Algebra & Number Theory

Volume 19
2025
No. 10

msp

Algebra & Number Theory

msp.org/ant

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

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

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 19:10 (2025)
https://doi.org/10.2140/ant.2025.19.1881

Smooth numbers are orthogonal to nilsequences

Lilian Matthiesen and Mengdi Wang

The aim of this paper is to study distributional properties of integers without large or small prime factors. Define an integer to be $[y', y]$-smooth if all of its prime factors belong to the interval $[y', y]$. We identify suitable weights $g_{[y',y]}(n)$ for the characteristic function of $[y', y]$-smooth numbers that allow us to establish strong asymptotic results on their distribution in short arithmetic progressions. Building on these equidistribution properties, we show that (a W-tricked version of) the function $g_{[y',y]}(n) - 1$ is orthogonal to nilsequences. Our results apply in the almost optimal range $(\log N)^K < y \leqslant N$ of the smoothness parameter y, where $K \geqslant 2$ is sufficiently large, and to any $y' < \min(\sqrt{y}, (\log N)^c)$.

As a first application, we establish for any $y > N^{1/\sqrt{\log_9 N}}$ asymptotic results on the frequency with which an arbitrary finite complexity system of shifted linear forms $\psi_j(\boldsymbol{n}) + a_j \in \mathbb{Z}[n_1, \dots, n_s]$, $1 \leqslant j \leqslant r$, simultaneously takes $[y', y]$-smooth values as the n_i vary over integers below N.

1. Introduction

Let $y > 0$ be a real number. A positive integer n is called y-smooth if its largest prime factor is at most y. The y-smooth numbers below N form a subset of the integers below N which is, in general, sparse but enjoys good equidistribution properties in arithmetic progressions and short intervals. These distributional properties turn y-smooth numbers into an important technical tool for many arithmetic questions. As an example for one of the striking applications of smooth numbers within analytic number theory, we mention [Vaughan 1989], which introduced smooth numbers in combination with a new iterative method to the

MSC2020: primary 11N37; secondary 11B30, 11D04, 11L15, 11N25.
Keywords: smooth numbers, multiplicative functions, nilsequences, Gowers uniformity norms, higher-order Fourier analysis.

study of bounds in Waring's problem. Wooley [1992] extended these methods and achieved substantial improvements on Waring's problem by working with smooth numbers. We refer to Granville's survey [2008] (see in particular Section 6) for a more comprehensive overview of applications of smooth numbers within number theory.

In the present paper, we prove new results on the equidistribution of smooth numbers in short intervals and arithmetic progressions. Our principal aim is to prove higher uniformity of y-smooth numbers in a sense that will be made precise below and for y ranging over an almost optimal range. In addition, we prove that provided y is not too small, the set of y-smooth numbers is sufficiently well distributed to guarantee the existence of nontrivial solutions to arbitrary finite complexity systems of linear equations.

In order to state our main results precisely, we first introduce a subset of the y-smooth numbers as well as a weighted version of its characteristic function that are both central to the rest of this paper. Given any real numbers $0 < y' \leqslant y$, we may consider the set of y-smooth numbers that are free from prime factors smaller than y'. We call such numbers $[y', y]$-smooth and denote their set by

$$S([y', y]) := \{n \in \mathbb{N} : p \mid n \Rightarrow p \in [y', y]\}.$$

Given any $x > 0$, the subset of $[y', y]$-smooth numbers $\leqslant x$ and its cardinality are denoted by

$$S(x, [y', y]) := S([y', y]) \cap [1, x] \quad \text{and} \quad \Psi(x, [y', y]) := |S(x, [y', y])|.$$

Our notation extends the following standard notation for y-smooth numbers:

$$S(y) := S([1, y]), \quad S(x, y) := S(x, [1, y]), \quad \text{and} \quad \Psi(x, y) = \Psi(x, [1, y]).$$

With this notation, we define the weighted characteristic function

$$g_{[y',y]}(n) = \frac{n}{\alpha(n, y)\Psi(n, [y', y])} \mathbf{1}_{S([y',y])}(n) \quad (n \in \mathbb{N})$$

of $[y', y]$-smooth numbers, where $\alpha(n, y)$ denotes the saddle point[1] associated to $S(n, y)$. If $1 \leqslant A < W$ are coprime integers, we further define a W*-tricked* version of $g_{[y',y]}$ by setting

$$g^{(W,A)}_{[y',y]}(m) = \frac{\phi(W)}{W} g_{[y',y]}(Wm + A) \quad (m \in \mathbb{N}).$$

In the definitions of the functions $g_{[y',y]}$ and $g^{(W,A)}_{[y',y]}$, the normalisations are chosen such that their average values are roughly 1.

Following these preparations, we are now ready to state the main result of this paper. The notation around nilsequences will be recalled in Section 8. Throughout this paper we write $\log_k x$ to denote the k-fold iterated logarithm of x.

Theorem 1.1 (higher uniformity). *Let N be a large positive parameter and let $K' \geqslant 1$, $K > 2K'$ and $d \geqslant 0$ be integers. Let $\frac{1}{2}\log_3 N \leqslant y' \leqslant (\log N)^{K'}$ and suppose that $(\log N)^K < y < N^\eta$ for some sufficiently*

[1] We recall the definition in Section 3.

small $\eta \in (0,1)$ *depending on the value of* d. *Let* $(G/\Gamma, G_\bullet)$ *be a filtered nilmanifold of complexity* Q_0 *and degree* d. *Finally, let* $w(N) = \frac{1}{2}\log_3 N$, $W = \prod_{p<w(N)} p$ *and define* $\delta(N) = \exp(-\sqrt{\log_4 N})$.

If K *is sufficiently large depending on the degree* d *of* $G_\bullet$, *then the estimate*

$$\left|\frac{W}{N}\sum_{n\leqslant (N-A)/W}(g^{(W,A)}_{[y',y]}(n)-1)F(g(n)\Gamma)\right| \ll_d (1+\|F\|_{\mathrm{Lip}})\delta(N)Q_0 + \frac{1}{\log w(N)} \tag{1-1}$$

holds uniformly for all $1 \leqslant A \leqslant W$ *with* $\gcd(A, W) = 1$, *all polynomial sequences* $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ *and all* 1-*bounded Lipschitz functions* $F : G/\Gamma \to \mathbb{C}$.

Remark 1.2. Since $\gcd(A, W) = 1$, we may, when working with $Wn + A \in S([y', y])$, restrict without loss of generality to the case $y' \geqslant w(N)$. In applications where y is not too small, the contribution from $S(y) \setminus S([w(N), y])$ can often be taken care of separately, leading to a result for $S(y)$ in the end.

Theorem 1.1 constitutes the first of two parts necessary in order to establish asymptotic results on the number of $[y', y]$-smooth solutions to finite complexity systems of linear forms using the nilpotent circle method. We emphasise that the setting of smooth numbers studied in the present paper is significantly more difficult than those settings considered in previous applications such as, e.g., [Green and Tao 2010], [Matthiesen 2018; 2020] or [Matthiesen 2012], which concern primes, a large class of multiplicative functions, and numbers representable by binary quadratic forms, respectively. The reason for this increase in difficulty lies partly in the sparsity of the set of y-smooth numbers and partly in the unavailability of sieve methods to study this set.

The parameter $w(N)$ is determined by the distribution of y-smooth numbers in arithmetic progressions and chosen in such a way that the subset of y-smooth numbers in any fixed reduced residue class modulo $W(N) = \prod_{p<w(N)} p$ is sufficiently well-distributed in arithmetic progressions. At the end of Section 2, we will discuss in more detail the need for applying such a "W-trick" when studying y-smooth numbers. In addition to the W-trick, the value y' may be used to further influence how well the resulting set of $[y', y]$-smooth numbers is distributed in progressions. Increasing the parameter y' beyond the value of $w(N)$ leads to better error terms for larger moduli in the distribution of $[y', y]$-smooth numbers in progressions. The proof of Theorem 1.1 relies on information on the distribution of $[y', y]$-smooth numbers in short intervals and in arithmetic progressions, which we establish in Sections 4 and 5.

From a technical perspective, our focus in Theorem 1.1 has been to establish a result in which the lower bound on the range of the smoothness parameter y is as small as possible in terms of N, while the W-trick (i.e., the value of W) is still independent of y. This allows one to combine this result in applications with inductive or recursive arguments in the y-parameter. We remark that when focussing on larger values of y, the function $w(N)$ in the statement can be chosen larger and consequently the bounds in this result improve.

As a first application, we estimate, for large values of y, the frequency with which an arbitrary finite complexity system of shifted linear forms $\psi_j(\boldsymbol{n}) + a_j \in \mathbb{Z}[n_1, \ldots, n_s]$, $1 \leqslant j \leqslant r$, simultaneously takes $[y', y]$-smooth values as the n_i vary over integers below N.

Theorem 1.3 (linear equations in smooth numbers). *Let $r, s \geqslant 2$ be integers and let $N > 2$ be a parameter. We consider the following setup*:

(i) *Let $\psi_1, \dots, \psi_r : \mathbb{Z}^s \to \mathbb{Z}$ be linear forms that are pairwise linearly independent over $\mathbb{Q}$, and let $a_1, \dots, a_r \in [-N, N] \cap \mathbb{Z}$ be integers. Let L denote the maximum of the absolute values of the coefficients of $\psi_1, \dots, \psi_r$.*

(ii) *Let $\mathfrak{K} \subseteq [-1, 1]^s$ be a fixed convex set of positive s-dimensional volume* $\mathrm{vol}\,\mathfrak{K} > 0$, *and suppose that the dilated copy*

$$N\mathfrak{K} = \{Nk \in \mathbb{R}^s : k \in \mathfrak{K}\}$$

satisfies $\psi_j(N\mathfrak{K}) + a_j \subseteq [1, N]$ for all $1 \leqslant j \leqslant r$.

(iii) *Let $\frac{1}{2}\log_3 N \leqslant y' \leqslant (\log N)^{K'}$ for some fixed $K' \geqslant 1$. Let $\eta \in (0, 1)$ be sufficiently small in terms of r and s and suppose that $y \leqslant N^{\eta}$.*

If $y \leqslant N^{\eta}$ is sufficiently large to ensure that $\Psi(N, y) > N/\log_8 N$, then

$$\sum_{\boldsymbol{n} \in \mathbb{Z}^s \cap N\mathfrak{K}} \prod_{j=1}^{r} g_{[y',y]}(\psi_j(\boldsymbol{n}) + a_j) = \mathrm{vol}(\mathfrak{K}) N^s \prod_{p<y'} \beta_p + o_{r,s,L}(N^s),$$

as $N \to \infty$ and where

$$\beta_p = \frac{1}{p^s} \sum_{\boldsymbol{u} \in (\mathbb{Z}/p\mathbb{Z})^s} \prod_{j=1}^{r} \frac{p}{p-1} \mathbf{1}_{\psi_j(\boldsymbol{u})+a_j \not\equiv 0 \ (\mathrm{mod}\ p)}.$$

We note as an aside that, as shown in [Green and Tao 2010, Section 4], this result implies an asymptotic count of the number of solutions in $[y', y]$-smooth numbers to systems of linear equations satisfying the non-degeneracy condition stated in [loc. cit., Theorem 1.8]. Theorem 1.3 is the first instance of a result of its kind for smooth numbers that applies in sparse situations where $\Psi(N, y) = o(N)$. We mainly include this result for illustration and remark that we have not tried to optimise the lower bound on $\Psi(N, y)/N$.[2] Theorem 1.3 follows by combining Theorem 1.1 with a "trivial" majorising function. Once suitable majorising functions are available on the full range of y on which Theorem 1.1 applies, a version of Theorem 1.3 will follow on that range of y, making Theorem 1.3 redundant. For this reason, we chose not to optimise the bounds.

The following unweighted version of Theorem 1.3 is an easy consequence of an asymptotic lower bound on the weight factor that appears in the definition of $g_{[y',y]}(n)$.

Corollary 1.4. *With the notation and under the assumptions of Theorem 1.3 the following holds. If $y \leqslant N^{\eta}$ is sufficiently large to ensure that $\Psi(N, y) > N/\log_8 N$ and $\frac{1}{2}\log_3 N \leqslant y' \leqslant (\log N)^{K'}$ for some fixed $K' \geqslant 1$, then the number $\mathcal{N}(N, \mathfrak{K})$ of $\boldsymbol{n} \in \mathbb{Z}^s \cap N\mathfrak{K}$ for which the given system of linear polynomials $(\psi_j(\boldsymbol{n}) + a_j)_{1 \leqslant j \leqslant r}$ takes simultaneous $[y', y]$-smooth values satisfies*

$$\mathcal{N}(N, \mathfrak{K}) := \sum_{\boldsymbol{n} \in \mathbb{Z}^s \cap N\mathfrak{K}} \prod_{j=1}^{r} \mathbf{1}_{S([y',y])}(\psi_j(\boldsymbol{n}) + a_j) \gg \mathrm{vol}(\mathfrak{K}) N^{s-r} \Psi(N, [y', y])^r \prod_{p} \beta_p$$

for all sufficiently large N.

[2] As this application only involves fairly large values in y, the bounds in Theorem 1.3 could be improved by establishing better bounds in Theorem 1.1 for large values of y.

Previous work and discussion. Before turning towards Theorem 1.1, we shall first discuss previous work in the direction of Theorem 1.3, partly because this work illustrates for which range of the smoothness parameter y one can hope to prove such a result.

In the dense case where $y > N^{\varepsilon}$ for any $\varepsilon > 0$, Theorem 1.3 has been proved by Lachand [2017] with $S([y', y])$ replaced by $S(y)$ and with $W = 1$, $A = 0$. The special case of Lachand's result where $s = 2$ and $a_1 = \cdots = a_r = 0$ also follows from work of Balog, Blomer, Dartyge and Tenenbaum [Balog et al. 2012]. The latter paper studies smooth values of binary forms and can be applied to $F(n_1, n_2) = \prod_{i=1}^{r} \psi_i(n_1, n_2)$. In a similar spirit, Fouvry [2010] investigated smooth values of absolutely irreducible polynomials $F(\boldsymbol{X}) \in \mathbb{Z}[X_1, \ldots, X_n]$ with $n \geqslant 2$ and $\deg F \geqslant 3$ and showed that

$$\sum_{\boldsymbol{n} \in \mathbb{Z}^n \cap [-N,N]} \mathbf{1}_{S(N^{d-\delta})}(F(\boldsymbol{n})) \gg N^n$$

for all $\delta < \frac{4}{3}$ and all $N > N_0(F)$, which also corresponds to the dense setting.

Concerning the sparse setting, i.e., smaller values of y, it follows from [Lagarias and Soundararajan 2011; 2012] that the lower threshold on y for which our counting function satisfies

$$\mathscr{N}(N, \mathfrak{K}) \to \infty, \quad \text{as } N \to \infty,$$

is $y > (\log N)^{\kappa}$ for some $\kappa \geqslant 1$. More precisely, Lagarias and Soundararajan studied primitive solutions to the equation $A+B=C$ in smooth numbers. Using our notation they proved in [Lagarias and Soundararajan 2012, Theorem 1.5] that GRH implies

$$\sum_{\substack{0<n_1+n_2\leqslant N \\ \gcd(n_1,n_2)=1}} \mathbf{1}_{S(y)}(n_1)\mathbf{1}_{S(y)}(n_2)\mathbf{1}_{S(y)}(n_1+n_2) \gg \frac{\Psi(N, y)^3}{N} \tag{1-2}$$

for any $y \geqslant (\log N)^{\kappa}$ and $\kappa > 8$. On the other hand, they showed in [loc. cit., Theorem 1.1] that the *abc* conjecture implies that the left-hand side above is bounded independent of N if $y = (\log N)^{\kappa}$ and $\kappa < 1$. In view of this latter result, and with applications in the spirit of Theorem 1.3 in mind, the range of y in which we prove Theorem 1.1 is optimal up to the size of the exponent K. Indeed, the left-hand side above is a special case of the expression $\mathscr{N}(N, \mathfrak{K})$ studied in Corollary 1.4 when ignoring the parameter y'. When combined with a simple majorising function, Theorem 1.3 and Corollary 1.4 follow from Theorem 1.1, the Green–Tao–Ziegler inverse theorem [Green et al. 2012] for the U^k-norms [3] and the generalised von Neumann theorem [Green and Tao 2010]. The existence of a suitable majorising function for $g_{[y',y]}$ on a larger range of y would imply an analogue of Theorem 1.3 on the intersection of that range and the range on which Theorem 1.1 holds. Such an analogue can only hold provided that $y > (\log N)^K$ for some sufficiently large K.

[3]To be precise, we require the quantitative version of the inverse theorem due to Manners [2018] in order to deduce explicit bounds from a very simple majorising function. A majorising function of the correct average order on a larger range of y would allow us to use the original qualitative inverse theorem [Green et al. 2012] instead. Since the first version of our paper appeared on the arXiv, a very substantial improvement to [Manners 2018] has been obtained by Leng, Sah and Sawhney [Leng et al. 2024]. We note that, in view of the bounds we obtain in Theorem 1.1, the range of y to which Theorem 1.3 and Corollary 1.4 apply would, without further optimisation, not significantly improve by exchanging in our proof the application of [Manners 2018] for [Leng et al. 2024].

Unconditional versions of the result (1-2) of [Lagarias and Soundararajan 2012, Theorem 1.5], but with more restrictive ranges for y, have been established by a number of authors. Harper [2016, Corollary 1] significantly improved those ranges and showed, almost matching the conditional range, that if $K > 1$ is sufficiently large then the asymptotic formula

$$\sum_{0<n_1+n_2\leqslant N} \mathbf{1}_{S(y)}(n_1)\mathbf{1}_{S(y)}(n_2)\mathbf{1}_{S(y)}(n_1+n_2) = \frac{\Psi(N,y)^3}{2N}\biggl(1+O\biggl(\frac{\log(u+1)}{\log y}\biggr)\biggr)$$

holds for all $y \geqslant (\log N)^K$, where $u = (\log N)/\log y$.

Concerning smooth solutions to Diophantine equations in many variables, there is a vast literature on Waring's problem in smooth numbers. Building on [Harper 2016] mentioned before, Drappeau and Shao [2016] studied Waring's problem in y-smooth numbers for $y \geqslant (\log N)^K$ and $K > 1$ sufficiently large, which is the currently largest possible range of y. We refer to that paper for references to previous work on this question.

Turning towards Theorem 1.1, the closest previous result is due to Lachand [2017] who proved, starting out from a Möbius inversion formula for $\mathbf{1}_{S(y)}$, that

$$\sum_{n\leqslant N}\biggl(\mathbf{1}_{S(y)} - \frac{\Psi(N,y)}{N}\biggr)F(g(n)\Gamma) = o((1+\|F\|_{\mathrm{Lip}})\Psi(N,y))$$

for any $\varepsilon > 0$ and uniformly for $x \geqslant y \geqslant \exp((\log N)/(\log\log N)^{1-\varepsilon})$. We shall explain at the end of Section 2 why, when restricting to large values of y, neither a W-trick nor the restriction to integers with prime factors in the interval $[y', y]$ is necessary.

As exponential phases $e(\theta n^k)$ for $\theta \in \mathbb{R}$, $k \in \mathbb{N}$, form a special case of the nilsequences $F(g(n)\Gamma)$ in Theorem 1.1, previous work on this question includes all work on exponential sum (and Weyl sum) estimates over smooth numbers. The latter is a well-studied subject. Our proof indeed builds on some of this previous work, in particular on that of [Harper 2016; Drappeau and Shao 2016; Wooley 1995].

2. Outline of the proof and overview of the paper

The proof of Theorem 1.1 occupies most of this paper. After recalling some of the background on smooth numbers in Section 3, we investigate in Section 4 the distribution of $[y', y]$-smooth numbers in short intervals and the results we prove here may be of independent interest. The results of this section motivate the definition of the weighted function $g_{[y',y]}(n)$ as well as an auxiliary weighted function $h_{[y',y]}(n)$, and we prove that these functions are equidistributed in *short intervals*. After extending in Section 5.1 work of Harper on the distribution of y-smooth numbers in progressions, we deduce in Section 5.2 that a W-tricked version of $g_{[y',y]}(n)$ is equidistributed in *short arithmetic progressions*. The work in Sections 4 and 5 (i.e., the fact that $g_{[y',y]}(n)$ is equidistributed in short arithmetic progressions) prepares the ground for the proof of the noncorrelation estimate (1-1). It allows us in Section 8 to reduce the task of proving (1-1) to the case in which the polynomial sequence g is "equidistributed". Using a Montgomery–Vaughan-type decomposition, this task is further reduced in Section 9.2 to a Type II sum estimate. More precisely, we reduce our task to bounding bilinear sums of the (slightly simplified) form

$$\sum_{m\in[y',y]} \sum_{N/m<n\leqslant(N+N_1)/m} \mathbf{1}_{S([y',y])}(n)\Lambda(m)F(g(mn)\Gamma).$$

Swapping the order of summation reduces this problem to that of bounding the correlation of the (W-tricked) von Mangoldt functions with equidistributed nilsequences, *provided* we can show an implication of the form

$$(g(m)\Gamma)_{m\leqslant N} \text{ is equidistributed} \quad \Longrightarrow \quad (g(mn)\Gamma)_{m\leqslant N/n} \text{ is equidistributed for most } n\in S(N,[y',y]).$$

A precise version of this implication will be established in Section 8.2. The proof of this implication builds on the material from Sections 5, 6 and 7. More precisely, it relies on a strong recurrence result for polynomial sequences over $[y', y]$-numbers that is proved in Section 7. The recurrence result in turn relies on Weyl sum estimates for $[y', y]$-smooth numbers that we establish in Section 6 by extending [Drappeau and Shao 2016]. As the Weyl sum estimates by themselves are in fact too weak for our purposes, we need to combine them with a bootstrapping argument, which requires good bounds on the distribution of smooth numbers in short arithmetic progressions (as established in Section 5) as input. The final section contains the proof of Theorem 1.3.

The most difficult part of this work is to ensure that y can be chosen as small as $(\log N)^K$ in our main result. To achieve this, we need to run the bootstrapping argument in Section 7 in four different stages of iterations, using the full scale of results proved in Sections 4 and 5, and we have to choose our parameters in Sections 8 and 9 very carefully. Similarly, working with y as small as $(\log N)^K$ requires us to use a Montgomery–Vaughan-type reduction to a Type II sum estimate in Section 9 instead of working with the intrinsic decomposition that was used, e.g., in [Harper 2016, Section 3] to establish minor arc estimates for exponential sums over smooth numbers. (The intrinsic decomposition appears in the proofs of Lemmas 6.2 and 6.4.) Harper's approach, which in principle generalises to the nilsequences setting, involves an application of the Cauchy–Schwarz inequality and reduces the minor arc estimate to bounds on degree-1 Weyl sums $\sum_{n\leqslant N} e(n\theta)$ for irrational θ. The losses of the Cauchy–Schwarz application in this approach can be compensated for by the strong bounds available for degree-1 Weyl sums for irrational θ. When working with an irrational (=equidistributed) nilsequence $n\mapsto F(g(n)\Gamma)$ instead of $n\mapsto e(\theta n)$ for irrational θ, the savings on the corresponding sum $\sum_{n\leqslant N} F(g(n)\Gamma)$ are usually much weaker, and in our case too weak in order to compensate for the loss of the Cauchy–Schwarz application when $y=(\log N)^K$.

Smooth numbers and the "W-trick". A result of the form of Theorem 1.1 requires as a necessary condition that the function it involves, here $g_{[y',y]}^{(W,A)}$, is equidistributed in arithmetic progressions to small moduli. The reason for this is that additive characters $n\mapsto e(an/q)$, where $a,q\in\mathbb{N}$ and $e(x)=e^{2\pi ix}$, form a special case of the nilsequences $n\mapsto F(g(n)\Gamma)$ that appear in the statement. The function $g_{[y',y]}^{(W,A)}$ is a weighted version of the characteristic function of $[y', y]$-smooth numbers, restricted to a reduced residue class A modulo W. Both the use of a W-trick, i.e., the restriction to integers of the form $n=Wm+A$, as well as the restriction to the subset of y-smooth numbers that are coprime to $P(y') := \prod_{p<y'} p$ are a means to ensure equidistribution in progressions. The weight is introduced in order to guarantee equidistribution in short intervals and progressions.

We proceed to explain why the W-trick is used. Combining work of Soundararajan [2008] and Harper [2012b] on the distribution of smooth numbers in progressions with work of de la Bretèche and Tenenbaum [2005] one may deduce that

$$\Psi(x, y; q, a) := \sum_{\substack{n \leqslant x \\ n \equiv a \pmod q}} \mathbf{1}_{S(y)}(n) \sim \frac{1}{\phi(q)} \sum_{\substack{n \leqslant x \\ \gcd(n,q)=1}} \mathbf{1}_{S(y)}(n) \sim \frac{\Psi(x, y)}{q} \prod_{p \mid q} \frac{1 - p^{-\alpha(x,y)}}{1 - p^{-1}}$$

for $(\log x)^2 < y \leqslant x,\ 2 \leqslant q \leqslant y^2$, and $(a, q) = 1$, as $\log x / \log q \to \infty$.

When $\alpha(x, y)$ is sufficiently close to 1, which happens when y is sufficiently close to x, the final product over prime divisors $p \mid q$ will be approximately 1 and the above asymptotic implies that $S(y)$ is equidistributed in *all* residue classes (reduced and nonreduced) modulo q for all small values of q. The necessary condition for the validity of Theorem 1.1 is met in this situation and no W-trick is required.

Once $\alpha(x, y)$ is no longer close to 1, the product over prime factors $p \mid q$ in the asymptotic formula above genuinely depends on q. In this case, $S(y)$ is seen to be equidistributed in the *reduced* residue classes modulo a fixed integer q. However, the density of $S(y)$ within a reduced class will differ from that in a nonreduced class modulo q, the latter being obtained by dividing out the common factor and applying the asymptotic formula with q replaced by a suitable divisor of q. To remove this discrepancy between reduced and nonreduced residue classes, one may use a W-trick with a slowly growing function $w(N)$. Fixing a reduced residue $A \pmod W$, one is then interested in the count of y-smooth integers of the form $n = W(qm + a) + A$, which satisfies

$$\Psi(x, y; Wq, Wa + A) \sim \frac{\Psi(x, y)}{Wq} \biggl(\prod_{p \mid W} \frac{1 - p^{-\alpha(x,y)}}{1 - p^{-1}} \biggr) \prod_{p' \mid q,\, p' \nmid W} \frac{1 - p'^{-\alpha(x,y)}}{1 - p'^{-1}}.$$

If $W = P(w(N))$ is the product of all primes $p < w(N)$, then all prime factors p' that appear in the final product satisfy $p' \geqslant w(N)$, i.e., are large. If q is not too large, this allows one to show that the final product is asymptotically equal to 1, with an error term that depends on $w(N)$. This ensures, within any fixed reduced residue class $A \pmod W$, that the necessary condition for the validity of Theorem 1.1 holds.

3. Smooth numbers

In this section we collect general properties of smooth numbers that we will frequently make use of within this paper. Recall the definitions of the sets $S(y)$, $S([y', y])$, $S(y, x)$ and $S([y', y], x)$ from the introduction. The relative quantity

$$u := \frac{\log x}{\log y} \tag{3-1}$$

frequently appears when describing properties of $S(x, y)$ and $S(x, [y', y])$.

3.1. ***The summatory function of*** $\mathbf{1}_{S(y)}$**.** Let

$$\Psi(x, y) = |S(x, y)| = \sum_{n \leqslant x} \mathbf{1}_{S(y)}(n)$$

denote the number of y-smooth numbers below x. As a function of x, $\Psi(x, y)$ may be viewed as the summatory function of the characteristic function $\mathbf{1}_{S(y)}$ of y-smooth numbers. The latter function is completely multiplicative and the associated Dirichlet series is given by

$$\zeta(s, y) = \sum_{n \in S(y)} n^{-s} = \prod_{p \leqslant y} (1 - p^{-s})^{-1} \quad (\Re(s) > 0).$$

Rankin's trick shows that $\Psi(x, y) \leqslant x^{\sigma} \zeta(\sigma, y)$ for all $\sigma > 0$. The (unique) saddle point of the function $\sigma \mapsto x^{\sigma} \zeta(\sigma, y)$ that appears here is usually denoted by $\alpha(x, y)$ and we have $\alpha(x, y) \in (0, 1)$. Hildebrand [1986a, Lemma 4] (see also [Hildebrand and Tenenbaum 1986, Lemma 2]) showed that

$$\alpha(x, y) = 1 - \frac{\log(u \log(u+1))}{\log y} + O\left(\frac{1}{\log y}\right) \quad (\log x < y \leqslant x). \tag{3-2}$$

This saddle point can be used in order to asymptotically describe $\Psi(x, y)$. More precisely, Hildebrand and Tenenbaum [1986, Theorems 1, 2] proved that, uniformly for $y / \log x \to \infty$,

$$\Psi(x, y) = \frac{x^{\alpha} \zeta(\alpha, y)}{\alpha \sqrt{2\pi \log x \log y}} \left(1 + O\left(\frac{1}{\log(u+1)} + \frac{1}{\log y}\right)\right), \tag{3-3}$$

where $\alpha = \alpha(x, y) > 0$ and $u = (\log x)/\log y$. In applications, it is frequently necessary to understand the relation between $\Psi(cx, y)$ and $\Psi(x, y)$ as $c > 0$ varies. In this direction, Theorem 3 of [Hildebrand and Tenenbaum 1986] shows that

$$\Psi(cx, y) = \Psi(x, y) c^{\alpha(x,y)} \left(1 + O\left(\frac{1}{u} + \frac{\log y}{u}\right)\right), \tag{3-4}$$

uniformly for $x \geqslant y \geqslant 2$ and $1 \leqslant c \leqslant y$.

The following related lemma is a consequence of (3-2).

Lemma 3.1. *Suppose that* $x > 2$ *is sufficiently large and that* $\log x < y \leqslant x$. *Then*

$$\alpha(cx, y) - \alpha(x, y) \ll 1/\log y$$

uniformly for all $c \in [1, 2]$.

Proof. We write $cx = x + x'$ and expand out the expression that (3-2) provides for $\alpha(x + x', y)$ when taking into account the definition (3-1) of u. The lemma follows provided we can show that $\log_2(x + x') = \log_2 x + O(1)$ and $\log_3(x + x') = \log_3 x + O(1)$. Both of these estimates may be deduced by repeated application of the identity $\log(t + t') = \log t + \log(1 + t'/t)$ combined with the observation that for $t > 2$ (in particular, for large t) the inequality $t \geqslant t'$ is preserved in the sense that $T > T'$ if $T := \log t$ and $T' := \log(1 + t'/t)$. Finally, the bound $\log(1 + t'/t) \ll 1$ provides the shape of the error term. □

3.2. ***The truncated Euler product.*** On the positive real line, the Dirichlet series $\zeta(\sigma, y)$ can be estimated and we will frequently use the following lemma which is [Montgomery and Vaughan 2007, Lemma 7.5].

