1}.$$

This proves the first equality.

For the second equality, the relations (3.3.2) imply

$$\begin{aligned} a_0^{p^2} &= a_1^p \cdot (t_1^p + t_1)^p - (t_0 t_1)^p a_0, \\ a_1^{p^2} &= a_0^p \cdot (t_0^p + t_0)^p - (t_0 t_1)^p a_1. \end{aligned}$$

Using the second of these, along with

$$\mathrm{ord}\big(a_0^p \cdot (t_0^p + t_0)^p\big) = \mathrm{ord}(b_1) + p \cdot \mathrm{ord}(t_0) = \frac{p^2}{p^2-1} - \frac{p^2 \cdot \mathrm{ord}(t_1)}{p+1} < 1 \le \mathrm{ord}\big((t_0 t_1)^p a_1\big),$$

we find that

$$\mathrm{ord}(a_1) = \frac{\mathrm{ord}\big(a_0^p \cdot (t_0^p + t_0)^p\big)}{p^2} = \frac{1}{p^2-1} - \frac{\mathrm{ord}(t_1)}{p+1}. \qquad \square$$

Now we can prove the main result of this subsection.

Proposition 3.3.4. *If we assume, as above, that $\Pi\Omega(G_k^\vee) = 0$ then*

$$\frac{1}{p+1} < \mathrm{ord}(\tau_0) < \frac{p}{p+1} \quad \textit{and} \quad \frac{1}{p+1} < \mathrm{ord}(\tau_1) < \frac{p}{p+1}.$$

Proof. First assume that

$$\frac{1}{p^2(p-1)} < \mathrm{ord}(t_1). \tag{3.3.4}$$

The discussion leading to (3.3.3) provides us with an $\mathcal{O}$-module isomorphism

$$\Omega_0(G^\vee)/p\Omega_0(G^\vee) \cong \mathcal{O}^\flat/(\varpi) \cong \mathcal{O}/(p),$$

and we fix any lift to an isomorphism $\Omega_0(G^\vee) \cong \mathcal{O}$.

It is easy to see from Lemmas 3.3.2 and 3.3.3 that $\mathrm{ord}(a_1)$ and $\mathrm{ord}(b_0)$ lie in the open interval $(0, 1/p)$, and so a_1^p and b_0^p have nonzero images in $\mathcal{O}^\flat/(\varpi)$. By (3.3.3) these images agree with the images of $\mathrm{HT}_0(\lambda)$ and $\mathrm{HT}_0(\Pi\lambda)$ under

$$\mathcal{O} \to \mathcal{O}/(p) \cong \mathcal{O}^\flat/(\varpi).$$

Thus

$$\mathrm{ord}(\mathrm{HT}_0(\lambda)) = \mathrm{ord}(a_1^p) = \frac{p}{p^2-1} - \frac{p \cdot \mathrm{ord}(t_1)}{p+1}$$

and

$$\mathrm{ord}(\mathrm{HT}_0(\Pi\lambda)) = \mathrm{ord}(b_0^p) = \frac{p}{p^2-1} + \frac{p^2 \cdot \mathrm{ord}(t_0)}{p+1}.$$

It follows that

$$\mathrm{ord}(\tau_0) = \mathrm{ord}(\mathrm{HT}_0(\Pi\lambda)) - \mathrm{ord}(\mathrm{HT}_0(\lambda)) = \frac{p}{p+1} - \frac{(p-1)}{p+1} \cdot \mathrm{ord}(t_1^p),$$

and so

$$\frac{1}{p+1} < \mathrm{ord}(\tau_0) < \frac{p}{p+1}.$$

The analogous inequalities for $\mathrm{ord}(\tau_1)$ follow from the relation $\tau_0\tau_1 = p$ of Proposition 3.1.2.

This proves Proposition 3.3.4 under the assumption (3.3.4). The proof when

$$\frac{1}{p^2(p-1)} < \mathrm{ord}(t_0) \tag{3.3.5}$$

is entirely similar.

Thus we are left to prove the claim under the assumption that both (3.3.4) and (3.3.5) fail. This assumption implies that

$$\frac{1}{p} = \mathrm{ord}(t_0) + \mathrm{ord}(t_1) \le \frac{2}{p^2(p-1)},$$

which implies that $p = 2$ and

$$\mathrm{ord}(t_0) = \tfrac{1}{4} = \mathrm{ord}(t_1).$$

In particular, Lemma 3.3.2 simplifies to

$$\mathrm{ord}(b_0) = \tfrac{1}{2} = \mathrm{ord}(b_1).$$

Consider the equality $a_0^2 = a_1 t_1^2 + b_1$ of (3.3.2). As in the proof of Lemma 3.3.3, the only way this can hold is if (at least) one of the relations

- $\mathrm{ord}(a_0) = \frac{1}{4}$
- $\mathrm{ord}(a_1) = 0$ and $\mathrm{ord}(a_0) \geq \frac{1}{4}$

holds. Similarly, the equality $a_1^2 = a_0 t_0^2 + b_0$ implies that (at least) one of the relations

- $\mathrm{ord}(a_1) = \frac{1}{4}$
- $\mathrm{ord}(a_0) = 0$ and $\mathrm{ord}(a_1) \geq \frac{1}{4}$

holds. Combining these shows that $\mathrm{ord}(a_0) = \frac{1}{4}$ and $\mathrm{ord}(a_1) = \frac{1}{4}$.

In particular, a_1^p has nonzero image in $\mathcal{O}^\flat/(\varpi)$, and

$$\mathrm{ord}(\mathrm{HT}_0(\lambda)) = \mathrm{ord}(a_1^p) = \tfrac{1}{2}.$$

On the other hand, b_0^p has trivial image in $\mathcal{O}^\flat/(\varpi)$, and so

$$\mathrm{ord}(\mathrm{HT}_0(\Pi\lambda)) \geq 1.$$

Therefore

$$\mathrm{ord}(\tau_0) = \mathrm{ord}(\mathrm{HT}_0(\Pi\lambda)) - \mathrm{ord}(\mathrm{HT}_0(\lambda)) \geq \tfrac{1}{2}.$$

The same reasoning shows that $\mathrm{ord}(\tau_1) \geq \frac{1}{2}$. As $\tau_0\tau_1 = p$ by Proposition 3.1.2, we must therefore have

$$\mathrm{ord}(\tau_0) = \tfrac{1}{2} = \mathrm{ord}(\tau_1),$$

completing the proof of Proposition 3.3.4. □

3.4. *The case* $\Pi\Omega_1(G_k^\vee) \neq 0$. We assume throughout Section 3.4 that

$$\Pi\Omega_1(G_k^\vee) \neq 0.$$

Once again, we will analyze the structure of $M^\flat = M/pM$, and use this to bound the Hodge–Tate periods of G. As in Section 3.3, the first step is to choose a convenient basis for $M^\flat$.

Lemma 3.4.1. *There are $\mathcal{O}^\flat$-bases $e_0, f_0 \in M_0^\flat$ and $e_1, f_1 \in M_1^\flat$ such that the operator $\Pi \in \Delta$ satisfies*

$$\Pi e_0 = 0, \qquad \Pi e_1 = 0, \quad \Pi f_0 = e_1, \quad \Pi f_1 = e_0, \tag{3.4.1}$$

and such that ψ satisfies

$$\psi(e_0) = e_1, \quad \psi(e_1) = te_0, \quad \psi(f_0) = tf_1, \quad \psi(f_1) = se_0 + f_0 \tag{3.4.2}$$

for some scalars $s, t \in \mathcal{O}^\flat$ with $\mathrm{ord}(t) = 1/p$. Moreover:

(1) *For any such basis, G_k is superspecial if and only if $\mathrm{ord}(s) > 0$.*

(2) *If G_k is not superspecial such a basis can be found with $s = 1$.*

Proof. Exactly as in the proof of Lemma 3.3.1, we may choose a basis such that (3.4.1) holds, and such that

$$\psi(e_0) = t_0 e_1, \quad \psi(e_1) = t_1 e_0, \quad \psi(f_0) = u_1 e_1 + t_1 f_1, \quad \psi(f_1) = u_0 e_0 + t_0 f_0$$

for some $u_0, u_1, t_0, t_1 \in \mathcal{O}^\flat$ with $\mathrm{ord}(t_0) + \mathrm{ord}(t_1) = 1/p$.

The Δ-module $\Omega(G_k^\vee)$ is identified with the image of

$$N^\flat \to (\sigma^* M^\flat)/\mathfrak{m}^\flat(\sigma^* M^\flat),$$

and this identifies $\Omega_1(G_k^\vee)$ with the (one-dimensional) k-span of the vectors

$$1 \otimes \psi(e_1) = t_1^p \otimes e_0, \quad 1 \otimes \psi(f_1) = u_0^p \otimes e_0 + t_0^p \otimes f_0$$

in $(\sigma^* M_0^\flat)/\mathfrak{m}^\flat(\sigma^* M^\flat)$. The assumption that Π does not annihilate $\Omega_1(G_k^\vee)$ implies that $\mathrm{ord}(t_0) = 0$, which allows us to rescale our basis vectors to make $t_0 = 1$, and then add a multiple of e_0 to f_0 to make $u_1 = 0$. Setting $t = t_1$ and $s = u_0$, the relations (3.4.2) now hold.

It follows from Proposition 3.2.3 and Theorem 2.2.2 that

$$G_k \text{ is superspecial} \iff V^2(D/pD) = 0 \iff \psi^2(M^\flat/\mathfrak{m}^\flat M^\flat) = 0 \iff \mathrm{ord}(s) > 0.$$

Finally, if $\mathrm{ord}(s) = 0$ it is an easy exercise in linear algebra to see that the given basis elements can be rescaled to make $s = 1$. □

As in Section 3.3, our fixed generator $\lambda \in T_p(G)$ determines an element

$$a_0 e_0 + a_1 e_1 + b_0 f_0 + b_1 f_1 \in M^\flat,$$

where the coefficients $a_0, a_1, b_0, b_1 \in \mathcal{O}^\flat$ satisfy

$$a_0^p = a_1 t^p + b_1 s^p, \quad a_1^p = a_0, \quad b_0^p = b_1, \quad b_1^p = b_0 t^p. \tag{3.4.3}$$

As in Section 3.3, we may identify

$$\Omega(G^\vee)/p\Omega(G^\vee) = N/(pN + \xi\sigma^* M) = N^\flat/\varpi\sigma^* M^\flat$$

with the direct summand of $\sigma^* M^\flat/\varpi\sigma^* M^\flat$ generated by the reductions of

$$1 \otimes \psi(e_0) = 1 \otimes e_1 \in \sigma^* M^\flat, \quad 1 \otimes \psi(f_1) = s^p \otimes e_0 + 1 \otimes f_0 \in \sigma^* M^\flat.$$

If we use this basis to identify

$$\Omega(G^\vee)/p\Omega(G^\vee) = N^\flat/\varpi\sigma^* M^\flat \cong \mathcal{O}^\flat/(\varpi) \oplus \mathcal{O}^\flat/(\varpi)$$

then, using Theorem 2.2.2, the partial Hodge–Tate morphisms

$$T_p(G)/pT_p(G) \xrightarrow{\mathrm{HT}_0} \Omega_0(G^\vee)/p\Omega_0(G^\vee) \cong \mathcal{O}^\flat/(\varpi),$$
$$T_p(G)/pT_p(G) \xrightarrow{\mathrm{HT}_1} \Omega_1(G^\vee)/p\Omega_1(G^\vee) \cong \mathcal{O}^\flat/(\varpi)$$

satisfy

$$\begin{aligned} \mathrm{HT}_0(\lambda) &= a_0, & \mathrm{HT}_0(\Pi\lambda) &= b_1, \\ \mathrm{HT}_1(\lambda) &= b_1, & \mathrm{HT}_1(\Pi\lambda) &= 0. \end{aligned} \tag{3.4.4}$$

Lemma 3.4.2. *We have*

$$\mathrm{ord}(b_0)=\frac{1}{p^2-1},\quad \mathrm{ord}(b_1)=\frac{p}{p^2-1}.$$

Moreover,

$$\mathrm{ord}(a_0)\geq\frac{1}{p^2-1},\quad \mathrm{ord}(a_1)\geq\frac{1}{p(p^2-1)},$$

and G_k is superspecial if and only if one (equivalently, both) of these inequalities is strict.

Proof. Exactly as in the proof of Lemma 3.3.2, both b_0 and b_1 are nonzero. The relations (3.4.3) therefore imply that

$$b_0^{p^2-1}=t^p,$$

from which the stated formulas for $\mathrm{ord}(b_0)$ and $\mathrm{ord}(b_1)=\mathrm{ord}(b_0^p)$ are clear.

The relations (3.4.3) imply that a_0 is a root of $x^{p^2}-xt^{p^2}-b_1^p s^{p^2}$, and by examination of the Newton polygon we see that

$$\mathrm{ord}(a_0)\geq\frac{1}{p^2-1}$$

with strict inequality if and only if $\mathrm{ord}(s)>0$. Combining this with $a_1^p=a_0$ completes the proof. □

Lemma 3.4.3. *If G_k is not superspecial then*

$$\varpi(a_0/b_1)^{p+1}\in(\mathcal{O}^\flat)^\times \quad \textit{and} \quad \varpi s^{p+1}/t^p\in(\mathcal{O}^\flat)^\times,$$

and these units have the same reduction to $k^\times$.

Proof. We have already noted that (3.4.3) implies $t^p=b_0^{p^2-1}$, from which one easily deduces the equality

$$\left(\frac{a_1}{b_0}\right)^{p^2}=\frac{a_1}{b_0}+\frac{s^p}{b_0^{p(p-1)}}$$

in the fraction field of $\mathcal{O}^\flat$. It follows from this and Lemma 3.4.2 that

$$\varpi^{p/(p+1)}\left(\frac{a_1}{b_0}\right)^{p^2} \quad \text{and} \quad \left(\frac{\varpi^{1/(p+1)}s}{b_0^{p-1}}\right)^p$$

are units in $\mathcal{O}^\flat$ with the same reduction to $k^\times$, hence the same is true after raising both to the power $(p+1)/p$. The lemma follows easily from this and the relations (3.4.3). □

Proposition 3.4.4. *If we assume, as above, that $\Pi\Omega_1(G_k^\vee)\neq 0$ then*

$$\frac{p}{p+1}\leq\mathrm{ord}(\tau_1) \tag{3.4.5}$$

with strict inequality if and only if G_k is superspecial. Moreover, if equality holds then

$$-\frac{p}{\tau_0^{p+1}}\in\mathcal{O}^\times \quad \textit{and} \quad \frac{\varpi s^{p+1}}{t^p}\in(\mathcal{O}^\flat)^\times$$

have the same reduction to $k^\times$.

Proof. Using (3.4.4) and Lemma 3.4.2, we find that

$$\mathrm{ord}(\mathrm{HT}_0(\Pi\lambda)) = \frac{p}{p^2-1},$$

and that

$$\mathrm{ord}(\mathrm{HT}_0(\lambda)) \geq \frac{1}{p^2-1}$$

with strict inequality if and only if G_k is superspecial. This implies that

$$\mathrm{ord}(\tau_0) = \mathrm{ord}(\mathrm{HT}_0(\Pi\lambda)) - \mathrm{ord}(\mathrm{HT}_0(\lambda)) \leq \frac{1}{p+1}$$

with strict inequality if and only if G_k is superspecial. The inequality (3.4.5) follows from this and the relation $\tau_0\tau_1 = p$ of Proposition 3.1.2, with strict inequality if and only if G_k is superspecial.

Suppose that equality holds in (3.4.5), so that G_k is not superspecial. Choose an $\alpha \in \mathcal{O}^\flat$ satisfying $\alpha^{p^2-1} = \varpi$. The construction of Section 2.1 determines an element $\alpha^\sharp \in \mathcal{O}$ whose image in $\mathcal{O}/(p) \cong \mathcal{O}^\flat/(\varpi)$ agrees with α.

Combining the relations (3.4.4) with Lemma 3.4.2 shows that

$$\frac{\mathrm{HT}_0(\Pi\lambda)}{(\alpha^\sharp)^p} \in \mathcal{O}^\times \quad \text{and} \quad \frac{b_1}{\alpha^p} \in (\mathcal{O}^\flat)^\times$$

have the same reduction to $k^\times$, as do

$$\frac{\mathrm{HT}_0(\lambda)}{\alpha^\sharp} \in \mathcal{O}^\times \quad \text{and} \quad \frac{a_0}{\alpha} \in (\mathcal{O}^\flat)^\times.$$

It follows that

$$\frac{\tau_0}{(\alpha^\sharp)^{p-1}} \in \mathcal{O}^\times \quad \text{and} \quad \frac{b_1}{a_0\alpha^{p-1}} \in (\mathcal{O}^\flat)^\times$$

have the same reduction to $k^\times$. Raising both to the power $p+1$ and applying Lemma 3.4.3 proves that

$$\frac{\varpi^\sharp}{\tau_0^{p+1}} \in \mathcal{O}^\times \quad \text{and} \quad \frac{\varpi s^{p+1}}{t^p} \in (\mathcal{O}^\flat)^\times$$

have the same reduction to $k^\times$. Now apply Lemma 2.1.1. □

4. The main results

We now formulate and prove our main results on the Ekedahl–Oort stratification of the Hodge–Tate period domain X defined by (1.2.1). Throughout Section 4 we assume that the conclusions of Theorem 2.2.2 hold. For example, it is enough to assume that $p > 2$.

4.1. ***The setup.*** Let T be a free Δ-module of rank one, and fix a generator $\lambda \in T$. Use the embeddings (1.3.1) to decompose

$$T \otimes_{\mathbb{Z}_p} C = T_{C,0} \oplus T_{C,1}$$

as a direct sum of 2-dimensional C-subspaces, in such a way that the action of $\mathbb{Z}_{p^2} \subset \Delta$ on the summands is through j_0 and j_1, respectively. Using the projection maps to the two factors, we obtain injective $\mathbb{Z}_p$-linear maps

$$q_0 : T \to T_{C,0}, \qquad q_1 : T \to T_{C,1}.$$

To each $\tau \in C \cup \{\infty\}$ we associate the Δ-stable plane

$$W_\tau \subset T \otimes_{\mathbb{Z}_p} C$$

spanned by the two vectors

$$\tau q_0(\lambda) - q_0(\Pi\lambda) \in T_{C,0}, \qquad p q_1(\lambda) - \tau q_1(\Pi\lambda) \in T_{C,1}.$$

The construction $\tau \mapsto W_\tau$ establishes a bijection

$$C \cup \{\infty\} \cong X(C).$$

Remark 4.1.1. It is not hard to see that the above bijection $\mathbb{P}^1(C) \cong X(C)$ arises from an isomorphism of schemes over $\mathbb{Q}_{p^2}$. The isomorphism cannot descend to $\mathbb{Q}_p$, for the simple reason that $X(\mathbb{Q}_p) = \varnothing$.

For the rest of Section 4.1 and Section 4.2 we hold $\tau \in C \cup \{\infty\}$ fixed, and let G be the p-divisible group over $\mathcal{O}$ determined by the pair (T, W_τ). Thus G comes equipped with an action of Δ, and Δ-linear identifications

$$\begin{array}{ccc} T_p(G) & \xrightarrow{\ \mathrm{HT}\ } & \Omega(G^\vee) \otimes_{\mathcal{O}} C \\ \| & & \| \\ T & \longrightarrow & (T \otimes_{\mathbb{Z}_p} C)/W_\tau \end{array}$$

In the notation of Section 3.1, the Hodge–Tate periods of G are

$$\tau_0 = \tau \quad \text{and} \quad \tau_1 = p/\tau. \tag{4.1.1}$$

4.2. ***Computing the reduction.*** Let G_k be the reduction of G to the residue field $k = \mathcal{O}/\mathfrak{m}$, and let (D, F, V) be its covariant Dieudonné module. We will show how to compute the isomorphism class of $G_k[p]$ from the Hodge–Tate periods (4.1.1).

Let $\mathbb{D} = \Delta \otimes_{\mathbb{Z}_p} k$ with its natural action of Δ by left multiplication. The embeddings (1.3.1) induce a decomposition

$$\mathbb{D} = \mathbb{D}_0 \oplus \mathbb{D}_1$$

in which $\mathbb{Z}_{p^2} \subset \Delta$ acts on $\mathbb{D}_i$ through the composition of $j_i : \mathbb{Z}_{p^2} \to \mathcal{O}$ with the reduction map $\mathcal{O} \to k$.

Choose k-bases

$$e_0,\, f_0 \in \mathbb{D}_0, \qquad e_1,\, f_1 \in \mathbb{D}_1$$

in such a way that $\Pi \in \Delta$ acts as

$$\Pi e_0 = 0, \quad \Pi e_1 = 0, \quad \Pi f_0 = e_1, \quad \Pi f_1 = e_0. \tag{4.2.1}$$

Theorem 4.2.1. *The inequalities*

$$\frac{1}{p+1} < \mathrm{ord}(\tau) < \frac{p}{p+1} \tag{4.2.2}$$

hold if and only if $\Pi\Omega(G_k^\vee) = 0$. *When these conditions hold, there is a* Δ*-linear isomorphism* $D/pD \cong \mathbb{D}$ *under which*

$$\begin{aligned} Fe_0 &= 0, & Ff_0 &= e_1, & Fe_1 &= 0, & Ff_1 &= e_0, \\ Ve_0 &= 0, & Vf_0 &= e_1, & Ve_1 &= 0, & Vf_1 &= e_0. \end{aligned}$$

Proof. If $\Pi\Omega(G_k^\vee) \neq 0$ then either $\Pi\Omega_1(G_k^\vee) \neq 0$ or $\Pi\Omega_0(G_k^\vee) \neq 0$. In the first case Proposition 3.4.4 implies

$$\frac{p}{p+1} \leq \mathrm{ord}(\tau_1).$$

In the second case the same proof, with indices 0 and 1 interchanged throughout, shows that

$$\frac{p}{p+1} \leq \mathrm{ord}(\tau_0).$$

In either case, these bounds imply that (4.2.2) fails.

Now assume that $\Pi\Omega(G_k^\vee) = 0$. We have already proved in Proposition 3.3.4 that (4.2.2) holds, and so it only remains to prove that D/pD admits an isomorphism to $\mathbb{D}$ with the prescribed properties.

Let $e_0,\, f_0 \in M_0^\flat$ and $e_1,\, f_1 \in M_1^\flat$ be the bases of Lemma 3.3.1. Using the formula for $\psi : M^\flat \to M^\flat$ prescribed in that lemma, and the relation $\phi \circ \psi = \varpi$, one can write down an explicit formula for ϕ, and then see that the induced operators on the reduction $M^\flat/\mathfrak{m}^\flat M^\flat$ are given by

$$\begin{aligned} \phi(e_0) &= 0, & \phi(f_0) &= ue_1, & \phi(e_1) &= 0, & \phi(f_1) &= ue_0, \\ \psi(e_0) &= 0, & \psi(f_0) &= e_1, & \psi(e_1) &= 0, & \psi(f_1) &= e_0, \end{aligned}$$

where $u^{-1} \in k^\times$ is the reduction of $-t_0^p t_1^p/\varpi \in (\mathcal{O}^\flat)^\times$.

The images of e_0, f_0, e_1, f_1 under the bijection

$$M^\flat/\mathfrak{m}^\flat M^\flat \xrightarrow{x \mapsto 1\otimes x} \sigma^*(M^\flat/\mathfrak{m}^\flat M^\flat) \cong M_{\mathrm{crys}}/pM_{\mathrm{crys}} \cong D/pD$$

provided by Theorem 2.2.2 form a k-basis of D/pD, denoted the same way, satisfying the relations (4.2.1) and

$$\begin{aligned} Fe_0 &= 0, & Ff_0 &= u^p e_1, & Fe_1 &= 0, & Ff_1 &= u^p e_0, \\ Ve_0 &= 0, & Vf_0 &= e_1, & Ve_1 &= 0, & Vf_1 &= e_0. \end{aligned}$$

It remains to prove that $u = 1$. The two embeddings (1.3.1) reduce to morphisms $j_0, j_1 : \mathbb{Z}_{p^2} \to k$, which then admit unique lifts to

$$j_0, j_1 : \mathbb{Z}_{p^2} \to W(k).$$

This allows us to decompose $D = D_0 \oplus D_1$ as W-modules, where $\mathbb{Z}_{p^2} \subset \Delta$ acts on the two summands via j_0 and j_1, respectively. Choose arbitrary lifts

$$\tilde{f}_0 \in D_0, \qquad \tilde{f}_1 \in D_1$$

of f_0 and f_1, and then define

$$\tilde{e}_0 = \Pi \tilde{f}_1 \in D_0, \qquad \tilde{e}_1 = \Pi \tilde{f}_0 \in D_1.$$

Using the fact that Π and V commute, we see that

$$\begin{aligned} V\tilde{e}_0 &= pb_1\tilde{e}_1 + pa_1\tilde{f}_1, & V\tilde{f}_0 &= a_1\tilde{e}_1 + pb_1\tilde{f}_1, \\ V\tilde{e}_1 &= pb_0\tilde{e}_0 + pa_0\tilde{f}_0, & V\tilde{f}_1 &= a_0\tilde{e}_0 + pb_0\tilde{f}_0, \end{aligned}$$

for scalars

$$a_0, a_1 \in 1 + pW(k), \qquad b_0, b_1 \in W(k).$$

Denote again by $\sigma : W(k) \to W(k)$ the lift of the Frobenius on k. Applying F to the expressions for $V\tilde{e}_1$ and $V\tilde{f}_1$ results in

$$p\tilde{e}_1 = \sigma(pb_0)F\tilde{e}_0 + \sigma(pa_0)F\tilde{f}_0, \qquad p\tilde{f}_1 = \sigma(a_0)F\tilde{e}_0 + \sigma(pb_0)F\tilde{f}_0,$$

from which one deduces

$$\big(\sigma(a_0)^2 - p\sigma(b_0)^2\big) \cdot F\tilde{f}_0 = \sigma(a_0)\tilde{e}_1 - p\sigma(b_0)\tilde{f}_1.$$

Reducing this modulo p proves that $Ff_0 = e_1$, and hence $u = 1$. $\square$

Theorem 4.2.2. *The inequality*

$$\operatorname{ord}(\tau) \le \frac{1}{p+1} \tag{4.2.3}$$

holds if and only if $\Pi\Omega_1(G_k^\vee) \neq 0$. *Moreover*:

(1) *If strict inequality holds in* (4.2.3), *there is a* Δ*-linear isomorphism* $D/pD \cong \mathbb{D}$ *under which*

$$\begin{aligned} Fe_0 &= e_1, & Ff_0 &= 0, & Fe_1 &= 0, & Ff_1 &= f_0, \\ Ve_0 &= e_1, & Vf_0 &= 0, & Ve_1 &= 0, & Vf_1 &= f_0. \end{aligned}$$

(2) *If equality holds in* (4.2.3), *there is a* Δ*-linear isomorphism* $D/pD \cong \mathbb{D}$ *under which*

$$\begin{aligned} Fe_0 &= u^p e_1, & Ff_0 &= -u^p e_1, & Fe_1 &= 0, & Ff_1 &= u^p f_0, \\ Ve_0 &= e_1, & Vf_0 &= 0, & Ve_1 &= 0, & Vf_1 &= e_0 + f_0, \end{aligned} \tag{4.2.4}$$

where u is the image of $-p/\tau_0^{p+1} = -p/\tau^{p+1}$ *under* $\mathcal{O}^\times \to k^\times$.

Proof. If (4.2.3) holds then Theorem 4.2.1 implies that $\Pi\Omega(G_k^\vee) \neq 0$, and so either

$$\Pi\Omega_0(G_k^\vee) \neq 0 \quad \text{or} \quad \Pi\Omega_1(G_k^\vee) \neq 0.$$

The first possibility cannot occur, as then the proof of Proposition 3.4.4, with the indices 0 and 1 reversed everywhere, would give the bound

$$\frac{p}{p+1} \leq \mathrm{ord}(\tau_0),$$

contradicting (4.2.3). Conversely, if $\Pi\Omega_1(G_k^\vee) \neq 0$ then (4.2.3) holds by Proposition 3.4.4.

Assume now that (4.2.3) holds, and that $\Pi\Omega_1(G_k^\vee) \neq 0$. Let $e_0, f_0 \in M_0^\flat$ and $e_1, f_1 \in M_1^\flat$ be the bases of Lemma 3.4.1. As in the proof of Theorem 4.2.1, the operator ϕ on $M^\flat$ can be computed from the formula for ψ given in the lemma. The induced operators on the reduction $M^\flat/\mathfrak{m}^\flat M^\flat$ are found to be

$$\begin{aligned} &\phi(e_0) = ue_1, && \phi(f_0) = -uv^p e_1, && \phi(e_1) = 0, && \phi(f_1) = uf_0,\\ &\psi(e_0) = e_1, && \psi(f_0) = 0, && \psi(e_1) = 0, && \psi(f_1) = ve_0 + f_0, \end{aligned}$$

where $u \in k^\times$ is the reduction of $\varpi/t^p \in (\mathcal{O}^\flat)^\times$, and $v \in k$ is the reduction of $s \in \mathcal{O}^\flat$. By the final claim of Lemma 3.4.1, we may further assume that

$$v = \begin{cases} 0 & \text{if } G_k \text{ is superspecial,} \\ 1 & \text{otherwise.} \end{cases}$$

Suppose that strict inequality holds in (4.2.3). Proposition 3.4.4 tells us that G_k is superspecial, and so $v = 0$. The images of e_0, f_0, e_1, f_1 under the bijection

$$M^\flat/\mathfrak{m}^\flat M^\flat \xrightarrow{x \mapsto 1\otimes x} \sigma^*(M^\flat/\mathfrak{m}^\flat M^\flat) \cong M_{\mathrm{crys}}/pM_{\mathrm{crys}} \cong D/pD$$

provided by Theorem 2.2.2 form a k-basis of D/pD, denoted the same way, satisfying the relations (4.2.1) and

$$\begin{aligned} &Fe_0 = u^p e_1, && Ff_0 = 0, && Fe_1 = 0, && Ff_1 = u^p f_0,\\ &Ve_0 = e_1, && Vf_0 = 0, && Ve_1 = 0, && Vf_1 = f_0. \end{aligned}$$

One can prove that $u = 1$ by lifting the basis elements to D and arguing exactly as in Theorem 4.2.1.

Suppose now that equality holds in (4.2.3). Proposition 3.4.4 implies that G_k is not superspecial, so $v = 1$, and that the reduction map $\mathcal{O}^\times \to k^\times$ sends $-p/\tau_0^{p+1} \mapsto u$. Arguing as in the previous paragraph, we obtain a k-basis e_0, f_0, e_1, f_1 of D/pD satisfying (4.2.1) and (4.2.4), completing the proof. □

Theorem 4.2.3. *The inequality*

$$\frac{p}{p+1} \leq \mathrm{ord}(\tau) \tag{4.2.5}$$

holds if and only if $\Pi\Omega_0(G_k^\vee) \neq 0$. *Moreover*:

(1) *If strict inequality holds in* (4.2.5), *there is a* Δ*-linear isomorphism* $D/pD \cong \mathbb{D}$ *under which*

$$\begin{aligned} Fe_0 &= 0, & Ff_0 &= f_1, & Fe_1 &= e_0, & Ff_1 &= 0, \\ Ve_0 &= 0, & Vf_0 &= f_1, & Ve_1 &= e_0, & Vf_1 &= 0. \end{aligned}$$

(2) *If equality holds in* (4.2.5), *there is a* Δ*-linear isomorphism* $D/pD \cong \mathbb{D}$ *under which*

$$\begin{aligned} Fe_0 &= 0, & Ff_0 &= u^p f_1, & Fe_1 &= u^p e_0, & Ff_1 &= -u^p e_0, \\ Ve_0 &= 0, & Vf_0 &= e_1 + f_1, & Ve_1 &= e_0, & Vf_1 &= 0. \end{aligned} \tag{4.2.6}$$

where u *is the image of* $-p/\tau_1^{p+1} = -\tau^{p+1}/p^p$ *under* $\mathcal{O}^\times \to k^\times$.

Proof. Recalling (4.1.1), the inequality (4.2.5) is equivalent to

$$\operatorname{ord}(\tau_1) \le \frac{1}{p+1}.$$

Using this observation, the proof is identical to that of Theorem 4.2.2, but with the indices 0 and 1 reversed everywhere. ☐

Corollary 4.2.4. *The* p*-divisible group* G_k *is superspecial if and only if*

$$\operatorname{ord}(\tau) \notin \left\{ \frac{1}{p+1}, \frac{p}{p+1} \right\}.$$

Moreover, the superspecial locus of $X(C)$ *is a union of three Ekedahl–Oort strata, characterized as follows*:

(1) *The subset of* $X(C)$ *defined by*

$$\frac{1}{p+1} < \operatorname{ord}(\tau) < \frac{p}{p+1}$$

is an Ekedahl–Oort stratum. On this stratum $\Pi\Omega(G_k^\vee) = 0$.

(2) *The subset of* $X(C)$ *defined by*

$$\operatorname{ord}(\tau) < \frac{1}{p+1},$$

is an Ekedahl–Oort stratum. On this stratum $\Pi\Omega_1(G_k^\vee) \neq 0$.

(3) *The subset of* $X(C)$ *defined by*

$$\operatorname{ord}(\tau) > \frac{p}{p+1}.$$

is an Ekedahl–Oort stratum. On this stratum $\Pi\Omega_0(G_k^\vee) \neq 0$.

Proof. Recall from Proposition 3.2.3 that G_k is superspecial if and only if V^2 annihilates D/pD. Given this, all parts of the claim are clear from Theorems 4.2.1, 4.2.2, and 4.2.3. ☐

Now consider the locus of points

$$\left\{\tau \in C : \mathrm{ord}(\tau) = \frac{1}{p+1}\right\} \cup \left\{\tau \in C : \mathrm{ord}(\tau) = \frac{p}{p+1}\right\} \subset X(C)$$

at which the corresponding p-divisible group does not have superspecial reduction. This set is a union of infinitely many Ekedahl–Oort strata.

Corollary 4.2.5. *The fibers of the composition*

$$\left\{\tau \in C : \mathrm{ord}(\tau) = \frac{1}{p+1}\right\} \xrightarrow{\tau \mapsto p/\tau^{p+1}} \mathcal{O}^\times \to k^\times$$

are Ekedahl–Oort strata, *as are the fibers of the composition*

$$\left\{\tau \in C : \mathrm{ord}(\tau) = \frac{p}{p+1}\right\} \xrightarrow{\tau \mapsto \tau^{p+1}/p^p} \mathcal{O}^\times \to k^\times.$$

Proof. For each $u \in k^\times$ let F_u and V_u be the operators on $\mathbb{D}$ defined by (4.2.4). Note that V_u is actually independent of u. We claim that the existence of a Δ-linear isomorphism

$$(\mathbb{D}, F_u, V_u) \xrightarrow{\phi} (\mathbb{D}, F_{u'}, V_{u'})$$

implies $u = u'$. To see this one checks that the first relation in

$$\phi \circ V_u = V_{u'} \circ \phi, \quad \phi \circ F_u = F_{u'} \circ \phi \tag{4.2.7}$$

implies that ϕ has the form

$$\phi(e_0) = ae_0, \quad \phi(e_1) = ae_1, \quad \phi(f_0) = af_0, \quad \phi(f_1) = af_1 + be_1$$

for some $a \in \mathbb{F}_p$ and $b \in k$. Using this, one checks that ϕ commutes with both F_u and $F_{u'}$. The second relation in (4.2.7) then implies that $F_u = F_{u'}$, and hence $u = u'$.

The same is true if we replace the operators of (4.2.4) with those of (4.2.6), and so the corollary follows from Theorems 4.2.2 and 4.2.3. □

Acknowledgements

The author thanks Keerthi Madapusi Pera for helpful conversations, and the anonymous referees for numerous suggestions and corrections.

References

[Andreatta and Goren 2003] F. Andreatta and E. Z. Goren, "Geometry of Hilbert modular varieties over totally ramified primes", *Int. Math. Res. Not.* **2003**:33 (2003), 1786–1835. MR Zbl

[Caraiani and Scholze 2017] A. Caraiani and P. Scholze, "On the generic part of the cohomology of compact unitary Shimura varieties", *Ann. of Math.* (2) **186**:3 (2017), 649–766. MR Zbl

[Fargues 2015] L. Fargues, "Quelques résultats et conjectures concernant la courbe", pp. 325–374 in *De la géométrie algébrique aux formes automorphes, I*, edited by J.-B. Bost et al., Astérisque **369**, Soc. Math. France, Paris, 2015. MR Zbl

[Kudla and Rapoport 2000] S. S. Kudla and M. Rapoport, "Height pairings on Shimura curves and p-adic uniformization", *Invent. Math.* **142**:1 (2000), 153–223. MR Zbl

[Lau 2010] E. Lau, "Frames and finite group schemes over complete regular local rings", *Doc. Math.* **15** (2010), 545–569. MR Zbl

[Lau 2018] E. Lau, "Dieudonné theory over semiperfect rings and perfectoid rings", *Compos. Math.* **154**:9 (2018), 1974–2004. MR Zbl

[Lau 2019] E. Lau, "Displayed equations for Galois representations", *Nagoya Math. J.* **235** (2019), 86–114. MR

[Moonen 2004] B. Moonen, "A dimension formula for Ekedahl–Oort strata", *Ann. Inst. Fourier* (*Grenoble*) **54**:3 (2004), 666–698. MR Zbl

[Oort 1975] F. Oort, "Which abelian surfaces are products of elliptic curves?", *Math. Ann.* **214** (1975), 35–47. MR Zbl

[Oort 2001] F. Oort, "A stratification of a moduli space of abelian varieties", pp. 345–416 in *Moduli of abelian varieties* (Texel, Netherlands, 1999), edited by C. Faber et al., Progr. Math. **195**, Birkhäuser, Basel, 2001. MR Zbl

[Scholze and Weinstein 2013] P. Scholze and J. Weinstein, "Moduli of p-divisible groups", *Camb. J. Math.* **1**:2 (2013), 145–237. MR Zbl

[Scholze and Weinstein 2020] P. Scholze and J. Weinstein, *Berkeley lectures on p-adic geometry*, Ann. of Math. Studies **207**, Princeton Univ. Press, 2020. Zbl

[Viehmann and Wedhorn 2013] E. Viehmann and T. Wedhorn, "Ekedahl–Oort and Newton strata for Shimura varieties of PEL type", *Math. Ann.* **356**:4 (2013), 1493–1550. MR Zbl

[Vignéras 1980] M.-F. Vignéras, *Arithmétique des algèbres de quaternions*, Lecture Notes in Math. **800**, Springer, 1980. MR Zbl

[Zhang 2018] C. Zhang, "Ekedahl–Oort strata for good reductions of Shimura varieties of Hodge type", *Canad. J. Math.* **70**:2 (2018), 451–480. MR Zbl

[Zink 2001] T. Zink, "Windows for displays of p-divisible groups", pp. 491–518 in *Moduli of abelian varieties* (Texel, Netherlands, 1999), edited by C. Faber et al., Progr. Math. **195**, Birkhäuser, Basel, 2001. MR Zbl

Communicated by Michael Rapoport
Received 2019-02-28 Revised 2020-01-03 Accepted 2020-02-06

howardbe@bc.edu *Department of Mathematics, Boston College, Chestnut Hill, MA, United States*

ALGEBRA AND NUMBER THEORY 14:4 (2020)
https://doi.org/10.2140/ant.2020.14.991

A moving lemma for relative 0-cycles

Amalendu Krishna and Jinhyun Park

We prove a moving lemma for the additive and ordinary higher Chow groups of relative 0-cycles of regular semilocal k-schemes essentially of finite type over an infinite perfect field. From this, we show that the cycle classes can be represented by cycles that possess certain finiteness, surjectivity, and smoothness properties. It plays a key role in showing that the crystalline cohomology of smooth varieties can be expressed in terms of algebraic cycles.

1. Introduction

Just as the classical Chow moving lemma played a fundamental role in studies of Chow groups of smooth algebraic varieties over a field, the moving lemma of Bloch [1986; 1994] played a significant role in studies of higher Chow groups of smooth algebraic varieties, i.e., the motivic cohomology. One limitation of those moving lemmas however is that they focus only on the proper intersection properties of the given cycles. Occasionally, the given circumstances require us to know more about the cycles beyond such proper intersection properties. For instance, we often need to know whether the given cycles are finite over the base scheme, and smooth, or, if not, whether they can be moved to such cycles. Such questions require more subtle treatments and may hold under special circumstances only.

The goal of this article is to prove a moving lemma of this sort for higher relative 0-cycles of a regular semilocal scheme essentially of finite type over an infinite perfect field k. Here, "essentially of finite type" means it is obtained by localizing a quasiprojective k-scheme at a finite set Σ of points. Achieving suitable finiteness and regularity of the cycles is the main characteristic of the moving lemma we seek.

MSC2010: primary 14C25; secondary 14F42, 19E15.
Keywords: algebraic cycles, moving lemma, higher Chow group, additive higher Chow group, linear projection, Grassmannian.

In the introduction, we state the main results, explain the motivation, and give an outline of the article.

1A. *The* sfs-*moving lemma.* Let k be an infinite perfect field. Let R be a regular semilocal k-algebra essentially of finite type. Let $V = \mathrm{Spec}(R)$ and let Σ denote the set of closed points of V. Let $\mathrm{Tz}^q(V, \bullet; m)$ be the nondegenerate additive cycle complex of V in codimension $q \geq 1$ and with modulus $m \geq 1$. Let $\mathrm{TCH}^q(V, n; m)$ denote the associated homology groups, called the additive higher Chow groups of V (see Section 2A).