Lemma 3.2. *Suppose that* $y \geqslant 2$. *If* $\max\{2/\log y, 1-4/\log y\} \leqslant \sigma \leqslant 1$, *then*

$$\prod_{p\leqslant y}(1-p^{-\sigma})^{-1} \asymp \log y.$$

If $2/\log y \leqslant \sigma \leqslant 1-4/\log y$, *then*

$$\prod_{p\leqslant y}(1-p^{-\sigma})^{-1} = \frac{1}{1-\sigma}\exp\biggl(\frac{y^{1-\sigma}}{(1-\sigma)\log y}\biggl\{1+O\biggl(\frac{1}{(1-\sigma)\log y}\biggr)+O(y^{-\sigma})\biggr\}\biggr).$$

3.3. ***The summatory function of*** $\mathbf{1}_{S([y',y])}$. De la Bretèche and Tenenbaum [2005] analysed the quantities

$$\Psi_m(x,y) = \sum_{\substack{n\in S(x,y)\\(n,m)=1}} 1 \quad\text{and}\quad \frac{\Psi_m(x/d,y)}{\Psi(x,y)}$$

and established asymptotic estimates for these expressions under certain assumptions on m and d. Observe that

$$\Psi_{P(y')}(x,y) = \Psi(x,[y',y])$$

for $P(y') = \prod_{p<y'} p$. We shall require versions of two of the results from [de la Bretèche and Tenenbaum 2005] that are restricted to the case $m = P(y')$ but come with sufficiently good explicit error terms. To state these results, define for any given integer $m \in \mathbb{N}$ the restricted Euler product

$$g_m(s) = \prod_{p\mid m}(1-p^{-s}) \quad (s\in\mathbb{C}). \tag{3-5}$$

The following lemma is a consequence of [de la Bretèche and Tenenbaum 2005, Théorème 2.1].

Lemma 3.3. *Let* $K'>0$ *and* $K>\max(2K',1)$ *be constants, let* $x\geqslant 2$ *and* $y'\leqslant(\log x)^{K'}$. *Then*

$$\Psi(x,[y',y]) = g_{P(y')}(\alpha(x,y))\Psi(x,y)\biggl(1+O\biggl(\frac{\log_2 x}{\log x}+\frac{1}{\log y}\biggr)\biggr)$$

uniformly for all $(\log x)^K < y \leqslant x$.

Proof. This result follows from [de la Bretèche and Tenenbaum 2005, Théorème 2.1 and Corollaire 2.2]. We need to verify that the error term is of the shape claimed in the statement above. Let $m = P(y') = \prod_{p<y'} p$. Since $y' < y^{1/2}$ and $\pi(y') \ll y^{1/2}/(\log y)$, we are either in the situation C_1 or C_2 of [loc. cit., Corollaire 2.2].

Observe that in our case $\bar{u} := \min(u, y/\log y) = u = \log x/\log y$ and

$$W_m = W_{P(y')} = \log p_{\pi(y')} \asymp \log y' \ll \log\log x.$$

Hence, $\vartheta_m \ll (\log\log x)/\log y$ $(\ll 1)$.

If C_1 holds, then the error term in the expression for $\Psi_{P(y')}(x,y)$ is bounded by

$$E^* := E^*_m(x,y) \ll \frac{\vartheta_m \log(u+2)}{u} \ll \frac{(\log_2 x)^2}{\log x}.$$

If condition C_2 holds, then $y^{1/\log(u+2)} \ll \omega(P(y')) < y' \leqslant (\log x)^{K'}$ and, hence,

$$\log y \ll (\log_2 x)\log(u+2) \ll (\log_2 x)^2.$$

This shows that $\log(u+2) \asymp \log_2 x$ and

$$E^* \ll \frac{\vartheta_{m_x}^2}{\log(u+2)} \ll \frac{(\log_2 x)^2}{(\log y)^2 \log(u+2)} \ll \frac{1}{\log y},$$

as required. □

By restricting [de la Bretèche and Tenenbaum 2005, Théorème 2.4] to the situation of the present paper, we obtain:

Lemma 3.4. (i) *If* $y' \leqslant (\log x)^{K'}$ *for some constant* $K' > 0$, *then we have*

$$\Psi_{P(y')}(x/d, y) \ll d^{-\alpha} \Psi_{P(y')}(x, y)$$

uniformly for all $\max(y'^2, (\log x)^2) < y \leqslant x,\ 1 \leqslant d \leqslant x/y$.

(ii) *If*, *moreover*, $1 \leqslant d \leqslant \exp((\log x)^{1/3})$, *the following more precise statement holds. We have*

$$\Psi_{P(y')}(x/d, y) = d^{-\alpha} \Psi_{P(y')}(x, y)\left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right)\right)$$

uniformly for all $\max(y'^2, (\log x)^2) < y \leqslant x$ *and* $1 \leqslant d \leqslant \exp((\log x)^{1/3})$, *where* $y' \leqslant (\log x)^{K'}$ *and* $K' > 0$ *as before.*

Proof. Part (i) is a direct application of [de la Bretèche and Tenenbaum 2005, Théorème 2.4(i)] combined with Lemma 3.3, when taking into account that $\delta = 0$ according to the remark following (2.8) of that paper, since $y > \log^2 x$.

Concerning part (ii), Lemma 3.3 implies that the asymptotic formula given in the statement is equivalent to

$$\Psi_{P(y')}(x/d, y) = d^{-\alpha} g_{P(y')}(\alpha(x, y)) \Psi(x, y)\left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right)\right).$$

The statement thus follows from [de la Bretèche and Tenenbaum 2005, Théorème 2.4(ii)] provided we can show that under our assumptions the error term in that result is in fact of the above-stated shape. In the notation of [loc. cit., Théorème 2.4(ii)] we therefore need to show that $h_m = o(1)$ and $(1 - t^2/(2u^2))^{bu} = 1 + o(1)$ as $x \to \infty$, where $o(1)$ needs to be made explicit in both cases. In the former case, we have

$$h_m = \frac{1}{u_y} + t\frac{(1+E_m)}{u} + E_m^*(x, y) \ll \frac{\log_2(2y)}{\log y} + \frac{1}{\log_2 x} + \frac{\log d}{\log x} + \frac{\log d}{\log y}\frac{E_m}{u},$$

where we used that $u_y = u + (\log y)/\log(u+2) \gg (\log y)/\log_2(2y)$ (cf. [de la Bretèche and Tenenbaum 2005, (2.18)]) and that $E_m^*(x, y) \ll 1/(\log_2 x)$ by the previous proof. For the term $(1+E_m)/u$, we obtain

$$E_m = \frac{\vartheta_m (u \log(u+2))^{\vartheta_m}}{1 + \vartheta_m \log(u+2)}.$$

Note that $\vartheta_m = (\log p_{\omega(m)})/\log y = (\log y')/\log y \leqslant \frac{1}{2}$ for $y \geqslant y'^2$. Further, $u \log(u+2) \geqslant 1$. Hence,

$$\frac{E_m}{u} \ll \frac{\vartheta_m \log^{1/2}(u+2)}{u^{1/2}(1 + \vartheta_m \log(u+2))}$$

If $\log y > (\log_2 x)^3$, then

$$\frac{\log d}{\log y}\frac{E_m}{u} \ll \frac{\log d}{\log y}\frac{\log^{1/2}(u+2)}{u^{1/2}(\log_2 x)^2} \ll \frac{(\log d)(\log_2 x)^{1/2}}{(\log x)^{1/2}(\log y)^{1/2}\log_2 x} \ll \frac{\log d}{(\log x)^{1/2}(\log_2 x)^2}.$$

If $\log y \leqslant (\log_2 x)^3$ and since $\log_2 x \ll \log y$, we have

$$t\frac{E_m}{u} \ll t\frac{\vartheta_m \log^{1/2}(u+2)}{u^{1/2}(1+\vartheta_m \log(u+2))} \ll t\frac{\vartheta_m(\log_2 x)^{1/2}(\log y)^{1/2}}{(\log x)^{1/2}\big(1+\frac{1}{(\log_2 x)^2}\log(u+2)\big)}$$
$$\ll t\frac{(\log_2 x)^2}{(\log x)^{1/2}} \ll \frac{\log d}{\log y}\frac{(\log_2 x)^2}{(\log x)^{1/2}} \ll \frac{(\log d)\log_2 x}{(\log x)^{1/2}}.$$

It remains to analyse $(1-t^2/(2u^2))^{bu}$. Since

$$\frac{t}{u} = \frac{\log d}{\log x} = o(1),$$

we have

$$\left(1-\frac{t^2}{2u^2}\right)^{bu} = \exp\left(-C\frac{t^2}{u}\right) = 1+O\left(\frac{1}{\log y}\right)$$

for some constant $C > 0$. The final step above follows since

$$\frac{t^2}{u} = \frac{(\log d)^2}{(\log y)\log x} \ll \frac{1}{\log y}$$

by our assumption on d. □

4. Smooth numbers in short intervals

Our aim in this section is to show that a suitably weighted version of the function $\mathbf{1}_{S([y',y])}$ is equidistributed in short intervals. More precisely, we will consider intervals contained in $[x, 2x]$ that are of length at least $x/(\log x)^c$. The error terms in Lemmas 3.3 and 3.4 are too weak in order to allow one to work with intervals as short as $x/(\log x)^c$. Better error terms are, however, available when considering the quantity

$$\frac{\Psi_{P(y')}(x/d, y)}{\Psi_{P(y')}(x, y)} = \frac{\Psi(x/d, [y', y])}{\Psi(x, [y', y])} \quad \text{instead of} \quad \frac{\Psi_{P(y')}(x/d, y)}{\Psi(x, y)}.$$

In the first subsection below we prove asymptotic formulas for the former quotient. In the second subsection we introduce a suitable smooth weight for $\mathbf{1}_{S([y',y])}$ and prove that the correspondingly weighted $[y', y]$-smooth numbers are equidistributed in short intervals of the above form.

4.1. ***The local behaviour of $\Psi(x, [y', y])$.*** In order to analyse the distribution of $S([y', y])$ in short intervals, we need to compare $\Psi(x, [y', y])$ to $\Psi(x(1+1/z), [y', y])$. In the notation of Lemma 3.4 this means that our d is very close to 1. Recall that $u = (\log x)/\log y$. We split our analysis below into two cases according to whether u is large or small.

Lemma 4.1 (local behaviour of $\Psi_{P(y')}$ for large u). *Let $x > 2$ be a parameter, let $K' > 0$ and $K \geqslant \max(2K', 2)$ be constants and suppose that $y' \leqslant (\log x)^{K'}$ and $(\log x)^K < y < \exp((\log x)^{1/3})$. Let $\alpha = \alpha(x, y)$ denote the saddle point associated to $S(x, y)$. Then*

$$\left|\Psi(x, [y', y]) - (1+z^{-1})^{-\alpha}\Psi\big((1+z^{-1})x, [y', y]\big)\right| \ll \Psi(x, [y', y])\left(\frac{\sqrt{\log y}}{z\log^{1/6} x} + \frac{O_A(1)}{\log^A x}\right)$$

for all $z > 1$.

Proof. We will follow the strategy via Perron's formula employed in [Hildebrand and Tenenbaum 1986, Lemma 8; de la Bretèche and Tenenbaum 2005, Section 4.2], and start by bounding Euler factors from the relevant Dirichlet series. Suppose that $s = \alpha + i\tau$ and $(\log x)^{-1/4} \leqslant \tau \leqslant (\log y)^{-1}$. Then[4] for any $y' \leqslant p \leqslant y$, we have

$$\frac{|1-p^{-\alpha}|}{|1-p^{-s}|} = \left(1 + \frac{2(1-\cos(\tau \log p))}{p^\alpha(1-p^{-\alpha})^2}\right)^{-1/2} \leqslant \exp\left(-c_1\frac{1-\cos(\tau\log p)}{p^\alpha(1-p^{-\alpha})^2}\right) \leqslant \exp\left(-c_2\frac{(\tau\log p)^2}{p^\alpha}\right),$$

where we took advantage of the fact that $1 - p^{-\alpha} \asymp 1$ for all $p \leqslant y$ since $\alpha > \frac{1}{2}$. Hence,

$$\begin{aligned}\prod_{y'\leqslant p\leqslant y}\frac{|1-p^{-\alpha}|}{|1-p^{-s}|} &\leqslant \exp\left(-c_2\frac{\tau^2}{y^\alpha}\sum_{y'\leqslant p\leqslant y}(\log p)^2\right) \leqslant \exp\left(-c_2\frac{\tau^2}{y^\alpha}\sum_{y/2\leqslant p\leqslant y}(\log p)^2\right)\\ &\leqslant \exp(-c_3\tau^2 y^{1-\alpha}\log y) \leqslant \exp\big(-c_4\tau^2 u\log(u+1)\log y\big)\\ &\leqslant \exp(-c_4\tau^2\log x). \end{aligned} \tag{4-1}$$

By invoking bounds of the above Euler factors in different regimes of τ, de la Bretèche and Tenenbaum [2005, (4.50)] deduced from Perron's formula that

$$\Psi(x, [y', y]) = \frac{1}{2\pi i}\int_{\alpha - i/\log y}^{\alpha + i/\log y}\zeta_m(s, y)x^s\,\frac{ds}{s} + O(x^\alpha\zeta_m(\alpha, y)R), \tag{4-2}$$

where

$$R = \exp\left(\frac{-c_5 u}{(\log 2u)^2}\right) + \exp(-(\log y)^{4/3}), \quad m = P(y') = \prod_{p<y'} p,$$

and

$$\zeta_m(s, y) = \prod_{p\leqslant y,\, p\nmid m}(1-p^{-s})^{-1}.$$

The approximation (4-2) applies in our situation since $u > (\log x)^{3/2} > (\log y)^{3/2}$ and $\pi(y') = \omega(P(y')) < (\log x)^{K'} < \sqrt{y}$, which ensures that the conditions [de la Bretèche and Tenenbaum 2005, (4.40)] are satisfied. Under our assumption that $\log x < y < \exp((\log x)^{1/3})$, we have

$$R \ll \exp(-c_5(\log x)^{1/2}) + \exp(-(\log\log x)^{4/3}) \ll_A (\log x)^{-A}.$$

[4] See the proof of [Hildebrand and Tenenbaum 1986, Lemma 8] for more details.

The bound (4-1) allows us to further truncate the integral in (4-2). More precisely,

$$\begin{aligned}\left|\int_{\alpha\pm i/(\log x)^{1/3}}^{\alpha\pm i/\log y} \zeta_m(s,y)x^s\,\frac{\mathrm{d}s}{s}\right| &\leqslant \zeta_m(\alpha,y)\frac{x^\alpha}{\alpha}\int_{1/(\log x)^{1/3}}^{1/\log y} \exp(-c_4\tau^2\log x)\,\mathrm{d}\tau\\ &\leqslant \frac{\zeta_m(\alpha,y)x^\alpha}{\alpha\log y}\exp(-c_4(\log x)^{-1/3})\\ &\leqslant \Psi(x,[y',y])\exp(-c_6(\log x)^{-1/3}),\end{aligned}$$

which implies that

$$\Psi(x,[y',y]) = \frac{1}{2\pi i}\int_{\alpha-i/(\log x)^{1/3}}^{\alpha+i/(\log x)^{1/3}} \zeta_m(s,y)x^s\,\frac{\mathrm{d}s}{s} + O_A\biggl(\frac{\Psi(x,[y',y])}{(\log x)^A}\biggr). \tag{4-3}$$

The latter approximation then yields

$$\begin{aligned}&\Psi(x,[y',y]) - (1+z^{-1})^{-\alpha}\Psi(x(1+z^{-1}),[y',y])\\ &\qquad= \frac{1}{2\pi i}\int_{\alpha-i/(\log x)^{1/3}}^{\alpha+i/(\log x)^{1/3}} \zeta_m(s,y)x^s\bigl(1-(1+z^{-1})^{s-\alpha}\bigr)\,\frac{\mathrm{d}s}{s} + O_A\biggl(\frac{\Psi(x,[y',y])}{(\log x)^A}\biggr)\\ &\qquad\ll \zeta_m(\alpha,y)x^\alpha\int_{-1/(\log x)^{1/3}}^{1/(\log x)^{1/3}} |1-(1+z^{-1})^{i\tau}|\,\frac{\mathrm{d}\tau}{\alpha} + O_A\biggl(\frac{\Psi(x,[y',y])}{(\log x)^A}\biggr)\\ &\qquad\ll \frac{\zeta_m(\alpha,y)x^\alpha}{z}\int_{-1/(\log x)^{1/3}}^{1/(\log x)^{1/3}} |\tau|\,\mathrm{d}\tau + O_A\biggl(\frac{\Psi(x,[y',y])}{(\log x)^A}\biggr)\\ &\qquad\ll \frac{\zeta_m(\alpha,y)x^\alpha}{z(\log x)^{2/3}} + O_A\biggl(\frac{\Psi(x,[y',y])}{(\log x)^A}\biggr)\\ &\qquad\ll \Psi(x,[y',y])\biggl(\frac{\sqrt{\log y}}{z(\log x)^{1/6}} + O_A((\log x)^{-A})\biggr),\end{aligned}$$

as claimed. □

The previous lemma applies to $y < \exp((\log x)^{1/3})$. Below, we establish an analogous result in the complementary range where $\exp((\log x)^{1/4}) \leqslant y \leqslant x$. In this case, $1-\alpha(x',y)$ is very small as x' ranges over $[x,2x]$, and the error terms in Lemma 3.3 are also very good. This allows us to reduce the problem of asymptotically evaluating $\Psi(x(1+z^{-1}),[y',y])$ to that of bounding $\Psi(x(1+z^{-1}),y)-\Psi(x,y)$. Bounds on the latter difference have been established by [Hildebrand 1986b, Theorem 3] in the case where $u=(\log x)/\log y$ is small, which applies to our current situation.[5]

Lemma 4.2 (local behaviour of $\Psi_{P(y')}$ for small u). *Let $x>2$ be a parameter, suppose that $y' \leqslant (\log x)^{K'}$ for some fixed $K'>0$ and that $\exp((\log x)^{1/4}) \leqslant y \leqslant x$, and let $\alpha=\alpha(x,y)$ denote the saddle point associated to $S(x,y)$. Then*

$$\bigl|\Psi(x(1+z^{-1}),[y',y]) - (1+z^{-1})^{\alpha}\Psi(x,[y',y])\bigr| \ll \Psi(x,[y',y])\frac{(\log_2 x)^2}{\log y} \tag{4-4}$$

holds uniformly for $1 \leqslant z \leqslant y^{5/12}$.

[5] Hildebrand's theorem has been extended by several authors and we refer to the survey paper [Granville 2008, Section 4.1] for a discussion of these extensions and references.

Proof. Under the assumptions on y it follows from (3-2) that the saddle point $\alpha(x, y)$ is very close to 1. More precisely, writing $u' = (\log x')/\log y$, we have

$$\alpha(x', y) = 1 - \frac{\log(u' \log(u'+1))}{\log y} + O\Big(\frac{1}{\log y}\Big) = 1 + O\Big(\frac{\log_2 x'}{\log y}\Big) = 1 + O\Big(\frac{\log_2 x}{(\log x)^{1/4}}\Big) \tag{4-5}$$

uniformly for all $x' \in [x, 2x]$ and $y \geqslant \exp((\log x)^{1/4})$. This implies, in particular, that

$$(1+z^{-1})^{\alpha} = (1+z^{-1})(1+z^{-1})^{O((\log_2 x/\log y)} = (1+z^{-1})\Big(1 + O\Big(\frac{\log_2 x}{z \log y}\Big)\Big)$$

for $\alpha = \alpha(x, y)$. Substituting this formula into (4-4) and rearranging the resulting expression reduces our task to that of establishing

$$\Psi\big(x(1+z^{-1}), [y', y]\big) - \Psi(x, [y', y]) = \frac{\Psi(x, [y', y])}{z}\Big(1 + O\Big(\frac{z(\log_2 x)^2}{\log y}\Big)\Big).$$

By Lemma 3.3, the difference on the left-hand side satisfies

$$\begin{aligned}&\Psi\big(x(1+z^{-1}), [y', y]\big) - \Psi(x, [y', y]) \\ &\quad = g_{P(y')}(\alpha_z)\Psi(x(1+z^{-1}), y) - g_{P(y')}(\alpha)\Psi(x, y) + O\Big(\Psi(x, [y', y])\Big(\frac{\log_2 x}{\log x} + \frac{1}{\log y}\Big)\Big),\end{aligned} \tag{4-6}$$

where $\alpha_z = \alpha(x(1+z^{-1}), y)$ is the saddle point associated to $S(x(1+z^{-1}), y)$. In order to simplify the main term above, we seek to relate $g_{P(y')}(\alpha_z)$ to $g_{P(y')}(\alpha)$. Following [de la Bretèche and Tenenbaum 2005, Section 3.6], define $\gamma_m(s) := \log\ g_m(s)$ and let $\gamma'_m(s)$ denote the derivative with respect to s. Since $\alpha_z \leqslant \alpha$, we then have

$$\log\Big(\frac{g_{P(y')}(\alpha_z)}{g_{P(y')}(\alpha)}\Big) = \int_{\alpha}^{\alpha_z} \gamma'_{P(y')}(\sigma)\ \mathrm{d}\sigma \leqslant (\alpha - \alpha_z) \sup_{\alpha_z \leqslant \sigma \leqslant \alpha} \gamma'_{P(y')}(\sigma)$$

and shall estimate the latter two factors in turn. Since $1 < 1+z^{-1} < 2$ for $z > 1$, it follows from (4-5) that

$$\alpha - \alpha_z \ll \frac{\log_2 x}{\log^{1/4} x}.$$

Concerning the second factor, we have

$$\gamma'_{P(y')}(\sigma) = \sum_{p \leqslant y'} \frac{\log p}{p^{\sigma} - 1} \ll \sum_{n \leqslant y'} \frac{\Lambda(n)}{n^{\sigma}} \ll {y'}^{1-\sigma} \sum_{n \leqslant y'} \frac{\Lambda(n)}{n} \ll \log y' \ll \log_2 x$$

since $\log y' \ll \log_2 x$ and $1 - \sigma \ll_{\varepsilon} \log^{-1/4+\varepsilon} x$ by (4-5). Hence

$$\frac{g_{P(y')}(\alpha_z)}{g_{P(y')}(\alpha)} = \exp\Big(\log\Big(\frac{g_{P(y')}(\alpha_z)}{g_{P(y')}(\alpha)}\Big)\Big) = \exp\Big(O\Big(\frac{(\log_2 x)^2}{\log^{1/4} x}\Big)\Big) = 1 + O\Big(\frac{(\log_2 x)^2}{\log^{1/4} x}\Big).$$

Substituting this expression into (4-6), we obtain

$$\begin{aligned}&\Psi(x(1+z^{-1}), [y', y]) - \Psi(x, [y', y]) \\ &\qquad = g_{P(y')}(\alpha)\big\{\Psi(x(1+z^{-1}), y) - \Psi(x, y)\big\} + O\Big(\Psi(x, [y', y])\frac{(\log_2 x)^2}{\log y}\Big),\end{aligned}$$

where we used Lemma 3.3 and (3-3) to simplify the error term. To the right-hand side, we may now apply Theorems 1 and 3 of [Hildebrand 1986b], which assert that

$$\Psi(x(1+z^{-1}), y) - \Psi(x, y) = \frac{\Psi(x, y)}{z}\left\{1 + O_\varepsilon\left(\frac{\log(u+1)}{\log y}\right)\right\}$$

holds uniformly for $\exp((\log_2 x)^{5/3+\varepsilon}) \leqslant y \leqslant x$ and $1 \leqslant z \leqslant y^{5/12}$. Thus, under our assumptions on z and y,

$$\begin{aligned}\Psi(x(1+z^{-1}), [y', y]) &- \Psi(x, [y', y]) \\ &= \frac{g_{P(y')}(\alpha)\Psi(x, y)}{z}\left\{1 + O\left(\frac{\log(u+1)}{\log y}\right)\right\} + O\left(\Psi(x, [y', y])\frac{(\log_2 x)^2}{\log y}\right).\end{aligned}$$

Another application of Lemma 3.3 to the right-hand side yields the desired result. □

4.2. *Equidistribution of weighted $[y', y]$-smooth numbers in short intervals.* The correction factors in Lemmas 4.1 and 4.2 suggest to consider the smoothly weighted version

$$n \mapsto n^{1-\alpha(x,y)}\mathbf{1}_{S([y',y])}(n)$$

of the characteristic function of $[y', y]$-smooth numbers. This choice of weight allows us to deduce from these lemmas that the new weighted function is equidistributed in all short intervals of length at least $x(\log x)^{-c}$, a property that is required in Section 7 and, indirectly, in Section 8. For simplicity, we further normalise our weighted function to mean value 1 in the following lemma.

Lemma 4.3 (equidistribution in short intervals). *Let $N > 2$ be a parameter and define the weighted and normalised function*

$$h_{[y',y]}(n) := \frac{N^\alpha}{\Psi(N, [y', y])}\frac{n^{1-\alpha}}{\alpha}\mathbf{1}_{S([y',y])}(n) \quad (N \leqslant n \leqslant 2N), \tag{4-7}$$

where $\alpha = \alpha(N, y)$. Let $y' \leqslant (\log N)^{K'}$ for some fixed $K' > 0$, let $K = \max(2, 2K')$, and suppose that $(\log N)^K < y \leqslant N$. Then

$$\sum_{N_0<n\leqslant N_0+N_1} h_{[y',y]}(n) = N_1\left\{1 + O\left(\frac{\log_3 N}{\log_2 N}\right)\right\} + O\left(\frac{N}{\log^{1/24} N}\right)$$

uniformly for all $N \leqslant N_0 < N_0 + N_1 \leqslant 2N$ such that $N_1 \gg N\exp(-(\log N)^{1/4}/4)$.

Proof. By partial summation,

$$\begin{aligned}\sum_{\substack{n\in S([y',y])\\ N_0<n\leqslant N_0+N_1}} n^{1-\alpha} &= (N_0+N_1)^{1-\alpha}\Psi(N_0+N_1, [y', y]) \\ &\quad - N_0^{1-\alpha}\Psi(N_0, [y', y]) - (1-\alpha)\int_{N_0}^{N_0+N_1}\frac{\Psi(t, [y', y])}{t^\alpha}\,\mathrm{d}t.\end{aligned} \tag{4-8}$$

We seek to bound the right-hand side with the help of Lemmas 4.1 and 4.2. To start with, suppose that $y < \exp((\log N)^{1/4})$. For $N \leqslant N_0 < t \leqslant N_0 + N_1 \leqslant 2N$, we have $t = N(1 + 1/z)$, where $z > 1$ and

$(t/N)^{-\alpha} \asymp 1$. Hence,

$$\Psi(t,[y',y]) = (t/N)^{\alpha}\Psi(N,[y',y])\biggl(1+O\biggl(\frac{\sqrt{\log y}}{z(\log N)^{1/6}}+O_A((\log N)^{-A})\biggr)\biggr)$$
$$= (t/N)^{\alpha}\Psi(N,[y',y])(1+O((\log N)^{-1/24}))$$

for $y < \exp((\log N)^{1/4})$ by Lemma 4.1. Thus,

$$\frac{N^{\alpha}}{\Psi(N,[y',y])}\sum_{\substack{n\in S([y',y])\\ N_0<n\leqslant N_0+N_1}} n^{1-\alpha} = (N_0+N_1)-N_0-(1-\alpha)\int_{N_0}^{N_0+N_1} 1\ \mathrm{d}t + O\biggl(\frac{N}{(\log N)^{1/24}}\biggr)$$
$$= \alpha N_1 + O\biggl(\frac{N}{(\log N)^{1/24}}\biggr),$$

which establishes the lemma when $y < \exp((\log N)^{1/4})$.

Suppose next that $\exp((\log N)^{1/4}) \leqslant y \leqslant N$. In view of the restriction on the size of z in Lemma 4.2, we cannot apply this lemma in the current situation with $x = N$ (which would correspond to the choice in the first part of the proof), but need to choose $x = N_0$ instead. For this purpose, let $\alpha' = \alpha(N_0, y)$ denote the saddle point associated to $S(N_0, y)$ and note that

$$|\alpha'-\alpha| = |\alpha(N_0,y)-\alpha(N,y)| \ll (\log y)^{-1} \ll (\log N)^{-1/4}$$

by Lemma 3.1 and the lower bound on y.

For $N \leqslant N_0 < t \leqslant N_0+N_1$ with $N\exp(-(\log N)^{1/4}/4) \leqslant N_1 \leqslant N$, we have

$$t = N_0(1+1/z), \quad \text{where } 1 < z \ll \exp((\log N)^{1/4}/4) = o(y^{5/12}),$$

and $(t/N_0)^{-\alpha'} \asymp 1$. Hence, Lemma 4.2 applies and yields

$$\Psi(t,[y',y]) = (t/N_0)^{\alpha'}\Psi(N_0,[y',y])\biggl(1+O\biggl(\frac{(\log_2 N)^2}{\log^{1/4} N}\biggr)\biggr). \tag{4-9}$$

Each of the three terms arising in the partial summation expression (4-8) involves a weighted count of the form $\Psi(t,[y',y])/t^{\alpha}$. With this and (4-9) in mind, observe that

$$(t/N_0)^{\alpha'}t^{-\alpha} = (t/N_0)^{\alpha'-\alpha}N_0^{-\alpha} = N_0^{-\alpha}(1+1/z)^{\alpha'-\alpha} = N_0^{-\alpha}(1+O(|\alpha-\alpha'|))$$
$$= N_0^{-\alpha}(1+O(\log N)^{-1/4}),$$

since

$$|(1+1/z)^a - 1^a| \leqslant \int_1^{1+1/z} |a|t^{a-1}\ \mathrm{d}t \leqslant \frac{|a|}{z} \leqslant |a|$$

for $0 < |a| < 1$, which we applied with $a = \alpha'-\alpha \ll (\log N)^{-1/4}$.

Hence, it follows from (4-8) and (4-9) that

$$\frac{N_0^{\alpha}}{\Psi(N_0,[y',y])}\sum_{\substack{n\in S([y',y])\\ N_0<n\leqslant N_0+N_1}} n^{1-\alpha} = (N_0+N_1)-N_0-(1-\alpha)\int_{N_0}^{N_0+N_1} 1\ \mathrm{d}t + O\biggl(\frac{N(\log_2 N)^2}{(\log N)^{1/4}}\biggr)$$
$$= \alpha(N,y)N_1 + O\biggl(\frac{N(\log_2 N)^2}{(\log N)^{1/4}}\biggr).$$

The lemma then follows since

$$\frac{(N_0/N)^{-\alpha}\Psi(N_0,[y',y])}{\Psi(N,[y',y])} = 1 + O\Big(\frac{\log_3 N}{\log_2 N}\Big)$$

by Lemma 3.3 and part (ii) of Lemma 3.4. □

Working with functions $h_{[y',y]} : [N, 2N] \to \mathbb{R}$ whose support is restricted to a dyadic interval is too restrictive for the purposes of our main theorem. The following lemma provides an intrinsically defined function $g_{[y',y]} : \mathbb{N} \to \mathbb{R}$ that approximates any function $h_{[y',y]}$ on the interval $[N, 2N]$ where the latter function is defined.

Lemma 4.4. *Define*

$$g_{[y',y]}(n) = \frac{n}{\alpha(n,y)\Psi(n,[y',y])}\mathbf{1}_{S([y',y])} \quad (n \in \mathbb{N}), \tag{4-10}$$

and suppose that $N \leqslant n < 2N$ *and that* $h_{[y',y]}$ *is defined on* $[N, 2N]$. *Then*

$$g_{[y',y]}(n) = h_{[y',y]}(n)\Big(1 + O\Big(\frac{\log_3 N}{\log_2 N}\Big)\Big),$$

provided that N *is sufficiently large.*

Proof. This follows immediately from Lemmas 3.1 and 3.4(ii). □

5. Major arc analysis: smooth numbers in arithmetic progressions and a W-trick

Extending previous work of Soundararajan [2008], Harper [2012b] proved that for $y \leqslant x$, $\varepsilon > 0$, $2 \leqslant q \leqslant y^{4\sqrt{e}-\varepsilon}$, and $(a,q) = 1$,

$$\Psi(x,y;q,a) := \sum_{\substack{n \leqslant x \\ n \equiv a \pmod q}} \mathbf{1}_{S(y)}(n) \sim \frac{1}{\phi(q)}\Psi_q(x,y)$$

as $\log x/\log q \to \infty$ provided that y is sufficiently large in terms of ε. This shows that, for $y \in (\log x, x]$, the y-smooth numbers up to x are equidistributed within the reduced residue classes to a given modulus q provided x is sufficiently large. Our aim in this section is to construct a subset of the y-smooth numbers with the property that its elements are equidistributed in the reduced residue classes of *any* modulus $q \leqslant (\log x)^c$. More precisely, we will show that, after applying a W-trick, the $[y', y]$-smooth numbers for any $y' \leqslant (\log N)^{K'}$, $K' > 0$, have this property. In addition, we will construct a suitable weight for this subset which allows us to retain equidistribution when we restrict any of the above progressions to a shorter interval of suitable length. These equidistribution properties will be required in order to reduce in Section 8 the task of establishing a noncorrelation estimate with nilsequences to the case where the nilsequence is highly equidistributed.