For $n \geq 1$, let $\mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$ denote the subgroup of sfs-cycles in $\mathrm{Tz}^n(V, n; m)$ (see Section 2E). Roughly speaking, an sfs-cycle is an element $\alpha \in \mathrm{Tz}^n(V, n; m)$ such that every irreducible component of α intersects $\Sigma \times F$ properly for every face $F \subset \square_k^{n-1}$, is finite and surjective over an irreducible component of V, and the image under every projection $V \times \square_k^{n-1} \to V \times \square_k^j$ $(0 \leq j \leq n-1)$ is a regular scheme. Those cycles have the trivial boundaries (see Lemma 2.21). Let $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m)$ denote the image of the canonical map $\mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n(V, n; m)$ (see Section 2F). The goal of this article is to prove the following result.

Theorem 1.1 (the sfs-moving lemma). *Let k be an infinite perfect field. Let $m, n \geq 1$ be integers. Let V be a smooth semilocal k-scheme essentially of finite type. Then the canonical map* $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n(V, n; m)$ *is an isomorphism.*

For the same V as above, let $z^q(V, \bullet)$ denote the cubical version of Bloch's cycle complex (see [Krishna and Levine 2008, Section 1]) and let $\mathrm{CH}^q(V, n)$ denote the associated higher Chow groups. We can define the subgroup $z^n_{\mathrm{sfs}}(V, n)$ of sfs-cycles and the higher Chow group $\mathrm{CH}^n_{\mathrm{sfs}}(V, n)$ of sfs-cycles analogous to the additive higher Chow group of sfs-cycles. There is a canonical map $\mathrm{CH}^n_{\mathrm{sfs}}(V, n) \to \mathrm{CH}^n(V, n)$. As a byproduct of the discussions toward the proof of Theorem 1.1, we can recover the following result, stated in [Elbaz-Vincent and Müller-Stach 2002].[1]

Theorem 1.2. *Let k be an infinite perfect field. Let $V = \mathrm{Spec}(R)$ be a smooth semilocal k-scheme essentially of finite type. Let $n \geq 1$ be an integer. Then the canonical map* $\mathrm{CH}^n_{\mathrm{sfs}}(V, n) \to \mathrm{CH}^n(V, n)$ *is an isomorphism.*

Theorem 1.1 provides the main geometric ground for the proof of the following result and a few of its consequences in the paper [Krishna and Park 2015a], discussed separately due to the huge size and complexities of the proofs of the current article. In particular, it allows one to describe the crystalline cohomology of a smooth scheme in positive characteristic in terms of algebraic cycles.

Theorem 1.3 [Krishna and Park 2015a]. *Let k be any field and let R be a smooth semilocal k-algebra essentially of finite type. Let $m, n \geq 1$ be integers. Then there is a natural isomorphism*

$$\tau_R : \mathbb{W}_m\Omega_R^{n-1} \xrightarrow{\cong} \mathrm{TCH}^n(R, n; m),$$

where $\mathbb{W}_m\Omega_R^\bullet$ is the big de Rham–Witt complex of Hesselholt and Madsen.

[1] In [Elbaz-Vincent and Müller-Stach 2002, Lemma 3.11], Theorem 1.2 is claimed for arbitrary fields, but we do not know if this can be achieved using the techniques of linear projections.

When R is a field, this was first proven by Rülling [2007]; the above is a higher dimensional generalization, but it also relies on the theorem of Rülling.

1B. ***The presentation lemma.*** We deduce Theorem 1.1 from the following general presentation lemma for residual cycles of linear projections. This has the flavor (hence the name) of Gabber's geometric presentation lemma (see [Colliot-Thélène et al. 1997]). Of course, our assertions are different and intricate.

Let k be an infinite perfect field. Given a finite map $h : Y' \to Y$ of k-schemes and a reduced closed subscheme $Z \subset Y'$, let $h^{+}(Z)$ be the closure of $h^{-1}(h(Z)) \setminus Z$ in Y' with the reduced induced closed subscheme structure. We call this the "residual scheme of Z" with respect to h.

Let $n \geq 1$ and let $\hat{A}_0, \dots, \hat{A}_{n-1}$ be smooth projective and geometrically integral k-schemes of positive dimensions. For $0 \leq j \leq n-1$, let $A_j \subset \hat{A}_j$ be a nonempty affine open subset. Set $C_0 := \operatorname{Spec}(k)$ and $C_j := \prod_{i=0}^{j-1} A_i$ for $j \geq 1$. Let $\pi_j : C_n \to C_j$ be the obvious projection. For any map $f : Y' \to Y$, let $f_j : Y' \times C_j \to Y \times C_j$ be the map $f \times \mathrm{id}_{C_j}$.

Let $\overline{X} \subset \mathbb{P}^m_k$ be a reduced closed subscheme of pure dimension $r \geq 1$ and let $X \subset \overline{X}$ be the complement of a hyperplane in $\mathbb{P}^m_k$ such that X is regular and integral. Let $\Sigma \subset X$ be a finite set of closed points. Let $Z \subset X \times C_n$ be an integral closed subscheme of dimension r such that the projection $Z \to C_n$ is not constant, and the projection $Z \to X$ is finite and surjective.

The presentation lemma for the residual schemes that we prove is the following.

Theorem 1.4. *Let k be an infinite perfect field. There exist a closed embedding $\overline{X} \hookrightarrow \mathbb{P}^N_k$, a hyperplane $H \subset \mathbb{P}^N_k$ with $X = \overline{X} \setminus H$, and a dense open subset $\mathcal{U} \subset \mathrm{Gr}(N-r-1, H)$ of the Grassmannian variety such that for each $L \in \mathcal{U}(k)$, the linear projection $\phi_L : \mathbb{P}^N_k \setminus L \to \mathbb{P}^r_k$ away from L defines a finite surjective morphism $\phi : \overline{X} \to \mathbb{P}^r_k$ such that the following hold*:

(1) *There exists a Cartesian square*:

$$\begin{array}{ccc} X & \hookrightarrow & \overline{X} \\ \downarrow{\scriptstyle\phi} & & \downarrow{\scriptstyle\phi} \\ \mathbb{A}^r_k & \hookrightarrow & \mathbb{P}^r_k \end{array}$$

(2) *ϕ is étale over an affine open neighborhood of $\phi(\Sigma)$.*

(3) *$\phi(x) \neq \phi(x')$ for each pair $x \neq x'$ of points in Σ.*

(4) *The map $k(\phi(x)) \to k(x)$ is an isomorphism for each $x \in \Sigma$.*

(5) *The induced map $Z \to \phi_n(Z)$ is birational.*

(6) *The map $\phi_n^{+}(Z) \to X$ is finite and surjective.*

(7) *$\pi_j(\phi_n^{+}(Z))$ is regular at all points lying over Σ for each $0 \leq j \leq n$.*

1C. ***Outline of proofs and remarks.*** We first remark that although V may be in general obtained by localizing a quasiprojective k-scheme at a finite set Σ of not necessarily closed points, for the proof of the sfs-moving lemma, we can easily reduce to the case of closed points. See Proposition 2.19. Then the proof the sfs-moving lemma can be broadly divided into two parts.

In the first part, we prove it when the underlying semilocal ring is the localization of an affine space $\mathbb{A}^r_k$ at a finite set of closed points. To solve this case, we rely on two key ingredients: the lemma of Bloch [1986, Lemma 1.2] and the moving lemma for cycles with modulus on affine spaces by Kai [2019]. (N.B., Part of what we need in this article from [loc. cit.] is also available in [Krishna and Park 2016].) The moving lemma of Kai allows us to ensure that our cycles can be made to intersect the closed points of the semilocal scheme V properly. After this, we apply an "spread out and specialize" type of argument using [Bloch 1986, Lemma 1.2] to achieve our goal.

Roughly speaking, we argue that we can equip the sfs-property to cycles after moving them via a certain kind of twisted translations by a general set of k-rational points of $\mathbb{A}^r_k$. This requires us to use that the ground field k is infinite. The rest of the argument is to construct a homotopy between the new and the original cycle. The plain translations by the rational points do not work and the twisted translations make the argument more involved than the classical case. This is done in Section 3.

In the second part, we prove the general case of the sfs-moving lemma by combining the affine space case and the presentation lemma (Theorem 1.4). The proof of the presentation lemma is an intricate application of the method of linear projections and moduli in algebraic geometry.

The reason for this intricacy lies in the fact that it is not sufficient for us to find enough linear projections which give finite and flat morphisms from a projective variety X to projective spaces. We need to invoke a more delicate linear projection in such a way that if we project a subvariety in some smooth family over X to a similar family over the projective space, the resulting residual scheme has certain desired geometric properties, e.g., regularity along a given set of fibers in the family. Even more, we need to ensure that if we project this smooth family over X to a smaller dimensional family via proper maps, then the images of the residual scheme continue to enjoy the good properties.

Showing that one can find enough such linear projections that do the above jobs lies at the heart of the argument. We see that the moduli spaces of linear subspaces that we encounter in the process are all rational, and we find enough rational lines in them. We then reduce the argument to studies of a family of linear subspaces parametrized by a rational line (pencil of linear subspaces). This simplifies the problem.

Along the proofs, we need to separate the cases of algebraically closed and general infinite perfect fields. We first prove the results over algebraically closed fields. Over a general infinite perfect field k, we argue that we can find enough linear subspaces after going to an algebraic closure $\bar{k}$ so that all desired properties are achieved (over $\bar{k}$) in such a generality that they remain to be satisfied for the original cycle over k after descent. One of these generalities we ensure over $\bar{k}$ is that the whole residual *scheme* is regular, and not just its irreducible components (even if the latter case suffices for the sfs-moving lemma). We then show that there are enough such linear subspaces defined over k. This is achieved using a Galois descent.

Carrying out this program rigorously takes up from Section 4 to Section 7. We combine them to prove the main results in Section 8.

We now make some remarks on our assumption on the ground field. We need k to be infinite to ensure that our moduli spaces have enough k-rational points. We need it to be perfect to achieve the regularity of various residual subvarieties. Although we only need the regularity of cycles, our argument at some stage uses the condition that some regular schemes that we encounter in the middle are actually smooth over k (e.g., see the last part of the proof of Proposition 7.8). The perfectness requirement is evident even in the proof of the sfs-moving lemma in affine space, where we need to use a specialization argument. To make sure that we do not destroy the regularity during the specialization, we need our over-field to be separably generated over k (e.g., see the proof of Lemma 3.11). This requires k to be perfect.

Recall that the moving lemma of Bloch and Chow hold over all fields. One proves this for infinite perfect fields first. The case of finite field reduces to the case of infinite perfect fields using the techniques of pro-ℓ-extensions and the push-pull operators on the Chow groups. However, we cannot use this technique in our case because the smoothness property of the sfs-cycles are not well-behaved under the push-forward operators. However, based on Theorem 1.1, we prove Theorem 1.3 in [Krishna and Park 2015a] over all base fields with different methods.

Finally, the reader may notice that our sfs-moving lemma is stated and proven in this paper for

$$\mathrm{TCH}^n(V, n; m) \quad \text{for } m \geq 1.$$

However, we remark that one does not miss out on anything by this assumption because it is shown in [Krishna and Park 2017, Theorem 1.5] that

$$\mathrm{TCH}^n(V, n; 0) = 0.$$

In particular, $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; 0) = 0$.

The main result of this article plays essential roles in [Gupta and Krishna 2019b; 2019a; Krishna and Park 2015a]. Apart from these applications, we hope that our presentation lemma through linear projection techniques as well as various results and ideas of manipulating locally closed subsets of the Grassmannian will be useful in the future to anyone in the mathematics community (in particular, those working with algebraic cycles) who uses the linear projection machines in the tool box.

1D. ***Conventions.*** Unless we specify otherwise, k is a fixed field. We shall assume later that k is infinite and perfect for our main results. A k-scheme is a separated scheme of finite type over k. An affine k-scheme is a k-scheme which is affine. A k-variety is an equidimensional reduced k-scheme. The product $X \times Y$ means $X \times_k Y$, unless we specify otherwise. We let $\mathbf{Sch}_k$ be the category of k-schemes and $\mathbf{Sm}_k$ of smooth k-schemes. A scheme essentially of finite type is a scheme obtained by localizing at a finite subset of (not necessarily closed) points of a quasiprojective subscheme of a finite type k-scheme. We include the case of not localizing at all. For $\mathcal{C} = \mathbf{Sch}_k, \mathbf{Sm}_k$, we let $\mathcal{C}^{\mathrm{ess}}$ be the extension of the category $\mathcal{C}$, whose objects are either those in $\mathcal{C}$ or those obtained by localizing an object of $\mathcal{C}$ at a finite subset.

Given $X \in \mathcal{C}^{\mathrm{ess}}$ and a finite set of points $\Sigma \subset X$, we write X_Σ for the localization of X along Σ. If $Y \subset X$ is an inclusion of a reduced locally closed subscheme, then the closure of Y is considered a closed subscheme of X with the reduced induced structure. The image of a reduced closed subset under a proper map is considered a closed subscheme of the target scheme with the reduced induced structure.

2. The fs and sfs-cycles

After recalling the definition of higher Chow groups and additive higher Chow groups, we define our main objects of study: the fs and sfs-cycles. We prove some preliminary results about these cycles.

2A. ***Higher Chow groups and additive higher Chow groups.*** Let k be a field. First recall (see [Bloch 1986]) the definition of higher Chow groups. Let $X \in \mathbf{Sch}_k^{\mathrm{ess}}$ be equidimensional. Let $\mathbb{P}_k^1 = \operatorname{Proj} k[Y_0, Y_1]$, and $\square^n = (\mathbb{P}_k^1 \setminus \{1\})^n$. Let $(y_1, \dots, y_n) \in \square^n$ be the coordinates. A *face* of $\square^n$ is a closed subscheme defined by a set of equations $\{y_{i_1} = \epsilon_1, \dots, y_{i_s} = \epsilon_s\}$, where $\epsilon_j \in \{0, \infty\}$. For $1 \le i \le n$ and $\epsilon = 0, \infty$, let $\iota_i^\epsilon : \square^{n-1} \to \square^n$ be the inclusion given by $(y_1, \dots, y_{n-1}) \mapsto (y_1, \dots, y_{i-1}, \epsilon, y_i, \dots, y_{n-1})$. Its image gives a codimension 1 face.

Let $q, n \ge 0$. When X is obtained by localizing at a nonclosed point, for closed subschemes in $X \times \square^n$, the notion of dimensions could be ambiguous but the codimensions are well-defined. So, we use dimensions only when there is no ambiguity.

Let $\underline{z}^q(X, n)$ be the free abelian group on the set of integral closed subschemes of $X \times \square^n$ of codimension q, that intersect properly with $X \times F$ for each face F of $\square^n$. We define the boundary map $\partial_i^\epsilon(Z) := [(\mathrm{Id}_X \times \iota_i^\epsilon)^*(Z)]$. This collection of data gives a cubical abelian group $(\underline{n} \mapsto \underline{z}^q(X, n))$ in the sense of [Krishna and Levine 2008, Section 1.1], and the groups $z^q(X, n) := \underline{z}^q(X, n)/\underline{z}^q(X, n)_{\mathrm{degn}}$ (in the notations of [loc. cit.]) give a complex of abelian groups, whose boundary map at level n is given by $\partial := \sum_{i=1}^n (-1)^i (\partial_i^\infty - \partial_i^0)$. The homology $\mathrm{CH}^q(X, n) := \mathrm{H}_n(z^q(X, \bullet), \partial)$ is called the higher Chow group of X.

We recall the definition of additive higher Chow groups from [Krishna and Park 2015b, Section 2] (see also [Park 2009]). Let $X \in \mathbf{Sch}_k^{\mathrm{ess}}$ be equidimensional. Let $\mathbb{A}_k^1 = \operatorname{Spec} k[t]$, $\mathbb{G}_m = \operatorname{Spec} k[t, t^{-1}]$, and $\overline{\square} = \mathbb{P}_k^1$. For $n \ge 1$, let $B_n = \mathbb{A}_k^1 \times \square^{n-1}$, $\overline{B}_n = \mathbb{A}_k^1 \times \overline{\square}^{n-1}$ and $\widehat{B}_n = \mathbb{P}_k^1 \times \overline{\square}^{n-1} \supset \overline{B}_n$. Let $(t, y_1, \dots, y_{n-1}) \in \overline{B}_n$ be the coordinates.

On $\overline{B}_n$, define the Cartier divisors $F_{n,i}^1 := \{y_i = 1\}$ for $1 \le i \le n-1$, $F_{n,0} := \{t = 0\}$, and let $F_n^1 := \sum_{i=1}^{n-1} F_{n,i}^1$. A *face* of B_n is a closed subscheme defined by a set of equations of the form $y_{i_1} = \epsilon_1, \dots, y_{i_s} = \epsilon_s$, where $\epsilon_j \in \{0, \infty\}$. For $1 \le i \le n-1$ and $\epsilon = 0, \infty$, let $\iota_{n,i}^\epsilon : B_{n-1} \to B_n$ be the inclusion $(t, y_1, \dots, y_{n-2}) \mapsto (t, y_1, \dots, y_{i-1}, \epsilon, y_i, \dots, y_{n-2})$. Its image is a codimension 1 face.

The additive higher Chow complex is defined similarly using the spaces B_n instead of $\square^n$, but together with proper intersections with all faces, we impose additional conditions called the *modulus conditions*, that control how the cycles should behave at "infinity": (see [Krishna and Park 2015b, Definition 2.1]) let X be a k-scheme, and let V be an integral closed subscheme of $X \times B_n$. Let $\overline{V}$ denote the Zariski closure of V in $X \times \overline{B}_n$ and let $\nu \colon \overline{V}^N \to \overline{V} \subset X \times \overline{B}_n$ be the normalization of $\overline{V}$. Let $m, n \ge 1$ be integers. We say that V satisfies the *modulus m condition* on $X \times B_n$, if as Weil divisors on $\overline{V}^N$ we have $(m+1)[\nu^*(F_{n,0})] \le [\nu^*(F_n^1)]$. When $n = 1$, we have $F_1^1 = \varnothing$, so it means $\nu^*(F_{1,0}) = 0$, or $\{t = 0\} \cap \overline{V} = \varnothing$.

If V is a cycle on $X \times B_n$, we say that V satisfies the modulus m condition if each of its irreducible components satisfies the modulus m condition. When m is understood, often we just say that V satisfies the modulus condition. Note that since $F_{n,0} = \{t = 0\} \subset \overline{B}_n$, replacing $\overline{B}_n$ by $\hat{B}_n$ in the definition does not change the nature of the modulus condition on V.

For an equidimensional $X \in \mathbf{Sch}_k^{\mathrm{ess}}$, and integers $m, n, q \geq 1$, we first define $\underline{\mathrm{Tz}}^q(X, 1; m)$ to be the free abelian group on integral closed subschemes Z of $X \times \mathbb{A}^1$ of codimension q, satisfying the modulus condition (see [Krishna and Park 2015b, Definition 2.5]). For $n > 1$, $\underline{\mathrm{Tz}}^q(X, n; m)$ is the free abelian group on integral closed subschemes Z of $X \times B_n$ of codimension q such that for each face F of B_n, Z intersects $X \times F$ properly on $X \times B_n$, and Z satisfies the modulus m condition on $X \times B_n$. For each $1 \leq i \leq n-1$ and $\epsilon = 0, \infty$, let $\partial_i^\epsilon(Z) := [(\mathrm{Id}_X \times \iota_{n,i}^\epsilon)^*(Z)]$. The proper intersection with faces ensures that $\partial_i^\epsilon(Z)$ are well-defined. The cycles in $\underline{\mathrm{Tz}}^q(X, n; m)$ are called the *admissible cycles* (or, often as *additive higher Chow cycles, or additive cycles*).

This gives the cubical abelian group $(\underline{n} \mapsto \underline{\mathrm{Tz}}^q(X, n+1; m))$ in the sense of [Krishna and Levine 2008, Section 1.1]. Using the containment lemma [Krishna and Park 2012, Proposition 2.4], that each face $\partial_i^\epsilon(Z)$ lies in $\underline{\mathrm{Tz}}^q(X, n-1; m)$ is implied from the defining conditions.

For a cycle $\sum_{i=1}^s n_i Z_i$,we let $|\alpha|$ be the closed subscheme $\bigcup_{i=1}^s Z_i$ with its reduced structure. This is called the support of α. If $f : Y \to X$ is flat and $\alpha \in \underline{\mathrm{Tz}}^q(X, n; m)$, we write $f^*(\alpha)$ often as α_Y. This shorthand is more evident when f is a localization morphism.

Definition 2.1 [Krishna and Park 2015b, Definition 2.6]. Let $X \in \mathbf{Sch}_k^{\mathrm{ess}}$ be equidimensional. The *additive higher Chow complex*, or just the *additive cycle complex*, $\mathrm{Tz}^q(X, \bullet; m)$ of X in codimension q with modulus m is the nondegenerate complex associated to the cubical abelian group $(\underline{n} \mapsto \underline{\mathrm{Tz}}^q(X, n+1; m))$, i.e., $\mathrm{Tz}^q(X, n; m)$ is the quotient $\underline{\mathrm{Tz}}^q(X, n; m)/\underline{\mathrm{Tz}}^q(X, n; m)_{\mathrm{degn}}$.

The boundary map of this complex at level n is given by $\partial := \sum_{i=1}^{n-1}(-1)^i(\partial_i^\infty - \partial_i^0)$, and it satisfies $\partial^2 = 0$. The homology $\mathrm{TCH}^q(X, n; m) := \mathrm{H}_n(\mathrm{Tz}^q(X, \bullet; m))$ for $n \geq 1$ is the *additive higher Chow group* of X with modulus m.

2B. ***Subcomplexes associated to some algebraic subsets.*** Let $X \in \mathbf{Sch}_k^{\mathrm{ess}}$ be a variety. Here are some subgroups of $\mathrm{Tz}^q(X, n; m)$ with a finer intersection property with a given finite set $\mathcal{W}$ of locally closed algebraic subsets of X:

Definition 2.2 [Krishna and Park 2012, Definition 4.2]. Define $\underline{\mathrm{Tz}}^q_{\mathcal{W}}(X, n; m)$ to be the subgroup of $\underline{\mathrm{Tz}}^q(X, n; m)$ generated by integral closed subschemes $Z \subset X \times B_n$ that additionally satisfy

$$\mathrm{codim}_{W \times F}(Z \cap (W \times F)) \geq q \text{ for all } W \in \mathcal{W} \text{ and all faces } F \subset B_n. \tag{2-1}$$

The groups $\underline{\mathrm{Tz}}^q_{\mathcal{W}}(X, n+1; m)$ for $n \geq 0$ form a cubical subgroup of $(\underline{n} \mapsto \underline{\mathrm{Tz}}^q(X, n+1; m))$ and they give the subcomplex $\mathrm{Tz}^q_{\mathcal{W}}(X, \bullet; m) \subset \mathrm{Tz}^q(X, \bullet; m)$ by modding out by the degenerate cycles. The homology groups are denoted by $\mathrm{TCH}^q_{\mathcal{W}}(X, n; m)$.

2C. ***Schemes with finite closed points.*** Recall that (see [Gabber et al. 2013, Section 2.2]) we say a scheme X is an FA-scheme if for any finite subset $\Sigma \subset X$, there exists an *affine* open neighborhood $U \subset X$ of Σ. We have the following [loc. cit.]:

Lemma 2.3. *Any quasiprojective k-scheme is FA. Any open subset of an FA-scheme is FA. Given any finite subset Σ of a quasiprojective k-scheme, and an open subset $U \subset X$ containing Σ, there exists an affine open neighborhood $W \subset U$ of Σ.*

Recall (Section 1D) that a semilocal k-scheme V is essentially of finite type if there is a quasiprojective k-scheme whose localization at a finite subset Σ of points gives V. By Lemma 2.3, we may obtain it by localizing an *affine* k-scheme of finite type.

Definition 2.4. For any semilocal k-scheme V essentially of finite type, a pair (X, Σ) consisting of an affine k-scheme X of finite type and a finite set Σ of points such that $V = \mathrm{Spec}(\mathcal{O}_{X,\Sigma})$, is called an *atlas* for V. A smooth (resp. regular) atlas (X, Σ) is an atlas such that X is smooth over k (resp. regular).

Lemma 2.5. *Let $V = \mathrm{Spec}(R)$ be a semilocal k-scheme obtained by localizing at a finite set Σ of points of a quasiprojective k-variety X. For a cycle α on $V \times B_n$, let $\bar{\alpha}$ be its Zariski closure in $X \times B_n$.*

Then $\alpha \in \mathrm{Tz}^q_\Sigma(V, n; m)$ if and only if there exists an affine open neighborhood $U \subset X$ of Σ such that $\bar{\alpha}_U \in \mathrm{Tz}^q_\Sigma(U, n; m)$.

Here, if $\partial(\alpha) = 0$, then we can assume that $\partial(\bar{\alpha}_U) = 0$. If α is a boundary, then we can assume $\bar{\alpha}_U$ is also a boundary. If V is smooth over k, then we may take (U, Σ) to be a smooth atlas.

Proof. The first three assertions were proven in [Krishna and Park 2016, Lemmas 4.13 and 4.14]. For the last one, choose any X of finite type using the first assertion. Since V is smooth, we have $X_{\mathrm{sing}} \cap V = \varnothing$ and $X_{\mathrm{sm}} = X \setminus X_{\mathrm{sing}} \supset \Sigma$. By Lemma 2.3, we can choose an affine open $U \subset X_{\mathrm{sm}}$ containing Σ. □

2D. ***The* fs*-cycles.*** Recall that for higher Chow groups of a semilocal k-scheme V in the Milnor range, [Elbaz-Vincent and Müller-Stach 2002, Lemma 3.11] used the notions called fs-cycles and sfs-cycles. An fs-cycle in [loc. cit.] is a cycle α on $V \times \square^n_k$ such that for each irreducible component Z, the morphism $Z \to V$ is finite and surjective. However, a moment's thought gives that it is not a good notion. For instance, if V is reducible, then one can almost never achieve the surjection part.

Even if we modify the definition a bit by requiring instead that the support $|\alpha| \to V$ is finite and surjective, still there is a problem when V is not irreducible: suppose $V = V_1 \cup V_2$ is a disjoint union of irreducible components. Suppose for $i = 1, 2$, we have an irreducible closed subscheme Z_i on $V \times \square^n_k$ such that $Z_i \to V_i$ is finite surjective. Then $W := Z_1 + Z_2$ and $W' := Z_1 + 2Z_2$ are both fs-cycles in this updated sense. But, then $W' - W = Z_2$ is still finite over V, while it is no longer surjective over V. As a result the set of fs-cycles in the above sense is not even closed under basic summation of cycles, thus they do not form a group.

The natural notion to work with is the following:

Definition 2.6. Let $X, Y \in \mathbf{Sch}_k^{\text{ess}}$. First suppose that Y is irreducible. In this case, we say that a morphism $Y \to X$ of k-schemes is *fs over* X (or an *fs-morphism*, or simply *fs* when X is understood) if it is finite and it is surjective to an irreducible component of X.

In case Y is not necessarily irreducible, we say $Y \to X$ is *fs over* X if for each irreducible component $Y_j \subset Y$, the induced map $Y_j \to X$ is fs over X.

We generalize it further: let $f : Y \to X$ be a morphism in $\mathbf{Sch}_k^{\text{ess}}$ and let $U \to X$ be a flat morphism. We say that $Y \to X$ is *fs over* U, if the fiber product $f' : Y \times_X U \to U$ is fs.

This notion coincides with the naïve notion mentioned above when X is irreducible. Unlike the naïve notion, this notion of fs-morphisms behaves well under base changes.

Lemma 2.7. *Let $f : Y \to X$ be an fs morphism in $\mathbf{Sch}_k^{\text{ess}}$. Let $U \to X$ be a flat morphism in $\mathbf{Sch}_k^{\text{ess}}$. Then the fiber product $f' : Y \times_X U \to U$ is fs.*

Proof. That the base change of a finite morphism is again finite is apparent. The remaining part on surjectivity over an irreducible component follows by [EGA IV$_2$ 1965, Proposition (2.3.7)(ii), page 16], where the dominance there is equivalent to surjectivity under finiteness. □

Lemma 2.8. *Let Z be a cycle on $Y \times B$ such that Z is fs over Y in the sense that each irreducible component of Z is fs over Y.*

Let $f : Y \to X$ be a finite surjective morphism in $\mathbf{Sch}_k^{\text{ess}}$ of irreducible schemes. Then the finite push-forward $f_*(Z)$ on $X \times B$ is fs over X.

Proof. We may assume Z is irreducible. Since $Z \to Y$ is finite surjective and $Y \to X$ is finite surjective, the composite $Z \to Y \to X$ is finite surjective. □

Here is one simple criterion on finiteness

Lemma 2.9 (finiteness criterion). *Let X be an equidimensional affine k-scheme essentially of finite type. Let $\hat{B}$ be a smooth projective geometrically integral k-scheme of finite type of dimension $n > 0$ and let $B \subset \hat{B}$ be a nonempty affine open subset.*

Let $Z \in z^n(X \times B)$ be an irreducible cycle. Then $Z \to X$ is fs over X if and only if Z is closed in $X \times \hat{B}$.

Proof. Let $f : Z \hookrightarrow X \times \hat{B} \to X$ be the composite map. Suppose f is fs over X. Since the second map is projective, by [Hartshorne 1977, Corollary II-4.8(e), Theorem II-4.9, pages 102–103], the first map is a closed immersion. This proves ($\Rightarrow$).

Conversely, suppose that Z is closed in $X \times \hat{B}$, i.e., the first map is a closed immersion (thus projective). Since the second map is projective, the composite f is projective. Hence, f is a projective morphism of affine schemes, so that it must be finite by [Hartshorne 1977, Exercise II-4.6, page 106]. Moreover, $Z \to X_i$ being a finite map of irreducible affine schemes of the same dimension, where X_i is the irreducible component that receives Z, this morphism must also be surjective. This proves ($\Leftarrow$). □

Lemma 2.10. *Let $V = \mathrm{Spec}(R)$ be a semilocal k-scheme essentially of finite type with the set of closed points Σ. Let $B \subset \hat{B}$ be as in Lemma 2.9. Let $F := \hat{B} \setminus B$. Let $Z \in z^n(V \times B)$ be an irreducible cycle and let $\hat{Z}$ be the Zariski closure of Z in $V \times \hat{B}$.*

Suppose that $\hat{Z} \cap (\Sigma \times F) = \varnothing$. Then given any affine atlas (X, Σ) for V, there exists an affine open subatlas (U, Σ) for V such that for the Zariski closure $\overline{Z}$ of Z in $X \times B$, the projection map $\overline{Z}_U \to U$ is fs over U.

If V is smooth over k from the first place, then we can choose (U, Σ) such that U is smooth over k as well.

Proof. Let (X, Σ) be a given atlas. Let $\hat{\overline{Z}}$ be the Zariski closure of $\overline{Z}$ in $X \times \hat{B}$ and let $\hat{f} : \hat{\overline{Z}} \hookrightarrow X \times \hat{B} \to X$ be the composition with the projection. Let $Y := \hat{f}(\hat{\overline{Z}} \cap (X \times F))$. Since $\hat{f}$ is projective and since $\hat{\overline{Z}} \cap (\Sigma \times F) = \hat{Z} \cap (\Sigma \times F) = \varnothing$, we see that $Y \subset X$ is a closed subset disjoint from Σ. Hence, $X \setminus Y$ is an open neighborhood of Σ such that $\hat{\overline{Z}} \cap ((X \setminus Y) \times F) = \varnothing$. By Lemma 2.3, we can find an affine open neighborhood U of Σ in $X \setminus Y$, so we have $\hat{\overline{Z}} \cap (U \times F) = \varnothing$. In particular, $\hat{\overline{Z}} \cap (U \times \hat{B}) = \overline{Z} \cap (U \times \hat{B})$. This means $\overline{Z}_U$ is closed in $U \times \hat{B}$. Hence, by Lemma 2.9, the map $\overline{Z}_U \to U$ is fs over U.

In case V is smooth, then by excising the singular locus of X, which is disjoint from Σ, we may assume that X is smooth. Then the open subset $U \subset X$ is also smooth. □

Let X be an equidimensional quasiprojective k-scheme and let $\Sigma \subset X$ be a finite set of points. By Lemma 2.3, we may replace X be an affine k-scheme. We have the following two notions of fs-cycles:

Definition 2.11. Let $V = X_\Sigma$. Let $m, n \geq 1$ be integers:

(1) A cycle $\alpha \in \mathrm{Tz}^n_\Sigma(X, n; m)$ is said to be an fs-*cycle along* Σ if there is an affine open neighborhood $U \subset X$ of Σ such that each irreducible component of α_U is fs over U. The group of fs-cycles along Σ is denoted by $\mathrm{Tz}^n_{\Sigma,\mathrm{fs}}(X, n; m)$.

(2) A cycle $\alpha \in \mathrm{Tz}^n_\Sigma(V, n; m)$ is said to be an fs-*cycle* if each irreducible component of α is fs over V. The group of fs-cycles is denoted by $\mathrm{Tz}^n_{\mathrm{fs}}(V, n; m)$.

These two notions are related as follows:

Corollary 2.12. *Let X be an equidimensional affine k-scheme and let $\Sigma \subset X$ be a finite set of points. Let $V = X_\Sigma$. Let $m, n \geq 1$ be integers. Then a cycle $\alpha \in \mathrm{Tz}^n_\Sigma(X, n; m)$ is an fs-cycle along Σ if and only if $\alpha_V \in \mathrm{Tz}^n_\Sigma(V, n; m)$ is an fs-cycle.*

Proof. ($\Rightarrow$) Since the localization map $V \to X$ is flat and it factors through any open neighborhood $U \subset X$ of Σ, one can pull-back by Lemma 2.7 to prove this direction.

($\Leftarrow$) By Lemma 2.5, there exists an affine open subatlas (U_1, Σ) of (X, Σ) for V such that the closure $\bar{\alpha}$ of Z in $U_1 \times B_n$ is in $\mathrm{Tz}^n_\Sigma(U_1, n; m)$.

For each irreducible component Z of α, let $\hat{Z}$ be its Zariski closure in $V \times \hat{B}$. Since Z is fs over V, by Lemma 2.9 Z is already closed in $V \times \hat{B}_n$, thus $Z = \hat{Z}$. In particular, $\hat{Z} \cap (\Sigma \times F_n) = \varnothing$. Hence by Lemma 2.10 there exists an affine open subatlas (U_Z, Σ) for V of (U_1, Σ) such that for the Zariski closure

$\overline{Z}$ of Z in $U_1 \times B_n$, the base change $\overline{Z}_{U_Z} \to U_Z$ is fs. By taking $U := \bigcap_Z U_Z$ where the intersection is taken over all (finitely many) irreducible components of α, we deduce that $\overline{Z}_U \to U$ is fs. This proves the corollary. □

We have the following a bit different characterization of the cycles centered around $\mathrm{Tz}^n_{\mathrm{fs}}(V, n; m)$:

Proposition 2.13. *Let $V = \mathrm{Spec}(R)$ be a semilocal k-scheme of geometric type with the set Σ of closed points. Let $m, n \geq 1$. Let $Z \in \mathrm{Tz}^n_{\Sigma}(V, n; m)$ be an irreducible cycle. Then Z is an* fs-*cycle if and only if there is an atlas (X, Σ) for V such that for the closures $\overline{Z}$ in $X \times B_n$ and $\hat{Z}$ in $V \times \hat{B}_n$, we have $\overline{Z} \in \mathrm{Tz}^n_{\Sigma}(X, n; m)$ and $\hat{Z} \cap (\Sigma \times F_n) = \varnothing$.*

Here, V is smooth over k if and only if we can choose (X, Σ) in the above such that X is smooth over k as well.

Proof. For the first assertion, suppose that Z is an fs-cycle. By Lemma 2.5, there is a affine atlas (X, Σ) for V such that $\overline{Z} \in \mathrm{Tz}^n_{\Sigma}(X, n; m)$. Since $Z \to V$ is fs over V, by Lemma 2.9, $\hat{Z} \cap (\Sigma \times F_n) = \varnothing$.

Conversely, suppose that for an atlas (X, Σ) and the closure $\overline{Z}$ in $X \times B_n$, we have $\overline{Z} \in \mathrm{Tz}^n_{\Sigma}(X, n; m)$ and $\hat{Z} \cap (\Sigma \times F_n) = \varnothing$. Then, by Lemma 2.10, we may shrink (X, Σ) to an affine open atlas (U, Σ) such that $\overline{Z}_U \to U$ fs over U. Hence $\overline{Z}_U \in \mathrm{Tz}^n_{\Sigma,\mathrm{fs}}(U, n; m)$. Now by Corollary 2.12, we have $Z \in \mathrm{Tz}^n_{\mathrm{fs}}(V, n; m)$.

For the second assertion, in case V was smooth, then we could have take X to be smooth here by the last assertion of Lemma 2.5. Conversely, a localization of a smooth scheme is smooth again, so that V is smooth over k. □

2E. *The* sfs-*cycles.* For $1 \leq j \leq n$, let $\pi_j : B_n \to B_j$ and $\hat{\pi}_j : \hat{B}_n \to \hat{B}_j$ be the projection maps. Let $X \in \mathbf{Sch}^{\mathrm{ess}}_k$ equidimensional. We shall often denote the maps $\mathrm{id}_X \times \pi_j : X \times B_n \to X \times B_j$ and $\mathrm{id}_X \times \hat{\pi}_j : X \times \hat{B}_n \to X \times \hat{B}_j$ simply by π_j and $\hat{\pi}_j$, respectively, if the scheme X is fixed in a given context.

For any reduced closed subscheme $Z \subset X \times B_n$ and $1 \leq j \leq n$, let $Z^{(j)} = (\mathrm{id}_X \times \pi_j)(Z)$ be the scheme-theoretic image of Z. Let $Z^{(0)}$ be the scheme-theoretic image of Z in X. Note that if the projection $Z \to X$ is proper, then $(\mathrm{id}_X \times \pi_j)(Z)$ is closed in $X \times B_j$ and, with its reduced induced closed subscheme structure, coincides with $Z^{(j)}$. The same holds for $Z^{(0)}$. We shall use $Z^{(j)}$ when $Z \to X$ is in fact finite.

Definition 2.14. Let $X \in \mathbf{Sch}^{\mathrm{ess}}_k$ be smooth over k and let $\Sigma \subset X$ be a finite set of points. Let $m, n \geq 1$ be integers. An integral cycle $[Z] \in \underline{\mathrm{Tz}}^n(X, n; m)$ is called an sfs-*cycle along* Σ, if $[Z] \in \underline{\mathrm{Tz}}^n_{\Sigma}(X, n; m)$, and there exists an affine neighborhood $U \subset X$ of Σ such that the following hold:

(1) Z_U is finite and surjective over an irreducible component of U, i.e., $Z_U \to U$ is an fs-morphism.

(2) The scheme $(Z^{(j)})_U$ is smooth over k for every $0 \leq j \leq n$.

A cycle $\alpha \in \underline{\mathrm{Tz}}^n(X, n; m)$ is called an *sfs-cycle along* Σ if every irreducible component of α is an sfs-cycle along Σ.

Lemma 2.15. *Let X be an equidimensional smooth affine k-scheme and let $\Sigma \subset X$ be a finite set of points. Let $V = X_\Sigma$. Let $m, n \geq 1$ be integers. Then $\alpha \in \underline{\mathrm{Tz}}^n_\Sigma(X, n; m)$ is an sfs-cycle along Σ if and only if $\alpha_V \in \underline{\mathrm{Tz}}^n_\Sigma(V, n; m)$ is an fs-cycle such that $Z^{(j)}$ is smooth over k for each $0 \leq j \leq n$ and for each irreducible component Z of α_V.*

Proof. Under Corollary 2.12, the ($\Rightarrow$) direction is obvious. We prove ($\Leftarrow$). By Corollary 2.12, together with Lemma 2.3, we can find an affine open neighborhood $U' \subset X$ of Σ such that the closure $\alpha_{U'} \in \underline{\mathrm{Tz}}^n_\Sigma(U', n; m)$ is an fs-cycle along Σ. Now let $Y \subset U'$ be the union of the images of the finite maps $(Z^{(j)}_{U'})_{\mathrm{sing}} \to U'$, where Z runs over all irreducible components of α and $0 \leq j \leq n$. Since $Z_{U'} \to U'$ is finite for each Z, this $Y \subset U'$ is a closed subset that does not meet Σ. By Lemma 2.3, we can choose an affine open neighborhood $U \subset U' \setminus Y$ of Σ. Then for each component Z of α and each $0 \leq j \leq n$, the scheme $Z^{(j)}_U$ is smooth over k. Note $(Z_U)^{(j)} = (Z^{(j)})_U$ naturally. This shows that α_U is an sfs-cycle along Σ. □

Another property that sfs-cycles enjoy is the following:

Lemma 2.16. *Let $\phi : X \to Y$ be an étale morphism of smooth affine k-schemes. Let $\Sigma \subset Y$ be a finite set of points and let $\Sigma' = \phi^{-1}(\Sigma)$. Let $Z \in \underline{\mathrm{Tz}}^n(Y, n; m)$ be an integral* sfs*-cycle along Σ. Then the flat pull-back $\phi^*(Z) \in \underline{\mathrm{Tz}}^n(X, n; m)$ is an sfs-cycle along Σ'.*

Proof. It is easy to see that $\phi^*(Z) \in \underline{\mathrm{Tz}}^n_{\Sigma'}(X, n; m)$. We now prove the other properties. We can shrink Y and assume that $Z \to Y$ is finite and surjective, and $Z^{(j)}$ is smooth over k for $0 \leq j \leq n$. Let $W := \phi^*(Z)$. It follows from Lemma 2.7 that W is an fs-cycle along Σ'. To prove that each $W^{(j)}$ is smooth over k, let $W_j := \phi^*(Z^{(j)})$ and consider the commutative diagram

$$\begin{array}{ccccccc} W & \twoheadrightarrow & W^{(j)} & \longrightarrow & W_j & \longrightarrow & X \\ \downarrow & & & & \downarrow & & \downarrow \phi \\ Z & & \longrightarrow\!\!\!\!\!\rightarrow & & Z^{(j)} & \longrightarrow & Y. \end{array} \tag{2-2}$$

Here, the map $W^{(j)} \to W_j$ exists uniquely since the right square is Cartesian. The outer big square is also Cartesian, and this implies that so is the left square. In particular, the vertical arrows are all étale, the horizontal arrows are all finite and surjective and all schemes in (2-2) are reduced. In particular, $W^{(j)} \twoheadrightarrow W_j$. On the other hand, as $W \to X$ is finite, $W^{(j)} = \pi_j(W)$ is a reduced closed subscheme of W_j. Thus $W^{(j)} = W_j$. Since Z and $Z^{(j)}$ are smooth over k and ϕ is étale, it follows that W and W_j are smooth over k. In particular, $W^{(j)} = W_j$ is smooth over k. This finishes the proof. □

2F. ***Additive higher Chow groups of* fs *and* sfs*-cycles.*** The goal of this paper is to prove the "sfs-moving lemma" which will show that the cycle class groups of sfs-cycles coincide with the additive higher Chow groups in the Milnor range for a smooth semilocal k-scheme essentially of finite type when k is an infinite perfect field.