5.1. ***Smooth numbers in arithmetic progressions.*** In this subsection we extend [Harper 2012b, Theorem 1] and show that numbers without small and large prime factors are equidistributed in progressions. More precisely, defining

$$\Psi(x, [y', y]; q, a) = \sum_{\substack{n \leqslant x \\ n \equiv a \pmod q}} \mathbf{1}_{S([y',y])}(n),$$

we show the following:

Theorem 5.1. *Let $K', K'' > 0$ and $K > \max(2K', 2)$ be constants, let $x \geqslant 2$ be a parameter and let $1 \leqslant y' \leqslant (\log x)^{K'}$ and $(\log x)^K < y \leqslant x$. If $q < (\log x)^{K''}$, $p \mid q \Rightarrow p < y'$ and $(a, q) = 1$, then*

$$\Psi(x, [y', y]; q, a) = \frac{\Psi(x, [y', y])}{\phi(q)}(1 + O(\log^{-1/5} x)),$$

provided that K and K/K'' are sufficiently large.

Remark. The theorem generalises to the case where $q \leqslant (\log x)^{K''}$ but $p \mid q \not\Rightarrow p < y'$. In this case, $\Psi(x, [y', y])$ needs to be replaced by $\Psi_q(x, [y', y])$.

As we require explicit error terms, we do not follow the proof strategy from [Harper 2012b], but instead start by establishing a version of the Perron-type bound given in [Harper 2012a, Proposition 1] that applies to

$$\Psi(x, [y', y], \chi) := \sum_{n \leqslant x} \chi(n) \mathbf{1}_{S([y',y])}.$$

Proposition 5.2. *There exist a small absolute constant $\varrho \in (0, 1)$ and a large absolute constant $C > 1$ such that the following is true. Let $K' > 0$ and $K > \max(2K', 2)$ be constants, let $1 \leqslant y' \leqslant (\log x)^{K'}$ and $(\log x)^K < y \leqslant x$, and suppose that x is large. Suppose that χ is a nonprincipal Dirichlet character with conductor $r \leqslant x^{\varrho}$, and to modulus $q \leqslant x$, such that $L(s, \chi) = \sum_n \chi(n) n^{-s}$ has no zeros in the region*

$$\Re(s) > 1 - \varepsilon, \quad |\Im(s)| \leqslant H,$$

where $C/\log y < \varepsilon \leqslant \min(\alpha/2, 1/(2K'))$ and $y^{0.9\varepsilon} \log^2 x \leqslant H \leqslant \min(x^{\varrho}, y^{5/6})$. Suppose, moreover, that at least one of the following holds:

(i) $y \geqslant (Hr)^C$.

(ii) $\varepsilon \geqslant 40 \log\log(qyH)/\log y$.

Then we have the bound

$$\Psi(x, \chi; [y', y]) \ll \Psi(x, [y', y])\big((\log x)^{-1/5} + \sqrt{\log x \log y}(H^{-1/2} + x^{-0.4\varepsilon} \log H)\big).$$

Remark 5.3. The $(\log x)^{-1/5}$ term in the bound can be omitted when $y < \exp(\log^{1/4} x)$, i.e., when the proof involves an application of Lemma 4.1 and avoids Lemma 4.2.

The proof of Proposition 5.2 presented below follows the proof of Harper's original result [2012a, Proposition 1] very closely. It involves an application of Perron's formula to relate $\Psi(x, [y', y]; \chi)$ to its Dirichlet L-function, defined as

$$L(s, \chi; [y', y]) = \sum_{n \in S([y',y])} \frac{\chi(n)}{n^s} = \prod_{y' \leqslant p \leqslant y} (1 - \chi(p) p^{-s})^{-1} \quad (\Re s > 0),$$

which is followed by a contour shift in the resulting contour integral. The following lemma will be used in order to estimate the integrals on the new contour.

Lemma 5.4. *Suppose that $\alpha = \alpha(x, y)$ is the saddle point associated to $S(x, y)$. Then under the assumptions of the proposition, we have*

$$\bigl|\log L(\sigma + it, \chi; [y', y]) - \log L(\alpha + it, \chi; [y', y])\bigr| \leqslant \frac{(\alpha - \sigma)\log x}{2}$$

for all $\alpha - 0.8\varepsilon \leqslant \sigma \leqslant \alpha$ and $|t| \leqslant H/2$.

Before proving this lemma, we complete the proof of the proposition.

Proof of Proposition 5.2 assuming Lemma 5.4. The truncated version of Perron's formula as given in [Montgomery and Vaughan 2007, Theorems 5.2 and 5.3] yields

$$\Psi(x, [y', y]; \chi) = \frac{1}{2\pi i}\int_{\alpha - iH/2}^{\alpha + iH/2} L(s, \chi; [y', y])\frac{x^s}{s}\,\mathrm{d}s + O\biggl(\frac{x^\alpha}{H}\sum_{n\in S([y',y])}\frac{\chi_0(n)}{n^\alpha} + \sum_{\substack{x/2<n<2x\\ n\in S([y',y])}}\chi_0(n)\min\Bigl\{1, \frac{x}{H|x-n|}\Bigr\}\biggr), \tag{5-1}$$

where $\alpha = \alpha(x, y)$. Since $x/(H|x-n|) \leqslant H^{-1/2}$ whenever $|x-n| > xH^{-1/2}$, we have

$$\min\Bigl\{1, \frac{x}{H|x-n|}\Bigr\} \leqslant \begin{cases} 1 & \text{if } |x-n| \leqslant xH^{-1/2}, \\ H^{-1/2} & \text{if } |x-n| > xH^{-1/2}. \end{cases}$$

Thus, the second error term in (5-1) is bounded above by

$$\sum_{\substack{n\in S([y',y])\\ x/2<n<2x}} H^{-1/2}\chi_0(n) + \sum_{\substack{|x-n|\leqslant xH^{-1/2}\\ n\in S([y',y])}} 1 \ll \frac{x^\alpha}{H^{1/2}}L(\alpha, \chi_0; [y', y]) + \bigl|\Psi(x + xH^{-1/2}, [y', y]) - \Psi(x - xH^{-1/2}, [y', y])\bigr|.$$

Since the first error term in (5-1) is smaller than the first term in the preceding line, it suffices to estimate the two terms in the bound above in order to bound the error in (5-1).

Concerning the first term, Lemma 3.3, combined with formulas (3-3) and (3-5), shows that

$$x^\alpha L(\alpha, \chi_0; [y', y]) \ll x^\alpha \sum_{n\in S([y',y])} n^{-\alpha} \ll x^\alpha \zeta(\alpha, y)\prod_{p<y'}(1 - p^{-\alpha}) \ll \Psi(x, [y', y])\sqrt{\log x \log y}. \tag{5-2}$$

Concerning the second term, applying Lemma 4.1 if $y < \exp(\log^{1/4} x)$ and Lemma 4.2 if $\exp(\log^{1/4} x) \leqslant y \leqslant x$ shows that

$$\begin{aligned}\Psi(x(1 + H^{-1/2}), [y', y]) - \Psi(x(1 - H^{-1/2}), [y', y]) &\ll \bigl((1 + H^{-1/2})^\alpha - (1 - H^{-1/2})^\alpha + H^{-1/2} + (\log x)^{-1/5}\bigr)\Psi(x, [y', y]) \\ &\ll (H^{-1/2} + (\log x)^{-1/5})\Psi(x, [y', y]),\end{aligned}$$

provided that $H \leqslant y^{5/6}$ so that $z = H^{1/2} \leqslant y^{5/12}$ in the application of Lemma 4.2, and where we used an integral to bound the difference. This shows that

$$\Psi(x, [y', y]; \chi) = \frac{1}{2\pi i} \int_{\alpha - iH/2}^{\alpha + iH/2} L(s, \chi; [y', y]) \frac{x^s}{s} \,\mathrm{d}s \\ + O\big(\Psi(x, [y', y])\big\{(\log x)^{-1/5} + H^{-1/2}\sqrt{\log x \log y}\big\}\big). \tag{5-3}$$

Since $L(s, \chi; [y', y])$ is defined and has no singularities in $\Re s > 0$, shifting the line of integration to $\Re s = \alpha - 0.8\varepsilon$ shows that the integral above is equal to

$$\int_{-H/2}^{H/2} L(\alpha - 0.8\varepsilon + it, \chi; [y', y]) \frac{x^{\alpha - 0.8\varepsilon + it}}{\alpha - 0.8\varepsilon + it} \,\mathrm{d}t \pm \int_{\alpha - 0.8\varepsilon}^{\alpha} L\Big(\sigma \pm \frac{iH}{2}, \chi; [y', y]\Big) \frac{x^{\sigma \pm iH/2}}{\sigma \pm \frac{iH}{2}} \,\mathrm{d}\sigma,$$

where the final expression indicates the sum over the two horizontal pieces. Multiplying each of their integrands by a trivial factor of the form $L(\alpha \pm iH/2, \chi; [y', y])^{+1-1}$ and applying Lemma 5.4, each of the horizontal integrals is seen to be bounded by

$$\frac{L(\alpha, \chi_0; [y', y])}{H} \int_{\alpha - 0.8\varepsilon}^{\alpha} x^{(\alpha - \sigma)/2} x^{\sigma} \,\mathrm{d}\sigma \ll \frac{x^{\alpha} L(\alpha, \chi_0; [y', y])}{H}.$$

A similar argument shows that the vertical integral is bounded by

$$L(\alpha, \chi_0; [y', y]) x^{\alpha - 0.8\varepsilon + 0.4\varepsilon} \int_{-H/2}^{H/2} \frac{\mathrm{d}t}{\alpha + |t|} \ll x^{\alpha} L(\alpha, \chi_0; [y', y]) x^{-0.4\varepsilon} \Big(\frac{1}{\alpha} + \log H\Big) \\ \ll \sqrt{\log x \log y}\, \Psi(x, [y', y]) x^{-0.4\varepsilon} \log H,$$

where we applied (5-2). Putting everything together,

$$\Psi(x, \chi; [y', y]) \ll \Psi(x, [y', y])\big((\log x)^{-1/5} + \sqrt{\log x \log y}(H^{-1/2} + x^{-0.4\varepsilon} \log H)\big),$$

as required. It remains to prove the lemma. □

Proof of Lemma 5.4. Reinterpreting the given difference as an integral shows that it is bounded above by

$$(\alpha - \sigma) \sup_{\sigma \leqslant \sigma' \leqslant \alpha} \left| \frac{L'(\sigma' + it, \chi; [y', y])}{L(\sigma' + it, \chi; [y', y])} \right| = (\alpha - \sigma) \sup_{\sigma \leqslant \sigma' \leqslant \alpha} \left| \sum_{n \in S([y', y])} \frac{\Lambda(n)\chi(n)}{n^{\sigma' + it}} \right|.$$

We may replace the summation condition in the final sum by $n \in [y', y]$ when bounding the contribution from proper prime powers $p^k \in [y', y]$ with $p < y'$ and $p^k > y$ with $p \leqslant y$ separately. Bounding this contribution from proper prime powers trivially, the expression above is seen to be

$$\ll (\alpha - \sigma) \sup_{\sigma \leqslant \sigma' \leqslant \alpha} \left| \sum_{y' \leqslant p \leqslant y} \frac{\log p\, \chi(n)}{p^{\sigma' + it}} \right| + (\alpha - \sigma) \sum_{p \leqslant y} \frac{\log p}{p^{2(\alpha - 0.8\varepsilon)}} \\ \ll (\alpha - \sigma) \left(\sup_{\alpha - 0.8\varepsilon \leqslant \sigma' \leqslant \alpha} \left| \sum_{n \leqslant y} \frac{\Lambda(n)\chi(n)}{n^{\sigma' + it}} \right| + \sup_{\alpha - 0.8\varepsilon \leqslant \sigma' \leqslant \alpha} \left| \sum_{n \leqslant y'} \frac{\Lambda(n)\chi(n)}{n^{\sigma' + it}} \right| + \frac{y^{1 - \alpha - 0.1\varepsilon}}{1 - \alpha} + \frac{1}{\varepsilon} \right),$$

as $2(\alpha - 0.8\varepsilon) \geqslant \alpha + 0.4\varepsilon$ if $\varepsilon \leqslant \alpha/2$. In the case that condition (i) of Proposition 5.2 holds, that is $y \geqslant (Hr)^C$, it follows from [Harper 2012a, Lemma 1] that

$$\sup_{\alpha-0.8\varepsilon\leqslant\sigma'\leqslant\alpha}\left|\sum_{n\leqslant y}\frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}}\right| \ll \frac{y^{1-\alpha-0.1\varepsilon}}{1-\alpha} + \log(rH) + \log^{0.9} q + \frac{1}{\varepsilon}$$

since $\log^2(qyH)/H \ll y^{-0.9\varepsilon}$ whenever $y^{0.9\varepsilon}\log^2 x \leqslant H$. Otherwise, condition (ii) of Proposition 5.2 holds by assumption, and we have $\log^2(qyH) < y^{2\varepsilon/40} = y^{0.05\varepsilon}$. Thus it follows from [Harper 2012a, Lemma 2] that the previous estimate holds in this case as well.

Concerning the sum over $n \leqslant y'$, we shall show that, for all sufficiently large x,

$$\sup_{\alpha-0.8\varepsilon\leqslant\sigma'\leqslant\alpha}\left|\sum_{n\leqslant y'}\frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}}\right| < \frac{\log x}{4}.$$

To see this, note that if $1 - 1/\log y' < \sigma' \leqslant 1$, then

$$\left|\sum_{n\leqslant y'}\frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}}\right| \leqslant \sum_{n\leqslant y'}\frac{\Lambda(n)}{n^{\sigma'}} \ll \log y' \ll K'\log_2 x = o(\log x).$$

If $\alpha - 0.8\varepsilon \leqslant \sigma' < 1 - 1/\log y'$, then partial summation and the prime number theorem imply that

$$\left|\sum_{n\leqslant y'}\frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}}\right| \ll \frac{y'^{1-\sigma'}}{1-\sigma'} \ll y'^{1-\sigma'}\log y' \ll y'^{1-\alpha}y'^{0.8\varepsilon}\log y' \ll y^{(1-\alpha)/2}(\log x)^{0.4}\log y',$$

where we also used that $\varepsilon \leqslant 1/(2K')$ and $y' < y^{1/2}$. On recalling the estimate (3-2), that is,

$$1 - \alpha = \frac{\log(u\log(u+1)) + O(1)}{\log y},$$

and noting that $\log(u+1) \ll \log_2 x$, the above is seen to be bounded by

$$\ll (u\log(u+1))^{1/2}(\log x)^{0.4}\log y' \ll \frac{K'(\log x)^{9/10}(\log_2 x)^{3/2}}{(\log y)^{1/2}} = o(\log x),$$

as required.

Collecting everything together, we conclude that

$$\begin{aligned}&\left|\log L(\sigma+it,\chi;[y',y]) - \log L(\alpha+it,\chi;[y',y])\right| \\ &\qquad \ll (\alpha-\sigma)\Big(\Big(y^{-0.1\varepsilon}+\frac{1}{4}\Big)\log x + \log(rH) + \log^{0.9} q + \frac{1}{\varepsilon}\Big).\end{aligned}$$

Since $\varepsilon > C/\log y$, $q < x$ and $r, H \leqslant x^{\varrho}$, we obtain

$$\begin{aligned}&\left|\log L(\sigma+it,\chi;[y',y]) - \log L(\alpha+it,\chi;[y',y])\right| \\ &\qquad \ll (\alpha-\sigma)\big((e^{-0.1C} + 4^{-1} + 2\varrho + C^{-1})\log x + \log^{0.9} x\big).\end{aligned}$$

The claimed bound thus follows provided $C > 1$ is sufficiently large, $0 < \varrho < 1$ sufficiently small and x is sufficiently large. □

Proof of Theorem 5.1. Since $(a, q) = 1$, character orthogonality implies that

$$\begin{aligned}\Psi(x, [y', y]; q, a) &= \frac{1}{\phi(q)} \sum_{\chi \pmod q} \bar\chi(a) \sum_{n \in S(x,[y',y])} \chi(n)\\ &= \frac{1}{\phi(q)} \Psi(x, [y', y]; \chi_0) + \frac{1}{\phi(q)} \sum_{\substack{\chi \pmod q\\ \chi \neq \chi_0}} \bar\chi(a) \Psi(x, [y', y]; \chi).\end{aligned}$$

If $p \mid q \Rightarrow p < y'$, then

$$\Psi(x, [y', y], \chi_0) = \sum_{n \leqslant x,\, (n,q)=1} 1_{S([y',y])}(n) = \Psi(x, [y', y]).$$

To complete the proof, it thus suffices to show that

$$\Psi(x, [y', y]; \chi) \ll \Psi(x, [y', y]) \log^{-1/5} x$$

uniformly for all nonprincipal characters $\chi \pmod q$.

The classical zero-free region (cf. [Montgomery and Vaughan 2007, Theorem 11.3]) implies that there is an absolute constant $0 < \kappa \leqslant 1$ such that $\prod_{\chi \neq \chi_0} L(\sigma + it, \chi)$ has at most one, necessarily simple, zero in the region

$$\{s : \sigma > 1 - \kappa/\log(qH), |t| < H\}.$$

If such an exceptional zero exists, let χ_{Siegel} denote the corresponding character. We consider the cases $\chi \neq \chi_{\mathrm{Siegel}}$ and $\chi = \chi_{\mathrm{Siegel}}$ separately.

Suppose first that $\chi \neq \chi_{\mathrm{Siegel}}$ and let $r \leqslant q \leqslant (\log x)^{K''}$ denote the conductor of χ. We claim that the conditions of Proposition 5.2 are satisfied if we choose

$$H = \min\{y^{\kappa/(2C)}, \exp(\log^{2/3} x), y^{5/6}, x^{\varrho}\} \quad \text{and} \quad \varepsilon = \kappa/\log(qH).$$

Suppose that x is sufficiently large to ensure that $\exp(\log^{2/3} x) < x^{\varrho}$ and suppose that $C > 1 \geqslant \kappa$. If $K > 2CK''/\kappa$ (or, in other words, if K/K'' is sufficiently large), then

$$q \leqslant (\log x)^{K''} < (\log x)^{\kappa K/(2C)} < y^{\kappa/(2C)}.$$

We thus have $qH < y^{\kappa/C}$ and, hence, $\varepsilon = \kappa/\log(qH) > C/\log y$. The upper bound $\varepsilon < \min(\alpha/2, 1/(2K''))$ holds as soon as x, and hence H, is sufficiently large. Concerning the conditions on H, suppose first that $H = \exp(\log^{2/3} x)$. In this case, $\varepsilon \asymp \log^{-2/3} x$ and $y^{0.9\varepsilon} \ll \exp(\log^{1/3} x)$. Hence, $H > y^{0.9\varepsilon} \log^2 x$ holds as soon as x is sufficiently large. Next, suppose that $H \neq \exp(\log^{2/3} x)$. In this case, $\min(y^{\kappa/(2C)}, y^{5/6}) \leqslant \exp(\log^{2/3} x)$ and $\varepsilon \asymp 1/\log y \gg 1/\log^{2/3} x$. From $\varepsilon \asymp 1/\log y$ we obtain $y^{0.9\varepsilon} \ll 1$, and thus $H > y^{0.9\varepsilon} \log^2 x$ holds provided that the exponent K in $y > (\log x)^K$ satisfies $\min(5K/6, K\kappa/(2C)) > 2$ and provided that x is sufficiently large. With the below application of Proposition 5.2 in mind, observe that the lower bound $\varepsilon \gg 1/\log^{2/3} x$, which holds in either of the above two cases, implies that

$$x^{-0.4\varepsilon} \ll \exp(-c \log^{1/3} x) \quad (c > 0).$$

Finally, observe that condition (i) of Proposition 5.2 is satisfied since $rH \leqslant qH \leqslant y^{\kappa/C} \leqslant y^{1/C}$ in view of $\kappa \leqslant 1$.

This shows that Proposition 5.2 applies to all $\chi \pmod q$, $\chi \neq \chi_{\mathrm{Siegel}}$ and yields

$$\begin{aligned}\Psi(x,[y',y];\chi) &\ll \Psi(x,[y',y])\big((\log x)^{-1/5}+\sqrt{\log x\log y}(H^{-1/2}+x^{-0.4\varepsilon}\log H)\big)\\ &\ll \Psi(x,[y',y])(\log x)^{-1/5}.\end{aligned}$$

To see that the latter bound holds, we shall show that the second and third terms in the bound above make a negligible contribution. Indeed, the third term is bounded as follows:

$$\begin{aligned}\sqrt{\log x\log y}\,x^{-0.4\varepsilon}\log H &\ll \sqrt{\log x\log y}\exp(-c\log^{1/3}x)(\log y+\log^{2/3}x)\\ &\ll \exp(-c'\log^{1/3}x),\end{aligned}$$

where $c, c' > 0$ are positive constants. Further, if $H=\exp(\log^{2/3}x)$, then the second term is bounded by

$$\sqrt{\log x\log y}\,H^{-1/2} \ll \exp(-(\log x)^{2/3}/2+\log_2 x),$$

which is negligible. If $H\neq\exp(\log^{2/3}x)$, then $\log y\ll\log^{2/3}x$ and $\varepsilon\asymp 1/(\log y)$, and the second term satisfies

$$\sqrt{\log x\log y}\,H^{-1/2}\ll\sqrt{\log x\log y}\,\frac{y^{-9\varepsilon/20}}{\log x}\ll\left(\frac{\log y}{\log x}\right)^{1/2}\ll(\log x)^{-1/3},$$

which is also negligible. This concludes the case of unexceptional characters.

We now consider the contribution of the potential exceptional character $\chi=\chi_{\mathrm{Siegel}}$. Following [Harper 2012a, Section 3.3], we split the analysis into two cases according to the size of y. Suppose first that $y'^2\leqslant y\leqslant x^{1/(\log\log x)^2}$. Applying the truncated Perron formula (5-3) with $H=y^{5/6}$, we obtain

$$\begin{aligned}\Psi(x,[y',y];\chi)=\frac{1}{2\pi i}\int_{\alpha-iy^{5/6}/2}^{\alpha+iy^{5/6}/2}L(s,\chi;[y',y])x^s\,\frac{\mathrm{d}s}{s}\\ +O\big(\Psi(x,[y',y])\big\{(\log x)^{-1/5}+y^{-5/12}\sqrt{\log x\log y}\big\}\big).\end{aligned}\tag{5-4}$$

As in [Harper 2012a], we proceed by bounding the integrand in absolute value from above. The argument used in Section 3.3 of that work is partly based on ideas from [Soundararajan 2008]. In our case, only small modifications are required. To start with, we have

$$\begin{aligned}\left|\frac{L(\alpha+it,\chi;[y',y])}{L(\alpha,\chi_0;[y',y])}\right| &= \prod_{y'\leqslant p\leqslant y}\left|\frac{1-\chi(p)p^{-\alpha-it}}{1-\chi_0(p)p^{-\alpha}}\right|^{-1}\leqslant\prod_{\substack{y'\leqslant p\leqslant y\\ p\nmid q}}\left|1+\sum_{k\geqslant1}\frac{1-\Re(\chi(p)p^{-it})}{p^{k\alpha}}\right|^{-1}\\ &\leqslant\exp\Big\{-\sum_{\substack{y'\leqslant p\leqslant y\\ p\nmid q}}\frac{1-\Re(\chi(p)p^{-it})}{p^\alpha}\Big\}\leqslant\exp\Big\{-\sum_{\substack{y^{1/2}\leqslant p\leqslant y\\ p\nmid q}}\frac{1-\Re(\chi(p)p^{-it})}{p^\alpha}\Big\},\end{aligned}$$

since $y'^2<y$. The next step is to show that in the final expression the argument of the exponential function satisfies

$$\sum_{\substack{y^{1/2}\leqslant p\leqslant y\\ p\nmid q}}\frac{1-\Re\chi(p)p^{-it}}{p^\alpha}\gg\frac{u}{\log^2(u+1)},$$

provided that $|t| \leqslant y/2$ and $y \leqslant x^{1/(\log\log x)^2}$. The proof of this bound proceeds by splitting the range of t into $|t| \leqslant 1/(2\log y)$ and $1/(2\log y) \leqslant |t| \leqslant y/2$. On the latter range, the proof of the second part of [Soundararajan 2008, Lemma 5.2] is employed. In that proof, all sums over primes can be restricted to the range $y^{1/2} \leqslant p \leqslant y,\ p \nmid q$. On the former range, Harper's argument [2012a, p. 17] applies directly. For sake of completeness, we summarise this argument here. Since $p \leqslant y$ and $t \leqslant 1/(2\log y)$, we have $|t \log p| \leqslant \frac{1}{2}$ and, by analysing the cases $\chi(p) = 1$ and $\chi(p) = -1$ for the quadratic character $\chi = \chi_{\mathrm{Siegel}}$ separately, one obtains

$$\begin{aligned}\sum_{y^{1/2}\leqslant p\leqslant y,\, p\nmid q} \frac{1-\Re\chi(p)p^{-it}}{p^{\alpha}} &\gg \sum_{y^{1/2}\leqslant p\leqslant y,\, p\nmid q} \frac{1-\chi(p)}{p^{\alpha}} \\ &\geqslant \frac{1}{\log y}\biggl(\sum_{y^{1/2}\leqslant p\leqslant y,\, p\nmid q} \frac{\log p}{p^{\alpha}} - \sum_{y^{1/2}\leqslant p\leqslant y,\, p\nmid q} \frac{\chi^*_{\mathrm{Siegel}}(p)\log p}{p^{\alpha}}\biggr),\end{aligned}$$

where χ^*_{Siegel} is the primitive character that induces χ_{Siegel}. Harper's argument is based on a combination of partial summation and the asymptotic evaluation of $\sum_{n\leqslant x} \Lambda(n)\chi(n)$ together with the observation that a Siegel zero has a negative contribution in that asymptotic evaluation. The error terms in the asymptotic evaluation in [Montgomery and Vaughan 2007, Theorem 11.16 and Exercise 11.3.1.2] are small when $q < (\log x)^{K''}$. Thanks to the minus sign in the final expression above, the contribution from the Siegel zero to this expression is positive and can therefore be ignored when seeking a lower bound. See [Harper 2012a, Section 3.3] for the remaining details.

Returning to (5-4), the above bounds and an application of (5-2) yield

$$\begin{aligned}\Psi(x,[y',y];\chi) &\ll L(\alpha,\chi_0;[y',y])e^{-c_3 u(\log u)^{-2}}x^{\alpha}\log y + \frac{\Psi(x,[y',y])}{\log^{1/5} x} \\ &\ll \Psi(x,[y',y])\bigl(\sqrt{\log x}(\log y)^{3/2}e^{-c_3 u(\log u)^{-2}} + \log^{-1/5} x\bigr).\end{aligned}$$

Since $u = (\log x)/\log y \gg \log\log^2 x$ when $y \leqslant \exp((\log x)/\log\log^2 x)$, the first term in the bound is $O_A(\Psi(x,[y',y])(\log x)^{-A})$ and we thus have

$$\Psi(x,[y',y];\chi) \ll (\log x)^{-1/5}\Psi(x,[y',y])$$

if $\chi = \chi_{\mathrm{Siegel}}$ and $y \leqslant x^{1/(\log\log x)^2}$.

It remains to analyse the range $x^{1/(\log\log x)^2} \leqslant y \leqslant x$ of y. In this case too, we follow the overall strategy used in [Harper 2012a] but need to make a number of technical changes arising from our slightly different situation. Observe that

$$\Psi(x,[y',y];\chi) = \sum_{n\in S(x,[y',y])} \chi(n) = \sum_{\substack{n\in S(x,y)\\ (n,P(y'))=1}} \chi(n) \leqslant \sum_{\substack{d\leqslant x\\ d\mid P(y')}} \biggl|\sum_{n\in S(x/d,y)} \chi(n)\biggr|,$$

where we applied Möbius inversion in order to remove the coprimality condition. Our next aim is to show that, at the expense of an acceptable error term, the sum over d can be further truncated in such a way that $\log(x/d) \asymp \log x$ uniformly for all values of d that remain in the sum. Recall that $\alpha(x,y) > 1-1/K+o(1)$

and that $\Psi(x, y') \leqslant x^{1-1/K'+o(1)}$. Let $K^* > K$ be sufficiently large so that $\alpha(x, y)(1-1/K^*) > 1-1/K'$. Then the $m = 1$ case of [de la Bretèche and Tenenbaum 2005, Theorem 2.4(i)], which applies uniformly for $1 \leqslant d \leqslant x$, implies that

$$\sum_{\substack{x^{1-1/K^*} \leqslant d \leqslant x \\ d \mid P(y')}} \Psi(x/d, y) \ll \Psi(x, y) x^{-(1-1/K^*)\alpha} \sum_{x^{1-1/K^*} \leqslant d \leqslant x} \mathbf{1}_{S(y')}(d) \\ \ll \Psi(x, y) x^{-(1-1/K^*)\alpha} x^{1-1/K'+o(1)} \ll \Psi(x, y) x^{-c_4} \quad (5\text{-}5)$$

for some positive constant $c_4 > 0$. Assuming for the moment that this bound is sufficiently good, it remains to analyse the expression

$$\sum_{\substack{d \leqslant x^{1-1/K^*} \\ d \mid P(y')}} \Bigg| \sum_{n \in S(x/d, y)} \chi(n) \Bigg|,$$

where $\log(x/d) \asymp \log x$. Fouvry and Tenenbaum [1996, Lemme 2.1] (see also their Lemme 2.2 and its deduction) studied bounds on character sums such as the inner sum above in a larger range for y than the one under current consideration. In view of the truncation of the sum over d and the bound $q \leqslant (\log x)^{K''}$ on the modulus of χ, [loc. cit., Lemme 2.1(i)] applies uniformly to the above character sum for each d. By invoking [de la Bretèche and Tenenbaum 2005, Theorem 2.4(i)] in the case $m = 1$, we obtain

$$\sum_{\substack{d \leqslant x^{1-1/K^*} \\ d \mid P(y')}} \Bigg| \sum_{n \in S(x/d, y)} \chi(n) \Bigg| \ll \sum_{\substack{d \leqslant x^{1-1/K^*} \\ d \mid P(y')}} d^{-\alpha} \Psi(x, y) \exp\bigl(-c\sqrt{\log y}\bigr) \\ \ll \exp\bigl(-c\sqrt{\log y}\bigr) \Psi(x, y) \prod_{p \leqslant y'} (1 - p^{-\alpha})^{-1}. \quad (5\text{-}6)$$

It remains to express the bounds (5-6) and (5-5) in terms of $\Psi(x, [y', y])$. By Lemma 3.3 and the first part of Lemma 3.2, we have

$$\Psi(x, y) \ll \Psi(x, [y', y]) \prod_{p \leqslant y'} (1 - p^{-\alpha})^{-1}.$$

Since $\log y' \asymp \log_2 x$ and $1 - \alpha \ll (\log_2 x)^{2+\varepsilon}/\log x$ for $x^{1/(\log\log x)^2} \leqslant y \leqslant x$, it follows from the first part of Lemma 3.2 (or direct computation) that $\prod_{p \leqslant y'} (1 - p^{-\alpha})^{-1} \ll \log y'$. Hence, it follows from (5-6) and (5-5) that

$$\Psi(x, [y', y]; \chi) \ll_A (\log x)^{-A} \Psi(x, [y', y])$$

if $\chi = \chi_{\mathrm{Siegel}}$ and $x^{1/(\log\log x)^2} \leqslant y \leqslant x$. Taking $A = \frac{1}{5}$ completes the proof. □

5.2. ***W-trick and equidistribution of weighted smooth numbers in short APs.*** The choice of our weight factor $n^{1-\alpha}$ ensures that the weighted version $h_{[y',y]}(n)$ of $\mathbf{1}_{S([y',y])}$ is equidistributed in short intervals and, by invoking Theorem 5.1, also in short progressions $n \equiv a \pmod q$, provided the residue class a is coprime to the modulus q of the progression and $q < y'$. Establishing noncorrelation with nilsequences later requires equidistribution in almost all residue classes $a \pmod q$. Given any $Q \in \mathbb{N}$ and $1 \leqslant A < Q$,

$\gcd(Q, A) = 1$, we consider the following renormalised restrictions:

$$h^{(Q,A)}_{[y',y]}(m) = \frac{\phi(Q)}{Q} h_{[y',y]}(Qm + A) \quad (m \sim (N - A)/Q), \tag{5-7}$$

$$g^{(Q,A)}_{[y',y]}(m) = \frac{\phi(Q)}{Q} g_{[y',y]}(Qm + A) \quad (m \in \mathbb{N}). \tag{5-8}$$

Let $w(N) = \frac{1}{2}\log_3 N$, let $W = W(N) = \prod_{p \leqslant w(N)} p$ and observe for later use that $w(N) = (\log_3 N')/2 + o(1)$ for any $N' \in (N/\log N, N]$. We will mainly be interested in the case where

- $Q = W(N)$, or
- $Q = \widetilde{W}$ for some $\widetilde{W} \leqslant (\log N)^{K''}$ such that $W(N) \mid \widetilde{W}$ and $p \mid \widetilde{W} \Rightarrow p \leqslant w(N)$

in (5-7) and (5-8).

The following lemma is an immediate consequence of Theorem 5.1 and Lemma 4.3.

Lemma 5.5 (equidistribution in short progression). *Let K', $K'' > 0$ and $K > \max(2K', 2)$ be constants and let $N \geqslant 2$ be a parameter. Suppose that $y' \leqslant (\log N)^{K'}$, and that $(\log N)^K < y \leqslant N$. Then, provided that K and K/K'' are sufficiently large, the following estimate holds uniformly for all N_0, N_1 such that*

$$N < N_0 < N_0 + N_1 \leqslant 2N$$

and $N_1 \gg N \exp(-(\log^{1/4} N)/4)$, for all $q \leqslant (\log N)^{K''}$ such that $p \mid q \Rightarrow p < y'$ and all $a \pmod q$ such that $\gcd(a, q) = 1$:

$$\sum_{\substack{N_0 < m \leqslant N_0 + N_1 \\ m \equiv a \ (\mathrm{mod}\ q)}} h_{[y',y]}(m) - \frac{N_1}{\phi(q)} \ll \frac{N_1}{\phi(q)} \frac{\log_3 N}{\log_2 N} + \frac{N}{\phi(q)\log^{1/24} N} + \frac{N}{\log^{1/5} N}.$$

Proof. Recall the definition (4-7) of $h_{[y',y]}(n)$. On omitting the weight factor $N^\alpha/(\alpha\Psi(N, [y', y]))$ for the moment, partial summation yields

$$\begin{aligned}&\sum_{N_0 < n \leqslant N_0 + N_1} n^{1-\alpha} \mathbf{1}_{n \equiv a \ (\mathrm{mod}\ q)} 1_{S([y',y])}(n) \\ &= (N_0 + N_1)^{1-\alpha}\Psi(N_0 + N_1, [y', y]; q, a) - (N_0)^{1-\alpha}\Psi(N_0, [y', y]; q, a) - (1-\alpha)\int_{N_0}^{N_0+N_1} \frac{\Psi(t, [y', y]; q, a)}{t^\alpha}\, dt.\end{aligned}$$

Theorem 5.1 implies

$$\Psi(t, [y', y]; q, a) = \frac{\Psi(t, [y', y])}{\phi(q)}(1 + O(\log^{-1/5} N)) \tag{5-9}$$

whenever $N \leqslant t \leqslant 2N$ and provided that K is sufficiently large. Inserting this expansion into the previous expression and recalling the partial summation application (4-8), we see that the contribution from the main term in (5-9) can be reinterpreted as a sum over $n^{1-\alpha}\mathbf{1}_{S([y',y])}$ that is not restricted to a progression. More precisely,

$$\begin{aligned}\phi(q) &\sum_{N_0 < n \leqslant N_0 + N_1} n^{1-\alpha} \mathbf{1}_{n \equiv a \ (\mathrm{mod}\ q)} \mathbf{1}_{S([y',y])}(n) \\ &= \sum_{N_0 < n \leqslant N_0 + N_1} n^{1-\alpha} \mathbf{1}_{S([y',y])}(n) + O\left(\frac{N^{1-\alpha}\Psi(N, [y', y])}{\log^{1/5} N}\right).\end{aligned} \tag{5-10}$$

The shape of the error term above follows by several applications of Lemmas 3.4(i) and 3.3. On reinserting the weights into (5-10) and invoking Lemma 4.3, we obtain

$$\sum_{\substack{N_0 \leqslant m \leqslant N_0+N_1 \\ m \equiv a \pmod q}} h_{[y',y]}(m) = \frac{1}{\phi(q)} \sum_{N_0 \leqslant n \leqslant N_0+N_1} h_{[y',y]}(n) + O\left(\frac{N}{\log^{1/5} N}\right)$$
$$= \frac{N_1}{\phi(q)}\left(1 + O\left(\frac{\log_3 N}{\log_2 N}\right)\right) + O\left(\frac{N}{\phi(q)(\log N)^{1/24}}\right) + O\left(\frac{N}{\log^{1/5} N}\right)$$

provided that $N_1 \gg N \exp(-(\log N)^{1/4}/4)$. □

By analysing the asymptotic order of the weight factor in $h_{[y',y]}$, we may deduce from the previous result an upper bound on the unweighted count of $[y', y]$-smooth numbers in short *intervals*. This bound will be integral in the application of an bootstrapping argument in Section 7.

Corollary 5.6 ($[y', y]$-smooth numbers in short intervals). *Let $K', K'' > 0$ and $K > \max(2K', 2)$ be constants and let $N \geqslant 2$ be a parameter. Suppose that $1 \leqslant y' \leqslant (\log N)^{K'}$ and that $(\log N)^K < y \leqslant N$. Let $P \subseteq [N, 2N]$ be any progression of length $|P| \geqslant N \exp(-(\log^{1/4} N)/4)$ and common difference $q \leqslant (\log N)^{K''}$ such that $p \mid q \Rightarrow p < y'$.*

Then the following assertions hold provided that K and K/K'' are sufficiently large.

(1) *Suppose that $|P| \geqslant L := N/(\log N)^{\ell}$ for some constant $\ell > 0$. Then*

$$|S(2N, [y', y]) \cap P| \ll \Delta \frac{\Psi(2N, [y', y]) \cdot |P|}{N},$$

where $\Delta = \log_3 N + (\log N)^{\ell - 1/24}$.

(2) *If, instead, $|P| \geqslant L := N \exp(-C\varpi(N))$ and $q \leqslant N/L \leqslant \exp(C\varpi(N)) = o(\log N)$ for some $\varpi(N) = o(\log_2 N)$, then*

$$|S(2N, [y', y]) \cap P| \ll \Delta \frac{\Psi(2N, [y', y]) \cdot |P|}{N},$$

where $\Delta = \log \varpi(N)$.

Proof. Note that if $n \in [N, 2N]$ is an element of $S([y', y])$, then $h_{[y',y]}(n) \neq 0$ and $N^{1-\alpha} \leqslant n^{1-\alpha} \leqslant (2N)^{1-\alpha}$, so that

$$h_{[y',y]}(n) = \frac{N^{\alpha}}{\Psi(N, [y', y])} \frac{n^{1-\alpha}}{\alpha} \asymp \frac{N}{\Psi(N, [y', y])} \quad (n \in [N, 2N] \cap S([y', y])). \tag{5-11}$$

We shall show that in either of the two cases in the statement

$$\frac{q}{\phi(q)} \ll \Delta,$$

which then leaves us with the task of deducing from Lemma 5.5 that

$$\sum_{n \in P} h_{[y',y]}(n) \ll \Delta |P|.$$

In order to bound $q/\phi(q)$, consider the decomposition $q = q_1 q_2$, where q_1 is composed of primes $p \leqslant \varpi$ and q_2 is composed of primes $p > \varpi$. Then

$$\prod_{p \mid q_1} (1-p^{-1})^{-1} \ll \log \varpi.$$

Concerning the product over prime divisors of q_2, we obtain

$$\prod_{p \mid q,\, p>\varpi} (1-p^{-1})^{-1} \ll \exp\biggl(\sum_{p \mid q,\, p>\varpi} p^{-1}\biggr) \ll 1 + O\biggl(\frac{\log q}{\varpi \log \varpi}\biggr), \tag{5-12}$$

where we used the bound $\omega(q_2) \leqslant (\log q)/\log \varpi$. Thus,

$$\frac{q}{\phi(q)} \ll \log \varpi + \frac{\log q}{\varpi} \ll \log_2 q,$$

where we chose $\varpi \log \varpi = \log q$. In case (2), we use this bound directly and obtain $q/\phi(q) \ll \log \varpi(N) = \Delta$. In case (1), we have $q/\phi(q) \ll \log_2 q \ll \log_3 N \ll \Delta$.

Turning towards the application of Lemma 5.5, note that for $|P| = N_1/q$ the bound in that lemma satisfies

$$\frac{N_1}{\phi(q)} \frac{\log_3 N}{\log_2 N} + \frac{N}{\phi(q) \log^{1/24} N} + \frac{N}{\log^{1/5} N} \ll |P| \biggl(\frac{q}{\phi(q)} \frac{\log_3 N}{\log_2 N} + (N/|P|)(\log N)^{-1/24}\biggr).$$

If $|P| > N(\log N)^{-\ell}$, we obtain

$$\frac{N_1}{\phi(q)} \frac{\log_3 N}{\log_2 N} + \frac{N}{\phi(q) \log^{1/24} N} + \frac{N}{\log^{1/5} N} \ll |P| \biggl(\frac{q}{\phi(q)} \frac{\log_3 N}{\log_2 N} + (\log N)^{\ell - 1/24}\biggr) \ll |P| \Delta,$$

provided $\Delta \gg (\log N)^{\ell-1/24}$, which holds under the assumptions of (1).

If, instead, $\log(N/|P|) \ll \varpi(N) = o(\log_2 N)$, then

$$\frac{N_1}{\phi(q)} \frac{\log_3 N}{\log_2 N} + \frac{N}{\phi(q) \log^{1/24} N} + \frac{N}{\log^{1/5} N} \ll |P| \biggl(\frac{q}{\phi(q)} \frac{\log_3 N}{\log_2 N} + o(1)\biggr) = o(|P|),$$

which concludes the proof. □

In addition to the above result, we shall need an analogous bound that is valid on much shorter intervals but may in return have a larger value of Δ. For y-smooth numbers, such a result appears as "Smooth Numbers Result 3" in [Harper 2016]; see also [Drappeau and Shao 2016, Lemma 3.3] for a slight extension. The lemma below will be used in Section 6, where we extend some of the results of [Drappeau and Shao 2016], as well as in Section 7, where it will find application in those situations where the progressions are too short for the previous corollary to be used.

Lemma 5.7. *Let $K' > 0$, $K \geqslant \max(2K', 2)$, $x > 2$, $1 \leqslant y' \leqslant (\log x)^{K'}$ and $(\log x)^K < y \leqslant x$. Suppose that $P \subseteq [x, 2x]$ is an arithmetic progression. Then, provided that x is sufficiently large, we have*

$$\#\{n \in P : n \in S([y', y])\} \ll (x/|P|)^{1-\alpha} \frac{\Psi(x, [y', y])|P|}{x} \log x.$$

Proof. Let $X \geqslant x$ denote the smallest element of P. This lemma is a generalisation of [Harper 2016, Smooth Numbers Result 3], which bounds

$$\sum_{\substack{X \leqslant n \leqslant X+Z \\ n \equiv a \pmod{q}}} \mathbf{1}_{S(y)}(n)$$

under the assumptions that $q \geqslant 1$ and $qy \leqslant Z \leqslant X$, i.e., concerns progressions of length $|P| \geqslant y$. The proof of [Harper 2016, Smooth Numbers Result 3] carries over directly to the situation where $S(y)$ is replaced by $S([y', y])$, with one small exception: The application of [de la Bretèche and Tenenbaum 2005, Théorème 2.4(i)] with $m = 1$ needs to be replaced by an application of Lemma 3.4(i), which is [loc. cit., Théorème 2.4(i)] with $m = P(y')$. While in the former case, the bound holds uniformly for $1 \leqslant d \leqslant x$, it is only available for $1 \leqslant d \leqslant x/y$ in our case.

Replacing the assumption that $qy \leqslant Z$ by $qy^2 \leqslant 2Z$ ensures that $d < X/y$ in all applications of [loc. cit., Théorème 2.4(i)]. More precisely, we have

$$\frac{Xqy}{Z2^{j+1}} \leqslant \frac{X}{y}$$

for all admissible values of j in Harper's proof.

It remains to consider the case where $|P| < y^2/2$, which corresponds to the case $|P| < y$ addressed in (the proof of) [Drappeau and Shao 2016, Lemma 3.3]. By bounding the left-hand side in our statement trivially by $|P|$, and observing that $|P|^{-1+\alpha} \gg y^{(-1+\alpha)/2}$, it suffices to show that

$$y^{\alpha-1} \frac{\Psi(x, [y', y])}{x^{\alpha}} \log x \gg 1.$$

Note that $y^{\alpha-1} = (u \log(u+1) + O(1))^{-1}$ in view of (3-2) and observe that

$$y^{\alpha-1} \frac{\Psi(x, [y', y])}{x^{\alpha}} \log x \gg y^{\alpha-1} \frac{\prod_{y' \leqslant p \leqslant y} (1 - p^{-\alpha})^{-1}}{\sqrt{\log x \log y}} \log x \gg y^{\alpha-1} \prod_{y' \leqslant p \leqslant y} (1 - p^{-\alpha})^{-1}$$

by Lemma 3.3 and (3-3). We split the analysis of the product over primes into two cases according to the size of y. When $\exp(\log x / \log_2 x) \leqslant y \leqslant x$, noting that $\alpha \leqslant 1$, we have

$$\prod_{y' \leqslant p \leqslant y} (1 - p^{-\alpha})^{-1} \gg \frac{\prod_{p \leqslant y} (1 - p^{-1})^{-1}}{\prod_{p \leqslant y'} (1 - p^{-1})^{-1}} \gg \frac{\log y}{1 + \log y'} \gg \frac{\log x}{(\log_2 x)^2}.$$

Hence,

$$y^{\alpha-1} \prod_{y' \leqslant p \leqslant y} (1 - p^{-\alpha})^{-1} \gg \frac{\log y}{\log x \log(u+1)} \frac{\log x}{(\log_2 x)^2} \gg \frac{\log x}{(\log_2 x)^3 \log_3 x} \gg 1.$$

When $y'^2 \leqslant y \leqslant \exp(\log x / \log_2 x)$, we have $u \geqslant \log_2 x$ and $\alpha \leqslant 1 - (\log_3 x)/\log y + O(1/\log y)$. Hence, the second conclusion of Lemma 3.2 shows that

$$\begin{aligned} y^{-1+\alpha} \prod_{y' \leqslant p \leqslant y} (1 - p^{-\alpha})^{-1} &\gg y^{-1+\alpha} \prod_{\sqrt{y} \leqslant p \leqslant y} (1 - p^{-\alpha})^{-1} \\ &\gg y^{-1+\alpha} \exp\biggl(\frac{y^{1-\alpha}(1+o(1)) - 2y^{(1-\alpha)/2}(1+o(1))}{(1-\alpha)\log y} \biggr) \\ &\gg \exp\biggl(\frac{y^{1-\alpha}(1+o(1)) - (\log y^{1-\alpha})^2}{\log y^{1-\alpha}} \biggr) \gg 1, \end{aligned}$$

where we used that $y^{1-\alpha} \asymp u \log u$, where $u \to \infty$ as $x \to \infty$, implies that $y^{(1-\alpha)/2} = o(y^{1-\alpha})$. □

6. Weyl sums for numbers without small and large prime factors

As technical input in the proof of the orthogonality of $(h^{(W,A)}_{[y',y]}(n)-1)$ to nilsequences, we shall require bounds on Weyl sums where the summation variable is restricted to integers in $S([y',y])$. Drappeau and Shao [2016, Section 5] established bounds on Weyl sums for smooth numbers $S(x,y)$ by extending the methods from Harper's minor arc analysis [2016]. Our aim in this section is to obtain analogous bounds for Weyl sums over $S(x,[y',y])$, that is, for Weyl sums over numbers free from small and large prime factors.

Following the notation in [Drappeau and Shao 2016], define for any given parameters $Q, x > 0$ and coprime integers $0 < a \leqslant q \leqslant Q$ the major arcs

$$\mathfrak{M}(q,a;Q,x)=\{\theta\in[0,1):|q\theta-a|\leqslant Qx^{-k}\}$$

and

$$\mathfrak{M}(Q,x)=\bigcup_{\substack{0\leqslant a<q\leqslant Q\\(a,q)=1}}\mathfrak{M}(q,a;Q,x). \tag{6-1}$$

Given any positive integer k, define the following smooth Weyl sum of degree k:

$$E_k(x,[y',y];\theta):=\sum_{n\in S(x,[y',y])}e(\theta n^k).$$

The main objective of this section is to establish the theorem below.

Theorem 6.1 (Weyl sums, x^η-smooth case). *Let $k>0$ be a fixed positive integer and suppose that $\eta\in(0,1/(4k)]$. Let $K>2K'>2$ and suppose that $y'\leqslant(\log x)^{K'}<(\log x)^K<y\leqslant x^\eta$ and that K is sufficiently large. Let $\alpha=\alpha(x,y)$ denote the saddle point associated to $S(x,y)$, and let $\theta\in\mathbb{T}:=\mathbb{R}/\mathbb{Z}$ be a frequency. Then*:

(1) *If x is sufficiently large and $\theta\notin\mathfrak{M}(x^{1/12},x)$, we have*

$$E_k(x,[y',y];\theta)\ll x^{1-c}$$

for some $c=c(k,\varepsilon)>0$.

(2) *If x is sufficiently large, $\theta\in\mathfrak{M}(x^{1/12},x)$ and if $0<a<q\leqslant x^{0.1}$ are coprime integers such that $|q\theta-a|\leqslant q^{-1}$, then*

$$E_k(x,[y',y];\theta)\ll\mathcal{Q}^{-c+2(1-\alpha)}(\log x)^5\Psi(x,[y',y])$$

for some $c=c(k,\varepsilon)>0$ and $\mathcal{Q}=q+x^k\|q\theta\|$.

We prove the two parts of this lemma in turn. The former case, in which θ is "highly irrational", will be handled by extending results of [Wooley 1995]. For this purpose, we first establish an auxiliary lemma generalising [loc. cit., Lemma 2.3], which in turn builds on an argument of [Vaughan 1989].

Lemma 6.2. *Suppose that $\theta\in\mathbb{T}$ is a frequency and $r\in\mathbb{N}$. If $1\leqslant y'<y\leqslant M<x$, then we have*

$$\sum_{\substack{n\in S(x,[y',y])\\(n,r)=1}}e(n^k\theta)\ll y(\log x)\max_{y'\leqslant p\leqslant y}\sum_{\substack{v\in\mathfrak{B}(M,p,y)\\(v,r)=1}}\sup_{\beta\in\mathbb{T}}\Bigg|\sum_{\substack{u\in S(x/M,[y',p])\\(u,r)=1}}e(u^kv^k\theta+u\beta)\Bigg|+M,$$

where $\mathfrak{B}(M,p,y)=\{M<v\leqslant Mp:p\mid v,\ v\in S([p,y])\}$.

Proof. We start by decomposing the elements $n \in S(x, [y', y])$ in a similar fashion as in [Vaughan 1989, Lemma 10.1]. By considering the prime factorisation $n = p_1^{k_1} \cdots p_s^{k_s}$, where the prime factors are ordered in *decreasing* order $y \geqslant p_1 > \cdots > p_s \geqslant y'$, it is immediate that, for every given $M \in [y, x)$ and every $n \in S(x, [y', y])$ with $n > M$, there is a unique triple (u, v, p) such that $n = uv$ and such that v is the smallest initial factor in the factorisation above that exceeds M. More precisely,

(1) $y' \leqslant p \leqslant y$ and $p \mid v$,

(2) $u \in S(x/v, [y', p])$,

(3) $M < v \leqslant Mp$ and $v \in S([p, y])$.

Using this factorisation, the exponential sum can be decomposed as

$$\sum_{\substack{n \in S(x,[y',y]) \\ (n,r)=1}} e(n^k\theta) = \sum_{y' \leqslant p \leqslant y} \sum_{\substack{v \in \mathfrak{B}(M,p,y) \\ (v,r)=1}} \sum_{\substack{u \in S(x/v,[y',p]) \\ (u,r)=1}} e(u^k v^k \theta) + O(M),$$

where the error term $O(M)$ bounds the contribution from all $n \in S(x, [y', y])$ with $n \leqslant M$. As in [Vaughan 1989, (10.9)], we can use the orthogonality principle to remove the dependence on v in the restriction $u \in S(x/v, [y', p])$. Since $v > M$, the inner sum above is equal to

$$\begin{aligned}
&\int_{\mathbb{T}} \Bigg(\sum_{\substack{u \in S(x/M,[y',p]) \\ (u,r)=1}} e(u^k v^k \theta + u\beta) \Bigg) \Bigg(\sum_{m \leqslant x/v} e(-m\beta) \Bigg) \,\mathrm{d}\beta \\
&\qquad \ll \sup_{\beta \in \mathbb{T}} \Bigg| \sum_{\substack{u \in S(x/M,[y',p]) \\ (u,r)=1}} e(u^k v^k \theta + u\beta) \Bigg| \int_{\mathbb{T}} \min\{x/v, \|\beta\|^{-1}\} \,\mathrm{d}\beta \\
&\qquad \ll \log x \sup_{\beta \in \mathbb{T}} \Bigg| \sum_{\substack{u \in S(x/M,[y',p]) \\ (u,r)=1}} e(u^k v^k \theta + u\beta) \Bigg|.
\end{aligned}$$

The lemma follows by combining this bound with the previous expression. ☐

The following lemma is a generalisation of the Weyl sum estimate for y-smooth numbers given by [Wooley 1995], and it relies on the observation that the conclusion of [loc. cit., Lemma 3.1] continues to hold when we restrict the y-smooth numbers to $[y', y]$-smooth numbers.

Lemma 6.3 (Weyl sum, θ is highly irrational). *Let $k \in \mathbb{N}$ be a fixed positive integer and suppose that $\sigma \in \left(0, \frac{1}{2}\right)$ and $\eta \in (0, \sigma/(4k)]$. Then there exists $c = c(k, \sigma) > 0$ such that the following statement is true. Let $1 \leqslant y' < y \leqslant x^\eta$ and $\theta \in (0, 1)$. Then, if $\theta \notin \mathfrak{M}(x^\sigma, x)$, we have*

$$E_k(x, [y', y]; \theta) \ll x^{1-c}.$$

Proof. We seek to apply a modified version of [Wooley 1995, Lemma 3.1] in a similar way as at the start of the proof of [loc. cit., Theorem 4.2]. Wooley's lemma produces bounds on

$$f(\theta; x, y) := E_k(x, [1, y]; \theta)$$

in terms of M and q, where $M = x^{\lambda} > y$ is a parameter such that $\lambda \in \left(\frac{1}{2}, 1\right)$ and q is any natural number such that $\|q\theta\| = |q\theta - a| < q^{-1}$ for some $a \in \mathbb{Z}$ with $(a, q) = 1$. As we shall explain below, one can deduce from the proof of the lemma that the same bounds hold for $E_k(x, [y', y]; \theta)$ in place of $f(\theta; x, y) = E_k(x, [1, y]; \theta)$, i.e., for the exponential sum relevant to us. In order for those bounds to be useful, additional assumptions on M and q are necessary, and our first aim is to show that, under the assumption of our lemma, we can choose M and q in such a way that

$$M > y, \quad q \leqslant 2(yM)^k, \quad |q\theta - a| < (yM)^{-k}/2 \quad \text{and} \quad q > (x/M)^k. \tag{6-2}$$

These conditions correspond to those in place at the start of the proof of [Wooley 1995, Theorem 4.2]. When combined[6] with the upper bound produced by [loc. cit., Lemma 3.1] it follows that

$$E_k(x, [y', y]; \theta) \ll x^{1-c(k,\sigma)}, \tag{6-3}$$

once we prove that the lemma can be extended to the case of $[y', y]$-smooth numbers.

Concerning the conditions (6-2), let $0 < \sigma^* < \sigma/2$. Then $\sigma^* + 2k\eta < \sigma < \frac{1}{2}$. By the Dirichlet approximation theorem, there is a positive integer $q \leqslant x^{k-\sigma^*}$ and some $a \in \mathbb{Z}$ with $(a, q) = 1$ such that

$$\|q\theta\| = |q\theta - a| \leqslant x^{\sigma^* - k} < x^{\sigma - k}.$$

Since $\theta \notin \mathfrak{M}(x^{\sigma}, x)$, this implies that $q > x^{\sigma}$. Let M be defined by the equation

$$x^{k-\sigma^*} = 2(yM)^k.$$

Then, since $\sigma^* + 2\eta < \frac{1}{2}$, we have

$$y < x^{\eta} < x^{1-\sigma^*/k-\eta} 2^{-1/k} \leqslant M = 2^{-1/k} x^{1-\sigma^*/k} y^{-1} \leqslant x^{1-\sigma^*/k}$$

as soon as x is sufficiently large. Hence, $y < M$. If we write $M = x^{\lambda}$, then the above line of inequalities also shows that $\lambda \in \left(\frac{1}{2}, 1\right)$ since $\sigma^*/k + \eta < \frac{1}{2}$. Observe further that $\sigma^* + k\eta < \sigma$ implies that

$$q > x^{\sigma} > 2x^{\sigma^* + k\eta} \geqslant (x/M)^k$$

as soon as x is sufficiently large. Hence, all conditions from (6-2) are satisfied and (6-3) therefore follows under the assumptions of our lemma.

It remains to show that [Wooley 1995, Lemma 3.1] can be extended to the case of exponential sums over $[y', y]$-smooth numbers. The proof strategy is to follow the original argument, applied to the exponential sum $E_k(x, [y', y]; \theta)$ instead of $E_k(x, [1, y]; \theta)$, and explain how all dependencies on y' that appear in any of the bounds on the way can be removed, so as to eventually arrive at an intermediate bound that agrees with the corresponding bound in the original proof. Following the original proof through to the end shows then that the upper bound that [loc. cit., Lemma 3.1] provides for $f(\theta; x, y) = E_k(x, [1, y]; \theta)$ is also an upper bound for $E_k(x, [y', y]; \theta)$.

[6] See the proof of [loc. cit., Theorem 4.2] for details.

Turning to the details, let $M = x^\lambda$ for any $\lambda \in \left(\frac{1}{2}, 1\right)$ and let $q \in \mathbb{N}$ be such that $|q\theta - a| < q^{-1}$ for some $a \in \mathbb{Z}$ with $(a, q) = 1$. The $r = 1$ case of Lemma 6.2 (which replaces [Wooley 1995, Lemma 2.3]) produces a prime $y' < p \leqslant y$ and some frequency $\gamma \in \mathbb{T}$ such that

$$E_k(x, [y', y]; \theta) \ll (\log x) y \sum_{v \in S(My,y)} \left| \sum_{u \in S(x/M,[y',p])} e(u^k v^k \theta + u\gamma) \right| + M.$$

Note carefully that the sum over v only involves y-smooth numbers. Following the argumentation of the start of the proof of [loc. cit., Lemma 3.1] (which involves expressing the absolute value of the inner sum as $\varepsilon(v, \theta) \sum_{u \in S(x/M,[y',p])} e(u^k v^k \theta + u\gamma)$ for some $\varepsilon(v, \theta) \in \mathbb{C}$ of unit modulus, reinterpreting certain exponential sums, as well as two applications of Hölder's inequality) shows that, for all positive integers $t, w \in \mathbb{N}$,

$$\left(\sum_{v \in S(My,y)} \left| \sum_{u \in S(x/M,[y',p])} y e(u^k v^k \theta + u\gamma) \right| \right)^{2tw} \ll (My)^{2w(t-1)} \left(\sum_c n_c \right)^{2w-2} \left(\sum_c n_c^2 \right) J_w(\theta),$$

where $J_w(\theta)$ is as in [loc. cit., (3.4)] and where n_c denotes the number of solutions to the equation $u_1^k + \cdots + u_t^k = c$ with $u_i \in S(x/M, [y', p])$. Since $p \leqslant y$, we have $S(x/M, [y', p]) \subset S(x/M, y)$, which implies that $\sum_c n_c^2$ is trivially bounded above by $\int_{\mathbb{T}} \left| \sum_{u \in S(x/M,y)} e(u^k \theta) \right|^{2t} \mathrm{d}\theta$. Using the trivial bound $\sum_c n_c \ll (x/M)^t$ for the remaining sum over c completes our task of removing all dependencies on y' from the upper bound on the given exponential sum $E_k(x, [y', y]; \theta)$. To summarise, we obtain

$$\left(\sum_{v \in S(My,y)} \left| \sum_{u \in S(x/M,[y',p])} e(u^k v^k \theta + u\gamma) \right| \right)^{2tw} \ll (My)^{2w(t-1)} (x/M)^{t(2w-2)} J_w(\theta) \int_{\mathbb{T}} \left| \sum_{u \in S(x/M,y)} e(u^k \theta) \right|^{2t} \mathrm{d}\theta,$$

corresponding to the bound in [loc. cit., (3.5)]. □

The following final lemma of this section is an easy generalisation of [Drappeau and Shao 2016, Proposition 5.7], which itself generalises the $k = 1$ case established in [Harper 2016, Theorem 1].

Lemma 6.4 (Weyl sum, θ is irrational). *Suppose that $\theta = a/q + \delta$ for some $q \in \mathbb{N}$, $a \in \mathbb{Z}$ and $\delta \in \mathbb{R}$ such that $(a, q) = 1$ and $|\delta| \leqslant 1/(2q)$. Further, let $x > 2$, $1 \leqslant y' \leqslant (\log x)^{K'}$ for some $K' \geqslant 1$ and let $y'^2 \leqslant y \leqslant x$. Write $\mathscr{Q} = q + x^k \|q\theta\|$, and assume that $4y^3 \mathscr{Q}^3 \leqslant x$. Then there exists some constant $c = c(k) > 0$ such that*

$$E_k(x, [y', y]; \theta) \ll \mathscr{Q}^{-c+2(1-\alpha)} (\log x)^5 \Psi(x, [y', y]).$$

Proof. The proof of [Drappeau and Shao 2016, Proposition 5.7] carries over almost directly when replacing any application of [loc. cit., Lemma 3.2] by one of Lemma 3.4(i) combined with Lemma 3.3. In doing so, we however have to ensure that the stronger condition that $1 \leqslant d \leqslant x/y$ of Lemma 3.4(i) is satisfied. For this reason we make the stronger assumption that $4y^3 \mathscr{Q}^3 \leqslant x$ instead of $4y^2 \mathscr{Q}^3 \leqslant x$. As only fairly straightforward changes are required, we explain below how to deduce our lemma from the proof of [Drappeau and Shao 2016, Proposition 5.7], instead of repeating the entire proof.