Let $m, n \geq 1$. Let X be a smooth affine k-scheme and let $\Sigma \subset X$ be a finite set of points. It follows from Definition 2.14 that $\mathrm{Tz}^n_{\Sigma,\mathrm{sfs}}(X, n; m)$ is a subgroup of $\mathrm{Tz}^n_{\Sigma,\mathrm{fs}}(X, n; m)$.

Definition 2.17. We let

$$\widetilde{\mathrm{TCH}}^n_{\Sigma}(X, n; m) = \frac{\ker(\partial : \mathrm{Tz}^n_{\Sigma}(X, n; m) \to \mathrm{Tz}^n(X, n-1; m))}{\mathrm{im}(\partial : \mathrm{Tz}^n(X, n+1; m) \to \mathrm{Tz}^n(X, n; m)) \cap \mathrm{Tz}^n_{\Sigma}(X, n; m)},$$
$$\mathrm{TCH}^n_{\Sigma,\mathrm{fs}}(X, n; m) = \frac{\ker(\partial : \mathrm{Tz}^n_{\Sigma,\mathrm{fs}}(X, n; m) \to \mathrm{Tz}^n(X, n-1; m))}{\mathrm{im}(\partial : \mathrm{Tz}^n(X, n+1; m) \to \mathrm{Tz}^n(X, n; m)) \cap \mathrm{Tz}^n_{\Sigma,\mathrm{fs}}(X, n; m)},$$
$$\mathrm{TCH}^n_{\Sigma,\mathrm{sfs}}(X, n; m) = \frac{\ker(\partial : \mathrm{Tz}^n_{\Sigma,\mathrm{sfs}}(X, n; m) \to \mathrm{Tz}^n(X, n-1; m))}{\mathrm{im}(\partial : \mathrm{Tz}^n(X, n+1; m) \to \mathrm{Tz}^n(X, n; m)) \cap \mathrm{Tz}^n_{\Sigma,\mathrm{sfs}}(X, n; m)}.$$

We similarly define $\widetilde{\mathrm{TCH}}^n_{\Sigma}(V, n; m)$, $\mathrm{TCH}^n_{\mathrm{fs}}(V, n; m)$, and $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m)$.

If X is not necessarily connected, note that the groups for X are obtained simply by taking the direct sums of the corresponding groups over all connected components of X.

In the above, the definition of the group $\widetilde{\mathrm{TCH}}^n_{\Sigma}(X, n; m)$ is slightly different from that of $\mathrm{TCH}^n_{\Sigma}(X, n; m)$ in Definition 2.2. However, we have:

Lemma 2.18. *The natural surjection* $\mathrm{TCH}^n_{\Sigma}(X, n; m) \twoheadrightarrow \widetilde{\mathrm{TCH}}^n_{\Sigma}(X, n; m)$ *is an isomorphism. Similarly,* $\mathrm{TCH}^n_{\Sigma}(V, n; m) \to \widetilde{\mathrm{TCH}}^n_{\Sigma}(V, n; m)$ *is an isomorphism.*

Proof. By the moving lemma for additive higher Chow groups of smooth affine schemes of W. Kai [2019] (see [Krishna and Park 2016, Theorem 4.1] for a sketch of its proof), the composition $\mathrm{TCH}^n_{\Sigma}(X, n; m) \twoheadrightarrow \widetilde{\mathrm{TCH}}^n_{\Sigma}(X, n; m) \to \mathrm{TCH}^n(X, n; m)$ is an isomorphism. Hence, the first arrow is injective. The proof for the second one is similar, except that we use [Krishna and Park 2016, Theorem 4.10]. □

We thus have canonical maps

$$\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n_{\mathrm{fs}}(V, n; m) \to \mathrm{TCH}^n_{\Sigma}(V, n; m) \to \mathrm{TCH}^n(V, n; m), \tag{2-3}$$

where the last map is an isomorphism by Lemma 2.18 and [Kai 2019]. Our goal is to show that all other maps are also isomorphisms.

2G. *Reduction to localization at closed points.* The semilocal k-schemes essentially of finite type we consider are obtained by localizing an affine k-scheme (see Lemma 2.3) at a finite set Σ of points which may not necessarily be closed. In Section 2G, we show that for the sfs-moving lemma, it is possible to reduce to the case when all points of Σ are actually closed. The following is the goal:

Proposition 2.19. *Suppose the natural map* $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n(V, n; m)$ *is an isomorphism for every smooth semilocal k-scheme V essentially of finite type, obtained by localizing at a finite set of closed points. Then the natural map* $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n(V, n; m)$ *is an isomorphism for every smooth semilocal k-scheme V essentially of finite type.*

We prove the following first:

Lemma 2.20. *Let V be a smooth semilocal k-scheme essentially of finite type, obtained by localizing an affine k-scheme X at a finite set Σ of, not necessarily closed, points. Let* $\alpha \in \underline{\mathrm{Tz}}^n(V, n; m)$.

Then there exist (1) *a smooth semilocal k-scheme V' essentially of finite type*, *obtained by localizing an affine k-scheme at a finite set Σ' of closed points with a flat localization map $V \to V'$ and* (2) *a cycle $\alpha' \in \underline{\mathrm{Tz}}^n(V', n; m)$ such that the flat pull-back map $\phi_V^{V'} : \underline{\mathrm{Tz}}^n(V', n; m) \to \underline{\mathrm{Tz}}^n(V, n; m)$ satisfies $\phi_V^{V'}(\alpha') = \alpha$. If $\partial\alpha = 0$, we can ensure $\partial\alpha' = 0$.*

Proof. By Lemma 2.3, we may assume that $V = X_\Sigma$, where X is a smooth affine k-scheme of finite type. For the cycle $\alpha \in \underline{\mathrm{Tz}}^n(V, n; m)$, by Lemma 2.5, there exists a smooth affine open neighborhood $U \subset X$ containing Σ such that the Zariski closure α_U of α in $U \times B_n$ is in $\underline{\mathrm{Tz}}^n(U, n; m)$. If $\partial\alpha = 0$, we can shrink U further (if necessary) so that $\partial\alpha_U = 0$.

For each $p \in \Sigma$, there exists a closed point $\mathfrak{m}_p \in U$ that is a specialization of p. (It exists by the basic fact in commutative algebra that any proper ideal of a commutative ring with unit is contained in a maximal ideal.) We choose it so that a distinct pair of points of Σ gives a distinct pair of points. Let $\Sigma' := \{\mathfrak{m}_p \mid p \in \Sigma\}$, and take $V' := U_{\Sigma'}$. Here, $\alpha_U \in \underline{\mathrm{Tz}}^n(U, n; m)$, and let $\alpha' \in \underline{\mathrm{Tz}}^n(V', n; m)$ be its flat pull-back via the localization map $V' \to U$. This satisfies $\partial\alpha' = 0$ if $\partial\alpha = 0$. By the construction of V', we also have the localization map $V \to V'$ and the flat pull-back map $\phi_V^{V'} : \underline{\mathrm{Tz}}^n(V', n; m) \to \underline{\mathrm{Tz}}^n(V, n, m)$. By the construction of α', we have $\phi_V^{V'}(\alpha') = \alpha$. This proves the lemma. □

We remark however that Lemma 2.20 does not say that the map $\phi_V^{V'}$ is surjective. It simply says that for each element α, there is some V' such that α can be an image of a cycle over V'.

Proof of Proposition 2.19. Since the map $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n(V, n; m)$ is automatically injective, it is enough to prove that this is surjective. Let $\alpha \in \mathrm{TCH}^n(V, n; m)$ be an arbitrary cycle class, and choose its cycle representative in $\underline{\mathrm{Tz}}^n(V, n; m)$, also denoted by α. Being a cycle representing a class in $\mathrm{TCH}^n(V, n; m)$, we have $\partial\alpha = 0$.

By Lemma 2.20, there exists now a smooth semilocal k-scheme (V', Σ') essentially of finite type, obtained by localizing at a finite set of closed points, a cycle class $\alpha' \in \mathrm{TCH}^n(V', n; m)$ and the localization map $\phi_V^{V'} : \mathrm{TCH}^n(V', n; m) \to \mathrm{TCH}^n(V, n; m)$ sends α' to α.

On the other hand, the localization map $\phi_V^{V'}$ sends the sfs-cycles over V' to the sfs-cycles over V. To see this, we first note that this map sends $\underline{\mathrm{Tz}}^n_{\Sigma'}(V', n; m)$ to $\underline{\mathrm{Tz}}^n_{\Sigma}(V, n; m)$ because the localization does not increase the dimensions of schemes, thus the proper intersection condition with Σ' implies the proper intersection condition with Σ. Now, the sfs-cycles are preserved under $\phi_V^{V'}$ because the localization (flat pull-back) of fs-morphisms are fs-morphisms by Lemma 2.7, while it is a basic fact in commutative algebra that a localization of a regular local ring is again a regular local ring. Hence, we have a commutative diagram:

$$\begin{array}{ccc} \mathrm{TCH}^n_{\mathrm{sfs}}(V', n; m) & \xrightarrow{\phi_{\mathrm{sfs}}} & \mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \\ \downarrow{\scriptstyle \mathrm{sfs}_{V'}} & & \downarrow{\scriptstyle \mathrm{sfs}_V} \\ \mathrm{TCH}^n(V', n; m) & \xrightarrow{\phi} & \mathrm{TCH}^n(V, n; m), \end{array} \tag{2-4}$$

where $\phi = \phi_V^{V'}$ and ϕ_{sfs} is the restriction of ϕ. By construction, we have $\phi(\alpha') = \alpha$. By the given assumption, we have that $\mathrm{sfs}_{V'}$ is surjective, so that there exists $\alpha'' \in \mathrm{TCH}^n_{\mathrm{sfs}}(V', n; m)$ such that $\mathrm{sfs}_{V'}(\alpha'') = \alpha'$. Hence $\alpha = \phi(\alpha') = \phi \circ \mathrm{sfs}_{V'}(\alpha'') =^\dagger \mathrm{sfs}_V \circ \phi_{\mathrm{sfs}}(\alpha'')$, where $\dagger$ holds by the commutativity of the diagram (2-4). In particular, $\alpha \in \mathrm{im}(\mathrm{sfs}_V)$. Since α was arbitrary in $\mathrm{TCH}^n(V, n; m)$, this shows that sfs_V is surjective, hence an isomorphism. □

We have one further result.

Lemma 2.21. *Let (V, Σ) be a smooth semilocal k-scheme essentially of finite type. Let $m, n \geq 1$ be integers. Let $\alpha \in \underline{\mathrm{Tz}}^n_\Sigma(V, n; m)$ be such that $|\alpha|$ is finite over V. Then α does not intersect $V \times F$ for any proper face $F \subset B_n$ at all. In particular, $\partial(\alpha) = 0$.*

Proof. We may assume that $\alpha = [Z]$ is an irreducible cycle and V is integral. We prove that $Z \cap (V \times F)$ is empty.

The composite $Z \cap (V \times F) \hookrightarrow Z \to V$ is finite by the given assumption. Hence, its image in V is closed and therefore must intersect Σ nontrivially if nonempty. It suffices therefore to show that the fiber product $\Sigma \times_V Z \times_{B_n} F = Z \cap (\Sigma \times F)$ is empty.

However, by the given assumption that $Z \in \underline{\mathrm{Tz}}^n_\Sigma(V, n; m)$, the proper intersection condition with Σ reads $\mathrm{codim}_{\Sigma \times F}\, Z \cap (\Sigma \times F)) \geq n$. Equivalently,

$$\dim Z \cap (\Sigma \times F) \leq \dim(\Sigma \times F) - n = \dim F - n < 0.$$

But this means $Z \cap (\Sigma \times F) = \varnothing$. This proves the lemma. □

Convention. Using Proposition 2.19, from now on, when we say a semilocal k-scheme essentially of finite type, it will mean that it is obtained by localizing at a finite set of closed points, unless we say otherwise.

3. The sfs-moving lemma in affine spaces

In this section, we prove a special case of Theorem 1.1 when the underlying semilocal scheme is a localization an affine space over k. This will be a ground for the general case of the theorem.

3A. ***The Set-up for affine spaces.*** We fix some notations that we shall use throughout this section.

Let k be an infinite perfect field. Let $m, n, r \geq 1$ be integers. We let $\Sigma \subset \mathbb{A}^r_k = \mathrm{Spec}(k[x_1, \dots, x_r])$ be a finite set of closed points. Let V be the localization of $\mathbb{A}^r_k$ at Σ. Let $j : V \to \mathbb{A}^r_k$ be the inclusion map. Let $p_n : \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1} \to \mathbb{A}^r_k \times \mathbb{A}^1_k$ and $q : \mathbb{A}^r_k \times \mathbb{A}^1_k \to \mathbb{A}^r_k$ denote the projection maps and let $q_n = q \circ p_n$. Using the automorphism $y \mapsto 1/(1-y)$ of $\mathbb{P}^1_k$, we replace $(\square, \infty, 0)$ by $(\mathbb{A}^1_k, 0, 1)$, and write $\square = \mathbb{A}^1_k$.

For any closed subset $Y \subset \mathbb{A}^r_k \times \mathbb{A}^1_k \times \square^{n-1}$, let $\bar{Y}$ be its closure in $\mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1}$. We let $Z \in \mathrm{Tz}^n_\Sigma(\mathbb{A}^r_k, n; m)$ be an irreducible cycle. For an integer $s \geq 0$ and a point $g \in \mathbb{A}^r_k$, we consider the map (see [Kai 2019])

$$\begin{aligned} \phi_{g,s} : \mathbb{A}^r_k \times \mathbb{A}^1_k \times \square \times \overline{\square}^{n-1} &\to \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1}; \\ (\underline{x}, t, y, y_1, \dots, y_{n-1}) &\mapsto (\underline{x} + yt^{s(m+1)}g, t, y_1, \dots, y_{n-1}). \end{aligned} \tag{3-1}$$

Note that $\phi_{g,s}$ is strictly speaking defined over the residue field of g, but to simplify notation we often won't make it explicit. If needed, one can take the scalar extension to the residue field of g to turn g into a rational point. For $a \in \square(k)$, we let $\phi_{g,s,a}$ be the composite map

$$\mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1} \hookrightarrow \mathbb{A}^r_k \times \mathbb{A}^1_k \times \square \times \overline{\square}^{n-1} \xrightarrow{\phi_{g,s}} \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1},$$

where the first arrow takes $(\underline{x}, t, \underline{y})$ to $(\underline{x}, t, a, \underline{y})$.

The evaluation of $\phi_{g,s}$ at $y = 1$ defines an isomorphism $\mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)} \to \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)}$, given by $\phi_{g,s,1}(\underline{x}, t) = (\underline{x} + t^{s(m+1)} g, t)$. Let $\phi^\sharp_{g,s,1} : k(g)[\underline{x}, t] \to k(g)[\underline{x}, t]$ be the corresponding $k(g)$-algebra isomorphism.

3B. ***Some properties of the twisted translations.*** Note that $\phi_{g,s}$ is a flat morphism. In particular, $\phi^*_{g,s}(Z)$ is an algebraic cycle on $\mathbb{A}^r_k \times \mathbb{A}^1_k \times \square^n$. In the next few lemmas, we verify some algebraic and geometric properties of $\phi^*_{g,s}(Z)$.

Lemma 3.1. *Let $f(\underline{x}, t) \in k[\underline{x}, t]$ be a nonzero polynomial. Then there is a nonempty open subset $U \subset \mathbb{A}^r_k$ such that for each $g \in U$ and sufficiently large $s \gg 0$ (not depending on g), the polynomial $\phi^\sharp_{g,s,1}(f)$ is monic in t over $k(g)[\underline{x}]$, i.e., integral over $k(g)[\underline{x}]$.*

Proof. Let $M := \deg_t f$ and write $f(\underline{x}, t) = \sum_{i=0}^M f_i(\underline{x}) t^{M-i}$ for some $f_i \in k[\underline{x}]$ and $M \geq 0$. Since $f \neq 0$, we have $f_0(\underline{x}) \neq 0$. Let $d_i = \deg_{\underline{x}}(f_i)$, which is the total degree in $\underline{x}$. We first consider the case $r = 1$ and take $U = \mathbb{A}^1_k \setminus \{0\}$. Let $c_i \in k$ be the coefficient of the highest degree term of $f_i(\underline{x})$. Since $f_0(\underline{x}) \neq 0$, we have $c_0 \in k^\times$. Then,

$$f(x + t^{s(m+1)} g, t) = \sum_{i=0}^M f_i(x + t^{s(m+1)} g) t^{M-i} = \sum_{i=0}^M c_i (g^{d_i} t^{d_i s(m+1) + M - i} + (\text{lower degree terms in } t)).$$

Let i_0 be the smallest integer such that $d_{i_0} = \max\{d_0, d_1, \dots, d_M\}$. Here, $c_{i_0} \in k^\times$ by definition.

If $d_{i_0} = 0$, then each $f_i(x)$ is a constant, so $f(x + t^{s(m+1)} g, t)$ gives an integral dependence in t as desired. Suppose $d_{i_0} > 0$. If $i_0 = 0$, then for each $i > 0$ and each $s > 0$, we have

$$d_0 s(m+1) + M \geq d_i s(m+1) + M > d_i s(m+1) + M - i.$$

Hence, the leading coefficient of the highest degree term in t is $c_0 g^{d_0} \in k(g)^\times$, so, after dividing by this unit $c_0 g^{d_0}$, we get a monic polynomial in t. Hence it is integral.

If $i_0 > 0$, then for each $i > i_0$ and each $s > 0$, we have

$$d_{i_0} s(m+1) + M - i_0 \geq d_i s(m+1) + M - i_0 > d_i s(m+1) + M - i,$$

while for $0 \leq i < i_0$, we have $d_i < d_{i_0}$ so that for every sufficiently large $s > 0$, we have

$$d_i s(m+1) + M - i < d_{i_0} s(m+1) + M - i_0.$$

Note that this choice of s depends only on f and not on g. Hence, for every sufficiently large $s > 0$ (not depending on g), again the leading coefficient of highest degree in t is $c_{i_0} g^{d_{i_0}} \in k(g)^\times$. Hence after dividing by this unit, it gives the desired integral dependence relation.

In case $r \geq 2$, the backbone of the proof is the same, but one problem is a possible cancellation of the highest degree terms in t, namely, if d_i is the total degree of $f_i(x_1, \dots, x_r)$, then possibly a multiple number of monomials in $\phi^\sharp_{g,s,1}(f)$ could have the same total degree d_i. However, such g's form a closed subscheme of $\mathbb{A}^r_k$ (depends on $f(\underline{x}, t)$), so for a general $g \in U$ for some nonempty open subset $U \subset \mathbb{A}^r_k$, we can avoid it. $\square$

W. Kai [2019, Proposition 2.3] (or see [Krishna and Park 2016, Claim of proof of Theorem 4.1]) defines a positive integer $s(Z)$ associated to Z, which plays a crucial role in proving the modulus condition for $\phi^*_{g,s}(Z)$.

Lemma 3.2. *Let $s \geq s(Z)$ be any integer. Then $\phi^*_{g,s}(Z) \in \mathrm{Tz}^n(\mathbb{A}^r_k, n+1; m)$ for any $g \in \mathbb{A}^r_k$.*

Proof. The modulus condition for $\phi^*_{g,s}(Z)$ follows from [Kai 2019, Proposition 2.3] (see also [Krishna and Park 2016, Proof of Theorem 4.1]). We show that $\phi^*_{g,s}(Z)$ intersects all faces of $\square^n$ properly. Let F be a face of $\square^n$. If $F = \{0\} \times F'$ for some face F' of $\square^{n-1}$, then the proper intersection follows directly from that of Z with F' since the map $\phi_{g,s,0}$ is identity. If $F = \{1\} \times F'$ for some face F' of $\square^{n-1}$, then the proper intersection also follows from that of Z with F' since the map $\phi_{g,s,1} : \mathbb{A}^r_k \times \mathbb{A}^1_k \times F' \to \mathbb{A}^r_k \times \mathbb{A}^1_k \times F'$ is an isomorphism. If $F = \square \times F'$ for some face F' of $\square^{n-1}$, then the map $\mathbb{A}^r_k \times \mathbb{A}^1_k \times \square \times F' \to \mathbb{A}^r_k \times \mathbb{A}^1_k \times F'$ is flat of relative dimension one and hence we get

$$\begin{aligned}
\dim(\phi^*_{g,s}(Z) \cap (\mathbb{A}^r_k \times \mathbb{A}^1_k \times \square \times F')) &= \dim(\phi^*_{g,s}(Z \cap F')) \\
&= \dim(Z \cap F') + 1 \leq \dim(Z) + 1 - \mathrm{codim}_{\square^{n-1}}(F') \\
&= \dim(\phi^*_{g,s}(Z)) - \mathrm{codim}_{\square^n}(\square \times F') \\
&= \dim(\phi^*_{g,s}(Z)) - \mathrm{codim}_{\square^n}(F).
\end{aligned}$$

This proves the desired proper intersection of $\phi^*_{g,s}(Z)$. $\square$

Lemma 3.3. *Assume that $n = 1$. For $g \in \mathbb{A}^r_k \setminus \{0\}$ and $s \gg 0$ as in Lemma 3.1, $\phi^*_{g,s,1}(Z)$ is finite and surjective over $\mathbb{A}^r_{k(g)}$.*

Proof. Since $\mathbb{A}^r_k \times \mathbb{A}^1_k$ is factorial, there exists an irreducible polynomial $f(\underline{x}, t) \in k[\underline{x}, t]$ such that $Z = \mathrm{Spec}(k[\underline{x}, t]/(f(\underline{x}, t)))$. The modulus condition mandates that this cycle does not intersect the divisor $\{t = 0\}$ in $\mathbb{A}^r_k \times \mathbb{A}^1_k$, so that after scaling f by a constant in $k^\times$, we must have $f = th - 1$ for some $h(\underline{x}, t) \in k[\underline{x}, t]$. By Lemma 3.1, $\phi^\sharp_{g,s}(th - 1)$ is monic in t for $g \in \mathbb{A}^r_k \setminus \{0\}$ and $s \gg 0$ up to scaling by a unit in $k(g)^\times$. This is equivalent to saying that $\phi^*_{g,s,1}(Z_{k(g)}) \to \mathbb{A}^r_{k(g)}$ is finite. As both have the same dimension and $\mathbb{A}^r_{k(g)}$ is integral, this morphism is automatically surjective. $\square$

3C. ***The three types of cycles.*** In order to generalize Lemma 3.3 to $n \geq 2$ case, we need to consider three types of cycles.

Lemma 3.4. *Suppose that the projection to the first factor $Z \to \mathbb{A}^r_k$ is dominant. Then there is a dense open subset $U \subset \mathbb{A}^r_k$ such that each $g \in U$ and integer $s > 0$, the projection to the first factor $\phi^*_{g,s,1}(Z_{k(g)}) \to \mathbb{A}^r_{k(g)}$ is still dominant.*

Proof. This is immediate from the definition of $\phi_{g,s}$. □

Lemma 3.5. *Assume that* (a) *the projection $q_n : Z \to \mathbb{A}^r_k$ is not dominant while* (b) *the projection* $\mathrm{pr}_2 : Z \to \mathbb{A}^1_k$ *is dominant. Then there is a dense open subset $U \subset \mathbb{A}^r_k$ such that for each $g \in U$ and $s > 0$, we have*

(1) $\dim(q_n(\phi^*_{g,s,1}(Z_{k(g)}))) = \dim(q_n(Z_{k(g)})) + 1$ *and*

(2) *the projection* $\mathrm{pr}_2 : \phi^*_{g,s,1}(Z_{k(g)}) \to \mathbb{A}^1_{k(g)}$ *is dominant.*

Proof. By (b), the map pr_2 is a dominant morphism to a regular curve, thus it is flat by [Hartshorne 1977, Proposition III-9.7, page 256]. In particular, $\mathrm{pr}_2(Z) \subset \mathbb{A}^1_k$ is a dense open subset. For each $g \in \mathbb{A}^r_k$ and $s > 0$, we have a surjection $\Phi : q_n(Z_{k(g)}) \times \mathrm{pr}_2(Z_{k(g)}) \to q_n(\phi^*_{g,s,1}(Z_{k(g)}))$, given by sending (x, t) to $x + t^{s(m+1)} g$. Thus, $\dim q_n(\phi^*_{g,s,1}(Z_{k(g)})) \le \dim q_n(Z_{k(g)}) + 1$.

On the other hand, for each fixed closed point $t_0 \in \mathrm{pr}_2(Z)$, the set $\Phi(q_n(Z_{k(g)}), t_0)$ has the same dimension as that of $q_n(Z_{k(g)})$, while it is an equidimensional proper closed subset of $q_n(\phi^*_{g,s,1}(Z_{k(g)}))$ when g is a general member, i.e., in an open subset of $\mathbb{A}^r_k$. Since $\mathrm{pr}_2(Z)$ is dense open in $\mathbb{A}^1_k$ and hence of positive dimension, we must have $\dim(q_n(\phi^*_{g,s,1}(Z_{k(g)})) > \dim(q_n(Z_{k(g)}))$. This proves (1). Property (2) is obvious because $\phi_{g,s}$ does not modify the $\mathbb{A}^1_k$-coordinate. □

Lemma 3.6. *Assume that neither of the projections $q_n : Z \to \mathbb{A}^r_k$ nor* $\mathrm{pr}_2 : Z \to \mathbb{A}^1_k$ *is dominant. Let $s \ge 1$ be any integer. Then there is a dense open subset $U \subset \mathbb{A}^r_k$ such that for each $g \in U$, there is an open neighborhood $\mathcal{W}_g \subset \mathbb{A}^r_{k(g)}$ of Σ such that $\phi^*_{g,s,1}(Z_{k(g)})$ restricted over $\mathcal{W}_g$ is empty.*

Proof. Since $\mathrm{pr}_2 : Z \to \mathbb{A}^1_k$ is not dominant and Z is irreducible, $\mathrm{pr}_2(Z)$ must be a singleton closed subset $\{t_0\}$. By the modulus condition that Z satisfies, we must have $t_0 \neq 0$ and $Z \subset \mathbb{A}^r_k \times \{t_0\} \times \square^{n-1}_k$. It is therefore sufficient to prove the lemma by replacing k by $k(t_0)$ and Σ by $\pi^{-1}_{t_0}(\Sigma)$, where $\pi_{t_0} : \mathrm{Spec}(k(t_0)) \to \mathrm{Spec}(k)$ is the base change. We can thus assume that $t_0 \in k^\times$. Consider the proper closed subset $\overline{q_n(Z)} \subset \mathbb{A}^r_k$ of dimension $< r$ and the dense open complement $\mathcal{U}_0 = \mathbb{A}^r_k \setminus \overline{q_n(Z)}$.

Because Z restricted over $\mathcal{U}_0$ is empty, we see that the translation $\phi^*_{g,s,1}(Z_{k(g)})$ restricted to the translation $\phi^*_{g,s,1}(\mathcal{U}_0)$ is empty for every $g \in \mathbb{A}^r_k$. Hence, it is enough to show that for an open subset $U \subset \mathbb{A}^r_k$, the set $\mathcal{W}_g := \phi^*_{g,s,1}(\mathcal{U}_0)$ contains Σ for each $g \in U$. However, this is evident because Σ is a finite set of closed point of $\mathbb{A}^r_k$ while $\mathcal{U}_0$ is a dense open subset of $\mathbb{A}^r_k$, and $\phi^*_{g,s,1}$ is translation by a nonzero constant factor $(t_0^{s(m+1)})$ of g. This proves the lemma. □

3D. ***Key lemmas.*** The key to our sfs-moving lemma for the localizations of $\mathbb{A}^r_k$ are the following two lemmas.

Let $W \subset \mathbb{A}^r_k \times B_n$ be a reduced closed subscheme and let $\overline{W}$ be its closure in $\mathbb{A}^r_k \times \overline{B}_n$ with reduced closed subscheme structure. We let $\overline{W}^o = \overline{W} \cap (\mathbb{A}^r_k \times \mathbb{G}_{m,k} \times \overline{\square}^{n-1})$. We fix a closed point $x \in \Sigma$ and integers $m, s \ge 1$.

Define

$$P_1 : \mathbb{A}^r_k \times \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1} \to \mathbb{A}^r_k,$$
$$P_2 : \mathbb{A}^r_k \times \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1} \to \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1}$$

to be the projection to the first factor, and the projection to the remaining factors. For a fixed $x \in \mathbb{A}^r_k$, define $\iota_x : \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1} \to \mathbb{A}^r_k \times \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1}$ to be the map $(g, t, \underline{y}) \mapsto (g, x + t^{s(m+1)}g, t, \underline{y})$. Let $\theta_x := P_2 \circ \iota_x$ and $\omega_{\overline{W},x} := (P_1 \circ \iota_x)|_{\theta_x^{-1}(\overline{W})}$, where $\theta_x^{-1}(\overline{W})$ is given its reduced induced closed subscheme structure. We then have the commutative diagram

$$\begin{array}{ccccccccccc}
 & & \theta_x^{-1}(\overline{W}^o) & \hookrightarrow & \mathbb{A}^r_k \times \mathbb{G}_{m,k} \times \square^{n-1} & \overset{\iota_x}{\hookrightarrow} & \mathbb{A}^r_k \times \mathbb{A}^r_k \times \mathbb{G}_{m,k} \times \square^{n-1} & \overset{P_2}{\longrightarrow} & \mathbb{A}^r_k \times \mathbb{G}_{m,k} \times \square^{n-1} & \hookleftarrow & \overline{W}^0 \\
 & \nearrow & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\theta_x^{-1}(W) & \hookrightarrow & \theta_x^{-1}(\overline{W}) & \hookrightarrow & \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1} & \overset{\iota_x}{\hookrightarrow} & \mathbb{A}^r_k \times \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1} & \overset{P_2}{\longrightarrow} & \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1} & \hookleftarrow & \overline{W} \\
 & & & & & & \downarrow P_1 & & & & \\
 & & & \omega_{\overline{W},x},\ \omega_{W,x} & & \searrow & \mathbb{A}^r_k, & & & &
\end{array} \tag{3-2}$$

where the top row's ι_x, P_2 are the restrictions of the second row, and $\omega_{W,x}$ is the natural composition. The vertical arrows are canonical open immersions. It is easy to check that ι_x is a closed immersion and θ_x is an isomorphism on the top row. Using (3-1) and (3-2), one immediately verifies the following observation which we shall use often.

Lemma 3.7. *Let $x \in \mathbb{A}^r_k$ be fixed. Then for each $g \in \mathbb{A}^r_k$, the map*

$$\omega^{-1}_{\overline{W},x}(g) \to \phi^*_{g,s,1}(\overline{W}),$$

$(g, t, \underline{y}) \mapsto (x, t, \underline{y})$, is an isomorphism. The same holds for W and $\overline{W}^o$ as well.

Another lemma we shall use is the following.

Lemma 3.8 [Bloch 1986, Lemma 1.2]. *Let X be an algebraic k-scheme and G a connected algebraic k-group acting on X. Let $A, B \subset X$ be closed subsets, and assume the fibers of the map $G \times A \to X$, $(g, a) \mapsto g \cdot a$ all have the same dimension, and that this map is dominant.*

Moreover, suppose that for an over-field $K \supset k$ and a K-morphism $\psi : X_K \to G_K$, there is a nonempty open subset $U \subset X$ such that for every $x \in U_K$, a scheme point, we have

$$\text{tr.}\deg_k k(\varphi \circ \psi(x), \pi(x)) \geq \dim(G),$$

where $\pi : X_K \to X_k$ and $\varphi : G_K \to G_k$ are the projection maps. Define $\phi : X_K \to X_K$ by $\phi(x) = \psi(x) \cdot x$ and suppose ϕ is an isomorphism. Then the intersection $\phi(A_K \cap U_K) \cap B_K$ is proper.

3E. ***Applications of the key lemmas.*** We apply the above two lemmas to our cycle Z and various other closed subsets associated to it. Let $\eta \in \mathbb{A}^r_k$ denote the generic point and let $K := k(\eta)$. We can regard $\eta \in \mathbb{A}^r_k(K)$. Apply Lemma 3.8 with

$$X = \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1}_k, \quad G = \mathbb{A}^r_k, \quad \psi(\underline{x}, t, \underline{y}) = (\eta)t^{s(m+1)}, \quad A = \Sigma \times \mathbb{A}^1_k \times \overline{\square}^{n-1}, \quad \text{and} \quad B = \overline{Z},$$

where G acts on $\mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1}$ by $g \cdot (\underline{x}, t, \underline{y}) = (g + \underline{x}, t, \underline{y})$. We let $\phi : X_K \to X_K$ be given by $\phi(\underline{x}, t, \underline{y}) = ((\eta)t^{s(m+1)} + \underline{x}, t, \underline{y})$. One checks immediately that the conditions of Lemma 3.8 are satisfied and we conclude that $\phi(A_K) \cap \overline{Z}_K$ has dimension at most zero. Comparing this with (3-2) and using Lemma 3.7, this is equivalent to saying that the generic fiber of $\omega_{\overline{Z},x}$ is finite for every $x \in \Sigma$.

It follows that if Z' is an irreducible component of $\theta_x^{-1}(\overline{Z})$, then either the map $\omega_{\overline{Z},x} : Z' \to \mathbb{A}^r_k$ is not dominant or it is dominant and generically quasifinite. In the dominant case, Chevalley's theorem on fiber dimensions (e.g., see [Hartshorne 1977, Exercise II-3.22, page 95]) tells us that we must have $\dim(Z') = r$ and $Z' \to \mathbb{A}^r_k$ is generically finite. In any case, it follows that there is a dense open subset of $\mathbb{A}^r_k$ over which $Z' \to \mathbb{A}^r_k$ is quasifinite (with possibly empty fibers).

By taking the finite intersection of such dense open subsets, running over all irreducible components of $\theta_x^{-1}(\overline{Z})$ and all $x \in \Sigma$, we conclude that there is a dense open subset $U \subset \mathbb{A}^r_k$ such that for each $x \in \Sigma$, the map $\omega^{-1}_{\overline{Z},x}(U) \to U$ is quasifinite. Using Lemma 3.7, equivalently we get:

Lemma 3.9. *For any integer $s \geq 1$, there is a dense open subset $U \subset \mathbb{A}^r_k$ such that for every $g \in U$, the set $(\Sigma \times \overline{B}_n)_{k(g)} \cap \phi^*_{g,s,1}(\overline{Z})_{k(g)} = (\Sigma \times \overline{B}_n)_{k(g)} \cap \overline{\phi^*_{g,s,1}(Z_{k(g)})}$ is finite.*

We can now show the following:

Lemma 3.10. *Let $s \gg 0$ be as in Lemma 3.1. Assume that Z is either dominant over $\mathbb{A}^r_k$ or restricts to zero on V. Then we can find a dense open $U \subset \mathbb{A}^r_k$ such that for $g \in U$, the scheme $\phi^*_{g,s,1}(Z)|_V$ is either empty or finite and surjective over V.*

Proof. We can assume $n \geq 2$ by Lemma 3.3. We let $U_1 \subset \mathbb{A}^r_k$ be the intersection of open subsets obtained in Lemmas 3.6 and 3.9. We can therefore assume that $\overline{\phi^*_{g,s,1}(Z_{k(g)})} \to \mathbb{A}^r_{k(g)}$ is dominant for all $g \in U_1$.

For $g \in U_1$, there is a commutative diagram

$$\begin{array}{ccc} \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)} \times \overline{\square}^{n-1}_{k(g)} & \xrightarrow{p_n} & \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)} \\ \downarrow{\scriptstyle \phi_{g,s,1}} & & \downarrow{\scriptstyle \phi_{g,s,1}} \\ \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)} \times \overline{\square}^{n-1}_{k(g)} & \xrightarrow{p_n} & \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)}, \end{array} \tag{3-3}$$

where the horizontal arrows are the projections.

If we let $W = p_n(\overline{Z}_{k(g)})$, it follows from Lemma 3.9 that the composite map $\overline{\phi^*_{g,s,1}(Z_{k(g)})} \to \phi^*_{g,s,1}(W) \to \mathbb{A}^r_{k(g)}$ is quasifinite over $\Sigma_{k(g)}$. Since $\overline{\phi^*_{g,s,1}(Z_{k(g)})} \to \mathbb{A}^r_{k(g)}$ is dominant by Lemma 3.4, it follows from Chevalley's theorem on fiber dimensions (see [Hartshorne 1977, Exercise II-3.22, page 95]) that there is an open neighborhood $U_g \subset \mathbb{A}^r_{k(g)}$ of $\Sigma_{k(g)}$ over which the map $\overline{\phi^*_{g,s,1}(Z_{k(g)})} \to \mathbb{A}^r_{k(g)}$ is quasifinite with nonempty fibers. We then get maps

$$\overline{\phi^*_{g,s,1}(Z_{k(g)})} \cap q_n^{-1}(U_g) \xrightarrow{p_n} \phi^*_{g,s,1}(W) \cap q^{-1}(U_g) \xrightarrow{q} U_g,$$

where the first map is projective and the composite map is quasifinite with nonempty fibers. This implies that the first map is also quasifinite, and hence, it is finite. Since $\bar{Z} \to W$ is dominant, so is the map $\overline{\phi^*_{g,s,1}(Z_{k(g)})} \to \phi^*_{g,s,1}(W)$ by (3-3). It follows that $\overline{\phi^*_{g,s,1}(Z_{k(g)})} \to \phi^*_{g,s,1}(W)$ is finite and surjective over U_g.

On the other hand, we have shown in Lemma 3.3 that $\phi^*_{g,s,1}(W) \to \mathbb{A}^r_{k(g)}$ is finite and surjective over $\mathbb{A}^r_k$ for our choice of $s \gg 0$ and $g \in \mathbb{A}^r_k \setminus \{0\}$. We conclude that there is an open neighborhood $U_g \subset \mathbb{A}^r_{k(g)}$ of $\Sigma_{k(g)}$ over which $\phi^*_{g,s,1}(\bar{Z}_{k(g)}) \to \mathbb{A}^r_{k(g)}$ is finite and surjective.