Extracting, as before, the greatest common divisor of $n \in S(x, [y', y])$ and q^∞, we obtain

$$E_k(x, [y', y]; \theta) = \sum_{\substack{d \mid q^\infty \\ d \leqslant x}} \mathbf{1}_{S([y',y])}(d) \sum_{\substack{n \leqslant x/d \\ (n,q)=1}} \mathbf{1}_{S([y',y])}(n) e(n^k d^k \theta).$$

We claim that the contribution from those terms with $d > \mathscr{Q}^{1/2}$ and those terms with $n \leqslant x/\mathscr{Q}$ is negligible, in the sense that

$$E_k(x, [y', y]; \theta) = \sum_{\substack{d \mid q^\infty \\ d \leqslant \mathscr{Q}^{1/2}}} \mathbf{1}_{S([y',y])}(d) \sum_{\substack{x/\mathscr{Q} \leqslant n \leqslant x/d \\ n \in S([y',y]) \\ (n,q)=1}} e(n^k d^k \theta) + O_\varepsilon\biggl(\frac{\Psi(x, [y', y])}{\mathscr{Q}^{\alpha/2-\varepsilon}}\biggr). \tag{6-4}$$

We start with the contribution from $\max(x/y, \mathscr{Q}^{1/2}) < d \leqslant x$. In this case $x/d < y < x^{1/3}$. Note that $q \leqslant \mathscr{Q} \leqslant x^{1/3}$ has $\omega(q) \leqslant \log x$ prime factors. Hence, the contribution to $E_k(x, [y', y]; \theta)$ is bounded by

$$\sum_{\substack{d \mid q^\infty \\ x/y < d \leqslant x}} y \leqslant x^{1/3} \Psi(x, \log x) \leqslant x^{1/3+o(1)}$$

by Erdős's bound stated in [Montgomery and Vaughan 2007, Section 7.1.1, Exercise 12]. Since $\alpha > \frac{1}{2}$ and $\mathscr{Q} < x^{1/3}$, it follows that

$$x^{1/3+o(1)} \ll \mathscr{Q}^{-\alpha/2} \Psi(x, [y', y]).$$

If $\mathscr{Q}^{1/2} < d < x/y$, then Lemma 3.4(i) applies and we obtain

$$\begin{aligned}
&\sum_{\substack{d \mid q^\infty \\ \mathscr{Q}^{1/2} < d \leqslant x/y}} \Psi(x/d, [y', y]) \mathbf{1}_{S([y',y])}(d) \\
&\qquad \ll \Psi(x, [y', y]) \sum_{\substack{\mathscr{Q}^{1/2} < d \leqslant x/y \\ d \mid q^\infty}} d^{-\alpha} \mathbf{1}_{S([y',y])}(d) \\
&\qquad \ll \Psi(x, [y', y]) \mathscr{Q}^{-\alpha/2} \sum_{d \mid q^\infty} d^{-\alpha/2} \mathbf{1}_{S([y',y])}(d) \ll \Psi(x, [y', y]) \mathscr{Q}^{-\alpha/2} \exp\biggl(\sum_{\substack{p > y' \\ p \mid q}} p^{-\alpha/2} + O(1)\biggr) \\
&\qquad \ll \Psi(x, [y', y]) \mathscr{Q}^{-\alpha/2} \exp\biggl(O\biggl(y'^{-\alpha/2} \frac{\log \mathscr{Q}}{\log y'}\biggr)\biggr) \ll_\varepsilon \Psi(x, [y', y]) \mathscr{Q}^{-\alpha/2+\varepsilon},
\end{aligned}$$

since $\alpha > \frac{1}{2}$.

To remove the contribution from $n \leqslant x/\mathscr{Q}$ we may now assume that $d < \mathscr{Q}^{1/2}$. This contribution is trivially bounded above by

$$\sum_{d < \mathscr{Q}^{1/2}} \sum_{\substack{n \leqslant x/\mathscr{Q} \\ (n,d)=1}} \mathbf{1}_{S([y',y])}(dn) \leqslant \Psi(x/\mathscr{Q}^{1/2}; [y', y]) \ll \mathscr{Q}^{\alpha/2} \Psi(x; [y', y]),$$

where we used Lemma 3.4(i) together with the bound $\mathscr{Q}^{1/2} < x/y$ which follows from the assumption that $4y^3\mathscr{Q}^3 \leqslant x$. This shows that (6-4) holds.

Following [Drappeau and Shao 2016], set $L = 4y\mathscr{Q}$ and decompose any element of $[x/\mathscr{Q}, x/d] \cap S([y', y])$ as the unique product mn for which $m \in [L, yL]$ and $P^+(m) \leqslant P^-(n)$, which is possible since

$y < Ly < x/\mathscr{Q}$ by the assumption that $4y^3\mathscr{Q}^3 \leqslant x$. Decomposing the range of m dyadically and extracting the largest prime factor of m implies that

$$E_k(x, [y', y]; \theta) \ll \log x \sup_{L \leqslant M \leqslant yL} \mathscr{E}(M) + \frac{\Psi(x, [y', y])}{\mathscr{Q}^{\alpha/4}},$$

where

$$\mathscr{E}(M) = \sum_{\substack{d \mid q^\infty \\ d \leqslant \mathscr{Q}^{1/2} \\ d \in S([y', y])}} \sum_{y' \leqslant p \leqslant y} \sum_{\substack{m \in S([y', p]) \\ M/p \leqslant m \leqslant \min\{2M, yL\}/p}} \sum_{\substack{n \in S([p, y]) \\ (pmn, q) = 1 \\ x/(pm\mathscr{Q}) \leqslant n \leqslant x/(pmd)}} e((pmnd)^k \theta).$$

The dependence of the summation condition of the inner sum on m can be removed with the help of the same trick as in the proof of Lemma 6.2, which leads to

$$\mathscr{E}(M) \ll (\log x) \sup_{\beta \in [0,1)} \sum_{\substack{d \mid q^\infty \\ d \leqslant \mathscr{Q}^{1/2} \\ d \in S([y', y])}} \sum_{y' \leqslant p \leqslant y} \sum_{\substack{m \in S([y', p]): \\ M/p \leqslant m \leqslant \min\{2M, yL\}/p}} \Bigg| \sum_{\substack{n \in S([p, y]) \\ (pmn, q) = 1 \\ x/(2M\mathscr{Q}) \leqslant n \leqslant x/(Md)}} e((pmnd)^k \theta + \beta n) \Bigg|.$$

An application of Cauchy–Schwarz, combined with the bound

$$\sum_{\substack{d \mid q^\infty \\ d \leqslant \mathscr{Q}^{1/2} \\ d \in S([y', y])}} 1 \leqslant \mathscr{Q}^{\varepsilon/2} \sum_{\substack{d \mid q^\infty \\ d \leqslant \mathscr{Q}^{1/2} \\ d \in S([y', y])}} d^{-\varepsilon} \leqslant \mathscr{Q}^{\varepsilon/2} \exp\Bigg(O(1) \sum_{p \mid q} p^{-\varepsilon}\Bigg) \ll \mathscr{Q}^{\varepsilon/2} \mathscr{Q}^{o(1)} \ll_\varepsilon \mathscr{Q}^{\varepsilon},$$

which follows from $\omega(d) \leqslant (\log \mathscr{Q})/\log y'$, we obtain

$$\mathscr{E}(M) \ll_\varepsilon (\log x) \mathscr{Q}^\varepsilon M^{1/2} \mathscr{S}_1(M)^{1/2},$$

where $\mathscr{S}_1(M)$ is defined as in [Drappeau and Shao 2016, display just below (5.3)] except that the sum over primes runs over $y' \leqslant p \leqslant y$:

$$\mathscr{S}_1(M) = \sum_{\substack{d \mid q^\infty \\ d \leqslant \mathscr{Q}}} \sum_{y' \leqslant p \leqslant y} \sum_{M/p < m \leqslant 2M/p} \Bigg| \sum_{\substack{n \in S([p, y]) \\ (pmn, q) = 1 \\ x/(2M\mathscr{Q}) \leqslant n \leqslant x/(Md)}} e((pmnd)^k \theta + \beta n) \Bigg|^2.$$

Here we relaxed the summation conditions on d (to match the corresponding sum in [loc. cit.]) and, following [Harper 2016; Drappeau and Shao 2016], on m, which allows one to later use a standard exponential sum estimate where the summation variable runs over a full interval. Expanding out the square, swapping the order of summation, applying the triangle inequality and relaxing some of the summation conditions in the outer sum leads to

$$\mathscr{S}_1(M) \ll \sum_{\substack{d \mid q^\infty \\ d \leqslant \mathscr{Q}}} \sum_{y' \leqslant p \leqslant y} \sum_{\substack{n_1, n_2 \in S([y', y]) \\ (n_1 n_2, q) = 1 \\ x/(2M\mathscr{Q}) < n_1 \leqslant n_2 \leqslant x/(Md)}} \Bigg| \sum_{M/p < m \leqslant 2M/p} e((pmd)^k \theta (n_1^k - n_2^k)) \Bigg|.$$

Note carefully that instead of relaxing the condition $n_1, n_2 \in S([p, y])$ to $n_1, n_2 \in S(y)$, as in [Drappeau and Shao 2016], we replaced it by $n_1, n_2 \in S([y', y])$. We now follow the analysis of $\mathscr{S}_1(M)$ from [loc. cit.], with the only change that in the definitions of $\mathscr{S}_j$, $j \in \{2, \dots, 5\}$, the variables n_1 and n_2 continue to be restricted to $S([y', y])$. Instead of bounding the quantity $\mathscr{S}_5(M; r', d; n_1, b; T)$ with the help of [loc. cit., Lemma 3.3], we use Lemma 5.7, which produces the analogous bound with $\Psi(x/(Md), y)$ replaced by $\Psi(x/(Md), [y', y])$. In the definition of $\mathscr{S}'_3$, we include the restriction to $n_1 \in S([y', y])$. Adapting the analysis of $\mathscr{S}_3$ requires the lower bound

$$\begin{aligned}\Psi(x/(Md), y) &\asymp \left(\frac{x}{Md}\right)^{\alpha'} (\log x \log y)^{-1/2} \prod_{y' \leqslant p \leqslant y} (1 - p^{-\alpha'})^{-1} \\ &\gg \left(\frac{x}{Md}\right)^{\alpha'} (\log x \log y)^{-1/2} \gg \left(\frac{x}{Md}\right)^{\alpha' + o(1)} \gg \left(\frac{x}{Md}\right)^{\alpha},\end{aligned}$$

where $\alpha' = \alpha(x/(Md), y)$, which follows from Lemma 3.3, the asymptotic expansion (3-3) as well as the bound $\alpha' = \alpha(x/(Md), y) > \alpha(x, y) = \alpha$ implied by the lower bound $M \leqslant L = 4y\mathscr{Q} > \log x$ and the estimate (3-2).

The remaining analysis in the proof of [loc. cit., Proposition 5.7] involves two applications of [loc. cit., Lemma 3.2], which may be replaced by applications of Lemma 3.4(i) combined with Lemma 3.3 since for $d \leqslant \mathscr{Q}$ and $M \leqslant yL$ we have

$$Md \leqslant yL\mathscr{Q} \leqslant 4y^2\mathscr{Q}^2 \leqslant x/(y\mathscr{Q}) \leqslant x/y,$$

thanks to our stronger assumption that $4y^3\mathscr{Q}^3 \leqslant x$. This ensures that $d \leqslant x/(yM)$ in the first application and that $M < x/y$ in the second application, i.e., that the conditions of Lemma 3.4(i) are satisfied. Our lemma thus follows from the proof of [loc. cit., Proposition 5.7]. □

Proof of Theorem 6.1. The first part of the result follows directly from Lemma 6.3 applied with $\sigma = \frac{1}{12}$. Concerning the second part, suppose that $\theta \in \mathfrak{M}(x^{1/12}, x)$ and recall that by definition (6-1) there exists a positive integer $q \leqslant x^{0.1}$ such that $\|q\theta\| \leqslant x^{1/12}x^{-k}$. For any such value of q, we have $\mathscr{Q} = q + \|q\theta\|x^k \leqslant 2x^{1/12}$ and $4y^3\mathscr{Q}^3 \leqslant x$ since $y < x^\eta$ and $\eta < \frac{1}{4}$. It then follows from Lemma 6.4 that

$$E_k(x, [y', y]; \theta) \ll \mathscr{Q}^{-c+2(1-\alpha)} (\log x)^5 \Psi(x, [y', y]). \qquad \square$$

7. Strongly recurrent polynomial sequences over smooth numbers

Our aim in this section is to show that if a sequence $\{\|\beta n^k\|\}_{n \in S(x,[y',y])}$ is strongly recurrent, then β is very close to a rational with small denominator q. More precisely:

Theorem 7.1. *Let $N > 2$ be a parameter, let $\theta \in \mathbb{R}$, and let $k \geqslant 1$ be a fixed integer. Let $K' > 0$, $K > \max(2, 2K')$, $1 \leqslant y' \leqslant (\log N)^{K'}$, and suppose that $(\log N)^K \leqslant y \leqslant N^\eta$ for some small constant $\eta = \eta(k) \in (0, 1)$. Let $\delta : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be a function of N that satisfies*

$$\log_5 N \ll \delta(N)^{-1} < \log_2 N$$

for all sufficiently large N. Finally suppose that $0 < \varepsilon \leqslant \delta/2$ *and that* $\varepsilon < \delta(N)^{-O(1)}(\log N)^{-C_1}$ *for some fixed constant* $C_1 \geqslant 1$. *Then the following assertion holds provided that* K *and* C_1 *are sufficiently large depending on* k.

If the bound $\|n^k\theta\| \leqslant \varepsilon$ *holds for at least* $\delta\Psi(N, [y', y])$ *elements* $n \in S(N, [y', y])$, *then there exists a positive integer* $0 < q \ll \delta^{-O(1)}$ *such that*

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}/N^k.$$

The corresponding problem for strongly recurrent *unrestricted* polynomial sequences $\{\|\beta n^k\|\}_{n\leqslant x}$ has been treated in [Green and Tao 2012a, Section 3] where the polynomial case was reduced to the linear case via an application of bounds in Waring's problem. While a suitable Waring-type result was proved in [Drappeau and Shao 2016, Theorem 2.4] for the set $S(x, y)$, no such result is currently available for the sparse subset $S(x, [y', y])$, although we expect such an analogue to hold.

For this reason, our approach proceeds instead via first reducing the problem via Fourier analysis to bounds on Weyl sums over $S(x, [y', y])$, which can then be combined with the results of Section 6. The Fourier analysis reduction is standard (see [Green and Tao 2012b, Proposition 3.1 and Lemma 3.2]). The bounds on the relevant Weyl sums obtained in Section 6 by themselves are, however, not strong enough in order to deduce the theorem above; they only provide part (1) of Lemma 7.4 below. In particular, those bounds on $\|q\theta\|$ are not strong enough in order to later analyse the correlation of $g_{[y',y]}(n)$ with nilsequences for very small values of y, that is, when $\log y \ll \log\log x$. To work around this problem, we will employ a bootstrapping argument in order to improve the bounds on $\|q\theta\|$. This argument is based on the combination of the following lemma due to [Drappeau and Shao 2016, Lemma 3.7] (which is a higher-dimensional version of the bootstrapping argument used in the proof of [Green and Tao 2012b, Lemma 3.2]) and the bounds we obtained in Section 5.2 on equidistribution of $S([y', y])$ in short progressions.

Lemma 7.2 [Drappeau and Shao 2016, bootstrapping lemma]. *Let* $k > 0$ *be a fixed integer and let* $\varepsilon', \delta \in (0, 1)$. *Let* $1 \leqslant L \leqslant x$ *be parameters and let* $\mathscr{A} \subset [x, 2x]$ *be a nonempty subset with the property that*

$$|\mathscr{A} \cap P| \leqslant \Delta\frac{|\mathscr{A}||P|}{x}$$

for some $\Delta \geqslant 1$ *and for any arithmetic progression* $P \subseteq [x, 2x]$ *of length at least* L *and common difference* $q = 1$. *Suppose that, for some* $\vartheta \in \mathbb{R}$ *with* $\|\vartheta\| \leqslant \varepsilon'/(Lx^{k-1})$, *there are at least* $\delta|\mathscr{A}|$ *elements* $m \in \mathscr{A}$ *satisfying* $\|m^k\vartheta\| \leqslant \varepsilon'$. *Then either* $\varepsilon' \gg \delta/\Delta$ *or*

$$\vartheta \ll \Delta\delta^{-1}\varepsilon'/x^k.$$

Remark 7.3. The original statement of [Drappeau and Shao 2016, Lemma 3.7] does not restrict the progressions P to have common difference $q = 1$. An inspection of the proof reveals, however, that the corresponding assumption is only required for progressions with common difference $q = 1$, i.e., for discrete intervals.

Theorem 7.1 is an immediate consequence of the following lemma, which will help us structure our proof.

Lemma 7.4. *Let* $\theta \in \mathbb{R}$, $k \geqslant 1$ *a fixed integer, and let* $K' > 0$ *and* $K > \max(2K', 2)$ *be constants. Let* $1 \leqslant y' \leqslant (\log x)^{K'}$ *and suppose that* $(\log x)^K \leqslant y \leqslant x^\eta$ *for some small constant* $\eta \in (0, 1)$. *Let* $\delta = \delta(x)$ *be such that* $\delta(x)^{-B} \ll_B x$ *for all* $B > 0$ *and suppose that* $0 < \varepsilon \leqslant \delta/2$. *Suppose that there are at least* $\delta\Psi(x, [y', y])$ *elements* $n \in S(x, [y', y])$ *for which* $\|n^k\theta\| \leqslant \varepsilon$. *Then*:

(1) *There is some integer* $0 < q \ll \delta^{-O(1)}$ *and a constant* $c > 0$ *depending on* k *such that*

$$\|q\theta\| \ll \delta^{-O(1)}(\log x)^{10/c}/x^k$$

provided K *is sufficiently large in terms of* c.

(2) *If, furthermore,* $\delta(x) > (\log_2 x)^{-1}$ *and* $\varepsilon < \delta(x)^{-O(1)}(\log x)^{-C_1}$ *for some fixed constant* $C_1 \geqslant 5+20/c$, *then the integer* $0 < q \ll \delta^{-O(1)}$ *from part* (1) *satisfies*

$$\|q\theta\| \ll \varepsilon\delta^{-O(1)}(\log x)^{C_1}/x^k,$$

provided that K *is sufficiently large in terms of* c.

(3) *Under the assumptions of part* (2) *the integer* $0 < q \ll \delta^{-O(1)}$ *from part* (1) *satisfies the bound*

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}(\log_3 x)/x^k,$$

provided that K *is sufficiently large in terms of* c.

(4) *If, in addition to all previous assumptions,* $\delta^{-1} \gg \log_5 x$, *then the integer* $0 < q \ll \delta^{-O(1)}$ *from part* (1) *is such that*

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}/x^k,$$

provided that K *is sufficiently large in terms of* c.

Proof of Lemma 7.4. Suppose that $I \subseteq [0, 1]$ is an interval of length $|I| = \varepsilon$, and that there are at least $\delta\Psi(x, [y', y])$ elements $n \in S(x, [y', y])$ such that $n^k\theta \pmod{\mathbb{Z}} \in I$. By approximating the characteristic function of the interval I by a suitable smooth Lipschitz function F, we obtain

$$\Bigg|\sum_{n\in S(x,[y',y])} F(n^k\theta \pmod{\mathbb{Z}})\Bigg| \geqslant \delta'\Psi(x, [y', y]) \tag{7-1}$$

for some $\delta' \in [\delta/2, \delta]$. We may assume that F is supported on a set of measure δ. Hence, on rescaling F and redefining δ' in the bound above as $\delta' = \delta^C$ for some $C \geqslant 1$, we may assume that $\|F\|_{\mathrm{Lip}} = 1$, which implies that, for every $k \in \mathbb{Z}$,

$$|\hat{F}(k)| = \int_{\mathbb{R}/\mathbb{Z}} F(\theta)e(-\theta k)\ \mathrm{d}\theta \leqslant \|F\|_\infty \leqslant \|F\|_{\mathrm{Lip}} = 1. \tag{7-2}$$

Following the proof of [Green and Tao 2012b, Lemma 3.1], our next step is to approximate F by a function whose Fourier transform is finitely supported, with a support defined in terms of δ'. For this purpose, consider the Fejér kernel $K(\theta) = \chi_Q * \chi_Q(\theta)$, where $\chi_Q(\theta) = (\delta'/16)\mathbf{1}_Q(\theta)$ is the normalised characteristic function of the short interval $Q = [-\delta'/16, \delta'/16]$. Since $\int_{\mathbb{R}/\mathbb{Z}} K = 1$ and by (7-2), the convolution $F_1 = F * K$ satisfies

$$\|F - F_1\|_\infty \leqslant \frac{\delta'}{4} \quad \text{and} \quad |\hat{F}_1(k)| \leqslant 1 \quad \text{for all } k \in \mathbb{Z}.$$

The rapid decay of the Fourier coefficients of K allows one to approximate F_1 by a finite truncation of its Fourier series as follows. If

$$F_2(\theta) := \sum_{0<|q|\ll\delta'^{-3}} \hat{F}_1(q)e(q\theta), \quad \text{then} \quad \|F_2 - F_1\|_\infty \leqslant \frac{\delta'}{4}.$$

When applying both of these approximations to (7-1) and then swapping the order of summation, it follows from the triangle inequality that

$$\frac{\delta'}{2}\Psi(x, [y', y]) \leqslant \sum_{0<|q|\ll\delta'^{-3}} |\hat{F}_1(q)| \Bigg| \sum_{n\in S(x,[y',y])} e(qn^k\theta) \Bigg| \leqslant \sum_{0<|q|\ll\delta'^{-3}} \Bigg| \sum_{n\in S(x,[y',y])} e(qn^k\theta) \Bigg|.$$

By the pigeonhole principle there therefore is some integer $0 < |q| \ll \delta'^{-3}$ such that

$$\delta'^4 \Psi(x, [y', y]) \ll \Bigg| \sum_{n\in S(x,[y',y])} e(qn^k\theta) \Bigg| = |E_k(x, [y', y]; q\theta)|. \tag{7-3}$$

Following the above reduction via Fourier analysis, we are now in the position to invoke the bounds on Weyl sums over smooth numbers from Section 6. To start with, the first part of Theorem 6.1 shows that if $q\theta \pmod{\mathbb{Z}}$ does not belong to $\mathfrak{M}(x^{1/12}, x)$, then the right-hand side of (7-3) is bounded by

$$E_k(x, [y', y]; q\theta) \ll x^{1-c} \ll x^{\alpha} x^{1-\alpha-c} \ll x^{-c'}\Psi(x, [y', y]),$$

where $c' := c - (1 - \alpha)$ and where the lower bound $\Psi(x, [y', y]) \gg x^{\alpha+o(1)}$ follows from (3-3) and Lemma 3.2. Observe that $c' > 0$ provided that K is sufficiently large in terms of c since $1 - \alpha(x, y) \leqslant 1/K + o(1)$ by (3-2). Hence, $\delta'^4 \ll x^{-c'}$ and, thus, $x \ll \delta'^{-4/c'} = \delta^{-4C/c'}$, contradicting our assumptions on δ, which implies $\delta^{-4C/c'} \ll_B x^{1/B}$ for all $B > 0$.

It follows that $q\theta \pmod{\mathbb{Z}} \in \mathfrak{M}(x^{1/12}, x)$. In this case, the second part of Theorem 6.1, applied with θ replaced by $q\theta$ (and applied in the special case where, in the statement of Theorem 6.1, $q = a = 1$), shows that

$$\sum_{n\in S(x,[y',y])} e(qn^k\theta) \ll (1 + x^k\|q\theta\|)^{-c+2(1-\alpha)}(\log x)^5\Psi(x, [y', y]).$$

Suppose that K is sufficiently large to ensure that $c - 2(1 - \alpha) > c/2$. Then it follows from the Fourier analysis bound (7-3) that

$$1 + x^k\|q\theta\| \ll \delta^{-8C/c}(\log x)^{10/c},$$

and hence

$$\|q\theta\| \ll \delta^{-8C/c}(\log x)^{10/c}x^{-k},$$

which proves part (1). For later use in the proofs of all remaining parts, we record the following consequence of part (1) and our assumptions.

Recurrence condition. Let $0 < q \ll \delta^{-O(1)}$ denote the integer produced by part (1). Then we have

$$\|n^k q\theta\| \leqslant q\|n^k\theta\| < \varepsilon' \tag{7-4}$$

for some $\varepsilon' \asymp q\varepsilon \ll \delta^{-O(1)}\varepsilon$ and for at least $\delta\Psi(x, [y', y])$ elements n of $S(x, [y', y])$.

To establish part (2), observe that in view of Lemma 5.7, the conditions of Lemma 7.2 are satisfied for the set $\mathscr{A} = S(2x, [y', y])$ and any lower bound $L \geqslant 1$ with a correction factor of the form $\Delta = (x/L)^{1-\alpha}\log x$. Moreover, we have

$$\|q\theta\| \ll \delta^{-8C/c}(\log x)^{10/c}x^{-k} = \varepsilon'/(Lx^{k-1})$$

if we set $L = \varepsilon'\delta^{8C/c}(\log x)^{-10/c}x$ and with ε' as in (7-4). In this case,

$$\begin{aligned}\varepsilon'\Delta &\leqslant \varepsilon'(\log x)(\varepsilon'\delta^{8C/c}(\log x)^{-10/c})^{\alpha-1} \ll (q\varepsilon)^{\alpha}\delta^{-O(1)}(\log x)^{1+5/c}\\ &\ll (\log x)^{-\alpha C_1}\delta^{-O(1)}(\log x)^{1+5/c} = o(\delta(x))\end{aligned}$$

provided that C_1 is sufficiently large depending on c. Here, we used that $\alpha > \frac{1}{2}$ if $K > 2$. To find a simple bound for Δ, note that

$$x/L \ll (\varepsilon q\delta^{8C/c}(\log x)^{-10/c})^{-1} \ll (\log x)^{C_1}\delta^{-O(1)}(\log x)^{10/c} \ll (\log x)^{3C_1/2},$$

provided that C_1 is sufficiently large (e.g., $C_1 \geqslant 5 + 20/c$), which implies that

$$\Delta = (x/L)^{1-\alpha}\log x \ll (\log x)^{1+3C_1(1-\alpha)/2} \ll (\log x)^{3C_1/4+1} \leqslant (\log x)^{C_1}.$$

Thus, in view of the recurrence condition above, the conclusion of part (2) follows from Lemma 7.2, applied with the given value of ε' and $\vartheta := q\theta$, provided that C_1 is sufficiently large in terms of c.

To prove part (3), we use the information from part (2), i.e., that the positive integer $q \ll \delta^{-O(1)}$ produced by part (1) satisfies

$$\|q\theta\| \ll \varepsilon\delta^{-O(1)}(\log x)^{C_1}/x^k. \tag{7-5}$$

We shall now apply the short intervals case of Corollary 5.6(1). This result and the recurrence condition imply that the conditions of the bootstrapping lemma are satisfied for the set $\mathscr{A} = S(x, [y', y])$, the lower bound $L = x/(\log x)^{\ell}$ and

$$\Delta = \log_3 x + (\log x)^{\ell-1/24}$$

for any constant $\ell > 0$. Observe that, since $\varepsilon < \delta^{-O(1)}(\log x)^{-C_1}$ and $\delta(x)^{-1} < \log_2 x$, a bound of the form $\ell \leqslant C_1$ implies

$$\begin{aligned}\varepsilon'\delta(x)^{-O(1)}\Delta &< \delta(x)^{-O(1)}(\log_3 x + (\log x)^{\ell-1/24})(\log x)^{-C_1}\\ &< (\log x)^{-C_1+o(1)} + (\log x)^{-1/24+o(1)} = o(\delta(x)).\end{aligned}$$

This ensures that the bootstrapping lemma produces bounds on $\|q\theta\|$ in all applications below.

Let $0 \leqslant j < 24C_1$ be an integer, and suppose inductively that

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}(\log x)^{C_1-j/24}/x^k,$$

the case $j = 0$ being (7-5). Let $\ell := C_1 - j/24$, and $L = x/(\log x)^\ell$. Then the bootstrapping lemma, applied with ε' replaced by $\varepsilon\delta^{-O_{C_1}(1)}$, implies that

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}(\log x)^{C_1-(j+1)/24}/x^k$$

if $j+1 < 24C_1$. Recalling the shape of Δ, we pick up a $(\log_3 x)$-factor when $j+1 \geqslant 24C_1$ and thus $\ell - 1/24 \leqslant 0$, and obtain

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}(\log_3 x)/x^k$$

as claimed.

It remains to establish part (4). In view of part (3), we have

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}(\log_3 x)/x^k$$

and there are $\delta\Psi(x, [y', y])$ elements $n \in S(x, [y', y])$ for which $\|n^k q\theta\| < \varepsilon'$ by (7-4). We shall now appeal to Corollary 5.6(2) in order to remove the $(\log_3 x)$-factor from the bound. For this purpose let $\varpi(x) = \log_4 x$ and recall the recurrence condition. Then Corollary 5.6(2) shows that the conditions of Lemma 7.2 are satisfied for the set $\mathscr{A} = S(x, [y', y])$, the lower bound $L = x/\log_3 x$ and $\Delta = \log\varpi(x) \ll \log_5 x \ll \delta^{-1}$. By the assumptions on δ and $\varpi(x)$ it follows easily that

$$\Delta\varepsilon'\delta^{-O_{C_1}(1)} \ll \delta^{-O_{C_1}(1)}(\log x)^{-C_1} = o(\delta).$$

Hence, Lemma 7.2 applied with ε' replaced by $\varepsilon\delta^{-O_{C_1}(1)}$ finally leads to the bound

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}/x^k. \qquad \square$$

8. Noncorrelation with nilsequences: a reduction and general lemmas

The aim of this section is to first establish an initial reduction of the noncorrelation estimate stated in Theorem 1.1 to the case where the nilsequence is equidistributed. In preparation for the proof of the reduced version we then show that most sequences in certain sparse families of subsequences of an equidistributed polynomial sequence are equidistributed. We start by recalling some notation around nilsequences.

Definition 8.1 (filtered nilmanifold). Let $d, m_G \geqslant 0$ be integers and let $M > 0$. We define a *filtered nilmanifold G/Γ of degree d, dimension m_G and complexity at most M* to be an s-step nilmanifold G/Γ, for some $1 \leqslant s \leqslant d$, of dimension m_G in the sense of [Green and Tao 2012b, Definition 1.1] such that G is equipped with a filtration $G_\bullet$ of degree $d \geqslant s$ in the sense of that definition and such that the Lie algebra $\mathfrak{g} = \log G$ is equipped with an M-rational Malcev basis adapted to $G_\bullet$ in the sense of [loc. cit., Definition 2.1].

A Malcev basis $\mathscr{X}$ gives rise to a metric $d_{\mathscr{X}}$ on G/Γ (see [loc. cit., Definition 2.2]). With respect to this metric, Lipschitz functions can be defined. More precisely, if $F : G/\Gamma \to \mathbb{C}$, we define (see [loc. cit., Definition 1.2]) the Lipschitz norm

$$\|F\|_{\mathrm{Lip}} := \|F\|_\infty + \sup_{x,y\in G/\Gamma,\, x\neq y} \frac{|F(x) - F(y)|}{d_{\mathscr{X}}(x, y)}$$

and call F a Lipschitz function if $\|F\|_{\mathrm{Lip}} < \infty$.

Definition 8.2 (polynomial sequence and nilsequence). Given a nilpotent Lie group G and a filtration

$$G_\bullet : G = G_0 = G_1 \supseteq G_2 \supseteq \cdots \supseteq G_d \supseteq G_{d+1} = \{\mathrm{id}_G\},$$

the set $\mathrm{poly}(\mathbb{Z}, G_\bullet)$ of *polynomial sequences* is defined as the set of all maps $g : \mathbb{Z} \to G$ such that, if $\partial_h g(n) := g(n+h)g(n)^{-1}$, the j-th discrete derivative $\partial_{h_j} \dots \partial_{h_1} g$ takes values in G_j for all $j \in \{1, \dots, d+1\}$ and all $h_1 \dots h_j \in \mathbb{Z}$. If $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ and $F : G/\Gamma \to \mathbb{C}$ is a Lipschitz function, then the sequence $\mathbb{Z} \to F(g(\cdot)\Gamma)$ is called a *nilsequence*.

With this notation in place, we restate the main result of our paper, which shows that $n \mapsto (g^{(W,A)}_{[y',y]}(n) - 1)$ is orthogonal to nilsequences.

Theorem 8.3 (noncorrelation with nilsequences). *Let N be a large positive parameter and let $K' \geqslant 1$, $K > 2K'$ and $d \geqslant 0$ be integers. Let $\frac{1}{2}\log_3 N \leqslant y' \leqslant (\log N)^{K'}$ and suppose that $(\log N)^K < y_0 < y < N^\eta$ for some sufficiently small $\eta \in (0,1)$ depending the value of d. Let $(G/\Gamma, G_\bullet)$ be a filtered nilmanifold of complexity Q_0 and degree d. Finally, let $w(N) = \frac{1}{2}\log_3 N + o(1)$, $W = P(w(N))$ and define $\delta(N) = \exp(-\sqrt{\log_4 N})$.*

If K is sufficiently large depending on the degree d of $G_\bullet$, then the estimate

$$\left| \frac{\widetilde{W}}{N} \sum_{n \leqslant (N-A)/\widetilde{W}} (g^{(\widetilde{W},A)}_{[y',y]}(n) - 1) F(g(n)\Gamma) \right| \ll_{d,C} (1 + \|F\|_{\mathrm{Lip}})\delta(N)Q_0 + \frac{1}{\log w(N)} \tag{8-1}$$

holds uniformly for all $\widetilde{W} = Wq$, where $q \leqslant (\log y_0)^C$ satisfies $p \mid q \Rightarrow p < w(N)$ and $C \geqslant 1$ is a fixed constant, for all $1 \leqslant A \leqslant \widetilde{W}$ with $\gcd(A, W) = 1$, all polynomial sequences $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ and all 1-bounded Lipschitz functions $F : G/\Gamma \to \mathbb{C}$.

By decomposing the summation range $1 \leqslant n \leqslant (N-A)/\widetilde{W}$ into dyadic intervals $n \sim (N'-A)/\widetilde{W}$ for $N' \in (N/\log N, N]$ and the initial segment $1 \leqslant n \ll N/(\widetilde{W}\log N)$, and approximating $g_{[y',y]}$ by $h_{[y',y]}$ on each dyadic interval, it follows from Lemma 4.4 that the task of proving Theorem 8.3 may be reduced to proving that the estimate

$$\left| \frac{\widetilde{W}}{N} \sum_{n \sim (N-A)/\widetilde{W}} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1) F(g(n)\Gamma) \right| \ll_{d,C} (1 + \|F\|_{\mathrm{Lip}})\delta(N)Q_0 + \frac{1}{\log w(N)} \tag{8-2}$$

holds under the assumptions of Theorem 8.3 and for the function $h^{(\widetilde{W},A)}_{[y',y]}$ that is defined on the interval $n \sim (N-A)/\widetilde{W}$.

8.1. *Reduction to noncorrelation with equidistributed nilsequences.*

Observe that the bound in the statement of Theorem 8.3 holds trivially unless

$$Q_0 \leqslant \delta(N)^{-1}.$$

This information can be used in combination with Green and Tao's factorisation theorem for polynomial sequences (which states that every polynomial sequence is the product of a slowly varying sequence, a highly equidistributed sequence and a periodic polynomial sequence) together with our results on the distribution of $n \mapsto h_{[y',y]}$ in short arithmetic progressions in order to reduce the statement to one in which the sequence g can be assumed to be equidistributed. Before stating the reduced version and proving the reduction, we recall the relevant definitions around equidistribution as well as the factorisation theorem.

Definition 8.4 (δ-equidistributed and totally δ-equidistributed sequence [Green and Tao 2012b, Definition 1.2]). Let G/Γ be a nilmanifold.

(1) Given a length $N > 0$ and an error tolerance $\delta > 0$, a finite sequence $(g(n)\Gamma)_{n\in[N]}$ is said to be δ-*equidistributed* if we have

$$\left|\mathbb{E}_{n\in[N]} F(g(n)\Gamma) - \int_{G/\Gamma} F\right| \leqslant \delta \|F\|_{\mathrm{Lip}}$$

for all Lipschitz functions $F : G/\Gamma \to \mathbb{C}$.

(2) A finite sequence $(g(n)\Gamma)_{n\in[N]}$ is said to be *totally* δ-*equidistributed* if we have

$$\left|\mathbb{E}_{n\in P} F(g(n)\Gamma) - \int_{G/\Gamma} F\right| \leqslant \delta \|F\|_{\mathrm{Lip}}$$

for all Lipschitz functions $F : G/\Gamma \to \mathbb{C}$ and all arithmetic progressions $P \subset [N]$ of length at least δN.

Definition 8.5 (rational sequence [Green and Tao 2012b, Definition 1.17]). Let G/Γ be a nilmanifold and let $Q > 0$ be a parameter. We say that $\gamma \in G$ is Q-*rational* if $\gamma^r \in \Gamma$ for some integer r, $0 < r \leqslant Q$. A Q-*rational point* is any point in G/Γ of the form $\gamma\Gamma$ for some Q-rational group element γ. A sequence $(\gamma(n))_{n\in\mathbb{Z}}$ is Q-*rational* if every element $\gamma(n)\Gamma$ in the sequence is a Q-rational point.

Definition 8.6 (smooth sequences [Green and Tao 2012b, Definition 1.18]). Let G/Γ be a nilmanifold with a Malcev basis $\mathscr{X}$. Let $(\varepsilon(n))_{n\in\mathbb{Z}}$ be a sequence in G, and let $M, N \geqslant 1$. We say that $(\varepsilon(n))_{n\in\mathbb{Z}}$ is (M, N)-*smooth* if we have $d(\varepsilon(n), \mathrm{id}_G) \leqslant M$ and $d(\varepsilon(n), \varepsilon(n-1)) \leqslant M/N$ for all $n \in [N]$, where id_G denotes the identity element of G.

The following proposition is Green and Tao's factorisation theorem [2012b, Theorem 1.19] for polynomial sequences, which asserts that any polynomial sequence can be decomposed into a product of a smooth, a highly equidistributed and a rational polynomial sequence.

Proposition 8.7 (Green–Tao factorisation theorem [2012b]; see also [Tao and Teräväinen 2023]). *Let $m, d \geqslant 1$, and let $M_0, N \geqslant 1$ and $A > 0$ be real numbers. Suppose that G/Γ is an m-dimensional nilmanifold together with a filtration $G_\bullet$ of degree d. Suppose that $\mathscr{X}$ is an M_0-rational Malcev basis adapted to $G_\bullet$ and that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$. Then there is an integer M with $M_0 \leqslant M \ll M_0^{O_{A,m,d}(1)}$, a rational subgroup $G' \subseteq G$, a Malcev basis $\mathscr{X}'$ for G'/Γ' in which each element is an M-rational combination of the elements of $\mathscr{X}$, and a decomposition $g = \varepsilon g' \gamma$ into polynomial sequences $\varepsilon, g', \gamma \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ with the following properties*:

(1) *$\varepsilon : \mathbb{Z} \to G$ is (M, N)-smooth.*

(2) *$g' : \mathbb{Z} \to G'$ takes values in G', and the finite sequence $(g'(n)\Gamma')_{n\in[N]}$ is totally $1/M^A$-equidistributed in G'/Γ', using the metric $d_{\mathscr{X}'}$ on G'/Γ'.*

(3) *$\gamma : \mathbb{Z} \to G$ is M-rational, and $(\gamma(n)\Gamma)_{n\in\mathbb{Z}}$ is periodic with period at most M.*

With the above notation in place, we may now state the equidistributed version of Theorem 8.3 and deduce this theorem from it.

Proposition 8.8 (orthogonality to equidistributed nilsequences). *Let N be a large positive parameter, let $K' \geqslant 1$, $K > 2K'$ and $d \geqslant 0$ be integers. Let $\frac{1}{2}\log_3 N \leqslant y' \leqslant (\log N)^{K'}$ and suppose that $(\log N)^K < y_0 < y \leqslant N^\eta$ for some $\eta \in (0,1)$ that is sufficiently small depending on d. Let $w(N) = \frac{1}{2}\log_3 N$, $W = P(w(N))$ and $\delta(N) = \exp(-C_0\sqrt{\log_4 N}))$ with $C_0 \in [1, (\log N)^{1/4}]$.*

Let G/Γ be a nilmanifold together with a filtration $G_\bullet$ of degree d, and suppose that $\mathscr{X}$ is a $\delta(N)^{-1}$-rational Malcev basis adapted to $G_\bullet$. Let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ be any polynomial sequence such that the finite sequence $(g(n)\Gamma)_{n\leqslant 2N/W}$ is totally $\delta(N)^{E_0}$-equidistributed for some $E_0 > 1$. Let $F : G/\Gamma \to \mathbb{C}$ be any 1-bounded Lipschitz function such that $\int_{G/\Gamma} F = 0$.

If $1 \leqslant q \leqslant (\log y_0)^C$ is $(w(N)-1)$-smooth, where $1 \leqslant C \ll 1$, if $0 \leqslant a < q$ and $0 < A < W$ are integers such that $(W, A) = 1$ (and thus $(Wa + A, Wq) = 1$), and if $0 < N_1 \leqslant N$, then we have

$$\left|\frac{Wq}{N} \sum_{\substack{m\in\mathbb{N}:\\ N<Wqm<N+N_1}} h^{(W,A)}_{[y',y]}(qm+a)F(g(m)\Gamma)\right| \ll_{d,\dim G,\|F\|_{\mathrm{Lip}},E_1} \delta(N)^{E_1} \tag{8-3}$$

for any given $E_1 \geqslant 1$, provided that E_0 is sufficiently large with respect to d, $\dim G$ and E_1, provided that K is sufficiently large depending on the degree d of $G_\bullet$, and provided that N is sufficiently large in terms of $\dim G$, d and E_0.

Proposition 8.8 implies Theorem 8.3. We shall prove that Proposition 8.8 implies (8-2), from which Theorem 8.3 follows. We may assume that $Q_0 \leqslant \delta(N)^{-1}$ as (8-2) is trivially true otherwise. Let $B > 1$ be a parameter. Then, by Proposition 8.7, applied with N replaced by $2N/\widetilde{W}$, there exists $Q_0 \leqslant Q \ll Q_0^{O_{B,\dim G,d}(1)}$ and a factorisation of the polynomial sequences g as $\varepsilon g'\gamma$ that satisfies properties (1)–(3) of that proposition. In particular, the polynomial sequence $\gamma : \mathbb{Z} \to G$ gives rise to a $\tilde{q}$-periodic function $\gamma(\cdot)\Gamma : \mathbb{Z} \to G/\Gamma$ for some period $1 \leqslant \tilde{q} \leqslant Q$, and the sequence $\varepsilon : \mathbb{Z} \to G$ is $(Q, 2N/\widetilde{W})$-smooth. The sequence $g' : \mathbb{Z} \to G'$ takes values in a Q-rational subgroup G' of G, it is a polynomial sequence with respect to the filtration $G'_\bullet := G_\bullet \cap G'$ and the finite sequence $(g'(n)\Gamma')_{n\leqslant 2N/\widetilde{W}}$ is Q^{-B}-equidistributed in G'/Γ', where $\Gamma' = \Gamma \cap G'$ and where equidistribution is defined with respect to the metric $d_{\mathscr{X}'}$ arising from a Malcev basis $\mathscr{X}'$ adapted to $G'_\bullet$. The existence of the Malcev basis $\mathscr{X}'$ is guaranteed by [Green and Tao 2012b, Lemma A.10], which also allows us to assume that each of its basis elements is a Q-rational combination of basis elements from $\mathscr{X}$.

In order to reduce the noncorrelation estimate to the case where the polynomial sequence is highly equidistributed, we seek to decompose the summation range of n in (8-2) into (short) subprogressions on which γ and ε are almost constant. Splitting the interval $(N/\widetilde{W}, 2N/\widetilde{W}]$ into arithmetic progressions with common difference $\tilde{q}$, we obtain

$$\sum_{n\sim N/\widetilde{W}} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)F(g(n)\Gamma) = \sum_{0\leqslant a<\tilde{q}} \sum_{\substack{n\sim N/\widetilde{W}\\ n\equiv a \pmod{\tilde{q}}}} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)F(\varepsilon(n)g'(n)\gamma_a\Gamma),$$

where $\gamma_a \in G$ is such that $\gamma(n)\Gamma = \gamma_a\Gamma$ whenever $n \equiv a \pmod{\tilde{q}}$.

Since F is a Lipschitz function and since $d_{\mathscr{X}}$ is right-invariant (see [Green and Tao 2012b, Appendix A]), we deduce that, for any $n_0, n \in \mathbb{Z}$,

$$\begin{aligned}\big|F(\varepsilon(n_0)g'(n)\gamma_a\Gamma) - F(\varepsilon(n)g'(n)\gamma_a\Gamma)\big| &\leqslant \|F\|_{\mathrm{Lip}}\, d_{\mathscr{X}}\big(\varepsilon(n_0)g'(n)\gamma_a, \varepsilon(n)g'(n)\gamma_a\big) \\ &= \|F\|_{\mathrm{Lip}}\, d_{\mathscr{X}}(\varepsilon(n_0), \varepsilon(n)).\end{aligned}$$

It then follows from the assumption ε is $(Q, 2N/\widetilde{W})$-smooth that

$$d_{\mathscr{X}}(\varepsilon(n_0), \varepsilon(n)) \leqslant \frac{Q|n_0 - n|}{N/\widetilde{W}}$$

whenever $n, n_0 \leqslant 2N/\widetilde{W}$, and therefore

$$\big|F(\varepsilon(n_0)g'(n)\gamma_a\Gamma) - F(\varepsilon(n)g'(n)\gamma_a\Gamma)\big| \ll_{\|F\|_{\mathrm{Lip}}} \log^{-1} Q \tag{8-4}$$

if, in addition, $|n_0 - n| \ll N/(\widetilde{W}Q\log Q)$. With this in mind, we refine the partition of our summation range and consider a subpartition

$$(N/\widetilde{W}, 2N/\widetilde{W}] = \bigcup_j P_j,$$

where each P_j is a "short" progression of common difference $\tilde{q}$, as before, but with diameter bounded by $O(N/(Q\widetilde{W}\log Q))$. The bound on the diameter ensures that ε is almost constant on each P_j. Note that the total number of short progressions P_j is $O(\tilde{q}Q\log Q)$. By fixing an element $\varepsilon_j \in \varepsilon(P_j)$ in the image of P_j under ε for each progression P_j, we obtain

$$\begin{aligned}&\sum_{n\sim N/\widetilde{W}} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)F(g(n)\Gamma) \\ &\quad = \sum_j \sum_{n\in P_j} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)F(\varepsilon_j g'(n)\gamma_a\Gamma) \\ &\qquad\qquad + O\bigg\{\sup_{P_j}\sup_{n\in P_j}\big|F(\varepsilon_j g'(n)\gamma_a\Gamma) - F(\varepsilon(n)g'(n)\gamma_a\Gamma)\big|\bigg(\sum_{n\sim N/\widetilde{W}} h^{(\widetilde{W},A)}_{[y',y]}(n) + \frac{N}{\widetilde{W}}\bigg)\bigg\}.\end{aligned} \tag{8-5}$$

By construction or, more precisely, by Lemma 5.5, we have

$$\sum_{n\sim N/\widetilde{W}} h^{(\widetilde{W},A)}_{[y',y]}(n) \ll N/\widetilde{W}.$$

Bounding the error in (8-5) with the help of this bound and (8-4), we obtain

$$\sum_{n\sim N/\widetilde{W}} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)F(g(n)\Gamma) = \sum_j \sum_{n\in P_j} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)F(\varepsilon_j g'(n)\gamma_a\Gamma) + O\bigg(\frac{N}{\widetilde{W}\log Q}\bigg).$$

Observe that in the argument of F, apart from two constant factors, only the sequence g' occurs, which has the property that $(g'(n)\Gamma)_{n\leqslant 2N/\widetilde{W}}$ is Q^{-B}-equidistributed. We aim to use this equidistribution property in combination with Proposition 8.8 in order to bound the correlations on the right-hand side above. For this purpose, we shall now first show that the sequence $n \mapsto \varepsilon_j g'(n)\gamma_a$ can be reinterpreted as a polynomial sequence $n \mapsto g^*(n)$ that is equidistributed on some filtered nilmanifold H/Λ. At the same time, we show that $F(\varepsilon_j g'(n)\gamma_a \Gamma) = \widetilde{F}(g^*(n)\Lambda)$ for a Lipschitz function $\widetilde{F} : H/\Lambda \to \mathbb{C}$. As a second step, we carry out a reduction that allows us to assume that $\int_{H/\Lambda} \widetilde{F} = 0$. And, thirdly and finally, we will apply Proposition 8.8 and complete the proof. We denote these steps (1)–(3) below.

(1) Define $g^*(n) := \gamma_a^{-1} g'(n)\gamma_a$. For an application of Proposition 8.8, it is necessary to verify that g^* is a polynomial sequence and that it inherits the equidistribution properties of g'. These questions have been addressed by Green and Tao [2012a, Section 2] and we follow their argument here. Let $H = \gamma_a^{-1} G' \gamma_a$ and define $H_\bullet = \gamma_a^{-1}(G')_\bullet \gamma_a$. Let $\Lambda = \Gamma \cap H$ and define $\widetilde{F} = \widetilde{F}_{a,j} : H/\Lambda \to \mathbb{R}$ via

$$\widetilde{F}(x\Lambda) = F(\varepsilon_j \gamma_a x \Gamma).$$

Then $g^* \in \mathrm{poly}(\mathbb{Z}, H_\bullet)$, we have $\widetilde{F}(g^*(n)\Lambda) = F(\varepsilon_j g'(n)\gamma_a \Gamma)$, and the correlation that we seek to bound takes the form

$$\sum_{n \in P_j} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)\widetilde{F}(g^*(n)\Lambda). \tag{8-6}$$

The "Claim" from the end of [Green and Tao 2012a, Section 2] guarantees the existence of a Malcev basis $\mathscr{Y}$ for H/Λ adapted to $H_\bullet$ such that each basis element Y_i is a $Q^{O(1)}$-rational combination of basis elements X_i from $\mathscr{X}$. Thus, there is $C' = O(1)$ such that $\mathscr{Y}$ is $Q^{C'}$-rational. Furthermore, the "Claim" implies that there is $c' > 0$, depending only on $\dim G$ and on the degree d of $G_\bullet$, such that whenever B is sufficiently large the sequence

$$(g^*(n)\Lambda)_{n \leqslant 2N/\widetilde{W}} \tag{8-7}$$

is totally $Q^{-c'B+O(1)}$-equidistributed in H/Λ, equipped with the metric $d_{\mathscr{Y}}$ induced by $\mathscr{Y}$. Taking B sufficiently large, we may assume that the sequence (8-7) is totally $M^{-c'B/2}$-equidistributed. Finally, the "Claim" also provides the bound $\|\widetilde{F}\|_{\mathrm{Lip}} \leqslant Q^{C''}\|F\|_{\mathrm{Lip}}$ for some $C'' = O(1)$. This shows that all conditions of Proposition 8.8 are satisfied except for $\int_{H/\Lambda} \widetilde{F} = 0$.

(2) Let $\mu(\widetilde{F}) = \int_{H/\Gamma} \widetilde{F}$ denote the mean value of $\widetilde{F}$ and observe that $\mu(\widetilde{F}) \ll 1$ since $\widetilde{F}$ is 1-bounded. Then $\int_{H/\Gamma} \bar{F} = 0$ if $\bar{F} := \widetilde{F} - \mu(\widetilde{F}) : H/\Gamma \to \mathbb{C}$ and

$$\begin{aligned}\sum_{n \in P_j} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)\widetilde{F}(g^*(n)\Lambda) &= \sum_{n \in P_j} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)\bar{F}(g^*(n)\Lambda) + O\biggl(\biggl|\sum_{n \in P_j} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)\biggr|\biggr) \\ &= \sum_{n \in P_j} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)\bar{F}(g^*(n)\Lambda) + O\biggl(\frac{|P_j|}{\log w(N)}\biggr)\end{aligned}$$

by Lemma 5.5. We may thus assume that $\int_{H/\Gamma} \widetilde{F} = 0$.

(3) By the previous two steps it remains to bound

$$\sum_j \sum_{n \in P_j} (h^{(\widetilde{W},A)}_{[y',y]}(n) - 1)\widetilde{F}(g^*(n)\Lambda),$$

where we may assume that $\int_{H/\Gamma} \widetilde{F} = 0$ and that $(g^*(n)\Lambda)_{n \leqslant 2N/\widetilde{W}}$ is $Q^{-c'B/2}$-equidistributed. Since P_j has common difference $\tilde{q} < Q$, the bound on the diameter of P_j implies that $|P_j| \gg N/(\tilde{q} Q \widetilde{W} \log Q)$.

Note further that $Q \ll Q_0^{O_{B,d,\dim G}(1)} \ll \delta(N)^{-C}$ for some $C = O_{B,d,\dim G}(1)$ since $Q_0 \leqslant \delta(N)^{-1}$. We may suppose that N is sufficiently large for $C \leqslant (\log N)^{1/4}$ to hold. Define $\delta'(N) = \delta(N)^{-C}$.

We may thus apply Proposition 8.8 with Wq replaced by $\widetilde{W}\tilde{q} = W(q\tilde{q})$, with δ replaced by δ', with $E_0 = c'B/2$, $g = g^*$, $G/\Gamma = H/\Lambda$, $\mathscr{X} = \mathscr{Y}$, and with $N_1 \gg N/(Q \log Q) > N\delta'(N)^{E_1/2}$ (assuming that $E_1 > 2$, which we may) to deduce that

$$\begin{aligned}\sum_j \sum_{n \in P_j} (h_{[y',y]}^{(\widetilde{W},A)}(n) - 1)\widetilde{F}(g^*(n)\Lambda) &\ll_d \sum_j (1 + \|\widetilde{F}\|_{\mathrm{Lip}})\delta'(N)^{E_1} \frac{N}{\widetilde{W}\tilde{q}} \\ &\ll_d (1 + Q^{C''}\|F\|_{\mathrm{Lip}})\delta'(N)^{E_1} \frac{\tilde{q} QN \log Q}{\widetilde{W}\tilde{q}} \\ &\ll_d (1 + \|F\|_{\mathrm{Lip}}) Q^{O(1)} \delta'(N)^{E_1} \frac{N}{\widetilde{W}} \\ &\ll_d (1 + \|F\|_{\mathrm{Lip}})\delta(N) Q_0 \frac{N}{\widetilde{W}},\end{aligned}$$

provided that B, and hence E_1, is sufficiently large to imply the final bound. This completes the proof of the deduction of (8-2) and, hence, Theorem 8.3. □

8.2. ***Sparse families of linear subsequences of equidistributed nilsequences.*** In Section 9.2 we will relate our task of bounding the one-parameter correlation (8-3) to that of bounding a bilinear sum. This reduction naturally leads to the problem of understanding equidistribution properties in families of linear subsequences of polynomial sequences that arise as follows. Let $(g(n)\Gamma)_{n \leqslant N}$ be a polynomial sequence and consider the family of sequences

$$\{n \mapsto (g(mn)\Gamma)_{n \leqslant N/m}\}_{m \in S([y',y]) \cap [M,2M)},$$

where the parameter $m \in [M, 2M)$ is further restricted to the sparse set $S([y', y])$, yielding a sparse family of subsequences. Our aim in this subsection is to show that if $(g(n)\Gamma)_{n \leqslant N}$ is δ-equidistributed for a suitable choice of δ, then almost all sequences in this family are $\delta^{1/C}$-equidistributed. Equidistribution properties in unrestricted families of linear subsequences have been studied by the first author in [Matthiesen 2018, Section 7]. In this section we show that, thanks to the sparse recurrence result from Section 7, the unrestricted result [loc. cit., Proposition 7.4] can indeed be extended to the sparse situation where $m \in S([y', y])$. Part of this section follows [loc. cit., Section 7] very closely.

The proof of Proposition 8.11 below uses the notion of a *horizontal character* [Green and Tao 2012b, Definition 1.5] on a nilmanifold G/Γ, which is defined to be a continuous additive homomorphism $\eta : G \to \mathbb{R}/\mathbb{Z}$ which annihilates Γ. The set of horizontal characters may be equipped with a height function $|\eta|$ as defined in [loc. cit., Definition 2.6]. This specific height function is called the *modulus* of η. All that is relevant to us in the present paper is that there are at most $Q^{O(1)}$ horizontal characters $\eta : G \to \mathbb{R}/\mathbb{Z}$ of modulus $|\eta| \leqslant Q$.

If $\eta : G \to \mathbb{R}/\mathbb{Z}$ is a horizontal character and $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$, where $G_\bullet$ is a filtration of degree d, then $\eta \circ g : \mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is a polynomial of degree at most d. For an arbitrary polynomial $P : \mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ of degree at most d, we may define two sets of coefficients, $\alpha_0, \dots, \alpha_d$ and $\beta_0, \dots, \beta_d$, in $\mathbb{R}/\mathbb{Z}$ via

$$P(n) = \alpha_0 + \alpha_1 \binom{n}{1} + \cdots + \alpha_d \binom{n}{d} = \beta_d n^d + \cdots + \beta_1 n + \beta_0.$$

The *smoothness norm* [Green and Tao 2012b, Definition 1.5] of P with respect to N is then defined to be

$$\|P\|_{C^\infty[N]} = \sup_{1 \leqslant j \leqslant d} N^j \|\alpha_j\|_{\mathbb{R}/\mathbb{Z}}$$

and we have

$$\|P\|_{C^\infty[N]} \ll_d \sup_{1 \leqslant j \leqslant d} N^j \|\beta_j\|_{\mathbb{R}/\mathbb{Z}} \quad \text{and} \quad \sup_{1 \leqslant j \leqslant d} N^j \|q\beta_j\|_{\mathbb{R}/\mathbb{Z}} \ll \|P\|_{C^\infty[N]}$$

for some positive integer $q \ll_d 1$ by [loc. cit., Lemma 3.2].

Smoothness norms and horizontal characters allow one to characterise δ-equidistributed polynomial sequences in the following sense:

Lemma 8.9 [Green and Tao 2012b, Theorem 2.9]. *Let m_G and d be nonnegative integers, let $0 < \delta < \frac{1}{2}$ and let $N \geqslant 1$. Suppose that G/Γ is an m_G-dimensional nilmanifold together with a filtration $G_\bullet$ of degree d and that $\mathscr{X}$ is a δ^{-1}-rational Malcev basis adapted to $G_\bullet$. Suppose $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$. If $(g(n)\Gamma)_{n \leqslant N}$ is not δ-equidistributed, then there exists a nontrivial horizontal character η with $0 < |\eta| \ll \delta^{-O_{d,m_G}(1)}$ such that*

$$\|\eta \circ g\|_{C^\infty[N]} \ll \delta^{-O_{d,m_G}(1)},$$

where C_d is a sufficiently large constant depending only on d.

In order to pass between the notions of equidistribution and total equidistribution for polynomial sequences, we shall use the following lemma, which is [Matthiesen 2018, Lemma 7.2].

Lemma 8.10. *Let N and A be positive integers and let $\delta : \mathbb{N} \to [0, 1]$ be a function that satisfies $\delta(x)^{-t} \ll_t x$ for all $t > 0$. Suppose that G has a $\delta(N)^{-1}$-rational Malcev basis adapted to the filtration $G_\bullet$. Then there is $1 \leqslant B \ll_{d,\dim G} 1$ such that the following holds provided N is sufficiently large. If $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ is a polynomial sequence such that $(g(n)\Gamma)_{n \leqslant N}$ is $\delta(N)^A$-equidistributed for some $A > B$, then $(g(n)\Gamma)_{n \leqslant N}$ is totally $\delta(N)^{A/B}$-equidistributed.*

With these preparations in place, we now turn towards the main result of this section.

Proposition 8.11 (equidistribution in sparse families of linear subsequences). *Let N be a large positive parameter, let $d \geqslant 0$, and let $K' > 0$ and $K > \max(2K', 2)$. Suppose that $1 \leqslant y' \leqslant (\log N)^{K'}$ and $(\log N)^K < y \leqslant N^\mu$, where $\mu = \mu(d) \in (0, 1)$ is sufficiently small depending on d. Let $\delta : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be a function of N that satisfies $\log_5 N \ll \delta(N)^{-1}$ and $\delta(N)^{-B} \ll_B \log_2 N$ for all $B > 0$. Suppose that $(G/\Gamma, G_\bullet)$ is a nilmanifold together with a filtration $G_\bullet$ of degree d and a $\delta(N)^{-1}$-rational Malcev basis adapted to it. Let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ be a polynomial sequence and suppose that the finite sequence $(g(n)\Gamma)_{n \leqslant N}$ is totally $\delta(N)^{E_1}$-equidistributed in G/Γ for some $E_1 \geqslant 1$. Then there is some $c_1 \in (0, 1)$ depending on d and $\dim G$ such that the following assertion holds for all integers*

$$M \in [N^{1/2}, N/y^{1/2}] \tag{8-8}$$

provided that K is sufficiently large depending on d, and that $c_1 E_1 \geqslant 1$.

Given any sequence $(A_m)_{m\in\mathbb{N}}$ of integers satisfying $|A_m| \leqslant m$, write $g_m(n) := g(mn + A_m)$ and let $\mathscr{B}_M$ denote the set of integers $m \in [M, 2M) \cap S([y', y])$ for which

$$(g_m(n)\Gamma)_{n\leqslant N/m}$$

fails to be totally $\delta(N)^{c_1 E_1}$-equidistributed. Then

$$\#\mathscr{B}_M \ll \Psi(2M, [y', y])\delta(N)^{c_1 E_1}.$$

Proof. Let M be a fixed integer in the range (8-8) and let $c_1 > 0$ to be determined in the course of the proof. Suppose that $E_1 > 1/c_1$. It follows from Lemma 8.10 that for every $m \in \mathscr{B}_M$, the sequence $(g_m(n)\Gamma)_{n\leqslant N/m}$ fails to be $\delta(N)^{c_1 E_1 B}$-equidistributed on G/Γ for some $1 \leqslant B \ll_{d,\dim G} 1$. By Lemma 8.9, there is a nontrivial horizontal character $\eta_m : G \to \mathbb{R}/\mathbb{Z}$ of magnitude $|\eta_m| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}$ such that

$$\|\eta_m \circ g_m\|_{C^\infty[N/M]} \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}. \tag{8-9}$$

For each nontrivial horizontal character $\eta : G \to \mathbb{R}/\mathbb{Z}$ we define the set

$$\mathscr{M}_\eta = \{m \in \mathscr{B}_M : \eta_m = \eta\}.$$

Note that this set is empty unless $|\eta| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}$. Suppose that

$$\#\mathscr{B}_M \geqslant \Psi(2M, [y', y])\delta(N)^{c_1 E_1}.$$

Since there are only $M^{O(1)}$ horizontal characters of modulus bounded by M, it follows from the pigeonhole principle that there is some η of modulus $|\eta| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}$ such that

$$\#\mathscr{M}_\eta \geqslant \Psi(2M, [y', y])\delta(N)^{c_1 E_1 C}$$

for some $C \asymp_{d,\dim G} 1$. Suppose

$$\eta \circ g(n) = \beta_d n^d + \cdots + \beta_1 n + \beta_0.$$

Then

$$\eta \circ g_m(n) = \eta \circ g(mn + A_m) = \alpha_d^{(m)} n^d + \cdots + \alpha_1^{(m)} n + \alpha_0^{(m)},$$

where

$$\alpha_j^{(m)} = m^j \sum_{i=j}^{d} \binom{i}{j} A_m^{i-j} \beta_i \quad (0 \leqslant j \leqslant d). \tag{8-10}$$

The bound (8-9) on the smoothness norm asserts that

$$\sup_{1\leqslant j\leqslant d} \frac{N^j}{M^j} \|\alpha_j^{(m)}\| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)},$$

which by downwards induction combined with (8-10) implies

$$\sup_{1\leqslant j\leqslant d} \frac{N^j}{M^j} \|\beta_j m^j\| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}.$$

Hence,

$$\|\beta_j m^j\| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)} (M/N)^j \quad (1 \leqslant j \leqslant d)$$

for every $m \in \mathscr{M}_\eta$.