To show this property for $\phi^*_{g,s,1}(Z_{k(g)})$, we fix $x \in \Sigma$ and use the diagram (3-2) where we take $\overline{W} = Y := \bar{Z} \setminus Z$. To understand the generic fiber of $\omega_{Y,x}$, we apply Lemma 3.8 with

$$X = \mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1}_k,\ G = \mathbb{A}^r_k,\ \psi(\underline{x}, t, \underline{y}) = (\eta)t^{s(m+1)},\ A = \Sigma \times \mathbb{A}^1_k \times \overline{\square}^{n-1}_k,\ B = Y, \tag{3-4}$$

where G acts on $\mathbb{A}^r_k \times \mathbb{A}^1_k \times \overline{\square}^{n-1}_k$ by $g \cdot (\underline{x}, t, \underline{y}) = (g + \underline{x}, t, \underline{y})$ as before. One checks immediately that the conditions of Lemma 3.8 are satisfied. It follows that the intersection $\phi_{\eta,s,1}(A_{k(\eta)}) \cap B_{k(\eta)}$ is proper. By a dimension counting, this means that $\phi_{\eta,s,1}(A_{k(\eta)}) \cap B_{k(\eta)} = \varnothing$. Equivalently, we have $A_{k(\eta)} \cap \phi^*_{\eta,s,1}(Y_{k(\eta)}) = \varnothing$. We conclude by Lemma 3.7 that for every $x \in \Sigma$, the map $\omega_{Y,x} : \theta_x^{-1}(Y) \to \mathbb{A}^r_k$ is not dominant. We can therefore find a dense open subset $U \subset U_1 \subset \mathbb{A}^r_k$ such that the fiber of $\omega_{Y,x} : \theta_x^{-1}(Y) \to \mathbb{A}^r_k$ is empty over U for every $x \in \Sigma$. In other words, for every $g \in U$, the intersection $\phi^*_{g,s,1}(Y_{k(g)}) \cap A_{k(g)} = (\overline{\phi^*_{g,s,1}(Z_{k(g)})} \setminus \phi^*_{g,s,1}(Z_{k(g)})) \cap A_{k(g)}$ is empty. But this means that the map $\phi^*_{g,s,1}(Z_{k(g)}) \to \mathbb{A}^r_{k(g)}$ is finite and surjective over an affine neighborhood of $\Sigma_{k(g)}$ (see Lemma 2.10). □

Lemma 3.11. *Assume that $Z \in \mathrm{Tz}^n_\Sigma(\mathbb{A}^r_k, n; m)$ is an irreducible cycle such that $Z \to \mathbb{A}^r_k$ is finite and surjective over an affine neighborhood of Σ. We can then find $s \gg 0$ and a dense open subset $U \subset \mathbb{A}^r_k$ such that for each $1 \le j \le n$ and for each $g \in U$, the scheme $(\phi^*_{g,s,1}(Z_{k(g)}))^{(j)}$ is regular over an affine neighborhood of $\Sigma_{k(g)}$.*

Proof. We take $W = Z_{\mathrm{sing}}$, the singular locus of Z, in (3-2) and consider the map $\omega_{Z_{\mathrm{sing}},x} : \theta_x^{-1}(Z_{\mathrm{sing}}) \to \mathbb{A}^r_k$ for $x \in \Sigma$. We had seen previously that the map θ_x on the top row of (3-2) is an isomorphism. In particular, the map $\theta_x : \theta_x^{-1}(Z_{\mathrm{sing}}) \to Z_{\mathrm{sing}}$ is an isomorphism. But this implies that $\dim(\theta_x^{-1}(Z_{\mathrm{sing}})) = \dim(Z_{\mathrm{sing}}) \le r - 1$. It follows that the map $\omega_{Z_{\mathrm{sing}},x} : \theta_x^{-1}(Z_{\mathrm{sing}}) \to \mathbb{A}^r_k$ is not dominant. We can therefore find a dense open subset $U \subset \mathbb{A}^r_k$ such that the fibers of ω_x over U are empty. By shrinking U further, we can assume that this holds for all $x \in \Sigma$.

It follows from Lemma 3.7 that for every $g \in U$, the closed subscheme $(\phi^*_{g,s,1}(Z_{k(g)}))_{\mathrm{sing}} = \phi^*_{g,s,1}((Z_{k(g)})_{\mathrm{sing}}) = \phi^*_{g,s,1}((Z_{\mathrm{sing}})_{k(g)})$ does not meet $(\Sigma \times B_n)_{k(g)}$. Here, the last equality uses the perfectness of k. But this means that $\phi^*_{g,s,1}(Z_{k(g)})$ is regular at all points lying over $\Sigma_{k(g)}$. By choosing $s \gg 0$ as in Lemma 3.1, shrinking U further, and using Lemma 3.10, we can assume that $\phi^*_{g,s,1}(Z_{k(g)})$ is finite and surjective over an affine neighborhood of $\Sigma_{k(g)}$. But then $\phi^*_{g,s,1}(Z_{k(g)})$ must be regular over an affine neighborhood of $\Sigma_{k(g)}$.

Let $Z^{(j)} \subset \mathbb{A}^r_k \times B_j$ be the projection of Z to B_j as in Section 2E for $1 \le j \le n$. Since $Z \to \mathbb{A}^r_k$ is finite and surjective over an affine neighborhood of Σ, each $Z^{(j)}$ is also finite and surjective over an affine neighborhood of Σ. We can therefore repeat the above process successively for each $Z^{(j)}$ by shrinking U further each time. In the end, we get a dense open subset $U \subset \mathbb{A}^r_k$ such that each $1 \le j \le n$ and for each $g \in U$, the scheme $\phi^*_{g,s,1}(Z^{(j)}_{k(g)})$ is regular over a common affine neighborhood of $\Sigma_{k(g)}$. Since the diagram

$$\begin{array}{ccc} \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)} \times \square^{n-1}_{k(g)} & \xrightarrow{\pi_j} & \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)} \times \square^{j-1}_{k(g)} \\ {\scriptstyle \phi_{g,s,1}} \downarrow & & \downarrow {\scriptstyle \phi_{g,s,1}} \\ \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)} \times \overline{\square}^{n-1}_{k(g)} & \xrightarrow{\pi_j} & \mathbb{A}^r_{k(g)} \times \mathbb{A}^1_{k(g)} \times \square^{j-1}_{k(g)} \end{array} \tag{3-5}$$

commutes and the vertical maps are isomorphisms, it follows that $\phi^*_{g,s,1}(Z^{(j)}_{k(g)}) = (\phi^*_{g,s,1}(Z_{k(g)}))^{(j)}$. We have therefore shown that there is a dense open subset $U \subset \mathbb{A}^r_k$ such that for every $g \in U$ and $1 \le j \le n$, the scheme $(\phi^*_{g,s,1}(Z_{k(g)}))^{(j)}$ is regular over a common affine neighborhood of $\Sigma_{k(g)}$. This finishes the proof. $\square$

Lemma 3.12. *For every integer $s \ge 1$, there is a dense open subset $U \subset \mathbb{A}^r_k$ such that for every $g \in U$, one has $\phi^*_{g,s,1}(Z_{k(g)}) \cap (\Sigma \times \mathbb{A}^1_k \times F)_{k(g)} = \varnothing$ for every proper face F of $\square^{n-1}$.*

Proof. We let F be a proper face of $\square^{n-1}$ and let $W = Z \cap (\mathbb{A}^r_k \times \mathbb{A}^1_k \times F)$. We fix a point $x \in \Sigma$ and consider the diagram (see (3-2)):

$$\begin{array}{ccccccc} \theta_x^{-1}(W) & \hookrightarrow & \mathbb{A}^r_k \times \mathbb{G}_{m,k} \times F & \overset{\iota_x}{\hookrightarrow} & \mathbb{A}^r_k \times \mathbb{A}^r_k \times \mathbb{G}_{m,k} \times F & \xrightarrow{P_2} & \mathbb{A}^r_k \times \mathbb{G}_{m,k} \times F \hookleftarrow W \\ & {\scriptstyle \omega_{W,x}} \searrow & & & \downarrow {\scriptstyle P_1} & & \\ & & & & \mathbb{A}^r_k. & & \end{array} \tag{3-6}$$

As in (3-2), the map $\theta_x = P_2 \circ \iota_x$ is an isomorphism. Note also that (see Lemma 3.7) for any $g \in \mathbb{A}^r_k$, the map $\omega^{-1}_{W,x}(g) \to \phi^*_{g,s,1}(Z) \cap (\{x\} \times \mathbb{A}^1_k \times F)$, which sends $(g, t, \underline{y})$ to $(x, t, \underline{y})$, is an isomorphism. It follows therefore that the map $\omega_{W,x}$ is not dominant. Equivalently, there exists a dense open $U \subset \mathbb{A}^r_k$ such that the fibers of $\omega_{W,x}$ over U are empty. Shrinking U further if necessary, we can assume that this happens for all $x \in \Sigma$. It is clear that for every $g \in U$, the set $\phi^*_{g,s,1}(Z_{k(g)}) \cap (\Sigma \times \mathbb{A}^1_k \times F)_{k(g)}$ is empty. This proves the lemma. $\square$

3F. ***The proof of the moving lemma for affine spaces.*** We can now prove the main result of this section, the sfs-moving lemma for the localizations of $\mathbb{A}^r_k$. We begin with the following intermediate modification step.

Lemma 3.13. *Let k be an infinite field and let $\alpha \in \mathrm{Tz}^n(\mathbb{A}^r_k, n; m)$. Let $V = \mathrm{Spec}(\mathcal{O}_{\mathbb{A}^r_k, \Sigma})$ for a finite subset $\Sigma \subset \mathbb{A}^r_k$ of closed points, with the localization map $j : V \to \mathbb{A}^r_k$. Assume that $\partial(j^*(\alpha)) = 0$. Then there are cycles $\beta \in \mathrm{Tz}^n(\mathbb{A}^r_k, n; m)$ and $\gamma \in \mathrm{Tz}^n(\mathbb{A}^r_k, n+1; m)$ with $\partial(j^*(\gamma)) = j^*(\alpha) - j^*(\beta)$ such that each component of β is either dominant over $\mathbb{A}^r_k$ or restricts to zero on V.*

Proof. We choose an integer $s \gg 0$ which is at least as large as the integer $s(Z)$ and the one chosen in Lemmas 3.5 and 3.6 for every irreducible component Z of α. It follows from Lemma 3.2 that $\phi^*_{g,s}(\alpha)$ intersects all faces of $\square^n$ properly. Taking the face $F = \{1\} \times \square^{n-1}$ (and using the containment lemma [Krishna and Park 2016, Proposition 2.2]), we see that $\phi^*_{g,s,1}(\alpha) \in \mathrm{Tz}^n(\mathbb{A}^r_{k(g)}, n; m)$ for all $g \in \mathbb{A}^r_k$. We can also assume that $s \gg 0$ is large enough so that Lemma 3.2 holds also for each boundary of each component of α.

We let $U \subset \mathbb{A}^r_k$ be any dense open which is contained in the intersection of the ones given by Lemmas 3.5 and 3.6 for all irreducible components of $|\alpha|$. We let $g \in U(k)$ be any element. It follows by our choice of g that if Z is a component of α, then $\phi^*_{g,s,1}(Z)$ is either dominant over $\mathbb{A}^r_k$, or it restricts to zero on V, or satisfies conditions (1) and (2) of Lemma 3.5.

We now compute

$$\begin{aligned}\phi^*_{g,s} \circ \partial(\alpha) &= \phi^*_{g,s}\Bigl(\sum_{i=1}^{n-1}(-1)^i(\partial^1_i - \partial^0_i)(\alpha)\Bigr)\\ &=^{\dagger} \sum_{i=1}^{n-1}(-1)^i(\partial^1_{i+1} - \partial^0_{i+1})(\phi^*_{g,s}(\alpha))\\ &= -\sum_{i=2}^{n}(-1)^i(\partial^1_i - \partial^0_i)(\phi^*_{g,s}(\alpha)),\end{aligned}$$

where $=^{\dagger}$ follows from (3-1). On the other hand, we have

$$\begin{aligned}\partial \circ \phi^*_{g,s}(\alpha) &= \sum_{i=1}^{n}(-1)^i(\partial^1_i - \partial^0_i)(\phi^*_{g,s}(\alpha))\\ &= (-1)(\partial^1_1 - \partial^0_1)(\phi^*_{g,s}(\alpha)) + \sum_{i=2}^{n}(-1)^i(\partial^1_i - \partial^0_i)(\phi^*_{g,s}(\alpha)).\end{aligned}$$

It follows that $\partial(\phi^*_{g,s}(\alpha)) + \phi^*_{g,s}(\partial(\alpha)) = (\partial^0_1 - \partial^1_1)(\phi^*_{g,s}(\alpha)) = \alpha - \phi^*_{g,s,1}(\alpha)$. Lemma 3.2 says that $\phi^*_{g,s}(\alpha) \in \mathrm{Tz}^n(\mathbb{A}^r_k, n+1; m)$. If we let $\gamma = \phi^*_{g,s}(\alpha)$ and $\beta = \phi^*_{g,s,1}(\alpha)$, we see that $\partial(j^*(\gamma)) = j^*(\alpha) - j^*(\beta)$. It also follows that $\partial(j^*(\beta)) = 0$.

We now replace α by β in $\mathrm{Tz}^n(\mathbb{A}^r_k, n; m)$ and repeat the above process. It follows from Lemmas 3.4, 3.5 and 3.6 that after finite steps, we arrive at new cycles $\beta \in \mathrm{Tz}^n(\mathbb{A}^r_k, n; m)$ and $\gamma \in \mathrm{Tz}^n(\mathbb{A}^r_k, n+1; m)$ such that $\partial(j^*(\gamma)) = j^*(\alpha) - j^*(\beta)$. Moreover, each component of β is either dominant over $\mathbb{A}^r_k$ or restricts to zero on V. □

Theorem 3.14. *Let k be an infinite perfect field and let $\alpha \in \mathrm{Tz}^n(\mathbb{A}^r_k, n; m)$. Let $V = \mathrm{Spec}(\mathcal{O}_{\mathbb{A}^r_k, \Sigma})$ for a finite subset $\Sigma \subset \mathbb{A}^r_k$ of closed points, with the localization map $j : V \to \mathbb{A}^r_k$. Assume that $\partial(j^*(\alpha)) = 0$. Then there are cycles $\beta \in \mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$ and $\gamma \in \mathrm{Tz}^n(V, n+1; m)$ such that $\partial(\gamma) = j^*(\alpha) - \beta$.*

Proof. By applying Lemma 3.13 and removing those components of the resulting new cycle α which restrict to zero on V, we can assume that every component of α is dominant over $\mathbb{A}^r_k$. Note that this does not change $\partial(j^*(\alpha))$.

We now choose an integer $s \gg 0$ which is at least as large as the integer $s(Z)$ and the one chosen in Lemmas 3.10 and 3.11 for every irreducible component Z of α. It follows from Lemma 3.2 that $\phi^*_{g,s}(\alpha)$ intersects all faces of $\square^n$ properly and $\phi^*_{g,s,1}(\alpha) \in \mathrm{Tz}^n(\mathbb{A}^r_{k(g)}, n; m)$ for all $g \in \mathbb{A}^r_k$ (see the proof of Lemma 3.13). We can also assume that $s \gg 0$ is large enough so that Lemma 3.2 holds also for each boundary of each component of α.

We let $U \subset \mathbb{A}^r_k$ be any dense open which is contained in the intersection of the ones given by Lemmas 3.10, 3.11 and 3.12 for all irreducible components of α. Since U is rational and k is infinite, $U(k)$ is a dense subset of U. We let $g \in U(k)$ be any element. We claim that $j^*(\phi^*_{g,s,1}(\alpha)) \in \mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$, where $\phi^*_{g,s}(-)$ is defined on $\mathrm{Tz}^n(\mathbb{A}^r_k, n; m)$ by the usual linear extension. By Lemmas 3.10 and 3.11, we only need to show that $\phi^*_{g,s,1}(\alpha) \in \mathrm{Tz}^n_{\Sigma}(\mathbb{A}^r_k, n; m)$. But this is equivalent to showing that $(\Sigma \times \mathbb{A}^1_k \times F) \cap |\phi^*_{g,s,1}(\alpha)| = \varnothing$ for every proper face F of $\square^{n-1}$, which in turn follows from Lemma 3.12. The claim is thus proven.

A computation identical to the one in the proof of Lemma 3.13 shows that

$$\partial(\phi^*_{g,s}(\alpha)) + \phi^*_{g,s}(\partial(\alpha)) = (\partial^0_1 - \partial^1_1)(\phi^*_{g,s}(\alpha)) = \alpha - \phi^*_{g,s,1}(\alpha).$$

Lemma 3.2 says that $\phi^*_{g,s}(\alpha) \in \mathrm{Tz}^n(\mathbb{A}^r_k, n+1; m)$. If $\partial(j^*(\alpha)) = 0$, we can set $\gamma = j^* \circ \phi^*_{g,s}(\alpha)$ and $\beta = j^*(\phi^*_{g,s,1}(\alpha))$. We get $\partial(\gamma) = j^*(\alpha) - \beta$ and we have shown above that $\beta \in \mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$. The theorem is now proven. □

Remark 3.15. The proof of Theorem 3.14 (where we take $n \geq 2$, replace B_n by $\square^{n-1}$ and take $s = 0$ everywhere in the proof) also shows that if $n \geq 1$ and $\alpha \in z^n(\mathbb{A}^r_k, n)$ is a higher Chow cycle with $\partial(j^*(\alpha)) = 0$, then we can find $\gamma \in z^n(V, n+1)$ and $\beta \in z^n_{\mathrm{sfs}}(V, n)$ such that $\partial(\gamma) = j^*(\alpha) - \beta$. Note that $n = 0$ case of this result is trivial.

4. The fs-property of residual cycles

Let k be an infinite perfect field. In this section, we discuss some results on linear projections in projective spaces, and show how these projections can be used to equip the residual cycle of a given cycle with certain finiteness properties over the base scheme. The main result of Section 4 is Theorem 4.15. It will be used later in proving the fs-moving lemma (see Lemma 8.7), a precursor to the final sfs-moving lemma.

For $0 \leq n < N$ and a linear subspace $H \subset \mathbb{P}^N_k$ defined over k, let $\mathrm{Gr}(n, H)$ be the Grassmannian scheme of n-dimensional linear subspaces of $\mathbb{P}^N_k$ contained in H. This is a homogeneous space of dimension

$(\dim(H)-n)(n+1)$. Unless we specify the field of definition, a linear subspace of $\mathbb{P}^N_k$ will mean a k-linear subspace.

Given two closed subschemes $Y, Y' \subset \mathbb{P}^N_k$, let $\mathrm{Sec}(Y, Y') \subset \mathbb{P}^N_k$ be the union of all lines $\ell_{yy'}$ joining distinct points $y \in Y$, $y' \in Y'$. In general, we have $\dim(\mathrm{Sec}(Y, Y')) \le \dim(Y)+\dim(Y')+1$. If $Y = Y'$, the scheme $\mathrm{Sec}(Y, Y') = \mathrm{Sec}(Y)$ is the secant variety of Y. If $Y' = L$ is a linear subspace, then $\mathrm{Sec}(Y, L) = C_L(Y)$ is the cone over Y with vertices in L.

4A. ***Containment and avoidance.*** Let $0 \le m \le n < N$ be integers and let $S, T \subset \mathbb{P}^N_k$ be two disjoint subsets.

Definition 4.1. We denote the set of n-dimensional linear subspaces of $\mathbb{P}^N_k$ containing S by $\mathrm{Gr}_S(n, \mathbb{P}^N_k)$. We write $\mathrm{Gr}_S(n, \mathbb{P}^N_k)$ as $\mathrm{Gr}_x(n, \mathbb{P}^N_k)$ if $S = \{x\}$ is a closed point. We denote the set of n-dimensional linear subspaces of $\mathbb{P}^N_k$ which do not intersect S by $\mathrm{Gr}(S, n, \mathbb{P}^N_k)$. If $S = \{x\}$, we write $\mathrm{Gr}(S, n, \mathbb{P}^N_k)$ as $\mathrm{Gr}(x, n, \mathbb{P}^N_k)$. We let $\mathrm{Gr}_S(T, n, \mathbb{P}^N_k) := \mathrm{Gr}_S(n, \mathbb{P}^N_k) \cap \mathrm{Gr}(T, n, \mathbb{P}^N_k)$. For any linear subspace $L \subset \mathbb{P}^N_k$, we define $\mathrm{Gr}_S(n, L)$ and $\mathrm{Gr}(T, n, L)$ similarly.

One checks that, when $M \subset \mathbb{P}^N_k$ is a linear subspace of dimension m, then $\mathrm{Gr}_M(n, \mathbb{P}^N_k)$ is a homogeneous space which is an irreducible closed subscheme of $\mathrm{Gr}(n, \mathbb{P}^N_k)$ of dimension $(N-n)(n-m)$. The following result is elementary. We leave the proof as an exercise.

Lemma 4.2. *Let $N > n$.* (1) *If $S' \subset S$, then* $\mathrm{Gr}(S, n, \mathbb{P}^N_k) \subset \mathrm{Gr}(S', n, \mathbb{P}^N_k)$. (2) *For any finite closed set $S \subset \mathbb{P}^N_k$, $\mathrm{Gr}(S, n, \mathbb{P}^N_k) \subset \mathrm{Gr}(n, \mathbb{P}^N_k)$ is a dense open subset.*

Lemma 4.3. *Let $X \subset \mathbb{P}^N_k$ be a closed subscheme of dimension $r \ge 1$ with $N \gg r$ and let $H \subset \mathbb{P}^N_k$ be a hyperplane, not containing any irreducible component of X. Then $\mathrm{Gr}(X, N-r-1, H)$ is a dense open subset of $\mathrm{Gr}(N-r-1, H)$.*

Proof. Consider the incidence scheme $S = \{(x, L) \in X \times \mathrm{Gr}(N-r-1, H) \mid x \in L\}$. We have the obvious projection maps $X \xleftarrow{\pi_1} S \xrightarrow{\pi_2} \mathrm{Gr}(N-r-1, H)$.

Each fiber of π_1 over $X \setminus (X \cap H)$ is empty. It is a smooth morphism over $X \cap H$ with its fiber over $x \in X \cap H$ to be $\mathrm{Gr}_x(N-r-1, H)$, whose dimension is $((N-1)-(N-r-1))(N-r-1-0) = r(N-r-1)$. It follows that $\dim(S) \le \dim(X \cap H) + \dim \mathrm{Gr}_x(N-r-1, H) = r-1+r(N-r-1) = r(N-r)-1$. Thus, $\pi_2(S)$ is a closed subscheme of $\mathrm{Gr}(N-r-1, H)$ of dimension $\le r(N-r)-1$ which is less than $\dim \mathrm{Gr}(N-r-1, H) = r(N-r)$. Hence, $\mathrm{Gr}(X, N-r-1, H) = \mathrm{Gr}(N-r-1, H) \setminus \pi_2(S)$ is a dense open subset. □

4B. ***Transverse intersection.*** For a reduced scheme X, let $X_{\mathrm{sing}} \subset X$ be the singular locus of X and let X_{sm} be its complement. For a closed subscheme $X \subset \mathbb{P}^N_k$, let $\mathrm{Gr}^{\mathrm{tr}}(X, n, \mathbb{P}^N_k)$ denote the set of n-dimensional linear subspaces which *do not* intersect X_{sing}, and whose intersection with X_{sm} is transverse (if not empty). We let

$$\mathrm{Gr}^{\mathrm{tr}}(X,S,n,\mathbb{P}^N_k) = \mathrm{Gr}(S,n,\mathbb{P}^N_k) \cap \mathrm{Gr}^{\mathrm{tr}}(X,n,\mathbb{P}^N_k) \quad \text{and} \quad \mathrm{Gr}^{\mathrm{tr}}_S(X,n,\mathbb{P}^N_k) = \mathrm{Gr}_S(n,\mathbb{P}^N_k) \cap \mathrm{Gr}^{\mathrm{tr}}(X,n,\mathbb{P}^N_k).$$

For a linear subspace $H \subset \mathbb{P}^N_k$, we define $\mathrm{Gr}^{\mathrm{tr}}(X, S, n; H)$ and $\mathrm{Gr}^{\mathrm{tr}}_S(X, n; H)$ similarly.

Lemma 4.4. *Let $r \geq 2$ be an integer and suppose $N \gg r$. Let $H \subset \mathbb{P}^N_k$ be a hyperplane. Let $L \subset \mathbb{P}^N_k$ be a linear subspace of dimension $N - r + 1$ intersecting H transversely and let $X \subset L$ be a curve (not necessarily connected) none of whose components is contained in H. Then the set of linear subspaces in* $\mathrm{Gr}^{\mathrm{tr}}(L, X, N-2, H)$ *is a dense open subset of* $\mathrm{Gr}(N-2, H)$.

Proof. Observe that $\mathrm{Gr}^{\mathrm{tr}}(L, N-2, H)$ is a dense open subset of $\mathrm{Gr}(N-2, H)$. Consider the map $\nu_L : \mathrm{Gr}^{\mathrm{tr}}(L, N-2, H) \to \mathrm{Gr}(N-r-1, L \cap H)$ given by $\nu_L(M) = L \cap M$. This ν_L is a smooth surjective morphism of relative dimension $2(r-1)$. It follows from Lemma 4.3 that $\mathrm{Gr}(X, N-r-1, L \cap H)$ is a dense open subset of $\mathrm{Gr}(N-r-1, L \cap H)$, so $\nu_L^{-1}(\mathrm{Gr}(X, N-r-1, L \cap H))$ is a dense open subset of $\mathrm{Gr}^{\mathrm{tr}}(L, N-2, H)$, and hence a dense open subset of $\mathrm{Gr}(N-2, H)$. □

4C. ***Affine Veronese embedding and linear projection.*** Recall that for positive integers $m, d \geq 1$, the Veronese embedding $\upsilon_{m,d} : \mathbb{P}^m_k \hookrightarrow \mathbb{P}^N_k$ is a closed embedding given by $\upsilon_{m,d}([\underline{x}]) = [M_0(\underline{x}), \dots, M_N(\underline{x})] = [M(\underline{x})]$, where $N = \binom{m+d}{m} - 1$ and $\{M_0, \dots, M_N\}$ are all monomials in $\{x_0, \dots, x_m\}$ of degree d, arranged in the lexicographic order.

If $[y_0, \dots, y_N] \in \mathbb{P}^N_k$ denotes the projective coordinates, it is clear that $\upsilon_{m,d}^{-1}(\{y_0 = 0\}) = \{x_0^d = 0\}$. In particular, the Veronese embedding yields Cartesian squares

$$\begin{array}{ccccc} \mathbb{A}^m_k & \longrightarrow & \mathbb{P}^m_k & \longleftarrow & dH_{m,0} \\ \downarrow{\scriptstyle \upsilon_{m,d}} & & \downarrow{\scriptstyle \upsilon_{m,d}} & & \downarrow{\scriptstyle \upsilon_{m,d}} \\ \mathbb{A}^N_k & \longrightarrow & \mathbb{P}^N_k & \longleftarrow & H_{N,0}, \end{array} \tag{4-1}$$

where $H_{m,0} \subset \mathbb{P}^m_k$ is the hyperplane $\{x_0 = 0\}$ and the vertical arrows are all closed embeddings. The closed embedding $\upsilon_{m,d} : \mathbb{A}^m_k \hookrightarrow \mathbb{A}^N_k$ is given by $\upsilon_{m,d}(y_1, \dots, y_m) = (M'_1, \dots, M'_N)$, where $\{M'_1, \dots, M'_N\}$ is the induced ordered set of all monomials in $\{y_1, \dots, y_m\}$ of degree bounded by d.

Let $1 \leq r < N$ be two integers. Recall (e.g., see [Krishna and Park 2012, Lemma 6.1]) that when $L \subset \mathbb{P}^N_k$ is a linear subspace of dimension $N - r - 1$, there is an associated projection map $\phi_L : \mathbb{P}^N_k \setminus L \to \mathbb{P}^r_k$, where $\mathbb{P}^r_k$ is a linear subspace of $\mathbb{P}^N_k$ such that $L \cap \mathbb{P}^r_k = \varnothing$. This map ϕ_L defines a vector bundle over $\mathbb{P}^r_k$ of rank $N - r$, whose fiber over a point $x \in \mathbb{P}^r_k$ is the affine space $C_x(L) \setminus L$, where $C_x(L) = \mathrm{Sec}(\{x\}, L)$.

Remark 4.5. The referee asked whether the above vector bundle $\phi_L : \mathbb{P}^N_k \setminus L \to \mathbb{P}^r_k$ is isomorphic to $\mathcal{O}_{\mathbb{P}^r_k}(1)^{\oplus(N-r)}$. Indeed, ϕ_L is (up to an isomorphism) the projection map of quotient stacks $\phi_L : [((\mathbb{A}^{r+1} \setminus \{0\}) \times_k V)/\mathbb{G}_m] \to [(\mathbb{A}^{r+1} \setminus \{0\})/\mathbb{G}_m]$, where $V = k^{N-r}$ and the $\mathbb{G}_m$-action everywhere is by scalar multiplication. Since $[((\mathbb{A}^{r+1} \setminus \{0\}) \times_k V)/\mathbb{G}_m] \cong [(\mathbb{A}^{r+1} \setminus \{0\})/\mathbb{G}_m] \times_{B\mathbb{G}_m} [V/\mathbb{G}_m]$, one identifies ϕ_L with the map $\mathbb{P}^r_k \times_{B\mathbb{G}_m} \pi^*(V(1)^{\oplus(N-r)}) \to \mathbb{P}^r_k$, where $V(1)$ is the line bundle on $B\mathbb{G}_m := [\mathrm{Spec}(k)/\mathbb{G}_m]$ associated to the 1-dimensional $\mathbb{G}_m$-representation given by the scalar multiplication on k, and $\pi : [(\mathbb{A}^{r+1} \setminus \{0\})/\mathbb{G}_m] \to B\mathbb{G}_m$ is the canonical projection.

Note that in general, if we let $\mathbb{G}_m$ act on k by weight $n \in \mathbb{Z}$ (i.e., $\lambda \cdot x = \lambda^n x$) and let $V(n)$ denote the corresponding line bundle on $B\mathbb{G}_m$, then $\pi^*(V(n))$ is isomorphic to $\mathcal{O}_{\mathbb{P}^r_k}(n)$. Hence the above $\pi^*(V(1)^{\oplus(N-r)})$ is isomorphic to $\mathcal{O}_{\mathbb{P}^r_k}(1)^{\oplus(N-r)}$, as wished.

Definition 4.6. Recall that if $X \subset \mathbb{P}^N_k$ is a closed subscheme with $X \cap L = \varnothing$, then ϕ_L restricted to X defines a projection $\phi_L : \phi_L|_X : X \to \mathbb{P}^r_k$. We call it the linear projection of X away from L. Since this is a morphism of projective schemes with affine fibers, it must be a finite morphism. In particular, $\dim(X) \le r$.

We shall use the following situation often: let $H \subset \mathbb{P}^N_k$ be a hyperplane containing L and $X \subset \mathbb{P}^N_k$ a closed subscheme with $X \cap L = \varnothing$ and $X \not\subset H$. Then ϕ_L defines the Cartesian squares of morphisms

$$\begin{array}{ccccc} X \setminus H & \longrightarrow & X & \longleftarrow & X \cap H \\ \downarrow & & \downarrow & & \downarrow \\ \mathbb{P}^r_k \setminus H & \longrightarrow & \mathbb{P}^r_k & \longleftarrow & \mathbb{P}^r_k \cap H. \end{array} \tag{4-2}$$

Together with (4-1), we deduce the following fact, which we use often:

Lemma 4.7. *Let $X \hookrightarrow \mathbb{A}^m_k$ be an affine scheme of dimension $r \ge 1$ and let $\overline{X} \hookrightarrow \mathbb{P}^m_k$ be its projective closure. Then, for $d \ge 1$, the Veronese embedding $\upsilon_{m,d} : \mathbb{P}^m_k \hookrightarrow \mathbb{P}^N_k$ and the linear projection away from $L \in \mathrm{Gr}(N-r-1, \mathbb{P}^N_k)(k)$ yield a Cartesian diagram with finite vertical maps*

$$\begin{array}{ccc} X & \longrightarrow & \overline{X} \\ {\scriptstyle\phi_L}\downarrow & & \downarrow{\scriptstyle\phi_L} \\ \mathbb{A}^r_k & \longrightarrow & \mathbb{P}^r_k \end{array} \tag{4-3}$$

if $L \in \mathrm{Gr}(\overline{X}, N-r-1, H_{N,0})(k)$, where $H_{N,0} = \{y_0 = 0\} \subset \mathbb{P}^N_k$ as in (4-1).

4D. ***The Set-up.*** Let k be an infinite perfect field. Here, we introduce the basic Set-up that will be used for most of the paper. This set of assumptions will be referred to as the Set-up of Section 4D.

(1) *The spaces*: Let X be an equidimensional reduced projective k-scheme of dimension $r \ge 1$ with a given closed embedding $\eta : X \hookrightarrow \mathbb{P}^N_k$ with $N \gg r$ and of degree $d+1 \gg 0$. We let $\hat{B}$ be a smooth projective geometrically integral k-scheme of positive dimension and let $B \subset \hat{B}$ be a nonempty affine open subset with $F := \hat{B} \setminus B$. Let $\Sigma \subset X_{\mathrm{sm}}$ be a finite set of closed points.

(2) *The linear projections*: Suppose that $H \subset \mathbb{P}^N_k$ is a hyperplane not meeting Σ, and that $X \setminus (X \cap H) \subset X_{\mathrm{sm}}$. For $L \in \mathrm{Gr}(X, N-r-1, H)(k)$, let $\phi_L : X \to \mathbb{P}^r_k$ be the linear projection away from L. If L is fixed in a given context, we often drop it from the notation of ϕ_L and write as $\phi : X \to \mathbb{P}^r_k$. We write $\hat{\phi} = \hat{\phi}_L = \phi_L \times \mathrm{id}_{\hat{B}} : X \times \hat{B} \to \mathbb{P}^r_k \times \hat{B}$.

(3) *The cycles*: Let $Z \subset X \times \hat{B}$ be a reduced closed subscheme with irreducible components $\{Z_1, \dots, Z_s\}$, each of dimension r. We suppose that both $X \times F$ and $H \times \hat{B}$ intersect properly with each irreducible component of Z. We let $\hat{f} : Z \to X$ and $\hat{g} : Z \to \hat{B}$ denote the restrictions of the projection maps. Let $E \subset \hat{B}$ be a closed subset containing F such that no component of Z is contained in $\hat{g}^{-1}(E)$. We suppose that each projection $Z_i \to \hat{B}$ is nonconstant.

(4) *The residual schemes and residual sets*: Let $L^+(Z)$ be the closure of $\hat{\phi}^{-1}(\hat{\phi}(Z)) \setminus Z$ in $X \times \hat{B}$ with the reduced closed subscheme structure. For any closed point $x \in \overline{X}$, we write $L^+(\{x\})$ as $L^+(x)$. We let $L^+(\Sigma) = \bigcup_{x \in \Sigma} L^+(x)$.

4E. ***A Nisnevich property of linear projections.*** The first result on "moving" our cycle Z is the following:

Lemma 4.8. *We are under the Set-up of Section 4D. After replacing the embedding $\eta : X \hookrightarrow \mathbb{P}^N_k$ by a bigger one via a Veronese embedding if necessary, there exists a dense open subset $\mathcal{U} \subset \mathrm{Gr}(X, N-r-1, H)$ such that each $L \in \mathcal{U}(k)$ satisfies the following*:

(1) *ϕ_L is étale at Σ.*

(2) *$\phi_L(x) \neq \phi_L(x')$ for each pair of distinct points $x \neq x' \in \Sigma$.*

(3) *$k(\phi_L(x)) \xrightarrow{\sim} k(x)$ for all $x \in \Sigma$.*

(4) *$L^+(x) \neq \varnothing$ for all $x \in \Sigma$.*

(5) *$L^+(x) \cap \hat{f}(\hat{g}^{-1}(E)) = \varnothing$ for all $x \in \Sigma$.*

(6) *$L^+(x) \cap \hat{f}(Z_i) = \varnothing$ for all $x \in \Sigma$ if $\hat{f} : Z_i \to X$ is not dominant over any irreducible component of X.*

Proof. Replacing the given embedding $X \hookrightarrow \mathbb{P}^N_k$ by its composition with a Veronese embedding, we may begin with a closed embedding $X \hookrightarrow \mathbb{P}^N_k$ such that $N \gg r$ and the degree of X in $\mathbb{P}^N_k$ is bigger than one.

Step 1. *First suppose that k is algebraically closed.* It follows from our assumption that $\dim(\hat{g}^{-1}(E)) \leq r-1$. Since $\hat{f}$ is projective, it follows that $\hat{f}(\hat{g}^{-1}(E))$ is a closed subset of X of dimension at most $r-1$. We let $W \subset X$ be the union of X_{sing}, $\hat{f}(\hat{g}^{-1}(E))$ and the images of all components of Z which are not dominant over X. This is a closed subset of X such that $\dim(W) \leq r-1$. In particular, $\dim(\mathrm{Sec}(D_1, W \cup D_2)) \leq r$ for any finite closed subsets $D_1, D_2 \subset X$. It follows from Lemma 4.3 that $\mathcal{U}_1 := \bigcap_{x \in \Sigma} \mathrm{Gr}(X \cup \mathrm{Sec}(\{x\}, W \cup (\Sigma \setminus \{x\})), N-r-1, H)$ is dense open in $\mathrm{Gr}(N-r-1, H)$. Furthermore, any $L \in \mathcal{U}_1(k)$ satisfies (5) and (6) by construction.

We continue the proof of the rest of the properties. Let $T_{\Sigma,X} \subset \mathbb{P}^N_k$ be the union of the tangent spaces to X at all points of Σ. Since $\Sigma \subset X_{\mathrm{sm}}$, we have $T_{\Sigma,X} = T_{\Sigma,X_{\mathrm{sm}}}$, which is a finite union of linear subspaces of dimension r. For each $x \in \Sigma$, the set $\mathcal{Z}_x = X \cup T_{\Sigma,X} \cup \mathrm{Sec}(\{x\}, X_{\mathrm{sing}} \cup (\Sigma \setminus \{x\}))$ is closed in $\mathbb{P}^N_k$ of dimension r. Therefore, the set $\mathcal{U} = \bigcap_{x \in \Sigma} \mathrm{Gr}(\mathcal{Z}_x, N-r-1, H) \cap \mathcal{U}_1$ is dense open in $\mathrm{Gr}(N-r-1, H)$ by Lemma 4.3. By construction, each $L \in \mathcal{U}(k)$ defines the finite surjective map $\phi_L : X \to \mathbb{P}^r_k$, which is unramified at Σ and separates the points of Σ. In particular, (2) holds.

Since X_{sm} is regular and dense in X, it follows that $\phi_L|_{X_{\mathrm{sm}}} : X_{\mathrm{sm}} \to \mathbb{P}^r_k$ is a dominant and quasifinite morphism between regular k-schemes. In particular, the map $\mathcal{O}_{\mathbb{P}^r_k, \phi_L(x)} \to \mathcal{O}_{X,x}$ is a local homomorphism of regular local rings with the finite closed fiber for each $x \in \Sigma$. It follows from [EGA IV$_2$ 1965, Proposition (6.1.5), page 136] (or [Matsumura 1986, Theorem 23.1, page 179]) that ϕ_L is flat at each point of Σ. Hence ϕ_L is étale at Σ, being flat and unramified, proving (1).

Since $k=\bar{k}$, the isomorphisms of the residue fields, (3) is evident. Property (4) follows because $\deg(\phi_L)>1$ by the assumptions on the chosen Veronese embedding of X. This proves the lemma in Step 1 when k is algebraically closed.

Step 2. *Now suppose that k is any infinite perfect field.* Let $\bar{k}$ be an algebraic closure. For any k-scheme A, let $\pi_A : A_{\bar{k}} \to A$ be the base change to $\bar{k}$. We have that $\Sigma_{\bar{k}} = \pi_X^{-1}(\Sigma)$ is still a finite closed set of the regular scheme $X_{\mathrm{sm},\bar{k}}$. By Step 1 applied to $X_{\bar{k}}$, $H_{\bar{k}}$ and $\Sigma_{\bar{k}}$, there exists a dense open $\mathcal{U}' \subset \mathrm{Gr}(N-r-1, H_{\bar{k}})$ where the mentioned properties (1)–(6) hold.

Since k is perfect, there exists a finite Galois extension $k \subset k'$ in $\bar{k}$ such that $\mathcal{U}'$ is defined over k'. Let $\mathcal{U} := \bigcap_{\sigma\in \mathrm{Gal}(k'/k)} \sigma\cdot\mathcal{U}'$. This is a nonempty open subset defined over the radical closure of k in k', but since k is perfect, this radical closure is equal to k. Hence $\mathcal{U} \subset \mathrm{Gr}(N-r-1, H)$ and it is defined over k (see [Colliot-Thélène et al. 1997, Lemma 3.4.3]). Here we have $\mathcal{U}_{\bar{k}} \subset \mathcal{U}'$. Now, for each $L \in \mathcal{U}(k)$, we have $X \cap L = \varnothing$ by our choice of the open set. So, we get a finite surjective map $\phi_L : X \to \mathbb{P}^r_k$ over k.

We prove that ϕ_L is étale at each point $x \in \Sigma$. Let $y := \phi_L(x)$. By the faithfully flat descent [EGA IV$_4$ 1967, Corollaire (17.7.3)(ii), page 72], the map $\phi_L : \mathrm{Spec}(\mathcal{O}_{X,x}) \to \mathrm{Spec}(\mathcal{O}_{\mathbb{P}^r_k,y})$ is étale if and only if its faithfully flat base change $\phi_{L,\bar{k}} : \mathrm{Spec}(\mathcal{O}_{X_{\bar{k}},x_{\bar{k}}}) \to \mathrm{Spec}(\mathcal{O}_{\mathbb{P}^r_{\bar{k}},y_{\bar{k}}})$ of the semilocal schemes via $\mathrm{Spec}(\bar{k}) \to \mathrm{Spec}(k)$ is étale. Here, $x_{\bar{k}} := \pi_X^{-1}(x)$ and $y_{\bar{k}} := \pi_{\mathbb{P}^r_k}^{-1}(y)$. But Step 1 shows that the latter map $\phi_{L,\bar{k}}$ is étale at each point of the set $x_{\bar{k}} \subset \Sigma_{\bar{k}}$, thus so is the former ϕ_L at x. This proves (1).

Since $\phi_{L,\bar{k}}$ separates the points of $\Sigma_{\bar{k}}$ by construction, (2) is obvious. Furthermore, this shows that for each $x \in \Sigma$, the map $\phi_{L,\bar{k}} : \pi_X^{-1}(x) \to \pi_{\mathbb{P}^r_k}^{-1}(y)$ is injective, where $y = \phi_L(x)$. Hence by Lemma 4.9 below, we have $k(x) = k(y)$, which proves (3). Property (4) is evident because $\deg(\phi_L) > 1$ and $k(\phi(x)) \simeq k(x)$ for each $x \in \Sigma$ by (3).