In view of the lower bound on $\#\mathscr{M}_\eta$, we seek to apply Theorem 7.1 with $k = j \leqslant d$ and with the cut-off parameter N in the theorem replaced by M. Observe that all assumptions in Theorem 7.1 about the relation between the parameters N, y, y', as well as those on the function δ and the size of ε are sufficiently flexible to allow for N to be replaced by any M with $\log M \asymp \log N$. Note that $N/M > y^{1/2} \geqslant (\log N)^{K/2} \gg (\log M)^{K/2}$, which allows us to choose

$$\varepsilon = \delta(N)^{-O_{d,\dim G}(c_1 E_1)}(M/N)^j < \delta(N)^{-O_{d,\dim G}(c_1 E_1)} y^{-j/2} < \delta(N)^{-O_{d,\dim G}(c_1 E_1)}(\log M)^{-jK/2}$$

and $C_1 = jK/2$ in the application of Theorem 7.1. Our assumptions on $\delta(N)$ imply that $\varepsilon < \delta^*/2$ for any δ^* of the form $\delta(N)^{c_1 E_1 C}$ with $C \asymp_{d,\dim G} 1$ as soon as N is sufficiently large. Hence, it follows from Theorem 7.1 that there exists an integer $1 \leqslant q_j \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}$ such that

$$\|q_j \beta_j\| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}(M/N)^{-j} M^{-j} = \delta(N)^{-O_{d,\dim G}(c_1 E_1)} N^{-j}.$$

Thus,

$$\beta_j = \frac{a_j}{\kappa_j} + \tilde{\beta}_j, \tag{8-11}$$

where $\kappa_j \mid q_j$, $\gcd(a_j, \kappa_j) = 1$ and

$$0 \leqslant \tilde{\beta}_j \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)} N^{-j}.$$

Hence,

$$\|\kappa_j \beta_j\| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)} N^{-j}. \tag{8-12}$$

Let $\kappa = \operatorname{lcm}(\kappa_1, \dots, \kappa_d)$ and set $\tilde{\eta} = \kappa\eta$. We proceed in a similar fashion as in [Green and Tao 2012a, Section 3]: The above implies that

$$\|\tilde{\eta} \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)} n/N \quad (n \leqslant N),$$

which is small provided n is not too large. Indeed, if $N' = \delta(N)^{c_1 E_1 C''} N$ for some sufficiently large constant $C'' \geqslant 1$ depending only on d and $\dim G$, and if $n \in \{1, \dots, N'\}$, then

$$\|\tilde{\eta} \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \leqslant \tfrac{1}{10}.$$

Let $\chi : \mathbb{R}/\mathbb{Z} \to [-1, 1]$ be a function of bounded Lipschitz norm that equals 1 on $\left[-\frac{1}{10}, \frac{1}{10}\right]$ and satisfies $\int_{\mathbb{R}/\mathbb{Z}} \chi(t)\, \mathrm{d}t = 0$. Then, by setting $F := \chi \circ \tilde{\eta}$, we obtain a Lipschitz function $F : G/\Gamma \to [-1, 1]$ that satisfies $\int_{G/\Gamma} F = 0$ and $\|F\|_{\mathrm{Lip}} \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}$. By choosing c_1 sufficiently small depending on d and $\dim G$ we may ensure that

$$\|F\|_{\mathrm{Lip}} < \delta(N)^{-E_1}$$

and, moreover, that

$$N' > \delta(N)^{E_1} N.$$

This choice of N', F and c_1 implies that, for all sufficiently large values of N, we have

$$\left| \frac{1}{N'} \sum_{1 \leqslant n \leqslant N'} F(g(n)\Gamma) \right| = 1 > \delta(N)^{E_1} \|F\|_{\mathrm{Lip}},$$

which contradicts our assumption that $(g(n)\Gamma)_{n \leqslant N}$ is totally $\delta(N)^{E_1}$-equidistributed. □

9. Noncorrelation with equidistributed nilsequences

In this section we prove Proposition 8.8, that is to say we show that the function $m \mapsto \mathbf{1}_{S([y',y])}(\widetilde{W}m + A')$ and its weighted version $m \mapsto h^{(W,A)}_{[y',y]}(qm+a)$ are orthogonal to highly equidistributed nilsequences for $w(N)$-smooth values of q and $0 \leqslant a < q$.

Both of these functions are W-tricked versions of *sparse multiplicative functions*. Building on a Montgomery–Vaughan-type decomposition, which allows one to replace twisted sums of multiplicative functions by bilinear expressions, it was proved in [Matthiesen 2018] that W-tricked and centralised versions of "*dense*" *multiplicative functions* are orthogonal to nilsequences. We shall show that this approach can be extended to the sparse setting we are looking at here in order to establish Proposition 8.8.

9.1. ***Removing the weight.*** We will prove Proposition 8.8 by a sequence of reductions. Recall that

$$h^{(W,A)}_{[y',y]}(m) = \frac{\phi(W)}{W} h_{[y',y]}(Wm+A) = \frac{\phi(W)}{W} \frac{N^{\alpha}(Wm+A)^{1-\alpha}}{\alpha \Psi(N,[y',y])} \mathbf{1}_{S([y',y])}(Wm+A),$$

where $\gcd(A, W) = 1$. Our first step is to remove the weight and reduce the correlation estimate for $h^{(W,A)}_{[y',y]}(qm+a)$ to one that only involves the characteristic function $\mathbf{1}_{S([y',y])}(\widetilde{W}m + A')$, where $\widetilde{W} = Wq$ and $A' = Wa + A$.

Lemma 9.1 (removing the weight). *Let $\widetilde{W} = Wq$ and $A' = Wa + A$ and define for any $x \in \mathbb{N}$ the quantity*

$$T(x) := \sum_{N/\widetilde{W} < m \leqslant x} \mathbf{1}_{S([y',y])}(\widetilde{W}m + A') F(g(m)\Gamma).$$

Then the conclusion (8-3) *of Proposition 8.8 holds provided that*

$$T(x) \ll_{d,\dim G,\|F\|_{\mathrm{Lip}},E_1} \delta(N)^{E_1} \frac{\Psi(N,[y',y])}{\phi(\widetilde{W})} \tag{9-1}$$

for all $x \in (N/\widetilde{W}, 2N/\widetilde{W}]$.

Proof. To start with, recall that

$$\phi(\widetilde{W})^{-1} = (\phi(W)q)^{-1} \prod_{p \mid q, p > w(N)} (1-p^{-1})^{-1} \asymp (\phi(W)q)^{-1}$$

by (5-12) and note that

$$\frac{N^{\alpha}(\widetilde{W}m + A')^{1-\alpha}}{\alpha N} \asymp 1$$

for $N/\widetilde{W} < m < (N+N_1)/\widetilde{W}$. With this in mind, partial summation shows that

$$\frac{\widetilde{W}}{N} \frac{\phi(W)}{W} \sum_{N/\widetilde{W} < m < (N+N_1)/\widetilde{W}} h_{[y',y]}(\widetilde{W}m + A') F(g(m)\Gamma)$$

$$= \frac{1}{\Psi(N,[y',y])} \frac{\widetilde{W}\phi(W)}{W} \sum_{N/\widetilde{W} < m < (N+N_1)/\widetilde{W}} \frac{N^{\alpha}(\widetilde{W}m + A')^{1-\alpha}}{\alpha N} \mathbf{1}_{S([y',y])}(\widetilde{W}m + A') F(g(m)\Gamma)$$

$$\ll \frac{|T((N+N_1)/\widetilde{W})|\phi(\widetilde{W})}{\Psi(N,[y',y])} + \frac{\phi(\widetilde{W})}{\Psi(N,[y',y])} \sum_{N/\widetilde{W}<m<(N+N_1)/\widetilde{W}} \frac{|T(m)|}{\alpha} \left| \left(\frac{\widetilde{W}(m+1)+A'}{N}\right)^{1-\alpha} - \left(\frac{\widetilde{W}m+A'}{N}\right)^{1-\alpha} \right|$$

$$\ll \frac{\phi(\widetilde{W})}{\Psi(N,[y',y])} \left\{ \left| T\left(\frac{N+N_1}{\widetilde{W}}\right)\right| + \sum_{N/\widetilde{W}<m<(N+N_1)/\widetilde{W}} \frac{|T(m)|(1-\alpha)}{\alpha N} \left| \int_{\widetilde{W}(m+1)+A'}^{\widetilde{W}m+A'} (N/t)^{\alpha}\, \mathrm{d}t \right| \right\}$$

$$\ll \frac{\phi(\widetilde{W})}{\Psi(N,[y',y])} \left\{ \left| T\left(\frac{N+N_1}{\widetilde{W}}\right)\right| + \frac{\widetilde{W}}{N} \sum_{N/\widetilde{W}<m<(N+N_1)/\widetilde{W}} |T(m)| \right\}$$

$$\ll_{d,\dim G,\|F\|_{\mathrm{Lip}},E_1} \delta(N)^{E_1},$$

provided that the stated bounds on $T(x)$ hold. □

9.2. ***Montgomery–Vaughan-type reduction to a Type II estimate.*** Our proof strategy for establishing (9-1) is to employ a Montgomery–Vaughan-type decomposition which replaces the given one-parameter correlation by a bilinear sum and eventually allows us to reduce (9-1) to a noncorrelation estimate of the von Mangoldt function with nilsequences. Since the parameter y can be very small and since the correlation involving the von Mangoldt function will have length y^c in the end, this approach requires a careful choice of the cut-off parameter $w(N)$ in the W-trick and the parameter $\delta(N)$ that controls the level of equidistribution in the nilsequence to ensure that uniform noncorrelation estimates for the von Mangoldt function, valid over the whole range produced by the Montgomery–Vaughan decomposition, can be deduced.

Proposition 9.2 (reduction to a bilinear correlation). *The bound* (9-1) *holds for all* $x \in (N/\widetilde{W}, 2N/\widetilde{W}]$ *provided that for any sequence* $(A_n)_{n\in\mathbb{N}}$ *with* $|A_n| \leqslant n$, *for any sequence* $(A'_n)_{n\in\mathbb{N}}$ *of integers* $0 < A'_n < \widetilde{W}$ *coprime to* $\widetilde{W}$ *and for any sufficiently large* $E_0 > 1$, *we have*

$$\frac{1}{\log N} \left| \sum_{1<n\leqslant N+N_1} \sum_{\substack{\widetilde{W}m+A'_n \in [y',y] \\ N<n(\widetilde{W}m+A'_n)\leqslant N+N_1}} \mathbf{1}_{S([y',y])}(n)\Lambda(\widetilde{W}m+A'_n)F(g(mn+A_n)\Gamma) \right| \ll \delta(N)^{E_1} \frac{\Psi(N,[y',y]))}{\phi(W)q}$$

for all $N_1 \in (0, N]$ *and all* $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ *such that* $(g(n)\Gamma)_{n\leqslant N}$ *is totally* $\delta(N)^{E_0}$*-equidistributed. The implied constant is allowed to depend on* d, $\dim G$, $\|F\|_{\mathrm{Lip}}$ *and* E_1.

Proposition 9.2 will be deduced from the following simple lemma, which is inspired by a bound from [Montgomery and Vaughan 1977], but does not involve a second moment.

Lemma 9.3. *Suppose that* $f : \mathbb{N} \to \mathbb{C}$, *let* N *be a positive parameter and let* $\mathscr{S}(N) \subset [N/2, N] \cap \mathbb{N}$ *denote a set. Then*

$$\sum_{n\in\mathscr{S}(N)} f(n) \leqslant \frac{1}{\log N} \sum_{n\in\mathscr{S}(N)} |f(n)| + \frac{1}{\log N} \left| \sum_{nm\in\mathscr{S}(N)} f(nm)\Lambda(m) \right|.$$

Proof. This is an immediate consequence of the bound

$$\sum_{N/2\leqslant n\leqslant N} \log(N/n) f(n) \leqslant \log 2 \sum_{N/2\leqslant n\leqslant N} |f(n)|. \qquad \square$$

Note that we have

$$T(x) = \sum_{N/\widetilde{W}<m\leqslant x} \mathbf{1}_{S([y',y])}(\widetilde{W}m+A')F(g(m)\Gamma) = \sum_{\substack{N<n\leqslant x\widetilde{W}\\ n\equiv A' \pmod{\widetilde{W}}}} \mathbf{1}_{S([y',y])}(n)F\left(g\left(\frac{n-A'}{\widetilde{W}}\right)\Gamma\right).$$

Writing $N_1 := x\widetilde{W} - N \in (0, N]$ and applying Lemma 9.3 to this expression yields

$$T(x) \ll \frac{1}{\log N}\left|\sum_{\substack{N<mn\leqslant N+N_1\\ mn\equiv A' \pmod{\widetilde{W}}}} \mathbf{1}_{S([y',y])}(mn)\Lambda(m)F\left(g\left(\frac{nm-A'}{\widetilde{W}}\right)\Gamma\right)\right| + O\left(\frac{\Psi(N,[y',y];\widetilde{W},A')}{\log N}\right). \quad (9\text{-}2)$$

If $A_n' \in \{0, \widetilde{W}-1\}$ is such that $nA_n' \equiv A' \pmod{\widetilde{W}}$, we may expand the congruence condition on m and replace m by $\widetilde{W}m + A_n'$. With this choice of A_n' we have

$$\frac{n(\widetilde{W}m+A_n')-A'}{\widetilde{W}} = nm + A_n$$

for some integer A_n with $|A_n| \leqslant n$. Hence the first term in the bound on $T(x)$ may be rewritten as

$$\mathscr{M}(x) := \frac{1}{\log N}\left|\sum_{1\leqslant n\leqslant N+N_1}\ \sum_{N/n<\widetilde{W}m+A_n'\leqslant (N+N_1)/n} \mathbf{1}_{S([y',y])}(n(\widetilde{W}m+A_n'))\Lambda(\widetilde{W}m+A_n')F(g(nm+A_n)\Gamma)\right|.$$

The error term in (9-2) is negligible in view of Theorem 5.1, Lemma 3.3 and (3-3), provided that $\log^{-1} N \ll_{E_1} \delta(N)^{E_1}$ which holds for the choice of $\delta(N)$ for Proposition 8.8. Thus (9-1) holds provided we can show that, under the assumptions of Proposition 8.8, the bound

$$\mathscr{M}(x) \ll_{d,\|F\|_{\mathrm{Lip}}} \delta(N)^{E_1}\frac{\Psi(N,[y',y]))}{\phi(\widetilde{W})} \quad (9\text{-}3)$$

holds for all $N_1 \in (0, N]$. To complete the proof of Proposition 9.2, it remains to show that we can replace the function $\Lambda(m)\mathbf{1}_{S([y',y])}(m)$ by $\Lambda(m)\mathbf{1}_{[y',y]}(m)$ in the condition (9-3). This is the content of the following lemma.

Lemma 9.4 (removing large prime powers). *We have*

$$\frac{1}{\log N}\sum_{\substack{N<nm\leqslant N+N_1\\ nm\equiv A' \pmod{\widetilde{W}}\\ \Omega(m)\geqslant 2,\ m>y}} \mathbf{1}_{S([y',y])}(nm)\Lambda(m) \ll_B \delta(N)^B\frac{\Psi(N,[y',y]))}{\phi(\widetilde{W})}.$$

Proof. Theorem 5.1 and Lemma 3.4(i) show that the left-hand side above is bounded by

$$\begin{aligned}&\ll \frac{\log y}{\log N}\sum_{p\leqslant y}\ \sum_{k\geqslant \max(2,(\log y)/\log p)}\ \max_{(A'',\widetilde{W})=1}\Psi\left(\frac{N}{p^k},[y',y];\widetilde{W},A''\right)\\ &\ll \frac{\log y}{\log N}\frac{\Psi(N,[y',y])}{\phi(\widetilde{W})}\sum_{p\leqslant y}\ \sum_{k\geqslant \max(2,(\log y)/\log p)} p^{-\alpha k} + O\left(\frac{\log y}{\log N}\frac{y^2}{\log y}\right),\end{aligned}$$

where the error term trivially bounds the contribution of those choices of p^k for which $N/y < p^k < (N+N_1)/y'$, i.e., to which Lemma 3.4(i) does not apply. This error term is acceptable since $y < N^\eta$ for some sufficiently small $\eta > 0$. The sum over primes in the main term satisfies

$$\sum_{p\leqslant y}\sum_{k\geqslant \max(2,(\log y)/\log p)} p^{-\alpha k} \ll \sum_{p\leqslant y^{1/2}} y^{-\alpha} + \sum_{\sqrt{y}<p\leqslant y} p^{-2\alpha} \ll \frac{y^{1/2-\alpha}}{\log y} + (y^{1/2})^{-2\alpha+1}$$
$$\ll y^{-1/2}u\log u \ll_\varepsilon \frac{\log N}{\log y}(\log N)^{-K/2+\varepsilon} \ll_B \delta(N)^B\frac{\log N}{\log y}$$

for all $B > 0$, since $\alpha > \frac{1}{2}$ and $\delta(N) = \exp(-C_0\sqrt{\log_4 N})$. □

9.3. *Explicit bounds for the correlation between Λ and nilsequences.* In view of the inner sum in the bilinear expression in Proposition 9.2, we may reduce the problem of establishing that bound to a noncorrelation estimate for the von Mangoldt function with equidistributed nilsequences. For this purpose, we require explicit bounds for correlations of length ξ of the W-tricked von Mangoldt function with nilsequences. A specific requirement on these bounds is that they work with the same W-trick (determined by $w(N)$ and independent of ξ) and are uniform over a large range of ξ. Taking into account that y can be as small as $(\log N)^K$, we shall prove a result that is applicable to $\xi \in [\log N, N]$ if $w(N)$ is suitably chosen (see Remarks 9.6(ii) below).

Theorem 9.5. *Let $N_0 > 2$ be a large constant and $N > N_0$ a parameter. Let $x_0, w, \delta, \kappa : \mathbb{N}_{>N_0} \to \mathbb{R}_{>0}$ be functions that are defined for all sufficiently large integers and satisfy the relations $\kappa(N) \geqslant 1$, $1 \leqslant w(N) \leqslant \frac{1}{2}\log_2 x_0(N)$ and $\kappa(N)^2 w(N)^{-1/\kappa(N)} \ll_B \delta(N)^B$ for all $N > N_0$ and all $B \geqslant 1$. Suppose that $w(N) \to \infty$ as $N \to \infty$.*

Suppose that G/Γ is a filtered nilmanifold of dimension $\dim G \leqslant \kappa(N)$ and complexity at most $\delta(N)^{-1}$, let $d \geqslant 0$ denote its degree, and let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$. Let $1 \leqslant q \leqslant (\log x_0(N))^E$, $1 \leqslant E \ll 1$, be an integer, let $W = P(w(N))$ and write $\widetilde{W} = Wq$. If $0 < A' < \widetilde{W}$ is an integer such that $(\widetilde{W}, A') = 1$, then

$$\frac{\widetilde{W}}{\xi}\sum_{n\leqslant \xi/\widetilde{W}}\left(\frac{\phi(\widetilde{W})}{\widetilde{W}}\Lambda(\widetilde{W}n + A') - 1\right)F(g(n)\Gamma) \ll_{d,B} (1 + \|F\|_{\mathrm{Lip}})\delta(N)^B$$

for all $\xi \in [x_0(N), N]$, all $B \geqslant 1$, and for all Lipschitz functions $F : G/\Gamma \to \mathbb{C}$.

Remarks 9.6. (i) Note that $\xi/\widetilde{W} = \xi^{1+o(1)}$ by the definition of W and the conditions on $w(N)$, $\delta(N)$ and q. More precisely, we have $\xi \geqslant x_0(N)$, $\widetilde{W} = Wq$, $W = \exp(w(N) + o(1)) = (\log x_0(N))^{1/2+o(1)}$, and $q \leqslant (\log x_0(N))^{O(1)}$.

(ii) The choices $x_0(N) := \log N$, $w(N) := \frac{1}{2}\log_3 N$, $\kappa(N) = (\log_5 N)^C$ with $1 \leqslant C = O(1)$ and $\delta(N) = \exp(-C_0\sqrt{\log_4 N})$ with $1 \leqslant C_0 \leqslant (\log_4 N)^{1/4}$ and $N > N_0$ for sufficiently large N_0 are permissible in the theorem above.

(iii) Observe that the parameter choices in (ii) are consistent with the assumptions of Proposition 8.11 as well as the assumptions of Theorem 7.1.

Proof. This proof is a small modification of the proof of [Matthiesen 2018, Lemma 9.5] given in the appendix to that paper. The starting point of that proof is the following decomposition of the von Mangoldt function as $\Lambda = \Lambda^{\flat} + \Lambda^{\sharp}$. Let $\mathrm{id}_{\mathbb{R}}(x) = x$ denote the identity function and let $\chi^{\flat} + \chi^{\sharp} = \mathrm{id}_{\mathbb{R}}$ be a smooth decomposition with the property that $\mathrm{supp}(\chi^{\sharp}) \subset (-1, 1)$ and $\mathrm{supp}(\chi^{\flat}) \subset \mathbb{R} \setminus \left[-\frac{1}{2}, \frac{1}{2}\right]$. Then, for any $\gamma \in (0, 1)$,

$$\frac{\phi(\widetilde{W})}{\widetilde{W}} \Lambda(\widetilde{W}n + A') - 1 = \frac{\phi(\widetilde{W})}{\widetilde{W}} \Lambda^{\flat}(\widetilde{W}n + A') + \left(\frac{\phi(\widetilde{W})}{\widetilde{W}} \Lambda^{\sharp}(\widetilde{W}n + A') - 1 \right),$$

where, see [Green and Tao 2010, (12.2)],

$$\Lambda^{\sharp}(n) = -\log \xi^{\gamma} \sum_{d \mid n} \mu(d) \chi^{\sharp}\left(\frac{\log d}{\log \xi^{\gamma}} \right) \quad (|t| \geqslant 1 \Rightarrow \chi^{\sharp}(t) = 0)$$

is a truncated divisor sum, where

$$\Lambda^{\flat}(n) = -\log \xi^{\gamma} \sum_{d \mid n} \mu(d) \chi^{\flat}\left(\frac{\log d}{\log \xi^{\gamma}} \right) \quad \left(|t| \leqslant \tfrac{1}{2} \Rightarrow \chi^{\flat}(t) = 0\right)$$

is an average of $\mu(d)$ running over large divisors of n.

It follows as in [Green and Tao 2010, Section 12] and [Matthiesen 2018, Appendix A] from the orthogonality of the Möbius function with nilsequences that

$$\frac{\widetilde{W}}{\xi} \sum_{n \leqslant \xi/\widetilde{W}} \frac{\phi(\widetilde{W})}{\widetilde{W}} \Lambda^{\flat}(\widetilde{W}n + A') F(g(n)\Gamma) \ll_{\|F\|_{\mathrm{Lip}}, G/\Gamma, B} (\log \xi)^{-B} \quad (B > 0).$$

Here, it is important that $\widetilde{W} \leqslant (\log x_0(N))^{O(1)} \leqslant (\log \xi)^{O(1)}$.

Concerning the contribution from $\Lambda^{\sharp}$, define $\lambda^{\sharp} : \mathbb{N} \to \mathbb{R}$,

$$\lambda^{\sharp}(n) := \frac{\phi(\widetilde{W})}{\widetilde{W}} \Lambda^{\sharp}(\widetilde{W}n + A') - 1.$$

Since $\mathscr{X}$ is $\delta(N)^{-1}$-rational, [Matthiesen 2018, Lemmas A.2 and A.3] imply that the following bound holds with $m = 2^d \dim G$ and for every $\varepsilon > 0$

$$\begin{aligned} \frac{\widetilde{W}}{\xi} \sum_{n \leqslant \xi/\widetilde{W}} \lambda^{\sharp}(n) F(g(n)\Gamma) &\ll \varepsilon(1 + \|F\|_{\mathrm{Lip}}) + \|\lambda^{\sharp}\|_{U^{d+1}[\xi/\widetilde{W}]} (m/\varepsilon)^{2m} \delta(N)^{-O(m)} \\ &\ll \varepsilon(1 + \|F\|_{\mathrm{Lip}}) + \frac{m^{2m} \|\lambda^{\sharp}\|_{U^{d+1}[\xi/\widetilde{W}]}}{(\varepsilon^2 \delta(N)^{O(1)})^m}. \end{aligned} \tag{9-4}$$

We shall show below that $\|\lambda^{\sharp}\|_{U^{d+1}[\xi/\widetilde{W}]} \ll_d w(N)^{-1/2^{d+1}}$. Choosing $\varepsilon = \delta(N)^B$ and recalling the assumption that $\kappa(N)^2 w(N)^{-1/\kappa(N)} \ll_B \delta(N)^B$ and that $m = 2^d \dim G \leqslant 2^d \kappa(N)$, it follows that the above is bounded by

$$\ll_{d,B} (1 + \|F\|_{\mathrm{Lip}}) \delta(N)^B + \delta(N)^{O_d(B\kappa(N))} \ll_{d,B} (1 + \|F\|_{\mathrm{Lip}}) \delta(N)^B,$$

as required.

The uniformity norm $\|\lambda^\sharp\|_{U^{d+1}[\xi/\widetilde{W}]}$, can, as in [Matthiesen 2018, Appendix A], be estimated using [Green and Tao 2010, Theorem D.3] since $w(N) \leqslant \frac{1}{2}\log_2 x_0(N)$ is sufficiently small (see the "important convention" in [loc. cit., Section 5]). In our case, the system of forms takes the shape

$$\Psi_{\mathscr{B}}(n, \boldsymbol{h}) = \big(\widetilde{W}(n + \boldsymbol{\omega}\cdot\boldsymbol{h}) + A'\big)_{\omega\in\mathscr{B}}, \quad (n, \boldsymbol{h}) \in \mathbb{Z}\times\mathbb{Z}^{d+1},$$

where $\mathscr{B} \subset \{0, 1\}^{d+1}$ is any nonempty subset. The corresponding set $\mathscr{P}_{\Psi_{\mathscr{B}}}$ of exceptional primes consists of those primes dividing $\widetilde{W} = Wq$, i.e., $|\mathscr{P}_{\Psi_{\mathscr{B}}}| = \pi(w(N)) \ll \log_2\xi/\log w(N)$. We further have

$$\prod_{p\notin\mathscr{P}_{\Psi_{\mathscr{B}}}} \beta_p^{(\mathscr{B})} = 1 + O_d\biggl(\frac{1}{w(N)}\biggr) \quad\text{and}\quad \prod_{p\in\mathscr{P}_{\Psi}} \beta_p^{(\mathscr{B})} = \biggl(\frac{\widetilde{W}}{\phi(\widetilde{W})}\biggr)^{|\mathscr{B}|}.$$

Provided the constant γ is chosen sufficiently small, it follows from [Green and Tao 2010, Theorem D.3] applied with $K_z = \{(n, \boldsymbol{h}) : 0 < n + \boldsymbol{\omega}\cdot\boldsymbol{h} \leqslant z \text{ for all } \boldsymbol{\omega} \in \{0, 1\}^{d+1}\}$, where $z = \xi/\widetilde{W}$, that

$$\begin{aligned}\|\lambda^\sharp\|_{U^{d+1}[z]}^{2^{d+1}} &= \frac{\mathrm{vol}(K_z)}{z^{d+2}} \sum_{\mathscr{B}\subseteq\{0,1\}^{d+1}} (-1)^{|\mathscr{B}|} \prod_{p\notin P_{\Psi_{\mathscr{B}}}} \beta_p^{(\mathscr{B})} + O_d\big((\log\xi^\gamma)^{-1/20}\exp(O_d(|\mathscr{P}_{\Psi_{\mathscr{B}}}|))\big)\\ &\ll_d \frac{\mathrm{vol}(K_z)}{z^{d+2}}\frac{1}{w(N)} + \exp\biggl(-\frac{\log_2\xi}{20} + O_d\biggl(\frac{\log_2\xi}{\log w(N)}\biggr)\biggr) \ll_d \frac{1}{w(N)}. \qquad\square\end{aligned}$$

As it provides a different approach which could prove useful for future generalisations, we include a second, alternative, proof for the special case of Theorem 9.5 in which $q = 1$. This proof is based on one of the main results of [Tao and Teräväinen 2023]. For this proof to work, $w(n)$ needs to be redefined as $w(n) = c'\log_3 N$ for some sufficiently small constant $c' > 0$.

Alternative proof for special case of Theorem 9.5. Assume for simplicity the choice of parameters from Remarks 9.6(ii) except for $w(N) = \frac{1}{2}\log_3 N$, which we replace by $w(n) = c'\log_3 N$ for some small $c' > 0$. Further, let $q = 1$ so that $\widetilde{W} = W$ and $A' = A$.

Define $\lambda_{W,A} = W^{-1}\phi(W)\Lambda(W\cdot{+A}) - 1$. By the proof of [loc. cit., Corollary 1.5] (see the end of Section 8 in that paper), we have

$$\|\lambda_{W,A}\|_{U^k[(\xi-A)/W]} \ll (\log_2\xi)^{-c} \tag{9-5}$$

provided that the constant $c' > 0$ in the definition of $w(N)$ is sufficiently small to ensure that the estimate [loc. cit., (8.19)] holds with $W = P(w(N))$ and $N = x_0(N)$.

Proceeding as in the proof of Theorem 9.5 above, we use [Matthiesen 2018, Lemmas A.2 and A.3] which show that the following bound holds with $m = 2^d \dim G$ and for every $\varepsilon > 0$:

$$\frac{W}{\xi}\sum_{n\leqslant(\xi-A)/W} \lambda_{W,A}(n)F(g(n)\Gamma) \ll \varepsilon(1 + \|F\|_{\mathrm{Lip}}) + \|\lambda_{W,A}\|_{U^{d+1}[(\xi-A)/W]}(m/\varepsilon)^{2m}\delta(N)^{-O(m)}.$$

Choosing $\varepsilon = \delta(N)^B$ and recalling that $\log_2\xi \gg \log_2 x_0(N) \gg \log_3 N$, the above is seen to be bounded by

$$\ll_d (1 + \|F\|_{\mathrm{Lip}})\delta(N)^B + \frac{\kappa(N)^2(\log_2\xi)^{-c}}{\delta(N)^{O_d(B\kappa(N))}} \ll_d (1 + \|F\|_{\mathrm{Lip}})\delta(N)^B,$$

where the constants c and the final implied constant depend on the exponent in the bound (9-5) from [Tao and Teräväinen 2023, Section 8] as well as on the degree d of the filtration. □

9.4. ***Conclusion of the proof of Proposition 8.8.*** In view of the reductions carried out in the previous subsections it remains to show that the condition of Proposition 9.2 is in fact valid in order to complete the proof of the equidistributed noncorrelation estimate stated in Proposition 8.8. Hence, the proof of Proposition 8.8 is complete once we have established the following lemma.