Conditions (5) and (6) are apparent for any $L \in \mathcal{U}(k)$ because for every $x \in \Sigma$, we have that $(L_{\bar{k}})^+(x') \cap \hat{f}_{\bar{k}}(\hat{g}_{\bar{k}}^{-1}(E_{\bar{k}})) = \varnothing = (L_{\bar{k}})^+(x') \cap \hat{f}_{\bar{k}}(Z_{i,\bar{k}})$ for all x' lying in the finite set $\pi_X^{-1}(x) \subset \Sigma_{\bar{k}}$. Note here that if Z_i is not dominant over a component of X, then no component of $Z_{i,\bar{k}}$ can be dominant over any component of $X_{\bar{k}}$. This finishes the proof of the lemma. □

We used the following in the middle of the proof of the above Lemma 4.8.

Lemma 4.9. *Let k be an infinite perfect field and let $\phi : X \to Y$ be a finite morphism of k-schemes. Consider the base change Cartesian square*:

$$\begin{array}{ccc} X_{\bar{k}} & \xrightarrow{\phi_{\bar{k}}} & Y_{\bar{k}} \\ \downarrow{\scriptstyle \pi_X} & & \downarrow{\scriptstyle \pi_Y} \\ X & \xrightarrow{\phi} & Y \end{array} \tag{4-4}$$

Let $x \in X$ be a closed point and let $y := \phi(x)$. Then one has $|\pi_Y^{-1}(y)| \le |\pi_X^{-1}(x)|$. The equality holds if and only if $[k(x):k(y)] = 1$. Furthermore, this equality holds when the map $\phi_{\bar{k}} : \pi_X^{-1}(x) \to \pi_Y^{-1}(y)$ is injective.

Proof. Since k is perfect, we have $|\pi_X^{-1}(x)| = [k(x):k]$ and $|\pi_Y^{-1}(y)| = [k(y):k]$. So, the field extensions $k \hookrightarrow k(y) \hookrightarrow k(x)$ imply the first and the second assertions. If the map $\phi_{L_{\bar{k}}} : \pi_X^{-1}(x) \to \pi_Y^{-1}(y)$ is injective, then $|\pi_Y^{-1}(y)| \geq |\pi_X^{-1}(x)|$. The last part of the lemma thus follows. □

4F. ***Some algebraic results.*** We discuss some algebraic results that will be needed.

Lemma 4.10. *Let $f : A \to B$ be an injective finite unramified local homomorphism of noetherian local rings that induces an isomorphism of the residue fields. Then f is an isomorphism.*

Proof. Let $\mathfrak{m}_A$ and $\mathfrak{m}_B$ be the maximal ideals of A and B, respectively. Since f is finite, to show that f is surjective it suffices to show that $A/\mathfrak{m}_A \to B/(\mathfrak{m}_A B)$ is surjective by Nakayama's lemma. But this follows because the map $A/\mathfrak{m}_A \to B/\mathfrak{m}_B$ is an isomorphism and so is the map $B/(\mathfrak{m}_A B) \to B/\mathfrak{m}_B$ as f is unramified. □

Lemma 4.11. *Let $f : Y' \to Y$ be a finite surjective morphism of regular k-schemes. Let $W \subset Y$ be an irreducible closed subset and let $y \in W$ be a closed point. Let $S = f^{-1}(y)$ and $W' = f^{-1}(W)$. Let $x \in S$ and let $Z \subset W'$ be an irreducible component passing through x. Suppose that f is étale at x and $k(y) \xrightarrow{\simeq} k(x)$. Then $Z \cap S = \{x\}$ if and only if Z is the only component of W' passing through x.*

Proof. We first observe that f must be a flat morphism (see [Hartshorne 1977, Exercise III-10.9, page 276]). We next note that any irreducible component of W' that passes through x will be in the connected component of Y' containing x. So, we may assume Y' is connected. On the other hand, $W \subset Y$ being irreducible, it must belong to a unique connected component of Y. Hence, we may also assume that Y is connected.

Now, first suppose $S = \{x\}$. We claim that f is an isomorphism locally around y, so that the lemma holds trivially. Indeed, it follows from Lemma 4.10 that the map $\mathcal{O}_{Y,y} \to \mathcal{O}_{Y',x}$ is an isomorphism. This implies that f is a finite and flat map with $[k(Y') : k(Y)] = 1$ (see [Liu 2002, Exercise 5.1.25, page 176]) and hence must be an isomorphism.

We now suppose $|S| > 1$. Consider the commutative diagram of semilocal rings

$$\begin{array}{ccccc}
\mathcal{O}_{Y,y} & \xrightarrow{\alpha_1} & \mathcal{O}_{Y',S} & \xrightarrow{\alpha_2} & \mathcal{O}_{Y',x} \\
\downarrow{\scriptstyle \beta_1} & & \downarrow{\scriptstyle \beta_2} & & \downarrow{\scriptstyle \beta_3} \\
\mathcal{O}_{W,y} & \xrightarrow{\alpha_3} & \mathcal{O}_{W',S} & \xrightarrow{\alpha_4} & \mathcal{O}_{W',x} \\
 & \searrow{\scriptstyle \gamma} & \downarrow{\scriptstyle \beta_4} & & \downarrow{\scriptstyle \beta_5} \\
 & & \mathcal{O}_{Z,S} & \xrightarrow{\alpha_5} & \mathcal{O}_{Z,x},
\end{array} \qquad \gamma' : \mathcal{O}_{W,y} \to \mathcal{O}_{Z,x} \tag{4-5}$$

where $\gamma := \beta_4 \circ \alpha_3$ and $\gamma' := \alpha_5 \circ \gamma$. Here, α_1 and α_3 are finite and flat, and $\alpha_2 \circ \alpha_1$ is étale. The lemma is equivalent to that α_5 is an isomorphism if and only if β_5 is.

Suppose α_5 is an isomorphism. Since β_4 is surjective and α_3 is finite, the map γ is finite. Thus, γ' is a finite map of local rings. Since $\alpha_2 \circ \alpha_1$ is étale, the map $\alpha_4 \circ \alpha_3$ is also étale. Since β_5 is surjective, we see that γ' is unramified. Thus, γ' is a finite and unramified map of local rings. Since $Z \to W$ is surjective and $k(y) \simeq k(x)$, the map γ' is an isomorphism by Lemma 4.10. In particular, $\alpha_4 \circ \alpha_3$ is an étale map of local rings such that $\beta_5 \circ \alpha_4 \circ \alpha_3$ is an isomorphism, in particular, étale. It follows that β_5 is étale, by [EGA IV$_4$ 1967, Proposition (17.3.4), page 62]. Thus, β_5 is a surjective étale map of local rings. But it can happen only if β_5 is an isomorphism.

Conversely, suppose that β_5 is an isomorphism. Let $\mathfrak{p}$ be the minimal prime of $\mathcal{O}_{W',S}$ such that $\mathcal{O}_{W',S}/\mathfrak{p} = \mathcal{O}_{Z,S}$ and let $\{\mathfrak{p}_1, \dots, \mathfrak{p}_m\}$ denote the set of distinct minimal primes of $\mathcal{O}_{W',S}$ different from $\mathfrak{p}$. To show that α_5 is an isomorphism, we need to show that $\mathfrak{p} + \mathfrak{p}_i = \mathcal{O}_{W',S}$ for all $1 \le i \le m$.

Claim 1. $$\mathfrak{p}_i \mathcal{O}_{W',x} = \mathcal{O}_{W',x} \quad \textit{for all } 1 \le i \le m.$$

Proof. Note that $\mathcal{O}_{W',x}$ is an integral domain because $\mathcal{O}_{Z,x}$ is an integral domain and β_5 is an isomorphism. Thus, we must have either $\mathfrak{p}_i \mathcal{O}_{W',x} = 0$ or $\mathfrak{p}_i \mathcal{O}_{W',x} = \mathcal{O}_{W',x}$. In the first case, we have $\mathfrak{p}_i \mathcal{O}_{Z,x} = 0$ as β_5 is an isomorphism. Equivalently, $\alpha_5 \circ \beta_4(\mathfrak{p}_i) = 0$. Since $\mathfrak{p}_i \neq \mathfrak{p}$, and $\mathfrak{p}_i, \mathfrak{p}$ are minimal, there is $a_i \in \mathfrak{p}_i \setminus \mathfrak{p}$ such that $\beta_4(a_i) \neq 0$. Hence, $\alpha_5 \circ \beta_4(a_i) \neq 0$, because α_5 is injective being a localization of an integral domain. This is a contradiction. Thus, we must have $\mathfrak{p}_i \mathcal{O}_{W',x} = \mathcal{O}_{W',x}$ for each i, proving Claim 1. □

Let $\mathfrak{m}$ be the maximal ideal of $\mathcal{O}_{W',S}$ defining the closed point x. By Claim 1, for any $1 \le i \le m$ there exists $a_i \in \mathfrak{p}_i \setminus \mathfrak{m}$ in $\mathcal{O}_{W',S}$ such that $\alpha_4(a_i)$ is invertible. Let $a = \prod_{i=1}^{m} a_i$. We see that there are nonzero elements $b, c \in \mathcal{O}_{Z',S}$ with $c \notin \mathfrak{m}$ such that $c(1 - ab) = 0$.

Claim 2. $$1 - ab \in \mathfrak{p}.$$

Proof. Let $v = 1 - ab$. Then, we have $cv = 0 \in \mathfrak{m}$ with $c \notin \mathfrak{m}$, so that $v \in \mathfrak{m}$ and $\alpha_4(v) = 0$. Toward contradiction, suppose $v \notin \mathfrak{p}$. Then $v \in \mathfrak{m} \setminus \mathfrak{p}$, so that $\beta_4(v) \neq 0$. Thus $\beta_5 \circ \alpha_4(v) = \alpha_5 \circ \beta_4(v) \neq 0$ because α_5 is injective. But this contradicts that $\alpha_4(v) = 0$. Hence, we have $v \in \mathfrak{p}$, proving Claim 2. □

By Claim 2, we have $v \in \mathfrak{p}$, $ab \in \mathfrak{p}_i$ for all i, while $v - ab = 1$. This shows that $\mathfrak{p} + \mathfrak{p}_i = \mathcal{O}_{W',S}$ for all $1 \le i \le m$. Thus, α_5 is an isomorphism. □

4G. ***Birationality under linear projections.*** Using Lemma 4.8, we shall show that the linear projections often give birational morphisms when restricted to a given integral closed subscheme. But first, we derive the following consequence of the results we proved in Section 4E and Section 4F. We continue to work with the Set-up of Section 4D.

We use a trick of "marking" irreducible components: for each $1 \le i \le s$, we fix a closed point $\alpha_i \in (Z_i)_{\text{sm}}$ such that (1) $\alpha_i \notin Z_j$ for $j \neq i$, (2) $x_i = \hat{f}(\alpha_i) \in X_{\text{sm}}$ but not in Σ, and (3) $b_i = \hat{g}(\alpha_i) \in B$. Note here that $\alpha_i \in (Z_i)_{\text{sm}}$ and $x_i \in X_{\text{sm}}$ can be achieved as follows: by the assumptions of the Set-up of Section 4D, each Z_i intersects $H \times \hat{B}$ properly and $X \setminus (X \cap H) \subset X_{\text{sm}}$. Then any choice of a point in $Z_i|_{(X \setminus (X \cap H)) \times B}$ maps to a point of X_{sm}. Moreover, perfectness of k implies that $(Z_i)_{\text{sm}} \cap Z_i|_{(X \setminus (X \cap H)) \times B} \neq \varnothing$. Let $\Xi = \{x_1, \dots, x_s\} \cup \Sigma$ and $E = \{b_1, \dots, b_s\} \cup F$. Since $Z_i \not\subset X \times F$ and $Z_i \to \hat{B}$ is nonconstant by the Set-up of Section 4D, no component of Z lies in $\hat{g}^{-1}(E)$.

Lemma 4.12. *After replacing the embedding $X \hookrightarrow \mathbb{P}^N_k$ by a bigger one via a Veronese embedding if necessary, there is a dense open subset $\mathcal{U} \subset \mathrm{Gr}(X, N-r-1, H)$ such that each $L \in \mathcal{U}(k)$ has the property that $Z_i \cap \hat{\phi}_L^{-1}(\hat{\phi}_L(\alpha_i)) = \{\alpha_i\}$ for all $1 \le i \le s$.*

Proof. We let $\pi : \mathrm{Spec}(\bar{k}) \to \mathrm{Spec}(k)$ denote the base change map. For any $A \in \mathbf{Sch}_k$, we shall write $\pi_A : A_{\bar{k}} \to A$ simply as π using a shorthand.

We fix i. Let $\beta_i := \hat{\phi}_L(\alpha_i)$. Let $\pi^{-1}(\alpha_i) = \{\alpha_{ij}\}_j$, which is a finite set of points, and let $x_{ij} := \hat{f}_{\bar{k}}(\alpha_{ij})$, $b_{ij} := \hat{g}_{\bar{k}}(\alpha_{ij})$. Note that all of α_{ij} and x_{ij} lie in the smooth loci of $(Z_i)_{\bar{k}}$ and $X_{\bar{k}}$, respectively.

We let $\Xi_i := \{x_{ij}\}_j \cup \Sigma_{\bar{k}}$ and $E_i := \{b_{ij}\}_j \cup F_{\bar{k}}$.

Applying Lemma 4.8 over $\bar{k}$ for the above Ξ_i (for Σ there) and E_i (for E there), we obtain a dense open subset $\mathcal{U}'_i \subset \mathrm{Gr}(X_{\bar{k}}, N-r-1, H_{\bar{k}})$ such that every $L \in \mathcal{U}'(k)$ satisfies the properties (1)–(6) there. Repeating the argument of Lemma 4.8 in Step 2, we obtain a dense open subset $\mathcal{U}_i \subset \mathrm{Gr}(X, N-r-1, H)$ such that for every $L \in \mathcal{U}_i(k)$, we have $L_{\bar{k}} \in \mathcal{U}'_i(\bar{k})$.

We show that the following map is bijective:

$$\hat{\phi}_{L,\bar{k}} : \pi^{-1}(\alpha_i) \to \pi^{-1}(\beta_i). \tag{4-6}$$

Suppose this is not injective, i.e., for some $j < j'$, we have $\hat{\phi}_{L,\bar{k}}(\alpha_{ij}) = \hat{\phi}_{L,\bar{k}}(\alpha_{ij'})$. Then $b_{ij} = \hat{g}_{\bar{k}}(\alpha_{ij}) = \hat{g}_{\bar{k}}(\alpha_{ij'}) = b_{ij'}$. Since $\bar{k}$ is algebraically closed, we can write $\alpha_{ij} = (x_{ij}, b_{ij})$ and $\alpha_{ij'} = (x_{ij'}, b_{ij'})$. Since $b_{ij} = b_{ij'}$ and $\alpha_{ij} \ne \alpha_{ij'}$, we must have $x_{ij} \ne x_{ij'}$.

But at the same time, we have

$$\hat{\phi}_{L,\bar{k}}(x_{ij}) = \hat{f}_{\bar{k}}(\hat{\phi}_{L,\bar{k}}(\alpha_{ij})) = \hat{f}_{\bar{k}}(\hat{\phi}_{L,\bar{k}}(\alpha_{ij'})) = \hat{\phi}_{L,\bar{k}}(x_{ij'}).$$

In particular, $x_{ij'} \in L^+_{\bar{k}}(x_{ij})$. Since $\hat{g}_{\bar{k}}(\alpha_{ij'}) = b_{ij'} \in E_i$, we thus have $\hat{f}_{\bar{k}}(\alpha_{ij'}) = x_{ij'} \in (L_{\bar{k}})^+(x_{ij}) \cap \hat{f}_{\bar{k}}(\hat{g}_{\bar{k}}^{-1}(E_i))$. But this contradicts property (5) of Lemma 4.8 satisfied by $L_{\bar{k}}$. Hence the map (4-6) is injective.

On the other hand, we have

$$\pi^{-1}(\beta_i) \times_{\mathrm{Spec}(k(\beta_i))} \mathrm{Spec}(k(\alpha_i)) = \mathrm{Spec}((\bar{k} \otimes_k k(\beta_i)) \otimes_{k(\beta_i)} k(\alpha_i)) = \mathrm{Spec}(\bar{k} \otimes_k k(\alpha_i)) = \pi^{-1}(\alpha_i)$$

so that it follows that the map (4-6) is surjective, as well, thus bijective.

Going back to the proof of the lemma, first note that we clearly have $Z_i \cap \hat{\phi}_L^{-1}(\beta_i) \supset \{\alpha_i\}$. For the inclusion in the other direction, toward contradiction suppose there is $\alpha' \in \hat{\phi}_L^{-1}(\beta_i) \setminus \{\alpha_i\}$ such that $\alpha' \in Z_i$. Clearly we have $\pi^{-1}(\alpha') \cap \pi^{-1}(\alpha_i) = \varnothing$. On the other hand, we have $\hat{\phi}_{L,\bar{k}}(\pi^{-1}(\alpha')) \subset \pi^{-1}(\beta_i) = \hat{\phi}_{L,\bar{k}}(\pi^{-1}(\alpha_i))$, where the second equality holds by the bijectivity of (4-6).

Hence there is some $\alpha'_j \in \pi^{-1}(\alpha')$ and $\alpha_{ij'} \in \pi^{-1}(\alpha_i)$ such that

$$\text{(a) } \alpha'_j \ne \alpha_{ij'}, \quad \text{while} \quad \text{(b) } \hat{\phi}_{L,\bar{k}}(\alpha'_j) = \hat{\phi}_{L,\bar{k}}(\alpha_{ij'}).$$

Property (b) implies that $\hat{g}_{\bar{k}}(\alpha'_j) = \hat{g}_{\bar{k}}(\alpha_{ij'}) = b_{ij'}$. Since $\bar{k}$ is algebraically closed, for $x' := \hat{f}_{\bar{k}}(\alpha'_j)$, we can express $\alpha'_j = (x', b_{ij'})$ and $\alpha_{ij'} = (x_{ij'}, b_{ij'})$. Because $\alpha'_j \not\alpha_{ij'}$ by (a), we must have $x' \ne x_{ij'} = \hat{f}_{\bar{k}}(\alpha_{ij'})$. In particular, $x' \in L^+_{\bar{k}}(x_{ij'})$. But $\hat{g}_{\bar{k}} = b_{ij'} \in E_i$ so that we obtain $x' \in L^+_{\bar{k}}(x_{ij'}) \cap \hat{f}_{\bar{k}}(\hat{g}_{\bar{k}}^{-1}(E_i))$. But, this contradicts property (5) of Lemma 4.8 satisfied by $L_{\bar{k}}$. Hence no such α' exists. Our proof then is over by taking $\mathcal{U} := \bigcap_{i=1}^{s} \mathcal{U}_i$. □

Combined with Lemma 4.11, we immediately have:

Corollary 4.13. *For each linear projection L as in Lemma 4.12 and each $1 \le i \le s$, one has that Z_i is the only irreducible component of $\hat{\phi}_L^{-1}(\hat{\phi}_L(Z_i))$ passing through a given marked point $\alpha_i \in Z_i \setminus \bigcup_{j \neq i} Z_j$.*

We can now prove the birationality of a given finite set of integral closed subschemes of $X \times \hat{B}_n$ under suitable linear projections.

Lemma 4.14. *For a suitable choice of the set E in the Set-up of Section 4D, after replacing the embedding $X \hookrightarrow \mathbb{P}_k^N$ by a bigger one via a Veronese embedding if necessary, there is a dense open subset $\mathcal{U} \subset \mathrm{Gr}(X, N-r-1, H)$ such that for each $L \in \mathcal{U}(k)$, the induced map $\hat{\phi}_L : Z_i \to \hat{\phi}_L(Z_i)$ is birational for all $1 \le i \le s$.*

Proof. We follow the choices of $\alpha_i \in Z_i$, Ξ and E that we made just before Lemma 4.12. We shall prove the lemma for this E. We let $\mathcal{U} \subset \mathrm{Gr}(X, N-r-1, H)$ be as given by Lemma 4.12 and fix $L \in \mathcal{U}(k)$. We let $T_i := \hat{\phi}_L(Z_i)$ and $\beta_i := \hat{\phi}_L(\alpha_i)$. To show that $\hat{\phi}_L : Z_i \to T_i$ is birational, we prove a stronger assertion that the map $\mathcal{O}_{T_i,\beta_i} \to \mathcal{O}_{Z_i,\beta_i}$ of semilocal rings is an isomorphism, where $\mathcal{O}_{Z_i,\beta_i} := \mathcal{O}_{Z_i, Z_i \cap \hat{\phi}_L^{-1}(\beta_i)}$. Consider the maps

$$\mathcal{O}_{T_i,\beta_i} \to \mathcal{O}_{Z_i,\beta_i} \to \mathcal{O}_{Z_i,\alpha_i}. \tag{4-7}$$

It follows from Lemma 4.12 that $Z_i \cap \hat{\phi}_L^{-1}(\beta_i) = \{\alpha_i\}$. In particular, the second map of (4-7) is an isomorphism, actually the identity map. By condition (1) of Lemma 4.8, the map ϕ_L is étale in an affine open neighborhood U' of Ξ, and thus ϕ_L is étale at α_i. In particular, the composite map in (4-7) is unramified. By condition (3) of Lemma 4.8, we have $k(\beta_i) \xrightarrow{\sim} k(\alpha_i)$. Hence, the first map of (4-7) is an injective finite unramified map of local rings, that induces an isomorphism of the residue fields. It is therefore an isomorphism by Lemma 4.10. This completes the proof. □

4H. ***A presentation lemma for moving to fs-cycles.*** The final result of Section 4 is the following Theorem 4.15, that will be used in the proof of the fs-moving lemma, specifically, in the proof of Lemma 8.7.

Theorem 4.15. *Under the Set-up of Section 4D, let $Z_i^0 := Z_i|_{X \times B}$ and $Z^0 := Z|_{X \times B}$.*

Then for a suitable choice of the set E in the Set-up, after replacing the embedding $X \hookrightarrow \mathbb{P}_k^N$ by its composition with a suitable Veronese embedding, there is a dense open subset $\mathcal{U} \subset \mathrm{Gr}(X, N-r-1, H)$ such that each $L \in \mathcal{U}(k)$ satisfies the following:

(1) *ϕ_L is étale at Σ.*

(2) *ϕ_L separates the points of Σ.*

(3) *$k(\phi_L(x)) \xrightarrow{\sim} k(x)$ for all $x \in \Sigma$.*

(4) *There exists an affine open neighborhood $U \subset X$ of Σ such that:*

(4a) *If Z_i^0 is an irreducible component of Z^0 that is dominant over an irreducible component of X, then for each component Z' of $L^+(Z_i^0)$, the map $Z'_U \to U$ is* fs *over U.*

(4b) *If Z_i^0 is an irreducible component of Z^0 that is not dominant over any irreducible component of X, then $L^+(Z_i^0)_U = 0$.*

Proof. For $1 \le i \le s$, choose closed points $\alpha_i \in Z_i^0 \setminus \big(\bigcup_{j \neq i} Z_j^0\big)$ such that

$$x_i := \hat{f}(\alpha_i) \in X_{\mathrm{sm}} \quad \text{and} \quad b_i := \hat{g}(\alpha_i) \in B,$$

as we did in Lemma 4.12. Let $\Xi := \{x_1, \dots, x_r\} \cup \Sigma \subset X_{\mathrm{sm}}$ and $E := \{b_1, \dots, b_r\} \cup F \subset \hat{B}$. Since $Z_i \not\subset X \times F$ and $Z_i \to \hat{B}$ is nonconstant, it is not contained in $\hat{g}^{-1}(E)$. We choose $\mathcal{U} \subset \mathrm{Gr}(X, N-r-1, H)$ as given by Lemma 4.12 and fix $L \in \mathcal{U}(k)$. In particular, all the properties of Lemma 4.8 holds, so that we have conditions (1)–(3) of the theorem.

To prove (4), first note that the irreducible components of $L^+(Z_i^0)$ are exactly the restrictions to $X \times B$ of the irreducible components of $L^+(Z_i)$. Let Z_i be an irreducible component of Z dominant over an irreducible component of X. Let Z' be an irreducible component of $L^+(Z_i)$. We prove that $Z' \cap (\Sigma \times F) = \varnothing$.

Suppose, on the contrary, that there is a closed point $\lambda \in Z' \cap (\Sigma \times F)$. This means that there is a closed point $\lambda' \in Z_i$ such that $\hat{\phi}_L(\lambda) = \hat{\phi}_L(\lambda')$. We claim in this case that

$$\{\lambda'\} = \hat{\phi}_L^{-1}(\hat{\phi}_L(\lambda)) \cap Z_i = \{\lambda\}. \tag{4-8}$$

Suppose we have shown that $\lambda' = \lambda$. Then we get $\lambda \in Z_i$ and (4-8) becomes equivalent to showing that $\hat{\phi}_L^{-1}(\hat{\phi}_L(\lambda)) \cap Z_i = \{\lambda\}$. But the proof of this equality is simply a repetition of the argument of Lemma 4.12. Hence, the claim is reduced to showing that $\lambda' = \lambda$.

Let's do it. First consider the case when k is algebraically closed. We can then uniquely write $\lambda = (x, b)$ for some closed points $x \in \Sigma$ and $b \in F$, and $\lambda' = (x', b)$, where $x' \in \hat{\phi}_L^{-1}(\hat{\phi}_L(x))$. If $x' \neq x$, then $x' \in L^+(x)$ and $x' \in \hat{f}(\hat{g}^{-1}(E))$, which contradicts condition (5) of Lemma 4.8. Hence, we must have $x' = x$ so that $\lambda' = \lambda$.

If k is not algebraically closed, we argue as in the proof of Lemma 4.12. Suppose again that $\lambda' \neq \lambda$. Then for the base change map $\pi : \mathrm{Spec}(\bar{k}) \to \mathrm{Spec}(k)$, we have $\pi^{-1}(\lambda') \cap \pi^{-1}(\lambda) = \varnothing$. Let $\beta' := \hat{\phi}_L(\lambda') = \hat{\phi}_L(\lambda)$. We show as in the argument of Lemma 4.12 that the map $\hat{\phi}_{L,\bar{k}} : \pi^{-1}(\lambda') \to \pi^{-1}(\beta')$ is bijective. Using this, we continue following the proof of Lemma 4.12, to get closed points $\tilde{\lambda} \in \pi^{-1}(\lambda)$ and $\tilde{\lambda'} \in \pi^{-1}(\lambda')$ such that $\hat{f}_{\bar{k}}(\tilde{\lambda}) \in (L_{\bar{k}})^+(\hat{f}_{\bar{k}}(\tilde{\lambda'})) \cap \hat{f}_{\bar{k}}(\hat{g}_{\bar{k}}^{-1}(E_i))$. But this contradicts property (5) of Lemma 4.8 for $L_{\bar{k}}$, which violates our choice of L. This proves (4-8).

Coming back to the proof of $Z' \cap (\Sigma \times F) = \varnothing$, we now note using Corollary 4.13 that $Z' \neq Z_i$. So, the two deductions $\lambda \in Z' \cap Z_i$ and $\hat{\phi}_L^{-1}(\hat{\phi}_L(\lambda)) \cap Z_i = \{\lambda\}$ from (4-8) together contradict Lemma 4.11. Hence, we must have $Z' \cap (\Sigma \times F) = \varnothing$, as desired.

Now, by Lemma 2.10, there is an affine open neighborhood $U_{i,Z'} \subset X_{\mathrm{sm}}$ of Σ such that $Z'_{U_{i,Z'}} \to U_{i,Z'}$ is fs. We take $U_1 := \bigcap U_{i,Z'}$ where the intersection is taken over all i such that Z_i dominant over a component of X and the irreducible components Z'. This open set U_1 works for (4a).

About property (4b), let Z_i be an irreducible component of Z which is not dominant over X. Let Z' be a component of $L^+(Z_i)$. In this case, we repeat the proof of (4a) above, where we now apply condition (6) of Lemma 4.8, to conclude that $Z' \cap (\Sigma \times \hat{B}) = \varnothing$.

It follows that $\hat{f}(L^+(Z_i))$ is a closed subset of X disjoint from Σ. Hence, we can apply Lemma 2.3 to obtain an affine open neighborhood U'_i of Σ in X such that $L^+(Z_i)_{U'_i} = \varnothing$. We take $U_2 := \bigcap U'_i$, where the intersection is taken over all i such that Z_i is not dominant over any component of X. This open set U_2 works for (4b). Taking $U := U_1 \cap U_2$, we have (4), and this concludes the proof of the theorem. □

5. Regularity of the original cycle over residual points

The focus of the remaining sections is to achieve the sfs-property of the residual cycle of Z along Σ via more refined linear projections. In order to achieve this, we first ensure that our original cycle Z is regular at all points lying over the residual set $L^+(\Sigma)$ of $\Sigma \subset X$. We later show that this regularity of Z at all points lying over $L^+(\Sigma)$ implies the regularity of the residual cycle of Z along Σ. The goal of this section is to achieve the first one when k is algebraically closed. The general case will be considered later.

5A. ***A basic algebraic result.*** We first discuss the following:

Lemma 5.1. *Let k be an algebraically closed field. Let $X \subset \mathbb{P}^N_k$ be a reduced closed subscheme of dimension* 1. *Suppose $N \gg 1$ and let $x \neq y$ be two closed points on X_{sm}. Let $\mathrm{Gr}_{x+2y}(N-1, \mathbb{P}^N_k) \subset \mathrm{Gr}(N-1, \mathbb{P}^N_k)$ be the set of hyperplanes containing $\{x, y\}$ that do not intersect X transversely at y. Then $\mathrm{Gr}_{\{x,y\}}(N-1, \mathbb{P}^N_k) \simeq \mathbb{P}^{N-2}_k$ and $\mathrm{Gr}_{x+2y}(N-1, \mathbb{P}^N_k) \simeq \mathbb{P}^{N-3}_k$.*

Proof. Recall that $\mathrm{Gr}_{\{x,y\}}(N-1, \mathbb{P}^N_k) \subset \mathrm{Gr}(N-1, \mathbb{P}^N_k)$ is the set of hyperplanes containing $\{x, y\}$. Since $x \neq y$, by elementary linear algebra on ranks of linear systems, we immediately have $\mathrm{Gr}_{\{x,y\}}(N-1, \mathbb{P}^N_k) \simeq \mathbb{P}^{N-2}_k$. We prove the second assertion. Since $N \gg 1$, we can find a linear form $s_1 \in W = H^0(\mathbb{P}^N_k, \mathcal{O}(1))$ which does not vanish anywhere in $\{x, y\}$. This yields a k-linear map $\alpha : W \to \mathcal{O}_{X,\{x,y\}}/\mathfrak{m}_x\mathfrak{m}_y^2 =: \mathcal{O}_{\{x+2y\}}$ given by $\alpha(s) = s/s_1$. Since k is algebraically closed, the ideal $\mathfrak{m}_y$ is generated by linear forms vanishing at y. Hence, the composite map $W \to \mathcal{O}_{X,\{x,y\}}/\mathfrak{m}_x\mathfrak{m}_y^2 \twoheadrightarrow \mathcal{O}_{X,y}/\mathfrak{m}_y^2 =: \mathcal{O}_{\{2y\}}$ is surjective and $\alpha^{-1}(\mathfrak{m}_y^2)$ is precisely the set of linear forms in W not transverse to X at y.

We first claim that α is surjective. Since x, y are two distinct regular closed points of X, the set $\mathrm{Gr}_y(x, N-1, \mathbb{P}^N_k)$ is nonempty and hence, $\mathfrak{m}_y/\mathfrak{m}_x\mathfrak{m}_y \xrightarrow{\sim} \mathcal{O}_{\{x\}}$ and there is a commutative diagram of short exact sequences:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \alpha^{-1}(\mathfrak{m}_x\mathfrak{m}_y) & \longrightarrow & \alpha^{-1}(\mathfrak{m}_y) & \longrightarrow & \mathcal{O}_{\{x\}} & \longrightarrow & 0 \\ & & \downarrow & & \downarrow\!\!\!\downarrow & & \| & & \\ 0 & \longrightarrow & \mathfrak{m}_x\mathfrak{m}_y/\mathfrak{m}_x\mathfrak{m}_y^2 & \longrightarrow & \mathfrak{m}_y/\mathfrak{m}_y^2 & \longrightarrow & \mathcal{O}_{\{x\}} & \longrightarrow & 0 \end{array} \tag{5-1}$$

In particular, the first vertical map is surjective. Since $\mathrm{Gr}_x(y, N-1, \mathbb{P}^N_k) \neq \varnothing$, we conclude that α is surjective.

To finish the proof, we look at the commutative diagram with exact rows:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \ker(\alpha) & \longrightarrow & W & \xrightarrow{\alpha} & \mathcal{O}_{\{x+2y\}} & \longrightarrow & 0 \\ & & \downarrow & & \| & & \downarrow & & \\ 0 & \longrightarrow & \alpha^{-1}(\mathfrak{m}_y^2) & \longrightarrow & W & \longrightarrow & \mathcal{O}_{\{2y\}} & \longrightarrow & 0 \end{array} \tag{5-2}$$

Since the last vertical arrow is surjective with one-dimensional kernel, by the snake lemma, the first vertical arrow is injective with one-dimensional cokernel. Since $\mathbb{P}(\alpha^{-1}(\mathfrak{m}_y^2)) \simeq \mathbb{P}^{N-2}_k$, we conclude that $\mathrm{Gr}_{x+2y}(N-1, \mathbb{P}^N_k) \simeq \mathbb{P}(\ker(\alpha)) \simeq \mathbb{P}^{N-3}_k$. □

5B. *The Set-up+*(fs). We suppose k is an infinite perfect field. The Set-up we now use repeatedly is the following situation, that we call the Set-up+(fs):

(1) *The Set-up*: We still suppose the Set-up of Section 4D, not necessarily specifying some closed subset $E \subset \hat{B}$.

(2) *The* fs-*property*: There exists an affine open neighborhood $X_{\mathrm{fs}} \subset X_{\mathrm{sm}}$ of Σ, that is dense open in X, such that the projection $Z \to X$ is fs over X_{fs}.

5C. *Regularity of the original cycle over residual points.* We now discuss two central results: Lemmas 5.2 and 5.10. Recall that X is equidimensional under the above assumptions.

Lemma 5.2. *Let k be an algebraically closed field. Suppose $r = 1$. We are under the Set-up+*(fs) *of Section 5B. Let $x \in X_{\mathrm{fs}}$ be a closed point and let $S \subset X \setminus \{x\}$ be another finite set of closed points.*

After replacing $\mathbb{P}^N_k$ by a bigger projective space via a Veronese embedding if necessary, there exists a dense open subset $\mathcal{U}_S \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ such that each $L \in \mathcal{U}_S(k)$ satisfies the following:

(1) *$L \cap ((X \setminus X_{\mathrm{fs}}) \cup S) = \varnothing$.*

(2) *L intersects X_{fs} transversely.*

(3) *$L \cap X$ consists of $(d+1)$-distinct closed points $c_0 = x, c_1, \dots, c_d$.*

(4) *Z is regular at all points lying over $\{c_1, \dots, c_d\}$. In particular, each component Z_i does not meet other irreducible components at points lying over $\{c_1, \dots, c_d\}$.*

Proof. Since $\dim(Z_{\mathrm{sing}}) = 0$, we see that $\hat{f}(Z_{\mathrm{sing}})$ is a finite closed subset of X. Since X_{fs} is dense in X, we have $|X \setminus X_{\mathrm{fs}}| < \infty$. Hence, $T := (\hat{f}(Z_{\mathrm{sing}}) \cup (X \setminus X_{\mathrm{fs}}) \cup S) \setminus \{x\}$ is a finite closed subset of X. Thus the hyperplanes disjoint from T form a dense open subset $\mathrm{Gr}(T, N-1, \mathbb{P}^N_k)$ of $\mathrm{Gr}(N-1, \mathbb{P}^N_k)$ by Lemma 4.2. The set $\mathcal{U}_1 := \mathrm{Gr}^{\mathrm{tr}}(X, N-1, \mathbb{P}^N_k) \cap \mathrm{Gr}(T, N-1, \mathbb{P}^N_k)$ is dense open in $\mathrm{Gr}(N-1, \mathbb{P}^N_k)$. If we show that $\mathcal{U}_S := \mathcal{U}_1 \cap \mathrm{Gr}_x(N-1, \mathbb{P}^N_k) \neq \varnothing$, then this set will be dense open in $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$. It is moreover clear that any $L \in \mathcal{U}_S(k)$ satisfies (1)–(4). It remains to show that $\mathrm{Gr}^{\mathrm{tr}}(X, N-1, \mathbb{P}^N_k) \cap \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ and $\mathrm{Gr}(T, N-1, \mathbb{P}^N_k) \cap \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ are both nonempty.

Let V be the set of linear forms in $H^0(\mathbb{P}_k^N, \mathcal{O}(1))$ that vanish at x. Note that $\dim|V| = N-1$ and that the maximal ideal $\mathfrak{m}_x \subset \mathcal{O}_{X,x}$ is generated by the members of V. Let $\mathcal{B} \subset X \times |V|$ be the incidence scheme consisting of pairs (y, L) such that L passes through y, but not transverse to X at y. We study the fiber of $\pi_1 : \mathcal{B} \to X$ over each $y \in X_{\mathrm{sm}} \setminus \{x\}$.

Choose $s_1 \in V$ such that $s_1(x) = 0$ but $s_1(y) \neq 0$. Consider the map $\beta : V \to \mathcal{O}_{X,y}/\mathfrak{m}_y^2$ given by $\beta(s) = s/s_1$. Since $\dim|V| = N-1$, while $\mathrm{Gr}_{\{x,y\}}(N-1, \mathbb{P}^N) \simeq \mathbb{P}^{N-2}$ and $\mathrm{Gr}_{\{x+2y\}}(N-1, \mathbb{P}^N) \simeq \mathbb{P}^{N-3}$ by at most $N-3$, because $\dim_k(\mathcal{O}_{X,y}/\mathfrak{m}_y^2) = 2$. Thus, $\dim(\mathcal{B}) \leq \dim X + \dim(\pi_1^{-1}(y)) \leq 1 + N - 3 = N-2$. Hence its image in $|V|$ under the projection $\pi_2 : X \times |V| \to |V|$ is a proper closed subset (note that X is projective). Since $N \gg 0$, its complement $\mathrm{Gr}_x^{\mathrm{tr}}(X, N-1, \mathbb{P}_k^N)$ in $\mathrm{Gr}_x(N-1, \mathbb{P}_k^N)$ is a dense open subset. Since $\dim(\mathrm{Gr}_x(N-1, \mathbb{P}_k^N)) = N-1$ and $T \subset X$ is a finite set of closed points different from x, the assertion that $\mathrm{Gr}(T, N-1, \mathbb{P}_k^N) \cap \mathrm{Gr}_x(N-1, \mathbb{P}_k^N)$ is nonempty follows from Lemma 5.1. We have therefore finished the proof. □

In Section 6A, we will obtain a slightly stronger version of Lemma 5.2. This is done in Lemma 5.9. The difference in the latter lemma from the former is that (following the notations of Lemma 5.2), after a possible reembedding, we may impose an additional property that for $L \cap X = \{c_0 = x, c_1, \ldots, c_d\}$, no three points of them are collinear.

At one bad extreme case, suppose X is contained in a 2-dimensional projective space. Then for any hyperplane L, which is a line, the hyperplane section $L \cap X$ is entirely collinear. This is an important obstacle to avoid. We will show in Lemma 5.7 that, after taking a Veronese reembedding for a high enough degree $d \geq 3$, we can always avoid it. It will be improved for the higher dimensional case in Lemma 5.8. These two are some technical grounds needed in Section 6A.

Once we can avoid the above extreme case using a Veronese reembedding, then one can employ the following well-known general result (see [Arbarello et al. 1985, Chapter III, page 109]):

Theorem 5.3 (general position theorem). *Let $N \geq 2$. Let $C \subset \mathbb{P}^N$ be an irreducible nondegenerate, possibly singular, curve of degree d. Then a general hyperplane meets C in d points, any N of which are linearly independent.*

Recall that a closed embedding $X \subset \mathbb{P}_k^n$ of an integral projective scheme X is said to be nondegenerate if no hyperplane of $\mathbb{P}_k^n$ contains X. We won't give the proof of Theorem 5.3 here. We mention that Theorem 5.3 for $N = 2$ is immediate, while, for $N \geq 3$ reduces to the following special case (see [loc. cit.]), that is more relevant to the paper:

Lemma 5.4. *Let $C \subset \mathbb{P}^N$ with $N \geq 3$ be an irreducible nondegenerate, possibly singular, curve of degree d. Then a general hyperplane meets C in d points, no three of which are collinear.*

Remark 5.5. To give a bit of the flavor of the proof of Lemma 5.4, we remark that with some efforts (see [Arbarello et al. 1985, pages 110–111] or imitate [Hartshorne 1977, Proposition IV-3.8, page 311]), one can argue that if Lemma 5.4 fails, then all tangent lines to C passes through a single fixed point $p \in C$. Then a linear projection from p would shrink the entire curve C to a point in $\mathbb{P}^{N-1}$. Since C is nondegenerate, we can argue this cannot happen.

Such a curve in $\mathbb{P}^N$ all of whose tangent lines pass through a fixed point is called *strange* (see [Hartshorne 1977, page 311]). We remark that in case C is nonsingular, it is known that the only nonsingular strange curves in any $\mathbb{P}^N$ are either a line or a conic in $\mathbb{P}^2$ in characteristic 2 (see [Samuel 1966, Theorem, Appendix to Chapter II, page 76] or [Hartshorne 1977, Theorem IV-3.9, page 312]).

We thank the referee for pointing to us that some technical part of our construction of the paper is relevant to noncollinearity of configurations of points and strange curves. □

Combined with the Bertini theorem ([Kleiman and Altman 1979, Theorem 1] or [Jouanolou 1983]), we immediately extend Lemma 5.4 to the following higher dimensional version, which we use:

Proposition 5.6 (linear general position theorem). *Let $X \subset \mathbb{P}^N$ with $N \geq 3$ be a nondegenerate, possibly singular, variety of degree d. Let $r = \dim X \geq 1$. Then for a general sequence of hyperplanes $H_1, \dots, H_r$ in $\mathbb{P}^N$, the intersection $X \cap H_1 \cap \cdots \cap H_r$ has d points, no three of which are collinear.*

Note that the above Proposition 5.6 holds for schemes that are nondegenerate in the projective spaces of dimension at least 3. This is another view of why we had a pathology about noncollinearity when X was contained in a 2-dimensional projective space in the paragraph before Theorem 5.3.

As said before, to avoid this problem, we need to replace the embedding by a bigger Veronese embedding. This is discussed now in the following:

Lemma 5.7. *Let $C \subset \mathbb{P}^n_k$ be a reduced projective curve. Suppose that there exists a 2-dimensional linear subspace $L \subset \mathbb{P}^n_k$ such that $C \subset L$. Let $\vartheta : \mathbb{P}^n_k \hookrightarrow \mathbb{P}^N_k$ be the d-uple Veronese embedding with $d \geq 3$. Then the image of each irreducible component of C via ϑ does not lie inside a 2-dimensional linear subspace of $\mathbb{P}^N_k$.*

Proof. We can assume C is an irreducible curve in order to prove the lemma. After a linear change of coordinates in $\mathbb{P}^n_k$, we may assume that $\mathbb{P}^n_k = \mathbb{P}(V)$ and $L = \mathbb{P}(W)$, where V is an $(n+1)$-dimensional k-vector space with a basis $\{x_0, \dots, x_n\}$ and $W = \mathrm{Span}_k\{x_0, x_1, x_2\}$ is a subspace of V. For any closed embedding $f : C \hookrightarrow \mathbb{P}^m_k$, we let $d_f(C)$ denote the degree of C under f.

Let $\iota : C \hookrightarrow L$ be the closed embedding as given in the assumption of the lemma. Let $d_0 := d_\iota(C) \geq 1$. Since L is linear in $\mathbb{P}^n_k$, the degree of C under the composite of the embeddings $C \hookrightarrow L \hookrightarrow \mathbb{P}^n_k$ is also d_0.

Toward contradiction, suppose that there is a 2-dimensional linear subspace $L' \subset \mathbb{P}^N_k$ such that $\vartheta(C) \subset L'$, where $\vartheta : \mathbb{P}^n_k \hookrightarrow \mathbb{P}^N_k$ is the d-uple Veronese embedding with $d \geq 3$. We denote the resulting embedding $C \hookrightarrow L'$ by $\vartheta|_C$.

By our choice of the embedding $L \hookrightarrow \mathbb{P}^n_k$, we have a commutative diagram

$$\begin{array}{ccccc} C & \xhookrightarrow{\ \iota\ } & L & \hookrightarrow & \mathbb{P}^n_k \\ & \searrow & \Big\downarrow{\scriptstyle \vartheta'} & & \Big\downarrow{\scriptstyle \vartheta} \\ & & M & \hookrightarrow & \mathbb{P}^N_k \end{array} \tag{5-3}$$

where $M \cong \mathbb{P}^r_k$ (with $r = (d+1)(d+2)/2 - 1$). The horizontal arrows in the right square are linear embeddings and the vertical arrows are the d-uple Veronese embeddings.

The linearity of the inclusion $L' \hookrightarrow \mathbb{P}^N_k$ implies that $d_{\vartheta|_C}(C)$ coincides with the degree of C under the composite closed embedding $C \hookrightarrow L' \hookrightarrow \mathbb{P}^N_k$. By the same argument, the degree of C for this composite embedding coincides with the degree of C for the composite embedding $C \hookrightarrow L \hookrightarrow M$. Since ϑ' is the d-uple Veronese embedding, it follows that the degree of C for the latter composite embedding is $d_0 d$. We conclude that $d_{\vartheta|_C}(C) = d_0 d$.

If we now apply the degree-genus adjunction formula for plane curves to the embedding ι, we get $g_a(C) = \frac{1}{2}(d_0 - 1)(d_0 - 2)$, where $g_a(C)$ is the arithmetic genus of C. The same formula for the embedding $\vartheta|_C$ yields $g_a(C) = \frac{1}{2}(d_0 d - 1)(d_0 d - 2)$.

Hence $(d_0 - 1)(d_0 - 2) = (d_0 d - 1)(d_0 d - 2)$, i.e., $d_0^2(d^2 - 1) - 3d_0(d - 1) = 0$. This factors into

$$d_0(d - 1)(d_0(d + 1) - 3) = 0. \tag{5-4}$$

Since $d_0 \geq 1$ and $d \geq 3$, the left hand side of (5-4) is $\geq 1 \cdot 2 \cdot (1 \cdot 4 - 3) > 0$, so that the equality of (5-4) cannot hold, thus a contradiction. This proves the lemma. □

An analogue of Lemma 5.7 in higher dimensions is the following.

Lemma 5.8. *Let $\iota : X \hookrightarrow \mathbb{P}^n_k$ be a reduced projective scheme of pure dimension $r \geq 2$. Assume that the degree of each irreducible component of X in $\mathbb{P}^n_k$ is at least two. Let $\Sigma \subset X$ be a finite set of closed points. For an integer $d \geq 1$, let $\vartheta : \mathbb{P}^n_k \hookrightarrow \mathbb{P}^N_k$ be the d-uple Veronese embedding.*

Then for all sufficiently large $d \geq 3$, (depending on X, Σ, n and the degrees of the irreducible components of X in $\mathbb{P}^n_k$), a general intersection $H_1 \cap \cdots \cap H_{r-1} \cap \vartheta(X)$ of X with hyperplanes H_i in $\mathrm{Gr}_\Sigma(N - 1, \mathbb{P}^N_k)(k)$ is a reduced curve, none of whose irreducible component is contained in a 2-dimensional linear subspace of $\mathbb{P}^N_k$.

Proof. By the Bertini theorems of Kleiman and Altman [1979, Theorem 1], an intersection of $\vartheta(X)$ with $(r-1)$ general hyperplanes containing Σ in a large enough d-uple Veronese embedding ϑ is a curve C, whose intersection with every irreducible component of $\vartheta(X)$ is again irreducible. Since k is perfect and X is reduced, it is actually geometrically reduced. It follows therefore from the Bertini theorem of Jouanolou [1983, Théorème 6.3] that C can be chosen to be reduced.

Let $X_1, \ldots, X_t$ be the irreducible components of X and let $C_1, \ldots, C_t$ denote the irreducible components of C.

Let s_i be the degree of X_i in $\mathbb{P}^n_k$ so that the degree of X in $\mathbb{P}^n_k$ is $s = \sum_{i=1}^r s_i$ (see [Hartshorne 1977, Proposition I-7.6, page 52]). Let $C = H_1 \cap \cdots \cap H_{r-1} \cap \vartheta(X)$ be as above. Let $d_\iota(C_i)$ denote the degree of C_i in $\mathbb{P}^n_k$ via the inclusion $\iota : C \hookrightarrow X \hookrightarrow \mathbb{P}^n_k$ and let $d_\vartheta(C_i)$ denote the degree of C_i in $\mathbb{P}^N_k$. Each of the hyperplanes $H_1, \ldots, H_{r-1} \subset \mathbb{P}^N_k$ restricts to a unique hypersurface of degree d in $\mathbb{P}^n_k$. Since these hyperplanes are sufficiently general, an elementary degree computation shows that $d_\iota(C_i) = d^{r-1} s_i$ and $d_\vartheta(C_i) = d^r s_i$ for each $1 \leq i \leq t$. We need to show that if d is sufficiently large, then each $C_i = H_1 \cap \cdots \cap H_{r-1} \cap \vartheta(X_i)$ is not contained in a 2-dimensional linear subspace of $\mathbb{P}^N_k$. To show this, we can assume that X and C are irreducible. In particular, $d_\iota(C) = sd^{r-1}$ and $d_\vartheta(C) = sd^r$.

We shall prove our assertion as an application of Castelnuovo's bound for the genus of curves. Let $3 \le n' \le n$ be the smallest integer such that $X \subset \mathbb{P}^{n'}_k \subset \mathbb{P}^n_k$, where the first embedding is nondegenerate and the second embedding is linear. Note that the lower bound on n' is forced by our assumption on the lower bounds of the dimension of X and its degree in $\mathbb{P}^n_k$.

Since $H_1, \dots, H_{r-1}$ restrict to general hypersurfaces of degree d in $\mathbb{P}^n_k$, we see that they restrict to hypersurfaces of the same degree in $\mathbb{P}^{n'}_k$. Since a hypersurface (of degree at least two) section of a nondegenerate closed subvariety of a projective space is necessarily nondegenerate (looking at the homogeneous coordinate rings), we conclude that the composite embedding $C \hookrightarrow X \hookrightarrow \mathbb{P}^{n'}_k$ is also nondegenerate. Furthermore, the degrees of X and C inside $\mathbb{P}^{n'}_k$ are the same as their respective degrees inside $\mathbb{P}^n_k$.

Let $m \ge 1$ and $0 \le \epsilon < n' - 1$ be two integers such that $sd^{r-1} - 1 = m(n' - 1) + \epsilon$. It follows from Castelnuovo's bound on the arithmetic genus (see [Harris 1982, Chapter 3; Arbarello et al. 1985, Chapter III, page 116] and see [Ballico 1989, Remark following Lemma 2.1] for singular curves) of C that

$$g_a(C) \le \frac{(n'-1)m(m-1)}{2} + m\epsilon. \tag{5-5}$$

Since $n' - 1 \ge 2$ and d is sufficiently large, we can assume $m < sd^{r-1} - 1$. We thus get

$$\begin{aligned} 2g_a(C) &\le (n'-1)m(m-1) + 2m\epsilon \\ &< (n'-1)m(m-1) + 2m(n'-1) \\ &= (n'-1)m(m+1) \\ &\le (sd^{r-1}-1)(sd^{r-1}-1) \\ &= (sd^{r-1}-1)^2. \end{aligned} \tag{5-6}$$

Now toward contradiction, suppose that inside $\mathbb{P}^N_k$, the curve C is contained in a 2-dimensional linear subspace $L \subset \mathbb{P}^N_k$. Since $d_\vartheta(C)$ is equal to the degree of C inside L, the degree-genus adjunction formula for the embedding $C \hookrightarrow L \cong \mathbb{P}^2_k$, yields $2g_a(C) = (sd^r - 1)(sd^r - 2)$. Note that if we let $e' := sd^{r-1} - 1$, then

$$\begin{aligned} 2g_a(C) &= (sd^r - 1)(sd^r - 2) \\ &= (d(sd^{r-1}-1) + d - 1)(d(sd^{r-1}-1) + d - 2) \\ &= d^2(e')^2 + (2d-3)d(e') + (d-1)(d-2), \end{aligned} \tag{5-7}$$

and because $d \ge 3$ and $s > 0$, we have $2g_a(C) > (e')^2 + e' + 0 \ge (e')^2$.

On the other hand, from (5-6) we had $2g_a(C) \le (e')^2$. This is a contradiction. □

We now present the aforementioned improvement of Lemma 5.2.

Lemma 5.9. *Let $X \hookrightarrow \mathbb{P}^N_k$ and $x \in X_{\mathrm{fs}}$ be as in Lemma 5.2. After replacing $\mathbb{P}^N_k$ by a bigger projective space via a Veronese embedding, there exists a dense open $\mathcal{U}_S \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ such that every $L \in \mathcal{U}_S(k)$ satisfies the following*:

(1) *Conditions* (1)–(4) *of Lemma 5.2.*

(2) *No three points of $L \cap X = \{x = c_0, c_1, \dots, c_d\}$ are collinear.*

Proof. Suppose first that X does not lie inside any 2-dimensional linear subspace of $\mathbb{P}^N_k$. In this case, we choose $\mathcal{U}_S$ just as in Lemma 5.2 so that (1) holds. Condition (2) holds by Proposition 5.6. Hence the lemma is proven in this case.

Suppose now that X lies inside a 2-dimensional linear space of $\mathbb{P}^N_k$. In this case, we choose a suitable Veronese embedding $\mathbb{P}^N_k \hookrightarrow \mathbb{P}^{N'}_k$ such that the image of each irreducible component of X does not lie in any 2-dimensional linear subspace of $\mathbb{P}^{N'}_k$ applying Lemma 5.7. Then after reembedding if necessary, we have a nonempty open subset $\mathcal{U}_S \subset \mathrm{Gr}_x(N'-1, \mathbb{P}^{N'}_k)$ such that conditions (1)–(4) of Lemma 5.2 hold.

In doing so, we can make sure that X is nondegenerate in a projective space of dimension at least 3. Then condition (1) holds by the choice of $\mathcal{U}_S$, while condition (2) holds by Proposition 5.6. This proves the lemma. □

The following result generalizes Lemma 5.9 to higher dimensional $r \geq 1$.

Lemma 5.10. *Let k be an algebraically closed field. Suppose $r \geq 1$. We are under the Set-up+*(fs) *of Section 5B. Let $x \in X_{\mathrm{fs}}$ be a closed point and let $S \subset X \setminus \{x\}$ be another finite set of closed points.*

After replacing $\mathbb{P}^N_k$ by a bigger projective space via a Veronese embedding if necessary, we have the following property: given any hyperplane $H_0 \subset \mathbb{P}^N_k$ disjoint from $S \cup \{x\}$ and a general $L_0 \in \mathrm{Gr}^{\mathrm{tr}}_{S\cup\{x\}}(H_0, N-r+1, \mathbb{P}^N_k)(k)$, there exists a dense open subset $\mathcal{U}_S \subset \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$ such that each $L \in \mathcal{U}_S(k)$ satisfies the following:

(1) *$L \cap L_0 \cap ((X \setminus X_{\mathrm{fs}}) \cup S) = \varnothing$.*

(2) *$L \cap L_0$ intersects X_{fs} transversely.*

(3) *$L \cap L_0 \cap X$ has $(d+1)$-distinct closed points $c_0 = x, c_1, \ldots, c_d$.*

(4) *Z is regular at all points lying over $\{c_1, \ldots, c_d\}$. In particular, each component Z_i does not meet other irreducible components at points lying over $\{c_1, \ldots, c_d\}$.*

(5) *$L_0 \cap X$ is an equidimensional reduced curve none of whose irreducible component lies inside a 2-dimensional linear subspace of $\mathbb{P}^N_k$.*

(6) *No three points of $L \cap L_0 \cap X = \{x = c_0, c_1, \ldots, c_d\}$ are collinear.*

Proof. In case $r = 1$, we have $\mathrm{Gr}(N-r+1, \mathbb{P}^N_k) = \mathrm{Gr}(N, \mathbb{P}^N_k) = \{\mathbb{P}^N_k\}$ so that $L_0 = \mathbb{P}^N_k$ and Lemma 5.10 follows from Lemmas 5.7 and 5.9. Hence we may assume $r \geq 2$. Let $X_1, \ldots, X_t$ be the irreducible components of X.

We saw in the proof of Lemma 5.8 that the Bertini theorems of Kleiman and Altman [1979, Theorem 1] and Jouanolou [1983] imply that an intersection of X with $(r-1)$ general hyperplanes containing $S \cup \{x\}$ in a large enough Veronese embedding of $\mathbb{P}^N_k$ is a reduced curve C whose intersection with every irreducible component of X is irreducible. This curve C contains $S \cup \{x\}$. We can also ensure that no component of C is contained in $\hat{f}(Z_{\mathrm{sing}}) \cup (X \setminus X_{\mathrm{fs}})$, it is regular at points away from X_{sing}, and for each component of $Z|_{C \times \hat{B}}$, its projection to $\hat{B}$ is nonconstant.

Hence, after replacing the embedding $\eta : X \hookrightarrow \mathbb{P}^N_k$ by its composition with a Veronese embedding of $\mathbb{P}^N_k$, we can find an $(r-1)$-tuple of general hyperplanes $(H_1, \dots, H_{r-1})$, each in $\mathrm{Gr}_{S\cup\{x\}}(N-1, \mathbb{P}^N_k)$, such that the linear subspace $L_0 = H_1 \cap \cdots \cap H_{r-1}$ has the following properties:

(a) L_0 is transverse to H_0.

(b) $C = L_0 \cap X$ is a reduced curve none of whose components lies in $\hat{f}(Z_{\mathrm{sing}}) \cup (X \setminus X_{\mathrm{fs}})$.

(c) $C \cap X_i$ is irreducible for each $1 \le i \le t$.

(d) C is regular at points away from X_{sing}.

(e) For each component of $Z|_{C\times\hat{B}}$, the projection to $\hat{B}$ is nonconstant.

Let $S' := (C \setminus \{x\}) \cap (\hat{f}(Z_{\mathrm{sing}}) \cup (X \setminus X_{\mathrm{fs}}) \cup S)$, which is a finite closed subset of C.

Note from the definition of the degree of the embedding $\eta : X \hookrightarrow \mathbb{P}^N_k$ that a general hyperplane inside L_0 will intersect C at $(d+1)$ distinct closed points. Applying Lemma 5.9 to the curve C, the finite set S', and $L_0 \simeq \mathbb{P}^{N-r+1}_k$ (which is regarded as the ambient projective space for C), there exists a dense open subset $\mathcal{U}_{C,S'} \subset \mathrm{Gr}_x(N-r, L_0)$ that satisfies the assertions (1)–(2) of Lemma 5.9. Note that as $N \gg r$, the subset $\mathrm{Gr}^{\mathrm{tr}}(L_0, N-1, \mathbb{P}^N_k)$ is dense open in $\mathrm{Gr}(N-1, \mathbb{P}^N_k)$.

Consider the regular map

$$\theta_{L_0} : \mathrm{Gr}^{\mathrm{tr}}(L_0, N-1, \mathbb{P}^N_k) \to \mathrm{Gr}(N-r, L_0), \tag{5-8}$$

given by $\theta_{L_0}(L) = L \cap L_0$.

One checks that θ_{L_0} is a surjective smooth morphism of relative dimension $r-1$. Since θ_{L_0} is a smooth and surjective morphism such that $\theta_{L_0}^{-1}(\mathrm{Gr}_x(N-r, L_0)) = \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$, we see that $\mathcal{U}_S := \theta_{L_0}^{-1}(\mathcal{U}_{C,S'})$ is a dense open subset of $\mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$.

We want to show that each $L \in \mathcal{U}_S(k)$ satisfies the desired conditions (1)–(4). This is a tautology, but let us write it in detail: suppose $L \in \mathcal{U}_S(k)$, i.e., $\theta_{L_0}(L) \cap S' = \varnothing$ and $\theta_{L_0}(L) = L \cap L_0$ satisfies (1)–(4) with Z replaced by $Z|_{C\times\hat{B}}$. Since $\theta_{L_0}(L) \cap ((X \setminus X_{\mathrm{fs}}) \cup S) = L \cap (L_0 \cap X) \cap ((X \setminus X_{\mathrm{fs}}) \cup S) = \theta_{L_0}(L) \cap C \cap ((X \setminus X_{\mathrm{fs}}) \cup S) \subset \theta_{L_0}(L) \cap S'$, and since $x \in X_{\mathrm{fs}}$, we see that $\theta_{L_0}(L) \cap ((X \setminus X_{\mathrm{fs}}) \cup S) = \varnothing$, proving (1).

Since L intersects L_0 transversely, which in turn intersects X transversely along X_{sm} by (b) and (d) above, we see that $\theta_{L_0}(L)$ intersects X transversely along X_{sm}, proving (2). Also, $\theta_{L_0}(L) \cap X = \theta_{L_0}(L) \cap C = \{x = c_0, c_1, \dots, c_d\}$ with $c_i \ne c_j$ for $i \ne j$, proving (3). Finally, since $(C \cap \hat{f}(Z_{\mathrm{sing}})) \setminus \{x\} \subset S'$ and since $\theta_{L_0}(L) \cap S' = \varnothing$, we see that Z is regular at all points lying over c_i for $1 \le i \le d$, proving (4).

We now prove (5). First of all, if the degree of any irreducible component of X inside $\mathbb{P}^N_k$ was less than or equal to two, before we do anything else, we first could have replaced $\mathbb{P}^N_k$ by its suitable Veronese embedding so as to ensure that the degree of any irreducible component of X is bigger than 2. In doing so, we see using Lemma 5.8 that the intersection L_0 of general $(r-1)$ hyperplanes $H_1, \dots, H_{r-1}$ lying in $\mathrm{Gr}_{S\cup\{x\}}(N-1, \mathbb{P}^N_k)$ will have the property that L_0 will satisfy the above (a)–(e), and $L_0 \cap X$ will be a reduced curve none of whose irreducible component is contained in a 2-dimensional linear subspace of $\mathbb{P}^N_k$. Note that since X is equidimensional and L_0 is general, the curve $L_0 \cap X$ will have this property too. This proves (5). The last property (6) is a direct consequence of (5), condition (2) of Lemma 5.9, which we already achieved from the beginning, and Proposition 5.6. □

Later, the set $\{c_1, \dots, c_d\}$ that we obtained in Lemma 5.10 will be taken to be $L^+(x)$ for $x \in \Sigma$, where Σ is the given set of finitely many closed regular points of X. This means the regularity of Z at points lying over the residual points $L^+(\Sigma)$. We will come back to this discussion, and it will be finished in Proposition 7.2.

6. Vertical separation of residual fibers

In this section, we prove some results which we shall need in order to prove the regularity of the residual cycle of Z along Σ. The main goal is to show that the distinct fibers, of the projection $Z \to X$ to the "horizontal axis" over the residual points of Σ (for a suitable linear projection) are mapped to disjoint sets under the projection $\hat{g} : Z \to \hat{B}$ to the "vertical axis". We call this property of linear projections, *the vertical separation of residual fibers*. We continue to use the Set-up+(fs) of Section 5B.

6A. *Separating residual fibers of Z along $\hat{B}$: the local case.* Let k be an algebraically closed field. In Lemma 5.2, under certain assumptions, we found a nonempty open subset of a Grassmannian such that each member L satisfies the properties (1)–(4) there. In Lemma 5.9, after choosing a Veronese reembedding into a bigger projective space, we achieved an additional noncollinearity of any three points of the hyperplane sections. It was generalized to Lemma 5.10 for $r \geq 1$.

In Section 6A, we want to further strengthen them, by constructing a nonempty open subset for which we have an additional separation property, which will be called property (I). This is eventually done in Proposition 6.5.

Up to Lemma 6.4, we assume the following. We suppose $r = 1$. We let $x \in X_{\mathrm{fs}}$ be a closed point and let $S \subset X \setminus \{x\}$ be another finite set of closed points. For any map $W \to X$ and a closed point $y \in X$, let W_y be the reduced fiber of W over y. We work under the Set-up of Lemma 5.9, which includes Lemma 5.2.

Since we want to prove a property called (I) by a kind of double induction argument on the pairs of numbers (m, n) with $0 \leq m \leq n \leq d-1$, we find it convenient to temporarily introduce some intermediate notations.

Definition 6.1. For $1 \leq n \leq d-1$ and $0 \leq m \leq n$, we say that a member $\underline{H} = (H, c_1, \dots, c_d) \in \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)(k) \times X^d$ is *(Z, x, m, n)-admissible*, if H satisfies the properties (1) and (2) of Lemma 5.9 with $H \cap X = \{x = c_0, c_1, \dots, c_d\}$, together with the additional property

$$(I)_{m,n} := \begin{cases} \hat{g}(Z_{c_i}) \cap \hat{g}(Z_{c_j}) = \varnothing & \text{for } 0 \leq i \neq j \leq n, \\ \hat{g}(Z_{c_i}) \cap \hat{g}(Z_{c_{n+1}}) = \varnothing & \text{for } 0 \leq i \leq m. \end{cases} \tag{6-1}$$

We remark that for $n = 0$ (thus we have just $(I)_{0,0}$), the first condition of (6-1) is empty.

Before anything else, we note the following elementary fact:

Lemma 6.2. *The projections $\hat{f} : Z \to X$ and $\hat{g} : Z \to \hat{B}$ are finite and the sets $\hat{g}(Z_x) \subset \hat{B}$ and $\hat{g}^{-1}(\hat{g}(Z_x)) \subset Z$ are finite subsets of closed points.*

Proof. Note that $\hat{f} : Z \to X$ is a projective morphism of reduced curves such that its restriction over the dense open subset X_{fs} of X is fs. Hence $\hat{f}$ is a projective quasifinite morphism, hence a finite morphism. Since $\hat{g}$ is a projective morphism from a curve which is nonconstant on each component of the source, it must also be finite. Since Z_x is a finite set, as Z is fs over X_{fs} and $x \in X_{\mathrm{fs}}$, the lemma now follows. □

Let $V_d \subset X^d$ be the nonempty open subset whose coordinates are all distinct from each other and distinct from x as well. More precisely, this is the complement of the union of all the small diagonals $\Delta_{i,j} \subset X^d$ defined by the equation $y_i = y_j$ for $1 \le i < j \le d$ as well as the subschemes given by $y_i = x$ for $1 \le i \le d$. Let $\pi : X^d \to \mathrm{Sym}^d(X) = X^d/\mathfrak{S}_d$ be the quotient map for the action by the symmetric group $\mathfrak{S}_d$ which permutes the coordinates. Since $\mathfrak{S}_d$ acts freely on $V_d \subset X^d$, the restriction $\pi : V_d \to \pi(V_d)$ is finite étale of degree $d!$.

Inside V_d, we consider the following subsets of "bad points" that do not satisfy the analogue of condition $(I)_{m,n}$ for $(y_1, \dots, y_d) \in V_d$. That is, for $y_0 := x$, let $D_0 := \varnothing$, while for $n \ge 1$, let $D_n \subset V_d$ be the subset of points $(y_1, \dots, y_d)$ such that $\hat{g}(Z_{y_i}) \cap \hat{g}(Z_{y_j}) \ne \varnothing$ for some $0 \le i \ne j \le n$ and $G^n_m \subset V_d$ be the subset of points such that $\hat{g}(Z_{y_i}) \cap \hat{g}(Z_{y_{n+1}}) \ne \varnothing$ for some $0 \le i \le m$.

Express $D_n = D_{n,1} \cup D_{n,2}$, where $D_{n,1}$ consists of the points $(y_1, \dots, y_d) \in V_d$ such that $\hat{g}(Z_{y_i}) \cap \hat{g}(Z_{y_j}) \ne \varnothing$ for some $1 \le i \ne j \le n$, while $D_{n,2}$ consists of the points $(y_1, \dots, y_d) \in V_d$ such that $\hat{g}(Z_{y_0}) \cap \hat{g}(Z_{y_i}) \ne \varnothing$ for some $1 \le i \le n$. We also write $G^n_m = \bigcup_{i=0}^m G^n_{m,i}$, where $G^n_{m,i}$ consists of the points $(y_1, \dots, y_d) \in V_d$ such that $\hat{g}(Z_{y_i}) \cap \hat{g}(Z_{y_{n+1}}) \ne \varnothing$ for $0 \le i \le m$. We check these "bad sets" are closed.

Lemma 6.3. *The subsets $D_{n,i}$ for $i = 1, 2$ and $G^n_{m,i}$ for $0 \le i \le m$ are closed subsets of V_d. In particular, D_n and G^n_m are closed subsets of V_d.*

Proof. Let $E_{n,1} \subset \hat{B}^d$ be the subset of points $(b_1, \dots, b_d)$ such that $b_i = b_j$ for some $1 \le i \ne j \le n$. Let $E_{n,2} \subset \hat{B}^d$ be the subset of points $(b_1, \dots, b_d)$ such that $b_i \in \hat{g}(Z_x)$ for some $1 \le i \le n$. The set $E_{n,1}$ is certainly closed in $\hat{B}^d$, while $E_{n,2}$ is closed in $\hat{B}^d$ because $\hat{g}(Z_x)$ is finite by Lemma 6.2. One checks that $D_{n,i} = \hat{f}^{\times d}((\hat{g}^{\times d})^{-1}(E_{n,i})) \cap V_d$ for $i = 1, 2$, where $\hat{f}^{\times d} : Z^d \to X^d$ and $\hat{g}^{\times d} : Z^d \to \hat{B}^d$ are the direct products of $\hat{f}$ and $\hat{g}$. Since $\hat{f}^{\times d}$ is finite by Lemma 6.2, this shows that $D_{n,i}$ is closed in V_d.

Similarly, let $J^n_{m,0} \subset \hat{B}^d$ be the subset of points $(b_1, \dots, b_d)$ such that $b_{n+1} \in \hat{g}(Z_x)$. This is closed since $\hat{g}(Z_x)$ is finite by Lemma 6.2. For $1 \le i \le m$, let $J^n_{m,i} \subset \hat{B}^d$ be the subset of points $(b_1, \dots, b_d)$ such that $b_i = b_{n+1}$. This is also closed. One checks that $G^n_{m,i} = \hat{f}^{\times d}((\hat{g}^{\times d})^{-1}(J^n_{m,i})) \cap V_d$, and this shows that $G^n_{m,i}$ is closed in V_d for $0 \le i \le m$. □

Coming back to the story, we let $\mathcal{U}_S \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ be the nonempty open set of Lemma 5.9. Let $\mathcal{U}_S \to \mathrm{Sym}^d(X)$ be the map given by $L \mapsto \sum_{i=1}^d [c_i]$, where $L \cap X = \{x = c_0, c_1, \dots, c_d\}$. By condition (3) of Lemma 5.2, its image is in $\pi(V_d)$. Define $\mathcal{V}_S$ by the Cartesian diagram

$$\begin{array}{ccc} \mathcal{V}_S & \xrightarrow{e} & V_d \\ {\scriptstyle\psi}\downarrow & & \downarrow{\scriptstyle\pi} \\ \mathcal{U}_S & \longrightarrow & \pi(V_d), \end{array} \tag{6-2}$$

so that ψ is a finite surjective étale map. The set $\mathcal{V}_S \setminus e^{-1}(D_n \cup G^n_m)$ is open in $\mathcal{V}_S$ by Lemma 6.3. Via the open map ψ, we define the open subset $\mathcal{U}^S_{m,n} := \psi(\mathcal{V}_S \setminus e^{-1}(D_n \cup G^n_m)) \subset \mathcal{U}_S$. This is open in $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$.

Lemma 6.4. *For $0 \le n \le d-1$ and $0 \le m \le n$, the subset $\mathcal{U}^S_{m,n} \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ is nonempty. In particular, it is a dense open subset of $\mathrm{Gr}_x(N-1, \mathbb{P}^n_k)$.*

Proof. Step 1. $\mathcal{U}^S_{0,0} \ne \varnothing$. Note that condition $(I)_{0,0}$ is independent of the choice of an x-fixing order on $L \cap X$. Let $T = S \cup (\hat{f}(\hat{g}^{-1}(\hat{g}(Z_x))) \setminus \{x\})$. This is a finite closed subset of X by Lemma 6.2. Applying Lemma 5.2 to T (in the place of S there), we obtain a dense open subset $\mathcal{U}_T$ of $\mathcal{U}_S \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$. On the other hand, condition (1) (in Lemma 5.2) for T implies that for each $L \in \mathcal{U}_T(k)$, we have $L \cap (\hat{f}(\hat{g}^{-1}(\hat{g}(Z_x))) \setminus \{x\}) = \varnothing$, which shows that $\hat{g}(Z_{c_0}) \cap \hat{g}(Z_{c_j}) = \varnothing$ for each $j \ne 0$ when $L \cap X = \{x = c_0, c_1, \ldots, c_d\}$, for every x-fixing order on $L \cap X$. Thus $(I)_{0,0}$ holds, and $\mathcal{U}_T \subset \mathcal{U}^S_{0,0}$, in particular $\mathcal{U}^S_{0,0} \ne \varnothing$.

Step 2. *For $0 \le n \le d-2$, if $\mathcal{U}^S_{n,n} \ne \varnothing$, then $\mathcal{U}^S_{0,n+1} \ne \varnothing$.* If $\mathcal{U}^S_{n,n} \ne \varnothing$, then it is a dense open subset of $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$. In particular, for the dense open subset $\mathcal{U}_T \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ of Step 1, the intersection $\mathcal{U}^S_{n,n} \cap \mathcal{U}_T$ is dense open in $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$. But, by definition, one notes that $\mathcal{U}^S_{n,n} \cap \mathcal{U}_T \subset \mathcal{U}^S_{0,n+1}$ so that $\mathcal{U}^S_{0,n+1} \ne \varnothing$.

Step 3. *For $0 \le n \le d-1$ and $0 \le m \le n-1$, if $\mathcal{U}^S_{m,n} \ne \varnothing$, then $\mathcal{U}^S_{m+1,n} \ne \varnothing$.* If $\mathcal{U}^S_{m,n} \ne \varnothing$, then it is dense open in $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$. For the dense open subset $\mathcal{U}_T \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ of Step 1, the intersection $\mathcal{U}^S_{m,n} \cap \mathcal{U}_T$ is therefore nonempty dense open in $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$.

Fix an element $L'_0 \in (\mathcal{U}^S_{m,n} \cap \mathcal{U}_T)(k)$ and let $L'_0 \cap X = \{x = c_0, c_1, \ldots, c_d\}$. Since every k-point of $\mathcal{U}^S_{m,n}$ satisfies condition (2) of Lemma 5.9, we know that no three points of $L'_0 \cap X$ are collinear. Thus $\{c_0, c_{m+1}, c_{n+1}\}$ are not collinear so that when $\ell = \mathrm{Sec}(\{c_0\}, \{c_{m+1}\})$ is the line joining c_0 and c_{m+1}, it does not pass through c_{n+1}.

We let $P = \mathrm{Sec}(\{c_{n+1}\}, \ell)$. The subspace $\mathrm{Gr}_\ell(N-1, \mathbb{P}^N_k)$ is of dimension $N-2$ and $\mathrm{Gr}_P(N-1, \mathbb{P}^N_k)$ is a closed subspace of $\mathrm{Gr}_\ell(N-1, \mathbb{P}^N_k)$ of dimension $N-3$ (see Lemma 5.1). Because we may assume $N \ge 3$, there is a one-parameter family (actually isomorphic to $\mathbb{P}^1_k$) $\mathcal{B}$ in $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ such that (i) $\{L'_0\} \in \mathcal{B}$, (ii) every member of the family $\mathcal{B}$ passes through both of $\{c_0, c_{m+1}\}$ and (iii) a general member does not pass through c_{n+1}. Since $\mathcal{U}^S_{m,n} \cap \mathcal{U}_T$ is dense open in $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ and $L'_0 \in \mathcal{U}^S_{m,n} \cap \mathcal{U}_T \cap \mathrm{Gr}_L(N-1, \mathbb{P}^N_k)$, the latter is dense open in $\mathrm{Gr}_\ell(N-1, \mathbb{P}^N_k)$. Hence, a general member of $\mathcal{B}$ is contained in $\mathcal{U}^S_{m,n} \cap \mathcal{U}_T$.

Let $W \subset \mathcal{B} \cap \mathcal{U}^S_{m,n} \cap \mathcal{U}_T$ be a smooth affine irreducible (rational) curve passing through $\{L'_0\}$. Consider again the quotient map $\pi : X^d \to \mathrm{Sym}^d(X) = X^d/\mathfrak{S}_d$, and the finite étale map $\pi : V_d \to \pi(V_d)$ for the open set V_d defined previously in (6-2). Consider the map $W \to \pi(V_d)$ given by $L \mapsto \sum_{i=1}^d [y_i]$, where

$L \cap X = \{x = y_0, y_1, \dots, y_d\}$. This yields the Cartesian product

$$\begin{array}{ccccc} W' & \hookrightarrow & \tilde{W} & \longrightarrow & V_d \\ & \searrow & \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \pi} \\ & & W & \longrightarrow & \pi(V_d) \end{array} \qquad e : W' \to V_d \tag{6-3}$$

so that ψ is finite and étale. Note also that the members of $\tilde{W}$ can be represented by $\underline{L} = (L, y_1, \dots, y_d) \in W \times V_d$ such that $L \cap X = \{x = y_0, y_1, \dots, y_d\}$. We let $W' \subset \tilde{W}$ be the component containing the point $(L'_0, c_1, \dots, c_d)$. For the "bad" closed subsets $D_n, G^n_{m+1} \subset V_d$ of Lemma 6.3, we have:

Claim. *$\mathcal{Y} := e^{-1}(D_n \cup G^n_{m+1})$ is a proper closed subset of W'.*

Proof. That this is a closed subset of W' follows by Lemma 6.3. We need to show that this is a proper subset. Note that $D_n = D_{n,1} \cup D_{n,2}$ and $G^n_{m+1} = \bigcup_{i=0}^{m+1} G^n_{m+1,i}$. We analyze each piece of them in what follows.

Case 1: We first show that $e^{-1}(D_{n,2}) = \varnothing$ and $e^{-1}(G^n_{m+1,0}) = \varnothing$.

Note that we had $W \subset \mathcal{B} \cap \mathcal{U}^S_{m,n} \cap \mathcal{U}_T$, where $\mathcal{U}_T$ is as in Lemma 5.2. Here, condition (1) of Lemma 5.2 (and S replaced by T) reads as '$L \cap ((X \setminus X_{\mathrm{fs}}) \cup T) = \varnothing$' for each $L \in \mathcal{U}_T(k)$. So, for every $L \in W(k)$, this is disjoint from $T = S \cup (\hat{f}(\hat{g}^{-1}(\hat{g}(Z_x))) \setminus \{x\})$. Hence, if $e^{-1}(D_{n,2}) \neq \varnothing$, then it gives an element $L \in W(k)$ such that $L \cap X = \{x = y_0, y_1, \dots, y_d\}$ satisfies $\hat{g}(Z_{y_0}) \cap \hat{g}(Z_{y_i}) \neq \varnothing$ for some $1 \le i \le n$, so that L intersects with a point of T, contradicting the above choice of W. Hence $e^{-1}(D_{n,2}) = \varnothing$. An identical argument shows that $e^{-1}(G^n_{m+1,0}) = \varnothing$.

Case 2: We now show that $e^{-1}(D_{n,1})$ and $e^{-1}(G^n_{m+1,i})$ for $1 \le i \le m$ are finite.

To do so, it is enough to show that these closed subsets are proper in W', as W' is an irreducible curve. Suppose $e^{-1}(D_{n,1}) = W'$. In particular $\underline{L}'_0 := (L'_0, c_1, \dots, c_d) \in e^{-1}(D_{n,1})$, so that $(c_1, \dots, c_d) \in D_{n,1}$, so $\hat{g}(Z_{c_i}) \cap \hat{g}(Z_{c_j}) \neq \varnothing$ for some $1 \le i \neq j \le n$. But, this contradicts that $L'_0 \in \mathcal{U}^S_{m,n}(k)$. Hence, $e^{-1}(D_{n,1})$ is proper closed in W'. By the same argument, we have $|e^{-1}(G^n_{m+1,i})| < \infty$.

Case 3: It remains to show that $|e^{-1}(G^n_{m+1,m+1})| < \infty$.

To do so, we will make use of our choice of W that $W \subset \mathcal{B}$. Recall that $\mathcal{B} \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ is a one-parameter family containing $\{L'_0\}$ such that every member of $\mathcal{B}$ passes through $\{c_0, c_{m+1}\}$, while a general member does not pass through c_{n+1}.

Consider the composite $q : W' \xrightarrow{e} V_d \to X^2$, where the last arrow takes $(y_1, \dots, y_d)$ to $(y_{m+1}, y_{n+1}) \in X^2$. Since every $L \in W(k) \subset \mathcal{B}(k)$ contains c_{m+1} by construction, the composition of q with the first projection $X^2 \to X$, taking (y_{m+1}, y_{n+1}) to y_{m+1}, is the constant map that takes all of W' to $c_{m+1} \in X$. On the other hand, the general member $L \in W(k)$ does not contain c_{n+1}. This implies that the composite of q with the second projection $X^2 \to X$, taking (y_{m+1}, y_{n+1}) to y_{n+1}, is nonconstant. Hence, the map q is nonconstant and the image $q(W')$ in X^2 is an irreducible curve contained in $\{c_{m+1}\} \times X \cong X$ (recall that k is assumed to be algebraically closed).

Write it as $W' \xrightarrow{u} q(W') \xrightarrow{v} X$, where u is induced by q and v is the projection to the coordinate y_{n+1}. Since both u and v are nonconstant morphisms of irreducible curves, they are dominant and quasifinite. In particular, the composite $v \circ u$ is quasifinite. Note that by definition, $e^{-1}(G^n_{m+1,m+1}) \subset \{(L, y_1, \ldots, y_d) \in W' \mid y_{n+1} \in S_1\} = (v \circ u)^{-1}(S_1)$, where $S_1 := f(\hat{g}^{-1}(\hat{g}(Z_{c_{m+1}})))$. Since $\hat{f}$ and $\hat{g}$ are finite by Lemma 6.2, the set S_1 is finite, thus $(v \circ u)^{-1}(S_1)$ is a finite set. Hence, we have $|e^{-1}(G^n_{m+1,m+1})| < \infty$, being a subset of a finite set. This finishes the proof of the Claim. □

Back to the proof of Step 3, since the set $\mathcal{Y}$ of the Claim is finite, the subset $W' \setminus \mathcal{Y} \subset W'$ is nonempty open. Since ψ is an open map and $W' \subset \tilde{W}$ is open subset such that $W' \to W$ is surjective, it follows that $\psi(W' \setminus \mathcal{Y}) \subset W$ is a nonempty (thus dense) open subset. By construction, $\psi(W' \setminus \mathcal{Y}) \subset \mathcal{U}^S_{m+1,n}$. In particular, we get $\mathcal{U}^S_{m+1,n} \neq \varnothing$. This proves Step 3.

Back to the proof of the lemma, by inductively applying the above three steps, we deduce that each $\mathcal{U}^S_{m,n}$ is a dense open subset of $\mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$. □

Now we allow $r \geq 1$. We can strengthen Lemma 5.10 as follows:

Proposition 6.5. *We follow the notations and the assumptions of Lemma 5.10. Let $r \geq 1$. After replacing $\mathbb{P}^N_k$ by a bigger projective space via Veronese if necessary, we have the following property: given any hyperplane $H_0 \subset \mathbb{P}^N_k$ disjoint from $S \cup \{x\}$ and a general $L_0 \in \mathrm{Gr}^{\mathrm{tr}}_{S\cup\{x\}}(H_0, N-r+1, \mathbb{P}^N_k)(k)$, there exists a dense open subset $\mathcal{U}^S_x \subset \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$ such that each $L \in \mathcal{U}^S_x(k)$ satisfies the properties (1)–(6) of Lemma 5.10 as well as the additional property $(I): \hat{g}(Z_{c_i}) \cap \hat{g}(Z_{c_j}) = \varnothing$ for each pair $0 \leq i \neq j \leq d$.*

Proof. The $r = 1$ case of the proposition follows from Lemma 6.4 with $(m, n) = (d-1, d-1)$. So we assume $r \geq 2$.

We use an argument of reduction to the $r = 1$ case as we did in Lemma 5.10. Using the notations there, choose a reembedding $\eta : X \hookrightarrow \mathbb{P}^N_k$, a general $L_0 \in \mathrm{Gr}(N-r+1, \mathbb{P}^N_k)(k)$ and $C = L_0 \cap X$ as in Lemma 5.10. Let $S' := (C \setminus \{x\}) \cap (\hat{f}(Z_{\mathrm{sing}}) \cup (X \setminus X_{\mathrm{fs}}) \cup S)$ and $W = Z|_{C \times \hat{B}}$. Applying the "$r = 1$" case of the proposition (proven in Lemma 6.4) to C, S' and W, with the identification $L_0 \simeq \mathbb{P}^{N-r+1}_k$, there is a dense open subset $\mathcal{U}' \subset \mathrm{Gr}_x(N-r, L_0)$ that satisfies the properties of Proposition 6.5 for $r = 1$ case. (In terms of the notations of Lemma 6.4, we have $\mathcal{U}' = \mathcal{U}^{S'}_{d-1,d-1}$.) Note that Lemma 6.4 is applicable to C by property (6) of Lemma 5.10.

Recall now that we had a smooth surjective morphism of varieties

$$\theta_{L_0} : \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k) \to \mathrm{Gr}_x(N-r, L_0)$$

from (5-8). So, the inverse image $\mathcal{U}^S_x := \theta^{-1}_{L_0}(\mathcal{U}')$ is a dense open subset of $\mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$. We claim that this $\mathcal{U}^S_x$ fulfills the requirements of the proposition for $r \geq 2$ case.

Indeed, since $W = Z|_{C \times \hat{B}}$, we see that $Z_y = W_y$ and hence $\hat{g}(Z_y) = \hat{g}(W_y)$ for any closed point $y \in C$. Hence, for $L \in \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)(k)$ with $\theta_{L_0}(L) \cap X = (L \cap L_0) \cap C = \{x = c_0, c_1, \ldots, c_d\}$, condition (I) is satisfied if and only if condition (I) is satisfied for $\theta_{L_0}(L)$ with X replaced by the curve C. This means $L \in \mathcal{U}^S_x(k)$ satisfies the proposition, as desired. □

6B. ***Separating residual fibers of Z along $\hat{B}$: the semilocal case.*** Note that in the statement of Proposition 6.5, the dense open subset that we found depends on the choice of a single regular closed point $x \in X$. We want to extend it to a finite subset Σ of regular points. This issue will be completely resolved in Proposition 7.2 by using the "cone admissibility" condition, which we develop as property (3) of the following Proposition 6.6. One further aspect on étaleness is studied in Section 6C.

Recall that when $M \subset \mathbb{P}^N_k$ is a linear subspace and $x \in \mathbb{P}^N_k$ is a closed point, after the base change $\mathrm{Spec}(k(x)) \to \mathrm{Spec}(k)$, the cone $C_x(M) = \mathrm{Sec}(\{x\}, M)$ is the smallest linear subspace containing both x and M. When $x \notin M$, we have $\dim(C_x(M)) = \dim(M) + 1$. In this article, we need to use the cones only when k is algebraically closed, so that no confusion will arise.

Proposition 6.6. *Let k be an algebraically closed field. We are under the Set-up+*(fs) *of Section 5B. After replacing the embedding $X \hookrightarrow \mathbb{P}^N_k$ by a bigger one via a Veronese embedding if necessary, we have the following: for the given hyperplane $H \subset \mathbb{P}^N_k$ disjoint from Σ and a general $L_0 \in \mathrm{Gr}^{\mathrm{tr}}(H, N-r+1, \mathbb{P}^N_k)(k)$, there exists a dense open subset $\mathcal{W} \subset \mathrm{Gr}(N-2, H)$ such that each $M \in \mathcal{W}(k)$ satisfies the following properties:*

(1) *M intersects L_0 transversely.*

(2) *$M \cap L_0 \cap X = \varnothing$.*

(3) *For each $x \in \Sigma$, the cone $C_x(M)$ lies in $\mathcal{U}_x^{\Sigma\setminus\{x\}}(k)$ for the open subset $\mathcal{U}_x^{\Sigma\setminus\{x\}} \subset \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$ of Proposition 6.5.*

Proof. Note that if $\Sigma = \{x_1, \dots, x_n\}$, then condition (3) consists of conditions $(3)_i : C_{x_i}(M) \in \mathcal{U}_{x_i}^{\Sigma\setminus\{x_i\}}(k)$ for $1 \le i \le n$. Suppose we proved the existence of a dense open subset $\mathcal{W}_i \subset \mathrm{Gr}(N-2, H)$ for which each member $M \in \mathcal{W}_i(k)$ satisfies conditions (1), (2), and $(3)_i$ for each $1 \le i \le n$. Then we can take $\mathcal{W} := \bigcap_{i=1}^n \mathcal{W}_i$, which is again a dense open subset of $\mathrm{Gr}(N-2, H)$. Hence, it is enough to prove the existence of those $\mathcal{W}_i$. Without loss of generality, we may assume $i = 1$. For notational simplicity, we let $x := x_1$ and $T := \Sigma \setminus \{x_1\}$. We note also that when $r = 1$, we have $\mathrm{Gr}^{\mathrm{tr}}(H, N-r+1, \mathbb{P}^N_k) = \mathrm{Gr}(N, \mathbb{P}^N_k) = \{\mathbb{P}^N_k\}$ so that the choice of L_0 plays no role. We prove the proposition for the cases of $r = 1$ and $r \ge 2$ separately.

Step 1. Suppose $r = 1$. Consider the affine morphism of schemes

$$\vartheta_x : \mathrm{Gr}(x, N-2, \mathbb{P}^N_k) \to \mathrm{Gr}_x(N-1, \mathbb{P}^N_k),\ L \mapsto C_x(L). \tag{6-4}$$

This is a smooth surjective morphism, and defines a vector bundle of rank $N-1$. For the closed irreducible subscheme $\mathrm{Gr}(N-2, H) \hookrightarrow \mathrm{Gr}(x, N-2, \mathbb{P}^N_k)$, the restriction $\vartheta_{x,H} : \mathrm{Gr}(N-2, H) \to \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ of ϑ_x, is an isomorphism.

Let $\mathcal{U}_x^T \subset \mathrm{Gr}_x(N-1, \mathbb{P}^N_k)$ be the dense open subset of Proposition 6.5, applied to x, T and $H_0 = H$ for $r = 1$. Then $\vartheta_{x,H}^{-1}(\mathcal{U}_x^T)$ is a dense open subset of $\mathrm{Gr}(N-2, H)$. Since $\mathrm{Gr}(X, N-2, H)$ is its dense open subset by Lemma 4.3, so is the intersection $\mathcal{W}_1 := \vartheta_{x,H}^{-1}(\mathcal{U}_x^T) \cap \mathrm{Gr}(X, N-2, H)$ in $\mathrm{Gr}(N-2, H)$. One checks that this satisfies the required conditions (1), (2), and $(3)_1$, proving the proposition for $r = 1$.

Step 2. Suppose now that $r \geq 2$. As we did previously in Lemma 5.10 with $H_0 = H$ via a Bertini argument of [Kleiman and Altman 1979], we choose a reembedding $\eta : X \hookrightarrow \mathbb{P}^N_k$, a general $L_0 \in \mathrm{Gr}^{\mathrm{tr}}(H, N-r+1, \mathbb{P}^N_k)(k)$, a curve $C = L_0 \cap X$, and $Z|_{C \times \hat{B}}$. Consider again the map in (6-4). When L_0 contains x, this ϑ_x induces a smooth surjective map $\vartheta_x^{L_0} : \mathrm{Gr}^{\mathrm{tr}}(L_0, x, N-2, \mathbb{P}^N_k) \to \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$, where we recall that $\mathrm{Gr}^{\mathrm{tr}}(L_0, x, n, \mathbb{P}^N_k) := \mathrm{Gr}^{\mathrm{tr}}(L_0, n, \mathbb{P}^N_k) \cap \mathrm{Gr}(x, n, \mathbb{P}^N_k)$. This restricts to give $\vartheta_{x,H} : \mathrm{Gr}^{\mathrm{tr}}(L_0, N-2, H) \to \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$. One checks that this map is an inclusion whose image is the dense open subset $\mathrm{Gr}^{\mathrm{tr}}_x(L_0 \cap H, N-1, \mathbb{P}^N_k)$. As $H \cap \{x\} = \varnothing$, we see that $\mathrm{Gr}^{\mathrm{tr}}_x(L_0 \cap H, N-1, \mathbb{P}^N_k)$ coincides with $\mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$. This implies that $\vartheta_{x,H}$ is an isomorphism.

Let $\mathcal{U}^T_x \subset \mathrm{Gr}^{\mathrm{tr}}_x(L_0, N-1, \mathbb{P}^N_k)$ be the dense open subset of Proposition 6.5 applied to x, T, and $H_0 = H$ for $r \geq 2$. Since $\vartheta_{x,H}$ is an isomorphism, $\vartheta^{-1}_{x,H}(\mathcal{U}^T_x)$ is dense open in $\mathrm{Gr}^{\mathrm{tr}}(L_0, N-2, H)$, thus dense open in $\mathrm{Gr}(N-2, H)$. Combining this with Lemma 4.3, we conclude that $\mathcal{W}_1 := \vartheta^{-1}_{x,H}(\mathcal{U}^T_x) \cap \mathrm{Gr}(C, N-2, H)$ is dense open in $\mathrm{Gr}(N-2, H)$. One checks that each $M \in \mathcal{W}(k)$ satisfies the required conditions (1), (2), and $(3)_1$. This finishes the proof. □

6C. *Étaleness of linear projections at $L^+(\Sigma)$.* Recall that we had obtained a linear projection $\phi_L : X \to \mathbb{P}^r_k$ that is étale at each point of Σ in condition (1) of Lemma 4.8. Unfortunately, this is not quite enough for us. We need to have L such that ϕ_L is étale at each point of $L^+(\Sigma)$ as well. We show that we can achieve this as a geometric consequence of condition (3) of Proposition 6.6.

Part of the requirement of Proposition 6.6 that $C_x(M)$ lies in $\mathcal{U}^{\Sigma \setminus \{x\}}_x(k)$ for the open set $\mathcal{U}^{\Sigma \setminus \{x\}}_x$ is that $C_x(M)$ intersects $X_{\mathrm{fs}} \subset X_{\mathrm{sm}}$ transversely. This comes from condition (2) of Lemma 5.10. Here is its geometric meaning.

Lemma 6.7. *Let k be an algebraically closed field and let $L \in \mathrm{Gr}(X, N-r-1, \mathbb{P}^N_k)(k)$. Let $\mathbb{P}^r_k$ be a linear subspace of $\mathbb{P}^N_k$ such that $L \cap \mathbb{P}^r_k = \varnothing$. Let $y \in \mathbb{P}^r_k$ be a closed point such that $C_y(L) \cap X_{\mathrm{sing}} = \varnothing$. Then $C_y(L)$ intersects X transversely if and only if the linear projection $\phi_L : X \to \mathbb{P}^r_k$ away from L is finite and étale over an affine neighborhood of y in $\mathbb{P}^r_k$.*

Proof. ($\Rightarrow$) Suppose that $C_y(L)$ intersects X transversely and let $E := C_y(L) \cap X$ be this scheme-theoretic intersection. Since k is perfect while $C_y(L)$ and X_{sm} have the complementary dimensions $N-r$ and r in $\mathbb{P}^N_k$, respectively, the transverse intersection is equivalent to saying that E is smooth, $|E| < \infty$, and each point of E is a simple regular point of X_{sm}. Because we are given that $X \cap L = \varnothing$ and $L \subset C_y(L)$, we see that $C_y(L) \cap X = (C_y(L) \setminus L) \cap X$, which is precisely the scheme-theoretic fiber $\phi_L^{-1}(y)$ over $y \in \mathbb{P}^r_k$.

Since $C_y(L) \cap X_{\mathrm{sing}} = \varnothing$, we see that $\phi_L^{-1}(y) \cap X_{\mathrm{sing}} = \varnothing$. Since ϕ_L is finite, $\phi_L(X_{\mathrm{sing}})$ is a closed subscheme of $\mathbb{P}^r_k$ not meeting y. Hence, there is an affine open $U \subset \mathbb{P}^r_k$ containing y such that $\phi_L^{-1}(U)$ is regular. We therefore get a Cartesian square

$$\begin{array}{ccc} E & \longrightarrow & \phi_L^{-1}(U) \\ \Big\downarrow{\scriptstyle \phi_L^y} & & \Big\downarrow{\scriptstyle \phi_L} \\ \mathrm{Spec}(k(y)) & \longrightarrow & U \end{array} \tag{6-5}$$

such that ϕ_L^y is smooth. Since ϕ_L is a finite map of regular affine schemes over k, it is flat by [Hartshorne 1977, Exercise III-10.9, page 276] (or [EGA IV$_2$ 1965, Proposition (6.1.5), page 136]). It follows therefore by [Hartshorne 1977, Exercise III-10.2, page 275] (or [EGA IV$_3$ 1966, Théorème (12.2.4)(iii), page 183]) that there is an affine neighborhood of y in U over which the restriction of the map ϕ_L is smooth, thus finite and étale.

($\Leftarrow$) If ϕ_L is étale over a neighborhood of y, then its base change to $\operatorname{Spec}(k(y))$, i.e., the map $\phi_L^y : E = C_y(L) \cap X \to \operatorname{Spec}(k(y))$ from the scheme-theoretic intersection is étale. Since $k = k(y)$, this means E is smooth over k so that the intersection is transverse. □

Corollary 6.8. *Let k be an algebraically closed field. Let $L \in \operatorname{Gr}(X, N-r-1, \mathbb{P}_k^N)(k)$ and realize the linear projection $\phi_L : X \to \mathbb{P}_k^r$ for a linear subspace $\mathbb{P}_k^r \subset \mathbb{P}_k^N$ such that $L \cap \mathbb{P}_k^r = \varnothing$. Suppose that for each $x \in \Sigma$, we have $C_x(L) \cap X_{\text{sing}} = \varnothing$ and $C_x(L)$ intersects X_{sm} transversely. Then there is an affine open neighborhood $U \subset \mathbb{A}_k^r$ of $\phi_L(\Sigma)$ such that $\phi_L : \phi_L^{-1}(U) \to U$ is finite and étale. In particular, $\phi_L : X \to \mathbb{P}_k^r$ is étale at every point of $\phi_L^{-1}(\phi_L(\Sigma))$.*

Proof. Note that for each $x \in \Sigma$, we have $C_x(L) = C_{\phi_L(x)}(L)$ since ϕ_L is given with a chosen internal linear subspace $\mathbb{P}_k^r \subset \mathbb{P}_k^N$. Since $C_{\phi_L(x)}(L) \cap X_{\text{sing}} = \varnothing$ and $C_{\phi_L(x)}(L)$ intersects X transversely, Lemma 6.7 says that there is an affine open neighborhood $U_x \subset \mathbb{P}_k^r$ of $\phi_L(x)$ such that $\phi_L : \phi_L^{-1}(U_x) \to U_x$ is finite and étale. Hence, for $U := \bigcup_{x \in \Sigma} U_x$, the restriction $\phi_L : \phi_L^{-1}(U) \to U$ is finite and étale. By Lemma 2.3, we may shrink this U into an affine open neighborhood of $\phi_L(\Sigma)$. This implies the corollary. □

7. Regularity of residual cycles over finite closed points

Our goal in Section 7 is to study the regularity of the residual cycles using the technique of vertical separation of residual fibers studied in Section 6. We continue to work with the Set-up+(fs) of Section 5B. In particular, for each irreducible component Z_i of Z, the projection $Z_i \to \hat{B}$ is nonconstant and the projection $Z_i \to X$ is fs over X_{fs}.

7A. ***Admissible sets.*** Property (I) in Proposition 6.5 encourages the following definition, that encodes a set of data needed to achieve the remaining properties of residual cycles.

Definition 7.1. Let k be an infinite perfect field. Let $x \in X_{\text{fs}}$ be a closed point. A finite subset $D \subset X_{\text{fs}}$ of distinct closed points is called (Z, x)-*admissible* if (1) $x \in D$, (2) Z is regular at all points lying over $D \setminus \{x\}$, and (3) $\hat{g}(Z_{x_1}) \cap \hat{g}(Z_{x_2}) = \varnothing$ for each distinct pair $x_1 \neq x_2$ in D.

The following application of Proposition 6.6 will be a basis for our proof of the regularity of the residual cycles along Σ. We study it for $k = \bar{k}$ case, but it will soon be generalized gradually.

Proposition 7.2. *Let k be an algebraically closed field. We are under the Set-up+*(fs) *of Section 5B. Let $Y \subset X$ be a closed subset of dimension at most $r-1$. After replacing the embedding $\eta : X \hookrightarrow \mathbb{P}_k^N$ by a bigger one via a Veronese embedding if necessary, we have the following: for the given hyperplane $H \subset \mathbb{P}_k^N$ disjoint from Σ, there is a dense open subset $\mathcal{U} \subset \operatorname{Gr}(N-r-1, H)$ such that for each $L \in \mathcal{U}(k)$,* *we have* $L \cap X = \varnothing$ *so that there is a finite and surjective linear projection* $\phi_L : X \to \mathbb{P}^r_k$. *Furthermore, it satisfies the following properties*:

(1) *The map* $\phi_L : \phi_L^{-1}(U) \to U$ *is étale for some affine open* $U \subset \mathbb{P}^r_k$ *containing* $\phi_L(\Sigma)$.

(2) $\phi_L(x) \neq \phi_L(x')$ *for each pair* $x \neq x' \in \Sigma$.

(3) $k(\phi_L(x)) \xrightarrow{\sim} k(x)$ *for each* $x \in \Sigma$.

(4) $L^+(x) \cap Y = \varnothing$ *for each* $x \in \Sigma$.

(5) $\phi_L^{-1}(\phi_L(x))$ *is* (Z, x)*-admissible for each* $x \in \Sigma$.

Proof. As in the proof of Proposition 6.5, we can choose a reembedding $\eta : X \hookrightarrow \mathbb{P}^N_k$ and a dense open subset $\mathcal{U}_1 \subset \mathrm{Gr}_\Sigma(N - r + 1, \mathbb{P}^N_k)$ such that each $L' \in \mathcal{U}_1(k)$ satisfies the condition that $L' \cap X$ is a reduced curve none of whose components is contained in Y, is regular away from X_{sing}, and for each component of $Z|_{X \times \hat{B}}$, the projection to $\hat{B}$ is nonconstant. Since $H \cap \Sigma = \varnothing$, we see that $\mathcal{U}_0 := \mathrm{Gr}_\Sigma(N-r+1, \mathbb{P}^N_k) \cap \mathrm{Gr}^{\mathrm{tr}}(H, N-r+1, \mathbb{P}^N_k) \neq \varnothing$. It follows that this intersection is dense open in $\mathrm{Gr}_\Sigma(N-r+1, \mathbb{P}^N_k)$. Letting $\mathcal{U}'_1 := \mathcal{U}_0 \cap \mathcal{U}_1$, we see that $\mathcal{U}'_1$, is a dense open subset of $\mathrm{Gr}_\Sigma(N-r+1, \mathbb{P}^N_k)$ such that each $L' \in \mathcal{U}'_1(k)$ intersects H transversely and $L' \cap X$ is a curve of the above type.

Choose $L_0 \in \mathcal{U}'_1(k)$. (N.B., When $r = 1$, there is a unique choice $L_0 = \mathbb{P}^N_k$ automatically, and we have $C = X$.) We now apply Proposition 6.6. It follows that there exists a dense open subset $\mathcal{W} \subset \mathrm{Gr}(N-2, H)$ such that each $M \in \mathcal{W}(k)$ satisfies conditions (1)–(3) of Proposition 6.6.

On the other hand, the subset $\mathrm{Gr}^{\mathrm{tr}}(L_0, \mathrm{Sec}(\Sigma, Y \cap C), N-2, H) \subset \mathrm{Gr}(N-2, H)$ is a dense open subset by Lemmas 4.3 and 4.4. Hence $\mathcal{V}' := \mathcal{W} \cap \mathrm{Gr}^{\mathrm{tr}}(L_0, \mathrm{Sec}(\Sigma, Y \cap C), N-2, H) \subset \mathrm{Gr}(N-2, H)$ is a dense open subset. Since L_0 intersects H transversely, the map $\mathrm{Gr}(N-2, H) \to \mathrm{Gr}(N-r-1, H)$, given by $M \mapsto L_0 \cap M$, is smooth and surjective (note that $N \gg r$). In particular, the image $\mathcal{U}_2 := \{L_0 \cap M \in \mathrm{Gr}(N-r-1, H) \mid M \in \mathcal{V}\}$ of $\mathcal{V}$ is a dense open subset of $\mathrm{Gr}(N-r-1, H)$. Let $\mathcal{U}'_2 \subset \mathrm{Gr}(X, N-r-1, H)$ be the dense open set of Lemma 4.8 so that $\mathcal{U} := \mathcal{U}_2 \cap \mathcal{U}'_2 \subset \mathrm{Gr}(X, N-r-1, H)$ is a dense open subset.

Claim. *Each* $L \in \mathcal{U}(k)$ *satisfies the properties* (1)–(5) *of the proposition.*

We ignore L from the notation of ϕ_L for simplicity. Before we prove the claim, we note that $\phi^{-1}(\phi(\Sigma)) \subset X_{\mathrm{fs}}$, as follows from condition (3) of Proposition 6.6 which includes condition (1) of Lemma 5.10.

Now, condition (3) of Proposition 6.6 also implies that, by Corollary 6.8, there is an affine neighborhood U of $\phi(\Sigma)$ such that $\phi^{-1}(U) \to U$ is finite étale. This proves (1).

Since our open set $\mathcal{U}$ is contained in the open set of Lemma 4.8, we can use the properties there, too. Condition (2) of Lemma 4.8 is that the map ϕ is injective on Σ, proving (2). Condition (3) is obvious because k is assumed to be algebraically closed. Condition (4) follows from our choice of M (thus of L) that it avoids the cone involving Y.

We now prove (5). We need to verify the three conditions of the (Z, x)-admissibility of Definition 7.1 for each $x \in \Sigma$. Condition (1) of Definition 7.1 that $x \in \phi^{-1}(\phi(x))$ is obvious.

We prove condition (2) of Definition 7.1. Condition (3) of Proposition 6.6 says that condition (4) of Lemma 5.10 applied to $C_x(L) \cap X$ holds. Note that the cone $C_x(L)$ plays the role of the linear space in the statement of Lemma 5.10. That is, each point of Z lying over a point of $(C_x(L) \cap X) \setminus \{x\}$ is a regular point of Z. This means that each point of Z lying over a point of $\phi^{-1}(\phi(x)) \setminus \{x\}$ is regular. This proves condition (2) of Definition 7.1 for $\phi^{-1}(\phi(x))$.

Condition (3) of Definition 7.1 for the (Z, x)-admissibility of $\phi^{-1}(\phi(x))$ for $x \in \Sigma$ follows from condition (I) of Proposition 6.5, which is part of condition (3) of Proposition 6.6. This proves (5). We have thus proven the claim, and hence, the proposition. □

7B. *Regularity of residual cycles: $k = \bar{k}$ case.* We now prove regularity of residual cycles at points lying over Σ using Proposition 7.2 when k is algebraically closed. Recall (Section 4D) that for a linear projection $\phi_L : X \to \mathbb{P}^r_k$, the residual scheme $L^+(Z)$ is the closure of $\hat{\phi}_L^{-1}(\hat{\phi}_L(Z)) \setminus Z$ in $X \times \hat{B}$ with the reduced induced closed subscheme structure.

We let $T := \hat{\phi}_L(Z) = \hat{\phi}_L(L^+(Z)) \subset \mathbb{P}^r_k \times \hat{B}$ with the reduced subscheme structure and let $\tilde{Z} := T \times_{(\mathbb{P}^r_k \times \hat{B})} (X \times \hat{B}) = \hat{\phi}_L^{-1}(T) = \hat{\phi}_L^{-1}(\hat{\phi}_L(Z))$ as a scheme. We first have:

Lemma 7.3. *We are under the Set-up+*(fs) *of Section 5B. Let $x \in X_{\rm fs}$ be a closed point. Suppose in addition that Z is irreducible.*

Let $\phi_L : X \to \mathbb{A}^r_k$ be a finite surjective morphism obtained by a linear projection as before such that $\phi_L^{-1}(\phi_L(x))$ satisfies condition (3) *of Definition 7.1 of (Z, x)-admissibility. Let $\alpha \in Z$ be a point lying over a point of $\phi_L^{-1}(\phi_L(x))$. Let $S = \hat{\phi}_L^{-1}\hat{\phi}_L(\alpha)$.*

Then $Z \cap S = \{\alpha\}$ and the natural map $\mathcal{O}_{Z,Z\cap S} \to \mathcal{O}_{Z,\alpha}$ is an isomorphism of local rings.

Proof. Suppose $\alpha \in Z$ lies over $x_1 \in \phi_L^{-1}(\phi_L(x))$. Toward contradiction, suppose there is a point $\alpha' \in Z$ lying over some $x_2 \in \phi_L^{-1}(\phi_L(x)) \setminus \{x_1\}$. Since $\hat{\phi}_L(\alpha) = \hat{\phi}_L(\alpha')$, we have $\hat{g}(\alpha) = \hat{g}(\alpha')$ in B, where $\hat{g} : X \times \hat{B} \to \hat{B}$ is the projection. Let b_0 be this common closed point. This $Z \to B$ is nonconstant and we have $\alpha \in Z_{x_1}$ and $\alpha' \in Z_{x_2}$ so that $\hat{g}(Z_{x_1}) \cap \hat{g}(Z_{x_2}) \ni b_0$, contradicting condition (3) of Definition 7.1 for the set $\phi_L^{-1}(\phi_L(x))$. □

Lemma 7.4. *Let k be algebraically closed. Let $L \in \mathcal{U}(k) \subset \mathrm{Gr}(N - r - 1, H)(k)$ be as in Proposition 7.2. Suppose Z is irreducible and let $\alpha = (a, b) \in Z$ be a closed point such that $a \in \phi_L^{-1}(\phi_L(\Sigma))$. Assume that Z is irreducible and $\alpha \in Z$. Then $\mathcal{O}_{\tilde{Z},\alpha} \to \mathcal{O}_{Z,\alpha}$ is an isomorphism. In particular, Z is the only irreducible component of $\tilde{Z}$ which passes through α, with multiplicity* 1, *and the cycle $[\tilde{Z}] - [Z]$ has no component equal to Z.*

Proof. We shall write ϕ_L simply as ϕ. Let $y = \phi(a)$ and $\beta = \hat{\phi}(\alpha) = (\phi(a), b) = (y, b)$. We let $x \in \Sigma$ be such that $y = \phi(x)$ and let $S = \phi^{-1}(y) \times \{b\} = \hat{\phi}^{-1}(\beta) \subset X \times \hat{B}$.

Let $U \subset \mathbb{P}^r_k$ be as in condition (1) of Proposition 7.2. Since $\hat{\phi}$ is finite and étale over $U \times \hat{B}$, it follows that the map $\tilde{Z} \to T$ is finite and étale over $T \cap (U \times \hat{B})$. In particular, the map of rings $\mathcal{O}_{T,\beta} \to \mathcal{O}_{\tilde{Z},S}$ is finite and étale. This in turn implies that the map $\mathcal{O}_{T,\beta} \to \mathcal{O}_{Z,Z\cap S}$ is finite and unramified.

On the other hand, by condition (5) of Proposition 7.2 that $\phi_L^{-1}(\phi_L(x))$ is (Z, x)-admissible, we deduce that for each $x \in \Sigma$, the map $\mathcal{O}_{Z,Z\cap S} \to \mathcal{O}_{Z,\alpha}$ is an isomorphism by Lemma 7.3. Hence, the map $\mathcal{O}_{T,\beta} \to \mathcal{O}_{Z,\alpha}$ is an injective (since $Z \twoheadrightarrow T$), finite and unramified map of local rings which induces isomorphism between the residue fields (as k is algebraically closed). Lemma 4.10 therefore says that the map $\mathcal{O}_{T,\beta} \to \mathcal{O}_{Z,\alpha}$ must be an isomorphism.

We next observe that as $\mathcal{O}_{T,\beta} \to \mathcal{O}_{\tilde{Z},S}$ is finite and étale, the map $\mathcal{O}_{T,\beta} \to \mathcal{O}_{\tilde{Z},\alpha}$ is étale. In particular, the map $\hat{\mathcal{O}}_{T,\beta} \to \hat{\mathcal{O}}_{\tilde{Z},\alpha}$ of completions is finite and étale. Since it induces an isomorphism between the residue fields, it follows again from Lemma 4.10 that $\hat{\mathcal{O}}_{T,\beta} \to \hat{\mathcal{O}}_{\tilde{Z},\alpha}$ is an isomorphism. Hence, there are local homomorphisms of complete local rings

$$\hat{\mathcal{O}}_{T,\beta} \to \hat{\mathcal{O}}_{\tilde{Z},\alpha} \twoheadrightarrow \hat{\mathcal{O}}_{Z,\alpha}, \tag{7-1}$$

where both the first map and the composite map are isomorphisms. Thus, the second map is an isomorphism too. The second map in (7-1) being *a priori* a surjection, the Krull intersection theorem [Matsumura 1986, Theorem 8.10, page 60] says that this map is an isomorphism if and only if $\mathcal{O}_{\tilde{Z},\alpha} \twoheadrightarrow \mathcal{O}_{Z,\alpha}$ (without completion) is an isomorphism. This in turn is equivalent to that Z is the only irreducible component of $\tilde{Z}$ passing through α, and Z has multiplicity 1 in $\tilde{Z}$. We have thus proven the lemma. □

Lemma 7.5. *Let k be algebraically closed and $L \in \mathcal{U}(k) \subset \mathrm{Gr}(N-r-1, H)(k)$ as in Proposition 7.2. Suppose that Z is irreducible. Then $L^+(Z)$ is regular at all points lying over Σ.*

Proof. We continue with the notations of the proof of Lemma 7.4. Let $\alpha = (x, b) \in X \times \hat{B}$ with $x \in \Sigma$ be such that $\alpha \in L^+(Z)$. Let $\beta = \hat{\phi}(\alpha) = (\phi(x), b) := (y, b)$. It follows from Lemma 7.4 that Z does not pass through α. This implies that the canonical map $\mathcal{O}_{\tilde{Z},\alpha} \to \mathcal{O}_{L^+(Z),\alpha}$ is an isomorphism. Therefore, it suffices therefore to show that $\mathcal{O}_{\tilde{Z},\alpha}$ is regular.

Since $\alpha \in L^+(Z)$, there must exist a closed point $\alpha' = (x', b) \in Z$ with $x' \in \phi^{-1}(y)$. As $\alpha \notin Z$, we must have $x' \neq x$. It follows again from Lemma 7.4 that $\mathcal{O}_{\tilde{Z},\alpha'} \xrightarrow{\cong} \mathcal{O}_{Z,\alpha'}$. We have also shown in the middle of the proof of Lemma 7.4 that the map $\hat{\mathcal{O}}_{T,\beta} \to \hat{\mathcal{O}}_{\tilde{Z},(a,b)}$ of completions in (7-1) is an isomorphism for every $a \in \phi^{-1}(y)$. We thus get the commutative diagram of local rings

$$\begin{array}{ccccccc}
\mathcal{O}_{\tilde{Z},\alpha} & \longleftarrow & \mathcal{O}_{T,\beta} & \longrightarrow & \mathcal{O}_{\tilde{Z},\alpha'} & \xrightarrow{\cong} & \mathcal{O}_{Z,\alpha'} \\
\downarrow & & \downarrow & & \downarrow & & \\
\hat{\mathcal{O}}_{\tilde{Z},\alpha} & \xleftarrow{\cong} & \hat{\mathcal{O}}_{T,\beta} & \xrightarrow{\cong} & \hat{\mathcal{O}}_{\tilde{Z},\alpha'}, & &
\end{array} \tag{7-2}$$

where the vertical arrows are completion maps.

Since $x' \neq x$, it follows from condition (2) in Definition 7.1 and condition (5) in Proposition 7.2 that $\mathcal{O}_{Z,\alpha'}$ is regular. It follows from (7-2) that all rings of the bottom of the diagram are regular, using a basic fact in commutative algebra that ($\star$) a noetherian local ring is regular if and only if its completion is a regular local ring [Matsumura 1986, proof of Theorem 19.5, page 157]. Equivalently, all rings of the top of the diagram are regular by ($\star$) again. In particular, $\mathcal{O}_{\tilde{Z},\alpha}$ is regular. This finishes the proof. □

To extend Lemma 7.5 to reducible subschemes Z in Lemma 7.7, we first consider the following:

Lemma 7.6. *Let k be algebraically closed. We are under the Set-up+*(fs) *of Section 5B. Here, Z is not necessarily irreducible. Then after replacing the embedding $X \hookrightarrow \mathbb{P}^N_k$ into a bigger space via a Veronese embedding if necessary, there is a dense open subset $\mathcal{U} \subset \mathrm{Gr}(X, N-r-1, H)$ such that for each $L \in \mathcal{U}(k)$, the induced map $\hat{\phi}_L$ takes distinct components of Z to distinct components of $\hat{\phi}_L(Z)$.*

Proof. As we did previously in Lemma 4.12, for each $1 \le i \le s$, choose a closed point $\alpha_i = (x_i, b_i) \in Z_i \setminus (\cup_{j \neq i} Z_j)$, so that $x_i := \hat{f}(\alpha_i) \in X_{\mathrm{sm}}$ and $b_i := \hat{g}(\alpha_i) \in B$. We observe that if $j \neq i$, then $Z_j \not\subset X \times \{b_i\}$, because $Z_j \to \hat{B}$ is nonconstant.

Let $A_i = \bigcup_{j \neq i} \hat{f}(Z_j \cap (X \times \{b_i\}))$. This is a closed subset of dimension $\le r-1$. In particular, $\dim(\mathrm{Sec}(A_i, \{x_i\})) \le r$. Note that $x_i \notin A_i$. Let

$$\mathcal{U} := \mathrm{Gr}(X, N-r-1, H) \cap \bigcap_{i=1}^{s} \mathrm{Gr}(\mathrm{Sec}(A_i, \{x_i\}), N-r-1, H).$$

This is dense open in $\mathrm{Gr}(N-r-1, H)$ by Lemma 4.3.

Suppose now that $\phi_L : X \to \mathbb{P}^r_k$ is the projection obtained by any $L \in \mathcal{U}(k)$. We fix an integer $1 \le i \le s$ and let $\beta_i := \hat{\phi}_L(\alpha_i)$. It is clear that $\beta_i \in \hat{\phi}_L(Z_i)$. We claim that $\beta_i \notin \hat{\phi}_L(Z_j)$ for $j \neq i$. To see this, note that $\beta_i \in \hat{\phi}_L(Z_j)$ if and only if $Z_j \cap (L^+(x_i) \times \{b_i\}) \neq \varnothing$. Equivalently, there exists a closed point $x'_j \neq x_i$ such that $\phi_L(x'_j) = \phi_L(x_i)$ and $x'_j \in A_i$. But this implies that $L \cap \mathrm{Sec}(A_i, \{x_i\}) \neq \varnothing$, which contradicts the choice of L. This proves the claim and hence the lemma. □

Lemma 7.7. *Let k be algebraically closed. We are under the Set-up+*(fs) *of Section 5B. Here, Z is not necessarily irreducible. Let $\mathcal{U} \subset \mathrm{Gr}(N-r-1, H)$ be the intersection of the dense open subsets of Proposition 7.2 and Lemma 7.6. Then for each $L \in \mathcal{U}(k)$, the residual scheme $L^+(Z)$ is regular at all points lying over Σ.*

Proof. For a choice of L, for simplicity write $\phi := \phi_L$. For $1 \le i \le s$, let $T_i = \hat{\phi}(Z_i)$ with the reduced closed subscheme structure and let $\tilde{Z}_i = T_i \times_{(\mathbb{P}^r_k \times \hat{B})} (X \times \hat{B}) = \hat{\phi}^{-1}(\hat{\phi}(Z_i))$ as a scheme.

The first claim is that $\tilde{Z}_i$ and $\tilde{Z}_j$ share no common component if $i \neq j$. Indeed, if they do share a common component, this would imply that $T_i = T_j$, which contradicts the choice of L as in Lemma 7.6.

Our second claim is that $L^+(Z_i)$ and $L^+(Z_j)$ do not meet at points lying over Σ if $i \neq j$. Suppose on the contrary that there is a closed point $\alpha = (x, b) \in L^+(Z_i) \cap L^+(Z_j)$ with $x \in \Sigma$. This implies that there are closed points $\alpha_i = (x_i, b) \in Z_i$ and $\alpha_j = (x_j, b) \in Z_j$ such that $x_i, x_j \in \phi^{-1}(y)$, where $y = \phi(x)$. It follows from Lemma 7.4 that $x_i, x_j \in \phi^{-1}(y) \setminus \{x\}$.

If $x_i = x_j$, then two components Z_i and Z_j of Z meet at $\alpha_i = \alpha_j$ that lies over $x_i = x_j$ in $\phi^{-1}(y) \setminus \{x\}$. In particular, Z is singular at a point lying over $x_i = x_j$ in $\phi^{-1}(y) \setminus \{x\}$, which contradicts condition (2) of Definition 7.1, which is part of condition (5) of Proposition 7.2. Hence we must have $x_i \neq x_j$. In this case, we get $b \in \hat{g}(Z_{x_i}) \cap \hat{g}(Z_{x_j}) \neq \varnothing$ for two distinct points $x_i, x_j \in \phi^{-1}(y) \setminus \{x\}$. This time, it contradicts condition (3) of Definition 7.1, which is part of condition (5) of Proposition 7.2. Hence, we proved the second claim.

It follows from the two claims that $L^+(Z)$ is regular at all points lying over Σ if and only if $L^+(Z_i)$ is so for every $1 \le i \le s$. Since we proved the latter holds in Lemma 7.5, we finished the proof of the lemma. □

7C. ***Regularity of residual cycles: general case.*** We can now generalize Proposition 7.2 to all infinite perfect field as follows. This includes the regularity of the residual cycle along Σ.

Proposition 7.8. *Let k be any infinite perfect field. We are under the Set-up+*(fs) *of Section 5B. Let $Y \subset X$ be a closed subset of dimension at most $r-1$. Then after replacing the embedding $\eta : X \hookrightarrow \mathbb{P}^N_k$ by a bigger one via a Veronese embedding if necessary, we have the following: for the given hyperplane $H \subset \mathbb{P}^N_k$ disjoint from Σ, there is a dense open subset $\mathcal{U} \subset \mathrm{Gr}(N-r-1, H)$ such that for each $L \in \mathcal{U}(k)$, we have $L \cap X = \varnothing$ so that there is a finite and surjective linear projection $\phi_L : X \to \mathbb{P}^r_k$. Moreover, it satisfies the following properties:*

1. *$\hat{\phi}_L(Z_i) \neq \hat{\phi}_L(Z_j)$ if $i \neq j$.*
2. *The map $\phi_L : \phi_L^{-1}(U) \to U$ is étale for some affine open $U \subset \mathbb{P}^r_k$ containing $\phi_L(\Sigma)$.*
3. *$\phi_L(x) \neq \phi_L(x')$ for each pair of distinct points $x \neq x' \in \Sigma$.*
4. *$k(\phi_L(x)) \xrightarrow{\sim} k(x)$ for each $x \in \Sigma$.*
5. *$L^+(x) \cap Y = \varnothing$ for each $x \in \Sigma$.*
6. *$L^+(Z)$ is regular at all points lying over Σ.*
7. *The map $\hat{\phi}_L : Z \to \hat{\phi}_L(Z)$ is birational.*

Proof. If k is algebraically closed, the proposition follows from Proposition 7.2 and Lemmas 7.6 and 7.7. In general, let $\bar{k}$ be an algebraic closure of k and let $\pi_X : X_{\bar{k}} \to X$ be the projection map from the base change to $\bar{k}$. We have $\Sigma_{\bar{k}} = \bigcup_{x \in \Sigma} \pi_X^{-1}(x)$. Choose a sufficiently large closed embedding $\eta : X \hookrightarrow \mathbb{P}^N_k$ so that for the induced embedding $X_{\bar{k}} \hookrightarrow \mathbb{P}^N_{\bar{k}}$, there exists a dense open subset $\tilde{\mathcal{U}} \subset \mathrm{Gr}(N-r-1, H_{\bar{k}})$ for which all assertions of Proposition 7.2 as well as Lemmas 7.6 and 7.7 applied to $X_{\bar{k}}$, $Z_{\bar{k}}$ and the set $\Sigma_{\bar{k}} \subset X_{\bar{k}}$ hold. (N.B., Under the base change to $\bar{k}$, the irreducible components Z_i of Z may decompose further into irreducible components Z_{ij} of $Z_{i,\bar{k}}$. At least $Z_{\bar{k}}$ and $Z_{i,\bar{k}}$ for all i stay reduced because the extension $\bar{k}$ over k is separable.)

Then we can argue via a Galois descent as in the Step 2 of the proof of Lemma 4.8 to find a dense open $\mathcal{U}_1 \subset \mathrm{Gr}(N-r-1, H)$ defined over k such that $(\mathcal{U}_1)_{\bar{k}} \subset \tilde{\mathcal{U}}$. We take $\mathcal{U} := \mathcal{U}_1 \cap \mathcal{U}_2$, where $\mathcal{U}_2$ is the open set in Lemma 4.8 so that we can also use the assertions of Lemma 4.8 as well.

Now, for each $L \in \mathcal{U}(k)$, we have $X \cap L = \varnothing$ by our choice of the open set. So, we get a finite linear projection map $\phi_L : X \to \mathbb{P}^r_k$ over k. We write this map as ϕ. Condition (1) is clear now by construction together with Lemma 7.6. Conditions (2), (3), (4) hold by conditions (2), (1), (3) of Lemma 4.8, respectively. Condition (5) follows immediately from condition (4) of Proposition 7.2.

To prove (6), as we did at the beginning of Section 7B, let $T := \hat{\phi}(Z) = \hat{\phi}(L^+(Z)) \subset \mathbb{P}^r_k \times \hat{B}$ with the reduced subscheme structure, and let $\tilde{Z} := T \times_{(\mathbb{P}^r_k \times \hat{B})} (X \times \hat{B}) = \hat{\phi}^{-1}\hat{\phi}(Z)$ as a scheme. Then we have the commutative diagram

$$\begin{array}{ccccc} Z_{\bar{k}} & \hookrightarrow & \tilde{Z}_{\bar{k}} & \xrightarrow{\hat{\phi}_{\bar{k}}} & T_{\bar{k}} \\ \downarrow & & \downarrow & & \downarrow \\ Z & \hookrightarrow & \tilde{Z} & \xrightarrow{\hat{\phi}} & T, \end{array} \tag{7-3}$$

where the vertical arrows are the base changes to $\bar{k}$. Note that the map $\hat{\phi} : Z \to T$ is surjective by definition. As both squares are Cartesian and the vertical maps are smooth, it follows that $L^+(Z_{\bar{k}}) \xrightarrow{\cong} L^+(Z)_{\bar{k}}$. By the choice of our open set $\mathcal{U}$, Lemma 7.5 shows that $L^+(Z_{\bar{k}})$ is regular at all points lying over $\Sigma_{\bar{k}}$, i.e., $L^+(Z)_{\bar{k}}$ is regular at all points lying over $\Sigma_{\bar{k}}$.

We replace Z by $Z|_V$ and consider the induced Cartesian squares

$$\begin{array}{ccccc} L^+(Z)_{\bar{k}} & \longrightarrow & V_{\bar{k}} & \longrightarrow & \mathrm{Spec}(\bar{k}) \\ \downarrow & & \downarrow & & \downarrow \\ L^+(Z) & \longrightarrow & V & \longrightarrow & \mathrm{Spec}(k), \end{array} \tag{7-4}$$

where the vertical arrows are the base changes to $\bar{k}$. Since $L^+(Z)_{\bar{k}}$ is regular and $\bar{k}$ is perfect, the top horizontal composite map is smooth. Hence, by the faithfully flat descent [EGA IV$_2$ 1965, Corollaire (17.7.3)(ii), page 72], the bottom horizontal composite map is smooth. In particular, $L^+(Z)$ is regular. This proves (6). Property (7) is a direct consequence of Lemma 4.14. □

Remark 7.9. Condition (4) of Proposition 7.2 or condition (5) of Proposition 7.8 that $L^+(x) \cap Y = \varnothing$ for each $x \in \Sigma$ is no longer needed in this version of the article toward the proof of the main theorems. However, we decided to keep them in this article because the property that the residual points of a projection can be made to avoid the given proper closed subscheme Y is nontrivial, and may be useful in an analysis of algebraic cycles in the future.

8. The main results

In this final section, we use various results of the previous sections to prove our main theorems: the presentation lemma and the sfs-moving lemma. The Set-up for the main results is as in Section 8A. This differs a bit from the Set-up of Section 4D and the Set-up+(fs) of Section 5B.

8A. ***The Set-up(⋆).*** Let k be an infinite perfect field and $n \geq 1$ an integer. We work under the following setting:

(1) *The box coordinates*: For $0 \leq i \leq n-1$, let $\hat{A}_i$ be a smooth projective geometrically integral k-scheme of positive dimension and let $A_i \subset \hat{A}_i$ be a nonempty affine open subset. Let $C_0 = \mathrm{Spec}(k) = \hat{C}_0$. For $1 \leq j \leq n$, we write $C_j = \prod_{i=0}^{j-1} A_i$ and $\hat{C}_j = \prod_{i=0}^{j-1} \hat{A}_i$. Let $\pi_j : \hat{C}_n \to \hat{C}_j$ be the projection map. We write $B = C_n$ and $\hat{B} = \hat{C}_n$. Let $F := \hat{B} \setminus B$.

(2) *The base scheme and the cycles*: Let $X \subset \mathbb{A}^m_k$ be an integral smooth affine closed subscheme of dimension $r \geq 1$ and let $\overline{X} \hookrightarrow \mathbb{P}^m_k$ be its closure with the reduced subscheme structure. Let $\Sigma \subset X$ be a finite set of closed points.

Let $Z \subset X \times B$ be a reduced closed subscheme of pure dimension r, and let $\{Z_1, \dots, Z_s\}$ be all of its irreducible components. Suppose $Z \to X$ is an fs-morphism, i.e., finite and surjective because X is integral. Let $E \subset \hat{B}$ be a closed subset containing F such that no irreducible component of Z is contained $X \times E$.

Let $\hat{Z} \subset \overline{X} \times \hat{B}$ denote the closure of Z in $\overline{X} \times \hat{B}$ with the reduced structure. Similarly, $\hat{Z}_i$ denotes the closure of Z_i in $\overline{X} \times \hat{B}$. We let $\hat{f} : \hat{Z} \to \overline{X}$ and $\hat{g} : \overline{Z} \to \hat{B}$ denote the projection maps.

For each $0 \leq j \leq n$, we define $Z^{(j)} = \pi_j(Z) := (\mathrm{id}_X \times \pi_j)(Z)$. Because $Z \to X$ is fs, this definition makes sense. Similarly we define $\hat{Z}^{(j)}$ for $0 \leq j \leq n$.

(3) *The linear projections*: Suppose we are given a Veronese embedding $\mathbb{P}^m_k \hookrightarrow \mathbb{P}^N_k$ with $N \gg m$. For $L \in \mathrm{Gr}(\overline{X}, N - r - 1, H)(k)$, where $H = \mathbb{P}^N_k \setminus \mathbb{A}^N_k$ as in Lemma 4.7, let $\phi_L : \overline{X} \to \mathbb{P}^r_k$ be the linear projection away from L which restricts to a finite map $\phi_L : X \to \mathbb{A}^r_k$. If L is fixed in a given context, we often drop it from ϕ_L and write ϕ.

For $0 \leq j \leq n$, let $\phi_j = \phi \times \mathrm{id}_{C_j} : X \times C_j \to \mathbb{A}^r_k \times C_j$, $\tilde{\phi}_j = \phi \times \mathrm{id}_{\hat{C}_j} : X \times \hat{C}_j \to \mathbb{A}^r_k \times \hat{C}_j$ and $\hat{\phi}_j = \phi \times \mathrm{id}_{\hat{C}_j} : \overline{X} \times \hat{C}_j \to \mathbb{P}^r_k \times \hat{C}_j$ be the induced maps. We let $L^+(Z)$ denote the closure of $\phi_n^{-1}(\phi_n(Z)) \setminus Z$ in $X \times B$ with the reduced structure. We define $L^+(\hat{Z})$ similarly.

8B. *The residual cycle.* For $L \in \mathrm{Gr}(\overline{X}, N - r - 1, H)(k)$ as in the Set-up ($\star$) of Section 8A, the morphism $\phi = \phi_L : X \to \mathbb{A}^r_k$ is a finite surjective morphism of affine k-schemes so that it is automatically flat by [Hartshorne 1977, Exercise III-10.9, page 276] (or [EGA IV$_2$ 1965, Proposition (6.1.5), page 136]). Hence, for algebraic cycles on $X \times C_j$, we have the proper push-forward ϕ_{j*} and the flat pull-back ϕ_j^* operations. (See [Fulton 1984, Sections 1.4 and 1.7].) For $X \times \hat{C}_j$, we have similar operations $\tilde{\phi}_{j*}$ and $\tilde{\phi}_j^*$.

Definition 8.1. If $Z \subset X \times C_j$ is an integral closed subscheme, the *residual cycle* by $\phi = \phi_L$ is defined to be

$$L^*([Z]) := \phi_j^* \phi_{j*}([Z]) - [Z].$$

We extend it $\mathbb{Z}$-linearly to all cycles on $X \times C_j$. Similarly, for cycles on $X \times \hat{C}_j$, we define the *residual cycle* by

$$L^*([Z]) := \tilde{\phi}_j^* \tilde{\phi}_{j*}([Z]) - [Z].$$

Note that by definition, $L^+([Z]) = |L^*([Z])|$.

Lemma 8.2. *We are under the Set-up*($\star$) *of Section 8A. In particular*, $Z \to X$ *is an* fs-*morphism. Suppose that* Z *is integral. Then for each* L *in the Set-up*($\star$), *the morphism* $L^+(Z) \to X$ *is also fs.*

Proof. Let $T = \phi_n(Z) \subset \mathbb{A}^r_k \times C_n$. Let $\tilde{Z} := T \times_{(\mathbb{A}^r_k \times C_n)} (X \times C_n) = \phi_n^{-1}(T)$ as a scheme. It suffices to show that the map $\tilde{Z} \to X$ is fs. Consider the commutative diagram

$$\begin{array}{ccccc} Z & \overset{\iota}{\hookrightarrow} & \tilde{Z} & \xrightarrow{\phi_n} & T \\ & \searrow & \downarrow{\scriptstyle \hat{f}} & & \downarrow{\scriptstyle \hat{f}'} \\ & & X & \xrightarrow{\phi} & \mathbb{A}^r_k, \end{array} \tag{8-1}$$

where the vertical arrows are the projection maps and the right square is Cartesian, and ι is the closed immersion.

Since $Z \to X$ is an fs-morphism and ϕ is an fs-morphism, the composite $(\phi \circ \hat{f})|_Z = \phi \circ \hat{f} \circ \iota$ is an fs-morphism. By the commutativity, this means $\hat{f}' \circ \phi_n \circ \iota$ is an fs-morphism. But since $Z \to T$ is surjective (as T being the image of Z under ϕ_n by definition), it follows that $\hat{f}'$ is finite (e.g., see [Liu 2002, Proposition 3.16(f), page 104]). Hence $\hat{f}'$ is an fs-morphism. Now, ϕ is flat, so by Lemma 2.7, the morphism $\hat{f}$ is an fs-morphism. □

Lemma 8.3. *We are under the Set-up($\star$) of Section 8A. In particular, $Z \to X$ is an* fs*-morphism. Suppose that Z is integral. Suppose that there is an integer $1 \le j \le n$ such that the projection map $Z^{(j)} \to C_j$ is nonconstant.*

Then for each $L \in \mathrm{Gr}(N - r - 1, H)(k)$ satisfying Lemma 4.14 and Proposition 7.8 for all $Z^{(i)}$ over $j \le i \le n$, we have the equalities

$$[L^+(Z^{(j)})] = L^*([Z^{(j)}]) \quad \textit{and} \quad \pi_{j*}(L^*([Z])) = m_j L^*([Z^{(j)}]),$$

where $m_j = [k(Z) : k(Z^{(j)})]$.

Proof. First of all, note that by Lemma 8.2, every component of $L^+(Z)$ is fs over X. In particular, by the finiteness criterion Lemma 2.9, each irreducible component of $L^+(Z)$ is closed in $X \times \hat{C}_n$. The push-forward $\pi_{j*}([L^*(Z)])$ is given by the projective map $\pi_j : X \times \hat{C}_n \to X \times \hat{C}_j$ is projective.

To prove the first equality, replacing Z by $Z^{(j)}$, we may assume $n = j$ and $Z^{(j)} = Z$. Let $T = \tilde{\phi}_n(Z) \subset \mathbb{A}^r_k \times \hat{C}_n$.

Note that the map $Z \to T$ is birational by Lemma 4.14. Hence, by the definition of the proper push-forward and flat pull-back of cycles, the first equality is equivalent to showing that $\tilde{Z} := T \times_{(\mathbb{A}^r_k \times \hat{C}_n)} (X \times \hat{C}_n) = \tilde{\phi}_n^{-1}(T)$ is a reduced scheme.

To show that $\tilde{Z}$ is reduced, let $U \subset \mathbb{A}^r_k$ be an affine open neighborhood of Σ as in condition (3) of Proposition 7.8. Since $Z \to X$ is finite and surjective, the open subset $T \cap (U \times \hat{C}_n)$ is dense in T. The map $\tilde{\phi}_n$ is étale over this dense open subset of T. Hence, $\tilde{Z} = \tilde{\phi}_n^{-1}(T)$ is reduced over this dense open subset of T. However, $\tilde{\phi}_n^{-1}(T) \to T$ is finite and flat everywhere, it means $\tilde{Z} = \tilde{\phi}_n^{-1}(T)$ is reduced. This proves the first equality.

For the second equality, consider the commutative diagram

$$\begin{array}{ccc} X \times \hat{C}_n & \xrightarrow{\tilde{\phi}_n} & \mathbb{A}^r_k \times \hat{C}_n \\ {\scriptstyle \pi_j}\downarrow & & \downarrow{\scriptstyle \pi_j} \\ X \times \hat{C}_j & \xrightarrow{\tilde{\phi}_j} & \mathbb{A}^r_k \times \hat{C}_j. \end{array} \tag{8-2}$$

This is a Cartesian square in which the vertical arrows are projective and the horizontal arrows are finite and flat. Hence, by [Fulton 1984, Proposition 1.7], we have

$$\begin{aligned}
\pi_{j*}(L^*([Z])) &= \pi_{j*}(\tilde{\phi}_n^* \circ \tilde{\phi}_{n*}([Z]) - [Z]) \\
&= \pi_{j*} \circ \tilde{\phi}_n^* \circ \tilde{\phi}_{n*}([Z]) - \pi_{j*}([Z]) \\
&= \tilde{\phi}_j^* \circ \pi_{j*} \circ \tilde{\phi}_{n*}([Z]) - \pi_{j*}([Z]) \\
&= \tilde{\phi}_j^* \circ \tilde{\phi}_{j*} \circ \pi_{j*}([Z]) - \pi_{j*}([Z]) \\
&= \tilde{\phi}_j^* \circ \tilde{\phi}_{j*}(m_j[Z^{(j)}]) - m_j[Z^{(j)}] \\
&= m_j(\tilde{\phi}_j^* \circ \tilde{\phi}_{j*}([Z^{(j)}]) - [Z^{(j)}]) \\
&= m_j L^*([Z^{(j)}]),
\end{aligned}$$

which proves the second equality. ☐

The following complements Lemma 8.3:

Lemma 8.4. *We are under the Set-up(⋆) of Section 8A. In particular, $Z \to X$ is an* fs-*morphism. Suppose that Z is integral such that $Z \to C_n$ is nonconstant. Suppose that for an integer $0 \le j \le n-1$, the projection $Z^{(j)} \to C_j$ is constant. Then for each $L \in \mathrm{Gr}(N-r-1, H)(k)$ satisfying the conditions of Proposition 7.8 for Z, we have the equality $\pi_j(Z') = \pi_j(Z)$ for each irreducible component Z' of $L^+(Z)$.*

Proof. Toward contradiction, suppose that there is an irreducible component Z' of $L^+(Z)$ such that $\pi_j(Z') \neq \pi_j(Z)$. In particular, this implies that $L^+(Z^{(j)}) \neq \varnothing$. On the other hand, we are given that $Z^{(j)} = \pi_j(Z) = X \times \{c_j\}$ for some closed point $c_j \in \hat{C}_j$. In this case, $\tilde{\phi}_j(Z^{(j)}) = \mathbb{A}^r_k \times \{c_j\}$ so that $\tilde{\phi}_j^{-1}\tilde{\phi}_j(Z^{(j)}) = X \times \{c_j\} = Z^{(j)}$. Hence, $L^+(Z^{(j)}) = \varnothing$. This is a contradiction. ☐

8C. ***The presentation lemma.*** We now prove the presentation lemma for residual cycles under linear projections. We are under the Set-up(⋆) in Section 8A.

Theorem 8.5. *Let k be an infinite perfect field. Let $Z \subset X \times C_n$ be an integral closed subscheme such that $Z \to X$ is finite surjective, and the projection $Z \to C_n$ is nonconstant.*

Then there exist an embedding $\eta : \overline{X} \hookrightarrow \mathbb{P}^N_k$ and a dense open subset $\mathcal{U} \subset \mathrm{Gr}(N-r-1, H)$, where $H := \mathbb{P}^N_k \setminus \mathbb{A}^N_k$, such that for each $L \in \mathcal{U}(k)$, the linear projection $\phi_L : \mathbb{P}^N_k \setminus L \to \mathbb{P}^r_k$ away from L defines a finite surjective morphism $\phi : \overline{X} \to \mathbb{P}^r_k$ satisfying the following properties:

(1) *There exists a Cartesian square*

$$\begin{array}{ccc} X & \hookrightarrow & \overline{X} \\ {\scriptstyle \phi}\downarrow & & \downarrow{\scriptstyle \phi} \\ \mathbb{A}^r_k & \hookrightarrow & \mathbb{P}^r_k. \end{array}$$

(2) ϕ *is étale over an affine open neighborhood of* $\phi(\Sigma)$.

(3) $\phi(x) \neq \phi(x')$ *for every pair* $x \neq x'$ *in* Σ.

(4) *The map* $k(\phi(x)) \to k(x)$ *is an isomorphism for each* $x \in \Sigma$.

(5) *The induced map* $Z \to \phi_n(Z)$ *is birational.*

(6) *The map* $L^+(Z) \to X$ *is finite surjective.*

(7) *For each* $0 \le j \le n$, *the scheme* $\pi_j(L^+(Z))$ *is regular at all points lying over* Σ.

Proof. Since $Z \to X$ is finite surjective, for each $0 \le j \le n$ the morphism $Z^{(j)} \to X$ is also finite surjective. Let $i_0 \in \{0, \dots, n\}$ be the largest integer i such that $Z^{(i)} \subset X \times \{b\}$ for some closed point $b \in C_i$. Note that $Z^{(0)} = X = X \times_k C_0$, so such i_0 exists. Note also that $i_0 \le n-1$ by our assumption.

Choose a large enough Veronese embedding $\mathbb{P}^m_k \hookrightarrow \mathbb{P}^N_k$ such that for the composite embedding $\eta : \overline{X} \hookrightarrow \mathbb{P}^N_k$, and the hyperplane $H = H_{N,0}$ as in Lemma 4.7, there are open dense subsets $\mathcal{U}_j \subset \mathrm{Gr}(N-r-1, H)$ such that each $L \in \mathcal{U}_j(k)$ satisfies Lemma 4.14 and conditions (1)–(6) of Proposition 7.8 for $Z^{(j)}$ over all $i_0+1 \le j \le n$. We let $\mathcal{U} = \bigcap_{j=i_0+1}^{n} \mathcal{U}_j$.

Condition (1) of the theorem automatically follows from our choice of H and Lemma 4.7. Conditions (2)–(4) follow directly from conditions (2)–(4) of Proposition 7.8. Condition (5) follows from Lemma 4.14. Condition (6) follows from Lemma 8.2.

We prove (7). We have to show that every irreducible component of $\pi_j(L^+(Z))$ is regular at all points lying over Σ and no two components of $\pi_j(L^+(Z))$ meet at points lying over Σ. We first assume that $i_0+1 \le j \le n$.

Let Z' be an irreducible component of $L^+(Z)$. Since $j > i_0$, Lemma 8.3 says that $\pi_j(Z')$ is a component of the effective cycle $\pi_{j*}(L^*([Z])) = m_j L^*([Z^{(j)}]) = m_j[L^+(Z^{(j)})]$ with $m_j \ge 1$. Since Z' was arbitrary, it follows that the irreducible components of $\pi_j(L^+(Z))$ are the same as those of $L^+(Z^{(j)})$. On the other hand, condition (6) of Proposition 7.8 (with our choice of L) says that $L^+(Z^{(j)})$ is regular at all points lying over Σ. It follows that each irreducible component of $\pi_j(L^+(Z))$ is regular at all points lying over Σ, and in particular no two components meet at points lying over Σ.

If $0 \le j \le i_0$, then Lemma 8.4 says that $\pi_j(L^+(Z))$ coincides with $\pi_j(Z)$, which in turn is of the form $X \times \{b\}$ for some closed point $b \in C_j$. In particular, $\pi_j(L^+(Z))$ is irreducible. As X is regular everywhere, in particular at all points lying over Σ, it follows that $\pi_j(L^+(Z))$ is regular at all points lying over Σ. This completes the proof of the theorem. □

8D. ***The* sfs*-moving lemma.*** We now prove the sfs-moving lemma for additive higher Chow groups of relative 0-cycles over semilocal k-schemes. A similar argument also proves the sfs-moving lemma for Bloch's higher Chow groups of relative 0-cycles over semilocal k-schemes.

Let k be an infinite perfect field. We apply Theorem 8.5 with $\hat{A}_i := \mathbb{P}^1_k$ for $0 \le i \le n-1$, while $A_0 := \mathbb{A}^1_k$ and $A_1 = \cdots = A_{n-1} = \square^1_k$ so that $C_j = B_j = \mathbb{A}^1_k \times \square^{j-1}_k$ for $j \ge 1$. The sfs-moving lemma for additive higher Chow groups of relative 0-cycles is the following:

Theorem 8.6. *Let R be a regular semilocal k-scheme essentially of finite type of dimension $r \ge 0$ over an infinite perfect field k. Let $V = \operatorname{Spec}(R)$ and let $m, n \ge 1$ be integers. Then the canonical map $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n(V, n; m)$ is an isomorphism.*

This theorem is proven in steps. Since R is regular, it is a product of regular semilocal k-domains, and each k-domain corresponds to a connected component of $\operatorname{Spec}(R)$. Thus we may reduce to the case when R is integral. We also remark that by Proposition 2.19, we may assume that R is obtained by localizing an integral smooth affine k-scheme at a finite set of closed points. Note that Theorem 8.6 is obvious for $r = 0$, so we may assume $r \ge 1$. We have injective maps (using Lemma 2.18),

$$\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n_{\mathrm{fs}}(V, n; m) \to \mathrm{TCH}^n_{\Sigma}(V, n; m) \to \mathrm{TCH}^n(V, n; m).$$

The last arrow is an isomorphism by [Krishna and Park 2016, Theorem 4.10]. We show that the middle arrow is an isomorphism, which we call the fs-*moving lemma*:

Lemma 8.7. *The map $\mathrm{TCH}^n_{\mathrm{fs}}(V, n; m) \to \mathrm{TCH}^n_{\Sigma}(V, n; m)$ is an isomorphism.*

Proof. By the above discussion, we assume V is integral. Since this map is injective, we only have to show that it is surjective. Let $\gamma \in \mathrm{Tz}^n_{\Sigma}(V, n; m)$ be a cycle with $\partial(\gamma) = 0$.

First suppose that there is an atlas $(\mathbb{A}^r_k, \Sigma)$ so that γ lifts to a cycle $\bar{\gamma} \in \mathrm{Tz}^n_{\Sigma}(\mathbb{A}^r_k, n; m)$. In this case, we can apply Theorem 3.14 and write $\gamma = \gamma_1 + \partial(\gamma_2)$, where $\gamma_1 \in \mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m) \subset \mathrm{Tz}^n_{\mathrm{fs}}(V, n; m)$ and $\gamma_2 \in \mathrm{Tz}^n(V, n+1; m)$. One immediately has $\partial(\gamma_1) = 0$, proving the desired surjectivity in this case.

In general, we write $\gamma = \alpha + \beta$, where no component of α is an fs-cycle and β is an fs-cycle. Lemma 2.5 says that there is a connected smooth affine atlas (X, Σ) for V, and cycles $\bar{\alpha}, \bar{\beta}, \bar{\gamma} \in \mathrm{Tz}^n_{\Sigma}(X, n; m)$ such that $\bar{\alpha}_V = \alpha$, $\bar{\beta}_V = \beta$, $\bar{\gamma}_V = \gamma$, $\bar{\gamma} = \bar{\alpha} + \bar{\beta}$ and $\partial(\bar{\gamma}) = 0$.

Since no component of α is fs over V, it follows that the projection of every component of $\bar{\alpha}$ to B_n must be nonconstant. We can therefore apply Theorem 4.15 to obtain a finite flat map $\phi : X \to \mathbb{A}^r_k$ such that α satisfies all the properties there. Let $\Sigma' = \phi(\Sigma)$, which consists of finitely many closed points of $\mathbb{A}^r_k$. Let $V' = \operatorname{Spec}(\mathcal{O}_{\mathbb{A}^r_k, \Sigma'})$ and $W := X \times_{\mathbb{A}^r} V'$. We have inclusions $\Sigma \subset V \subset W \subset X$, and a finite flat morphism $\phi : W \to V'$.

Write $\bar{\alpha} = \bar{\alpha}_1 + \bar{\alpha}_2$, where each component of $\bar{\alpha}_1$ is dominant over X and no component of $\bar{\alpha}_2$ is dominant over X. As β is an fs-cycle over V, after shrinking X if needed, $\bar{\beta}$ is an fs-cycle over X along Σ by Corollary 2.12.

We now have

$$\bar{\gamma} = \bar{\alpha}_1 + \bar{\alpha}_2 + \bar{\beta} = (\bar{\alpha}_1 - \phi_n^* \phi_{n*}(\bar{\alpha}_1)) + (\bar{\alpha}_2 - \phi_n^* \phi_{n*}(\bar{\alpha}_2)) + (\bar{\beta} - \phi_n^* \phi_{n*}(\bar{\beta})) + \phi_n^* \phi_{n*}(\bar{\gamma}).$$

Let $\bar{\alpha}'_i := \bar{\alpha}_i - \phi_n^* \phi_{n*}(\bar{\alpha}_i)$ for $i = 1, 2$, and $\bar{\beta}' := \bar{\beta} - \phi_n^* \phi_{n*}(\bar{\beta})$. Since $\bar{\beta}$ is an fs-cycle on X along Σ and ϕ is finite, $\phi_{n*}(\bar{\beta})$ is an fs-cycle over $\mathbb{A}_k^r$ by Lemma 2.8. Since $X \to \mathbb{A}_k^r$ is flat, by Lemma 2.7 $\phi_n^* \phi_{n*}(\bar{\beta})$ is an fs-cycle over X along Σ. In particular, $\bar{\beta}' \in \mathrm{Tz}^n_{\Sigma,\mathrm{fs}}(X, n; m)$. On the other hand, by Theorem 4.15, we have $(\bar{\alpha}'_2)_V = 0$ and $(\bar{\alpha}'_1)_V \in \mathrm{Tz}^n_{\Sigma,\mathrm{fs}}(V, n; m)$.

Since $\bar{\gamma} \in \mathrm{Tz}^n_\Sigma(X, n; m)$ with $\partial(\bar{\gamma}) = 0$, it follows that $\phi_*(\bar{\gamma}) \in \mathrm{Tz}^n_{\Sigma'}(\mathbb{A}_k^r, n; m)$ with $\partial(\phi_{n*}(\bar{\gamma})) = 0$. By the previous case, there are cycles $\eta_1 \in \mathrm{Tz}^n_{\mathrm{fs}}(V', n; m)$, and $\eta_2 \in \mathrm{Tz}^n(V', n+1; m)$ such that $j^*(\phi_{n*}(\bar{\gamma})) = \eta_1 + \partial\eta_2$. Equivalently, $\phi_{n*}(\bar{\gamma}_W) = \eta_1 + \partial\eta_2$. Hence, $\phi_n^* \phi_{n*}(\bar{\gamma}_W) = \phi^*(\eta_1) + \phi_n^*(\partial\eta_2) = \phi_n^*(\eta_1) + \partial(\phi_n^*(\eta_2))$. Moreover, $\phi_n^*(\eta_1)$ is an fs-cycle by Lemma 2.7. Combining these, we have

$$\gamma = (\bar{\gamma})_V = (\bar{\alpha}'_1)_V + \bar{\beta}'_V + (\phi_n^*(\eta_1))_V + \partial((\phi_n^*(\eta_2))_V) = \gamma_1 + \partial((\phi_n^*(\eta_2))_V),$$

where $\gamma_1 := (\bar{\alpha}'_1)_V + \bar{\beta}'_V + (\phi_n^*(\eta_1))_V \in \mathrm{Tz}^n_{\mathrm{fs}}(V, n; m)$. Since $\partial\gamma = 0$, we also deduce that $\partial\gamma_1 = 0$. This completes the proof of the lemma. □

Proof of Theorem 8.6. We may assume that V is integral. Using Lemma 8.7, it suffices to show that the map $\mathrm{TCH}^n_{\mathrm{sfs}}(V, n; m) \to \mathrm{TCH}^n_{\mathrm{fs}}(V, n; m)$ is surjective. Let $\alpha \in \mathrm{Tz}^n_{\mathrm{fs}}(V, n; m)$ be an fs-cycle, which always has $\partial(\alpha) = 0$ by Lemma 2.21. Write $\alpha = \alpha_1 + \alpha_2$, where $\alpha_2 \in \mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$, while $\alpha_1 \in \mathrm{Tz}^n_{\mathrm{fs}}(V, n; m)$ but no component of α_1 lies in $\mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$. Note that $\partial(\alpha_i) = 0$ for $i = 1, 2$ by Lemma 2.21 again. It is enough to prove that α_1 is equivalent to a cycle in $\mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$. Replacing α by α_1, we may therefore assume that no component of α lies in $\mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$.

Apply Lemma 2.5 to choose a connected smooth affine atlas (X, Σ) for V and a cycle $\bar{\alpha} \in \mathrm{Tz}^n_\Sigma(X, n; m)$ such that $\partial(\bar{\alpha}) = 0$. If $X \simeq \mathbb{A}_k^r$, we can apply Theorem 3.14 to write $\alpha = \beta + \partial(\gamma)$, where $\beta \in \mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m) \subset \mathrm{Tz}^n_{\mathrm{fs}}(V, n; m)$ and $\gamma \in \mathrm{Tz}^n(V, n+1; m)$. This solves the problem in this case.

Suppose that X is not an affine space. If Z is a component of α whose projection to B_n is constant, then Z is already an sfs-cycle. But, we supposed no component of α is an sfs-cycle. Hence, $Z \to B_n$ is nonconstant for each irreducible component Z. It follows that Lemma 8.3 and Theorem 8.5 apply to every component of α. Let $\phi : X \to \mathbb{A}_k^r$ be the finite and flat map as in Theorem 8.5 and let $\Sigma' = \phi(\Sigma)$. By shrinking $\mathcal{U} \subset \mathrm{Gr}(N - r - 1, H)$ if necessary, we can assume that conditions (1)–(7) of Theorem 8.5 hold for each $L \in \mathcal{U}(k)$ and for each component of α.

Let $V' = \mathrm{Spec}(\mathcal{O}_{\mathbb{A}_k^r, \Sigma'})$ and let $W = X \times_{\mathbb{A}_k^r} V'$. We have inclusions $\Sigma \subset V \subset W \subset X$ and a finite and flat morphism $\phi_\Sigma : W \to V'$ of smooth semilocal k-schemes. Let $j : V \to W$ be the localization map.

We can write $\bar{\alpha}_W = (\bar{\alpha}_W - \phi_n^* \phi_{n*}(\bar{\alpha}_W)) + \phi_n^* \phi_{n*}(\bar{\alpha}_W)$. We have $\partial(\phi_{n*}(\bar{\alpha}_W)) = \phi_{n*}(\partial(\bar{\alpha}_W)) = 0$. By the previous case of affine space atlas, we can write $\phi_{n*}(\bar{\alpha}_W) = \eta_1 + \partial(\eta_2)$, where $\eta_1 \in \mathrm{Tz}^n_{\mathrm{sfs}}(V', n; m)$ and $\eta_2 \in \mathrm{Tz}^n(V', n+1; m)$. This yields $\phi_n^* \phi_{n*}(\bar{\alpha}_W) = \phi_n^*(\eta_1) + \partial(\phi_n^*(\eta_2))$. Since $\phi : W \to V'$ is finite and étale, it follows by Lemmas 2.7 and 2.16 that $\phi_n^*(\eta_1) \in \mathrm{Tz}^n_{\mathrm{sfs}}(W, n; m)$.

It follows from Lemma 8.3 and Theorem 8.5 that $j^*(\bar{\alpha}_W - \phi_n^* \phi_{n*}(\bar{\alpha}_W)) \in \mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$. Let $\beta = j^*(\bar{\alpha}_W - \phi_n^* \phi_{n*}(\bar{\alpha}_W)) + j^*(\phi_n^*(\eta_1)) \in \mathrm{Tz}^n_{\mathrm{sfs}}(V, n; m)$ and $\gamma = j^*(\phi_n^*(\eta_2)) \in \mathrm{Tz}^n(V, n+1; m)$. Then, we get

$$\begin{aligned}\alpha = j^*(\bar{\alpha}_W) &= j^*(\bar{\alpha}_W - \phi_n^*\phi_{n*}(\bar{\alpha}_W)) + j^*\phi_n^*(\eta_1) + j^*(\partial(\phi_n^*(\eta_2)))\\ &= j^*(\bar{\alpha}_W - \phi_n^*\phi_{n*}(\bar{\alpha}_W)) + j^*\phi_n^*(\eta_1) + \partial(j^*\phi_n^*(\eta_2))\\ &= \beta + \partial(\gamma).\end{aligned}$$

Since $\partial(\alpha) = 0$, we must have $\partial(\beta) = 0$ as well. This proves the theorem. □

Proof of Theorem 1.2. We take $n \geq 2$, $A_0 = \hat{A}_0 = \mathrm{Spec}(k)$, $A_i = \square_k$ and $\hat{A}_i := \mathbb{P}_k^1$ for $1 \leq i \leq n-1$ in Theorem 8.5. We now repeat the proof of Theorem 8.6 verbatim using Remark 3.15. □

Acknowledgments

We thank Spencer Bloch, Hélène Esnault, Najmuddin Fakhruddin, Bruno Kahn, Marc Levine, and Kay Rülling for some conversations related to the question studied here throughout the time we were working on it. We thank the handling editors and, the anonymous referee of Algebra & Number Theory, who read the article with thoroughness and patience, and supplied numerous suggestions, which helped the authors in improving the article.

We acknowledge that part of this work was done during Park's visits to TIFR and Krishna's visits to KAIST, and we thank both the institutions. Park would like to thank Juya and Damy for his peace of mind at home through the time this paper was written. Park was partially supported by the National Research Foundation of Korea (NRF) grant (2015R1A2A2A01004120 and 2018R1A2B6002287) funded by the Korean government (MSIP), and TJ Park Junior Faculty Fellowship funded by POSCO TJ Park Foundation.

References

[Arbarello et al. 1985] E. Arbarello, M. Cornalba, P. A. Griffiths, and J. Harris, *Geometry of algebraic curves, I*, Grundlehren der Math. Wissenschaften **267**, Springer, 1985. MR Zbl

[Ballico 1989] E. Ballico, "On singular curves in the case of positive characteristic", *Math. Nachr.* **141** (1989), 267–273. MR Zbl

[Bloch 1986] S. Bloch, "Algebraic cycles and higher K-theory", *Adv. Math.* **61**:3 (1986), 267–304. MR Zbl

[Bloch 1994] S. Bloch, "The moving lemma for higher Chow groups", *J. Algebraic Geom.* **3**:3 (1994), 537–568. MR Zbl

[Colliot-Thélène et al. 1997] J.-L. Colliot-Thélène, R. T. Hoobler, and B. Kahn, "The Bloch–Ogus–Gabber theorem", pp. 31–94 in *Algebraic K-theory* (Toronto, 1996), edited by V. P. Snaith, Fields Inst. Commun. **16**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[EGA IV$_2$ 1965] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR Zbl

[EGA IV$_3$ 1966] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, III", *Inst. Hautes Études Sci. Publ. Math.* **28** (1966), 5–255. MR Zbl

[EGA IV$_4$ 1967] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, IV", *Inst. Hautes Études Sci. Publ. Math.* **32** (1967), 5–361. MR Zbl

[Elbaz-Vincent and Müller-Stach 2002] P. Elbaz-Vincent and S. Müller-Stach, "Milnor K-theory of rings, higher Chow groups and applications", *Invent. Math.* **148**:1 (2002), 177–206. MR Zbl

[Fulton 1984] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik (3) **2**, Springer, 1984. MR Zbl

[Gabber et al. 2013] O. Gabber, Q. Liu, and D. Lorenzini, "The index of an algebraic variety", *Invent. Math.* **192**:3 (2013), 567–626. MR Zbl

[Gupta and Krishna 2019a] R. Gupta and A. Krishna, "Relative K-theory via 0-cycles in finite characteristic", preprint, 2019. arXiv

[Gupta and Krishna 2019b] R. Gupta and A. Krishna, "Zero-cycles with modulus and relative K-theory", preprint, 2019. arXiv

[Harris 1982] J. Harris, *Curves in projective space*, Sémin. Math. Sup. **85**, Univ. Montreal Press, 1982. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Math. **52**, Springer, 1977. MR Zbl

[Jouanolou 1983] J.-P. Jouanolou, *Théorèmes de Bertini et applications*, Prog. Math. **42**, Birkhäuser, Boston, 1983. MR Zbl

[Kai 2019] W. Kai, "A moving lemma for algebraic cycles with modulus and contravariance", *Int. Math. Res. Not.* (online publication May 2019).

[Kleiman and Altman 1979] S. L. Kleiman and A. B. Altman, "Bertini theorems for hypersurface sections containing a subscheme", *Comm. Algebra* **7**:8 (1979), 775–790. MR Zbl

[Krishna and Levine 2008] A. Krishna and M. Levine, "Additive higher Chow groups of schemes", *J. Reine Angew. Math.* **619** (2008), 75–140. MR Zbl

[Krishna and Park 2012] A. Krishna and J. Park, "Moving lemma for additive higher Chow groups", *Algebra Number Theory* **6**:2 (2012), 293–326. MR Zbl

[Krishna and Park 2015a] A. Krishna and J. Park, "de Rham-Witt sheaves via algebraic cycles", preprint, 2015. arXiv

[Krishna and Park 2015b] A. Krishna and J. Park, "DGA-structure on additive higher Chow groups", *Int. Math. Res. Not.* **2015**:1 (2015), 1–54. MR Zbl

[Krishna and Park 2016] A. Krishna and J. Park, "On additive higher Chow groups of affine schemes", *Doc. Math.* **21** (2016), 49–89. MR Zbl

[Krishna and Park 2017] A. Krishna and J. Park, "A moving lemma for cycles with very ample modulus", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **17**:4 (2017), 1521–1549. MR Zbl

[Liu 2002] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Grad. Texts Math. **6**, Oxford Univ. Press, 2002. MR Zbl

[Matsumura 1986] H. Matsumura, *Commutative ring theory*, Cambridge Stud. Adv. Math. **8**, Cambridge Univ. Press, 1986. MR Zbl

[Park 2009] J. Park, "Regulators on additive higher Chow groups", *Amer. J. Math.* **131**:1 (2009), 257–276. MR Zbl

[Rülling 2007] K. Rülling, "The generalized de Rham–Witt complex over a field is a complex of zero-cycles", *J. Algebraic Geom.* **16**:1 (2007), 109–169. MR Zbl

[Samuel 1966] P. Samuel, *Lectures on old and new results on algebraic curves*, Tata Inst. Fund. Res. Lect. Math. **36**, Tata Inst. Fund. Res., Bombay, 1966. MR Zbl

Communicated by Hélène Esnault
Received 2019-04-13 Revised 2019-11-18 Accepted 2019-12-16

amal@math.tifr.res.in *School of Mathematics, Tata Institute of Fundamental Research, Mumbai, India*

jinhyun@kaist.edu *Department of Mathematical Sciences, KAIST, Daejeon, South Korea*

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 14 No. 4 2020

1937-0652(2020)14:4;1-Y