Lemma 9.7. *Let $E_1 \geqslant 1$. Under the assumptions of Proposition 8.8 and with the notation $\widetilde{W} = Wq$, we have*

$$\frac{1}{\log N}\left|\sum_{n\leqslant N}\sum_{\substack{\widetilde{W}m+A'_n\in\\ [y',\min(y,N/n)]}} \mathbf{1}_{S([y',y])}(n)\Lambda(\widetilde{W}m+A'_n)F(g(mn+A_n)\Gamma)\right| \ll \delta(N)^{E_1}\frac{\Psi(N,[y',y]))}{\phi(\widetilde{W})}$$

for all sequences $(A_n)_{n\in\mathbb{N}}$, $(A'_n)_{n\in\mathbb{N}}$ of integers such that $|A_n| \leqslant n$ and $\gcd(A'_n, \widetilde{W}) = 1$, and for all $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ such that the finite sequence $(g(n)\Gamma)_{n\leqslant N}$ is totally $\delta(N)^{E_0}$-equidistributed for some sufficiently large $E_0 > 1$. The implied constant may depend on the degree d of $G_\bullet$, $\dim G$, $\|F\|_{\mathrm{Lip}}$ and E_1.

Proof. Splitting the summation range of the outer sum into three intervals and abbreviating

$$G(m,n) = \mathbf{1}_{S([y',y])}(n)\Lambda(\widetilde{W}m+A'_n)F(g(mn+A_n)\Gamma),$$

we obtain

$$\begin{aligned}\sum_{n\leqslant N}\sum_{\widetilde{W}m+A'_n\in[y',\min(y,N/n)]} G(m,n) &= \sum_{n<\delta^{2E_1}N/y}\sum_{y'\leqslant\widetilde{W}m+A'_n\leqslant y} G(m,n)\\ &\quad+\sum_{\delta^{2E_1}N/y<n\leqslant N/y^{2/3}}\sum_{y'\leqslant\widetilde{W}m+A'_n\leqslant\min(N/n,y)} G(m,n)\\ &\quad+\sum_{0\leqslant k\leqslant K}\sum_{N/y^{2/3}<n\leqslant N/(y'2^k)}\sum_{\widetilde{W}m+A'_n\in[y'2^k,\min(y'2^{k+1},N/n)]} G(m,n),\end{aligned} \tag{9-6}$$

where $K = \lceil\log(y^{2/3}/y')/\log 2\rceil$ and in particular $y'2^K \asymp y^{2/3}$.

The decomposition above has been chosen in such a way that the contributions from the initial and final segment are negligible while in the middle segment the summation range of the inner sum is guaranteed to be sufficiently long and $\min(y, N/n)$ is not too small compared to N/n. More precisely, by Lemma 3.4(i), the final term is trivially bounded above by

$$\begin{aligned}\sum_{0\leqslant k\leqslant K}\frac{y'2^k}{\phi(\widetilde{W})}\Psi\left(\frac{N}{y'2^k},[y',y]\right) &\ll \frac{\Psi(N,[y',y])}{\phi(\widetilde{W})}\sum_{0\leqslant k\leqslant K}(y'2^k)^{(1-\alpha)} \ll \frac{\Psi(N,[y',y])}{\phi(\widetilde{W})}(y'2^K)^{(1-\alpha)}\\ &\ll y^{2(1-\alpha)/3}\frac{\Psi(N,[y',y])}{\phi(\widetilde{W})} \ll (\log N)^{2/3}\frac{\Psi(N,[y',y])}{\phi(\widetilde{W})},\end{aligned}$$

which is acceptable in view of the $(\log N)^{-1}$ factor in the statement.

Similarly, the first term is trivially bounded above by

$$\frac{y}{\phi(\widetilde{W})}\Psi\Big(\delta(N)^{2E_1}\frac{N}{y},[y',y]\Big) \ll \delta(N)^{2E_1\alpha}y^{1-\alpha}\frac{\Psi(N,[y',y])}{\phi(\widetilde{W})} \ll (\log N)\delta(N)^{E_1}\frac{\Psi(N,[y',y])}{\phi(\widetilde{W})},$$

since $\alpha > \frac{1}{2}$, which is also acceptable.

It remains to analyse the middle range and we start by observing that the summation range of the inner sum over $\widetilde{W}m + A'_n$ is always sufficiently long, that is, of length $\gg \delta(N)^{E_1}m_1$ if m_1 denotes the upper endpoint of the range. In fact, $m_1 = \min(y, N/n) \geqslant y^{2/3} > y^{1/2} > y'$, which shows that $|\min(y, N/n) - y'| \sim \min(y, N/n)$ and moreover $y' = o(\delta(N)^B \min(y, N/n)/\phi(\widetilde{W}))$ for any $B \asymp 1$.

In order to apply Proposition 8.11 on the equidistribution of linear subsequences of an equidistributed nilsequence, we dyadically decompose the sum over n in the middle range of (9-6) into $O(\log N)$ intervals $(M_j, 2M_j]$, where $\delta(N)^{2E_1}N/y \leqslant M_j < N/y^{2/3}$, and one potentially shorter interval. Note that $N^{1/2} < Ny^{-2} < N\delta(N)^{2E_1}y^{-1} \leqslant M_j$, assuming that $\eta < \frac{1}{4}$ (so that $y < N^{1/4}$) and that N is sufficiently large. For each M_j as above consider the sum

$$\sum_{\substack{n\sim M_j\\ n\leqslant N/y^{2/3}}} \mathbf{1}_{S([y',y])}(n) \sum_{\widetilde{W}m+A'_n\in[y',\min(y,N/n)]} \Lambda(\widetilde{W}m + A'_n)F(g(mn + A_n)\Gamma). \tag{9-7}$$

Let $E_2 > 1$ be a parameter to be chosen later. Then, by Proposition 8.11, all but $O(\delta(N)^{E_2}\Psi(M_j, [y', y]))$ of the finite sequences

$$(g_n(m)\Gamma)_{m\leqslant N/n}, \quad n \in (M_j, 2M_j] \cap S([y', y]),$$

where $g_n(m) := g(mn + A_n)$, are totally $\delta(N)^{E_2}$-equidistributed in G/Γ provided that the parameter $E_0 > 1$ from our assumptions is sufficiently large. We bound the contribution to (9-7) from all exceptional values of $n \in (M_j, 2M_j] \cap S([y', y])$ trivially by

$$O(\delta(N)^{E_2}\Psi(M_j, [y', y]))\frac{\min(y, N/M_j)}{\phi(\widetilde{W})}.$$

For all other values of $n \in (M_j, 2M_j] \in S([y', y])$, we seek to apply Theorem 9.5 with the choice of parameters given in Remarks 9.6(ii). The lower bound $M_j \geqslant \delta(N)^{2E_1}N/y$ implies that $y \geqslant \delta^{2E_1}N/M_j \geqslant \delta^{E_2/2}N/n$ for all $n \in (M_j, 2M_j]$ provided that $E_2 > 4E_1$. Thus $\min(N/n, y) \geqslant \delta^{E_2/2}N/n$ and the sequence

$$(g_n(m)\Gamma)_{m\leqslant\min(N/n,y)}$$

is totally $\delta^{E_2/2}$-equidistributed in G/Γ whenever n is nonexceptional, and $\xi = \min(y, N/n)$ satisfies $\xi \in [\log N, N]$. Finally, the upper bound on y' implies that, for any $B \asymp 1$, $y' = o(\delta(N)^B \min(y, N/M_j)/\phi(q))$. For nonexceptional n as above, it thus follows from Theorem 9.5, applied with $B = E_2/4$, that

$$\begin{aligned}\frac{\phi(\widetilde{W})}{\min(y, N/M_j)} &\sum_{\widetilde{W}m+A'_n\in[y',\min(y,N/n)]} \Lambda(\widetilde{W}m + A'_n)F(g(mn + A_n)\Gamma)\\ &= \sum_{\widetilde{W}m+A'_n\in[y',\min(y,N/n)]} F(g(mn + A_n)\Gamma) + O_{d,E_1,\|F\|_{\mathrm{Lip}}}(\delta(N)^{E_2/4}) \ll_{d,E_1,\|F\|_{\mathrm{Lip}}} \delta(N)^{E_2/4},\end{aligned}$$

where we used that $p < w(N)$ for every $p \mid q$.

Taking the contributions from both exceptional and nonexceptional $n \in (M_j, 2M_j] \in S([y', y])$ into account, the expression (9-7) is thus bounded by

$$\ll_{d,E_1,\|F\|_{\mathrm{Lip}}} (\delta(N)^{E_2/4} + \delta(N)^{E_1}) \min(y, N/M_j) \frac{\Psi(M_j, [y', y])}{\phi(\widetilde{W})}.$$

Summing over all j, the contribution of the middle segment to (9-6) is therefore bounded above by

$$\begin{aligned} &\ll_{d,E_1,\|F\|_{\mathrm{Lip}}} (\delta(N)^{E_2/4} + \delta(N)^{E_1}) \sum_j \min(y, N/M_j) \frac{\Psi(M_j, [y', y])}{\phi(\widetilde{W})} \\ &\ll_{d,E_1,\|F\|_{\mathrm{Lip}}} \frac{\delta(N)^{E_1}}{\phi(\widetilde{W})} \sum_j \sum_{n \sim M_j} \mathbf{1}_{S([y',y])}(n) \min(y, N/n). \end{aligned} \tag{9-8}$$

We shall now make use of the fact that the inner sum in the middle range is guaranteed to be sufficiently long in order to deduce that we make an $\delta(N)^{E_1}$-saving on average on each inner sum provided E_0 and, hence, E_2 are sufficiently large. Reversing the steps in the Montgomery–Vaughan-type decomposition will then complete the proof.

Turning towards the details, recall that for any $n \sim M_j$ the interval $[y', \min(y, N/n)]$ has length $|\min(y, N/n) - y'| \sim \min(y, N/n)$, so that

$$\min(y, N/n) \ll |\min(y, N/n) - y'| \ll \sum_{y' \leqslant m \leqslant \min(y, N/n)} \Lambda(m) \quad (n \sim M_j),$$

by the prime number theorem. By combining this estimate with the bound (9-8), the contribution of the middle segment to (9-6) is seen to be bounded above by

$$\begin{aligned} &\ll \frac{\delta(N)^{E_1}}{\phi(\widetilde{W})} \sum_j \sum_{n \sim M_j} \mathbf{1}_{S([y',y])}(n) \min(y, N/n) \\ &\ll \frac{\delta(N)^{E_1}}{\phi(\widetilde{W})} \sum_j \sum_{n \sim M_j} \mathbf{1}_{S([y',y])}(n) \sum_{y' \leqslant m \leqslant \min(y, N/n)} \Lambda(m) \mathbf{1}_{S([y',y])}(m) \\ &\ll \frac{\delta(N)^{E_1}}{\phi(\widetilde{W})} \sum_{\delta(N)^{2E_1} N/y \leqslant n < 2N/y^{2/3}} \sum_{y' \leqslant m \leqslant \min(y, N/n)} \Lambda(m) \mathbf{1}_{S([y',y])}(mn) \\ &\ll \frac{\delta(N)^{E_1}}{\phi(\widetilde{W})} \sum_{n < 2N} \mathbf{1}_{S([y',y])}(n) \log n \ll (\log N) \delta(N)^{E_1} \frac{\Psi(N, [y', y])}{\phi(\widetilde{W})}, \end{aligned}$$

provided that E_0 is sufficiently large (to ensure that $E_2 > 4E_1$), and where the implied constant may depend on d, $\dim G$, E_1 and $\|F\|_{\mathrm{Lip}}$. The lemma follows when taking into account the $(\log N)^{-1}$-factor in the statement. □

10. The proof of Theorem 1.3

We will use the transferred generalised von Neumann theorem [Green and Tao 2010, Proposition 7.1] combined with a simple majorising function in order deduce Theorem 1.3 from Theorem 1.1. There are two choices of majorants for $h_{[y',y]}$ readily available in the case where $y \geqslant N^{\varepsilon}$ for any fixed $\varepsilon \in (0, 1)$. These can still be used for somewhat smaller values of y. We state both majorant constructions below.

10.1. ***GPY-type sieve majorant.*** Consider the following GPY-type sieve majorant for numbers free from prime factors $p < y'$:

$$\Lambda_{\chi,y'}(n) = \log y' \Bigl(\sum_{d \mid n} \mu(d)\chi\Bigl(\frac{\log d}{\log y'}\Bigr)\Bigr)^2,$$

where $\chi : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ is a smooth function with support in $[-1, 1]$ which takes the value $\chi(x) = 1$ for all $x \in \bigl[-\frac{1}{2}, \frac{1}{2}\bigr]$. In particular,

$$\Lambda_{\chi,y'}(n) = \log y' \Bigl(\sum_{\substack{d \mid n \\ d \leqslant y'}} \mu(d)\chi\Bigl(\frac{\log d}{\log y'}\Bigr)\Bigr)^2,$$

which implies that $\Lambda_{\chi,y'}(n) = \log y'$ for all integers n that are free from prime factors $p \leqslant y'$. Since $\alpha = 1 - O((\varepsilon \log N)^{-1} \log_2 N)$ for $y = N^{\varepsilon}$, Lemma 3.2 shows that $\zeta(\alpha, y) \asymp \log y \asymp \log N$ and $g_{P(y')} \asymp (\log y')^{-1}$. Thus, by (3-3) and Lemma 3.3, we have

$$h_{[y',N^{\varepsilon}]}(n) \leqslant \frac{(N/n)^{\alpha} n}{\Psi(N, [y', N^{\varepsilon}])} \ll \frac{N \log N}{N g_{P(y')}(\alpha)\zeta(\alpha, N^{\varepsilon})} \ll \log y'.$$

This shows that

$$h_{[y',N^{\varepsilon}]}(n) \ll \Lambda_{\chi,y'}(n) \quad (N < n \leqslant 2N).$$

Hence, $\Lambda_{\chi,y'}(n)$ satisfies the majorisation property (1) from above. It follows from [Green and Tao 2010, Theorem D.3] that the W-tricked version of $\Lambda_{\chi,y'}(n)$ satisfies the *linear forms condition* as stated in [loc. cit., Definition 6.2]. We omit the proof here as it is essentially contained in [loc. cit., Appendix D] and only mention that it relies on the fact that $\Lambda_{\chi,y'}$ carries the structure of a truncate divisor sum.

10.2. ***The normalised characteristic function $\mathbf{1}_{S([y',N])}$.*** An alternative majorant function in the dense setting is given by the function

$$\nu(n) = \frac{P(y')}{\phi(P(y'))} \mathbf{1}_{S([y',2N])}(n) \quad (n \leqslant 2N).$$

Since

$$\mathbf{1}_{S([y',2N])}(n) = \mathbf{1}_{(n,P(y'))=1}(n)$$

for $n \leqslant 2N$, this function corresponds to the *Cramér model* for the von Mangoldt function that was studied in [Tao and Teräväinen 2023]: we have

$$\nu(n) = \Lambda_{\mathrm{Cramér},y'}(n) := \frac{P(y')}{\phi(P(y'))} \mathbf{1}_{(n,P(y'))=1}(n)$$

for all $n \leqslant 2N$.

Lemma 10.1. *Let* $\varepsilon \in (0, 1)$. *Then we have*

$$h_{[y',y]}(n) \ll \Lambda_{Cramér,y'}(n)$$

uniformly for all $y > N^{\varepsilon}$, *all* $n \in (N, 2N]$ *and all sufficiently large* N.

Proof. Since $\mathbf{1}_{(n,P(y'))=1} = \mathbf{1}_{S([y',2N])}(n)$ for $n \leqslant 2N$, it remains to note that

$$h_{[y',y]}(n) \leqslant \frac{n(N/n)^{\alpha}}{\Psi(N,[y',N^{\varepsilon}])} \ll \frac{N\log N}{N g_{P(y')}(\alpha)\zeta(\alpha,N^{\varepsilon})} \ll \log y' \asymp \frac{P(y')}{\phi(P(y'))},$$

when $n \sim N$ and $y > N^{\varepsilon}$. □

Let $2 < w(N) < y' = (\log N)^{K'}$, $W = P(w(N))$, $0 < A < W$ and $(A, W) = 1$. Then, by [Tao and Teräväinen 2023, Corollary 5.3], we have

$$\left\| \frac{\phi(W)}{W}\nu(W\cdot+A) - 1 \right\|_{U^k[\frac{N-A}{W}]} \ll_k w(N)^{-c} \tag{10-1}$$

for some constant $c > 0$. We observe that, using $\Lambda_{\chi,y'}(n)$ as a majorant for ν, it follows from this uniformity norm estimate and the generalised von Neumann theorem [Green and Tao 2010, Proposition 7.1] that the W-tricked version of ν satisfies the linear forms condition from [Green and Tao 2008, Definition 6.2].

10.3. ***Proof of Theorem 1.3.*** Let $y_0 \leqslant y \leqslant N^{\eta}$, where $\eta \in (0, 1)$ is sufficiently small depending on d and y_0 will be determined in the course of the proof, and let $w(N) \leqslant y' \leqslant (\log N)^{K'}$ for some fixed $K' \geqslant 1$. The transferred version of the quantitative inverse theorem for U^k-norms of [Manners 2018] proved in [Tao and Teräväinen 2023, Theorem 8.3] allows us to deduce explicit U^k-norm estimates from Theorem 1.3, which we will later combine with the transferred generalised von Neumann theorem [Green and Tao 2010, Proposition 7.1]. More precisely, Theorem 1.1 implies that if $\Psi(N,[y',y])P(y')/(N\phi(P(y'))) \geqslant (\log_4 N)^{-1/2}$, then

$$\begin{aligned}\left| \frac{W}{N} \sum_{N/W<n\leqslant 2N/W} (g^{(W,A)}_{[y',y]}(n) - 1)F(g(n)\Gamma) \right| &\ll (1+\|F\|_{\mathrm{Lip}})(\log_4 N)^{-1} \\ &\ll \frac{1+\|F\|_{\mathrm{Lip}}}{\sqrt{\log_4 N}} \frac{N\phi(P(y'))}{\Psi(N,[y',y])P(y')}\end{aligned}$$

for all filtered nilmanifolds of dimension $O((\log_5 N)^{O(1)})$ and complexity Q_0 that is bounded by

$$Q_0 < \exp\bigl(\tfrac{1}{2}\sqrt{\log_4 N}\bigr),$$

and all nilsequences attached to it. Observe that

$$(\log_4 N)^{-1/2} = o(\exp(-\exp(C_1/\tilde{\delta}^{C_2})))$$

and

$$\exp\exp(C_1/\tilde{\delta}^{C_2}) = o\bigl(\exp\bigl(\tfrac{1}{2}\sqrt{\log_4 N}\bigr)\bigr)$$

for all constants $C_1, C_2 > 1$ if $\tilde{\delta} = 1/\log_7 N$. Note further that the function from Section 10.2 may be used as a trivial majorising function for the following rescaled version of $g^{(W,A)}_{[y',y]}(n)$, which essentially corresponds to $\mathbf{1}_{S([y',y])}(n)$. We have

$$\frac{P(y')}{\phi(P(y'))}\frac{\Psi(N,[y',y])}{N} g^{(W,A)}_{[y',y]}(n) \ll \nu(n) \quad (n \leqslant (N-A)/W).$$

By (10-1), we have

$$\|\nu - 1\|_{U^k[(N-A)/W]} \ll_k (\log_3 N)^{-c} \ll_B \tilde{\delta}^B$$

for all $B \geqslant 1$. Hence the transferred quantitative U^k inverse theorem of [Manners 2018], as stated in [Tao and Teräväinen 2023, Theorem 8.3], implies that uniformly for all $0 < A < W$ with $(W, A) = 1$ we have

$$\|g_{[y',y]}^{(W,A)} - 1\|_{U^k([(N-A)/W])} \ll \frac{\tilde{\delta} N \phi(P(y'))}{\Psi(N, [y', y]) P(y')},$$

which is small compared to the mean value 1 of $g_{[y',y]}^{(W,A)}$ provided that

$$\Psi(N, [y', y]) P(y')/(N\phi(P(y'))) = o(\log_7 N).$$

In order to apply the generalised von Neumann theorem [Green and Tao 2010, Proposition 7.1] in our situation, observe that the $o(1)$ error term in [loc. cit., Proposition 7.1′] corresponds to the error term in the (D, D, D)-linear forms condition for ν. It follows from [Tao and Teräväinen 2023, Proposition 5.2] that any function of the form

$$\tilde{\nu}(n) = \frac{1}{t} \sum_{i=1}^{t} \frac{\phi(W)}{W} \nu^{(N)}(Wn + A_i),$$

with $0 < A_i < W,\ \gcd(A_i, W) = 1$ for all $1 \leqslant i \leqslant t$, satisfies the (D, D, D)-linear forms condition from [Green and Tao 2010, Definition 6.2] for any given $D = O_{r,s,L}(1)$ with an error term of the form

$$w(N)^{-c} = (\log_3 N)^{-c}.$$

Applying [loc. cit., Proposition 7.1] with

$$f_i(n) = \frac{P(y')}{\phi(P(y'))} \frac{\Psi(N, [y', y])}{N} g_{[y',y]}^{(W,A)}(n)$$

and with $\tilde{\nu}$ as above, we obtain

$$\sum_{\boldsymbol{n} \in \mathbb{Z}^s \cap (N\mathfrak{K})/W} \prod_{j=1}^{r} g_{[y',y]}(W\psi_j(\boldsymbol{n}) + A_j) = \left(\frac{W}{\phi(W)}\right)^r \left\{ \frac{N^s \operatorname{vol} \mathfrak{K}}{W^s} + o_{s,r,\|\Psi\|}\left(\left(\frac{N}{W}\right)^s\right) \right\}$$

provided that

$$\|g_{[y',y]}^{(W,A)} - 1\|_{U^k([(N-A)/W])} = o\left(\left(\frac{\Psi(N, [y', y]) P(y')}{N\phi(P(y'))}\right)^{r-1}\right)$$

for $k = O_{r,s}(1)$, as well as

$$w(N)^{-c} = (\text{linear forms condition error term}) = o\left(\left(\frac{\Psi(N, [y', y]) P(y')}{N\phi(P(y'))}\right)^{r}\right).$$

These conditions are certainly satisfied if

$$\frac{\Psi(N, [y', y]) P(y')}{N\phi(P(y'))} \gg (\log_8 N)^{-1}.$$

Since

$$\frac{\Psi(N, [y', y]) P(y')}{N\phi(P(y'))} \asymp \frac{\Psi(N, y)}{N} \prod_{p \leqslant y'} \frac{1 - p^{-\alpha(N,y)}}{1 - p^{-1}} \leqslant \frac{\Psi(N, y)}{N} = u^{-u+o(u)}$$

by Lemma 3.3, it follows that $N^{-1}\Psi(N, y) \gg (\log_8 N)^{-1}$. Let y_0 be such that

$$\frac{\Psi(N, y_0)}{N} = (\log_8 N)^{-1}$$

and write $u_0 = (\log N)/\log y_0$. Then

$$\frac{\log N}{\log y_0} = u_0 < u_0^{u_0+o(u_0)} = \log_8 N$$

and, if $y > y_0$, then

$$1-\alpha(N, y) \ll \frac{\log(u_0 \log(u_0+2))}{\log y_0} \ll \frac{(\log_8 N)^2}{\log N}.$$

Hence, the first part of Lemma 3.2 applies to $g_{P(y')}(\alpha(N, y))$ and yields

$$g_{P(y')}(\alpha(N, y)) \asymp \frac{1}{\log y'} \asymp \frac{\phi(P(y'))}{P(y')}$$

for $y \geqslant y_0$ and $y' = (\log N)^{K'}$. Thus

$$\frac{\Psi(N, [y', y])P(y')}{N\phi(P(y'))} \asymp \frac{\Psi(N, y)}{N} \gg (\log_8 N)^{-1}$$

for $y \geqslant y_0$. From

$$\exp(u_0^2) > u_0^{u_0+o(u_0)} = \log_8 N,$$

which holds for sufficiently large N, we deduce that $y_0 < N^{1/\sqrt{\log_9 N}}$.

Suppose now that $y_0 \leqslant y \leqslant N^\eta$. Then we obtain

$$\begin{aligned}
&\sum_{\boldsymbol{n}\in\mathbb{Z}^s\cap N\mathfrak{K}} \prod_{j=1}^{r} g_{[y',y]}(\psi_j(\boldsymbol{n})+a_j) \\
&\qquad = \sum_{\boldsymbol{A}\in\{0,\dots,W-1\}^s} \sum_{\substack{W\boldsymbol{n}+\boldsymbol{A}\\ \in\mathbb{Z}^s\cap N\mathfrak{K}}} \prod_{j=1}^{r} g_{[y',y]}(\psi_j(W\boldsymbol{n}+\boldsymbol{A})+a_j) \\
&\qquad = \sum_{\boldsymbol{A}\in\{0,\dots,W-1\}^s} \sum_{\boldsymbol{n}\in\mathbb{Z}^s\cap(N\mathfrak{K}-\boldsymbol{A})/W} \prod_{j=1}^{r} g_{[y',y]}(W\psi_j(\boldsymbol{n})+\psi_j(\boldsymbol{A})+a_j) \\
&\qquad = \left(\frac{W}{\phi(W)}\right)^r \left\{\frac{N^s \operatorname{vol}\mathfrak{K}}{W^s} + o_{s,r,\|\Psi\|}\left(\left(\frac{N}{W}\right)^s\right)\right\} \sum_{\boldsymbol{A}\in\{0,\dots,W-1\}^s} \prod_{j=1}^{r} \mathbf{1}_{\gcd(W,\psi_j(\boldsymbol{A})+a_j)=1} \\
&\qquad = \{N^s \operatorname{vol}\mathfrak{K} + o_{s,r,\|\Psi\|}(N^s)\} \prod_{p<w(N)} \beta_p,
\end{aligned}$$

where

$$\beta_p = \frac{1}{p^s} \sum_{\boldsymbol{u}\in(\mathbb{Z}/p\mathbb{Z})^s} \prod_{j=1}^{r} \frac{p}{p-1} \mathbf{1}_{\psi_j(\boldsymbol{u})+a_j\not\equiv 0 \pmod p}.$$

Theorem 1.3 now follows on recalling[7] from [Green and Tao 2010, Lemma 1.3] that the condition that the forms ψ_i be pairwise linearly independent over $\mathbb{Q}$ implies $\beta_p = 1 + O_{s,r,L}(p^{-2})$ for all primes $p > w(N)$.

[7] Observe that our local factors are identical to those defined in [Green and Tao 2010, (1.6)]

Acknowledgements

We would like to thank Adam Harper for very helpful comments on a previous version of this paper and the anonymous referee for their astute questions and suggestions, leading us, in particular, to rework our arguments and remove the previous restriction to large values of y' ($\asymp (\log N)^C$). While working on this paper, the authors were supported by the Swedish Research Council (VR) grant 2021-04526 and the Göran Gustafsson Foundation.

References

[Balog et al. 2012] A. Balog, V. Blomer, C. Dartyge, and G. Tenenbaum, "Friable values of binary forms", *Comment. Math. Helv.* **87**:3 (2012), 639–667. MR

[de la Bretèche and Tenenbaum 2005] R. de la Bretèche and G. Tenenbaum, "Propriétés statistiques des entiers friables", *Ramanujan J.* **9**:1-2 (2005), 139–202. MR

[Drappeau and Shao 2016] S. Drappeau and X. Shao, "Weyl sums, mean value estimates, and Waring's problem with friable numbers", *Acta Arith.* **176**:3 (2016), 249–299. MR

[Fouvry 2010] É. Fouvry, "Friabilité des valeurs d'un polynôme", *Arch. Math.* (*Basel*) **95**:5 (2010), 411–421. MR

[Fouvry and Tenenbaum 1996] É. Fouvry and G. Tenenbaum, "Répartition statistique des entiers sans grand facteur premier dans les progressions arithmétiques", *Proc. Lond. Math. Soc.* (3) **72**:3 (1996), 481–514. MR

[Granville 2008] A. Granville, "Smooth numbers: computational number theory and beyond", pp. 267–323 in *Algorithmic number theory*: *lattices*, *number fields*, *curves and cryptography*, edited by J. P. Buhler and P. Stevenhagen, Math. Sci. Res. Inst. Publ. **44**, Cambridge Univ. Press, 2008. MR

[Green and Tao 2008] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions", *Ann. of Math.* (2) **167**:2 (2008), 481–547. MR

[Green and Tao 2010] B. Green and T. Tao, "Linear equations in primes", *Ann. of Math.* (2) **171**:3 (2010), 1753–1850. MR

[Green and Tao 2012a] B. Green and T. Tao, "The Möbius function is strongly orthogonal to nilsequences", *Ann. of Math.* (2) **175**:2 (2012), 541–566. MR

[Green and Tao 2012b] B. Green and T. Tao, "The quantitative behaviour of polynomial orbits on nilmanifolds", *Ann. of Math.* (2) **175**:2 (2012), 465–540. MR

[Green et al. 2012] B. Green, T. Tao, and T. Ziegler, "An inverse theorem for the Gowers $U^{s+1}[N]$-norm", *Ann. of Math.* (2) **176**:2 (2012), 1231–1372. MR

[Harper 2012a] A. J. Harper, "Bombieri–Vinogradov and Barban–Davenport–Halberstam type theorems for smooth numbers", preprint, 2012. arXiv 1208.5992

[Harper 2012b] A. J. Harper, "On a paper of K. Soundararajan on smooth numbers in arithmetic progressions", *J. Number Theory* **132**:1 (2012), 182–199. MR

[Harper 2016] A. J. Harper, "Minor arcs, mean values, and restriction theory for exponential sums over smooth numbers", *Compos. Math.* **152**:6 (2016), 1121–1158. MR

[Hildebrand 1986a] A. Hildebrand, "On the local behavior of $\Psi(x, y)$", *Trans. Amer. Math. Soc.* **297**:2 (1986), 729–751. MR

[Hildebrand 1986b] A. Hildebrand, "On the number of positive integers $\leqslant x$ and free of prime factors $> y$", *J. Number Theory* **22**:3 (1986), 289–307. MR

[Hildebrand and Tenenbaum 1986] A. Hildebrand and G. Tenenbaum, "On integers free of large prime factors", *Trans. Amer. Math. Soc.* **296**:1 (1986), 265–290. MR

[Lachand 2017] A. Lachand, "On the representation of friable integers by linear forms", *Acta Arith.* **181**:2 (2017), 97–109. MR

[Lagarias and Soundararajan 2011] J. C. Lagarias and K. Soundararajan, "Smooth solutions to the *abc* equation: the *xyz* conjecture", *J. Théor. Nombres Bordeaux* **23**:1 (2011), 209–234. MR

[Lagarias and Soundararajan 2012] J. C. Lagarias and K. Soundararajan, "Counting smooth solutions to the equation $A+B=C$", *Proc. Lond. Math. Soc.* (3) **104**:4 (2012), 770–798. MR

[Leng et al. 2024] J. Leng, A. Sah, and M. Sawhney, "Quasipolynomial bounds on the inverse theorem for the Gowers $U^{s+1}[N]$-norm", preprint, 2024. arXiv 2402.17994

[Manners 2018] F. Manners, "Quantitative bounds in the inverse theorem for the Gowers U^{s+1}-norms over cyclic groups", preprint, 2018. arXiv 1811.00718

[Matthiesen 2012] L. Matthiesen, "Linear correlations amongst numbers represented by positive definite binary quadratic forms", *Acta Arith.* **154**:3 (2012), 235–306. MR

[Matthiesen 2018] L. Matthiesen, "Generalized Fourier coefficients of multiplicative functions", *Algebra Number Theory* **12**:6 (2018), 1311–1400. MR

[Matthiesen 2020] L. Matthiesen, "Linear correlations of multiplicative functions", *Proc. Lond. Math. Soc.* (3) **121**:2 (2020), 372–425. MR

[Montgomery and Vaughan 1977] H. L. Montgomery and R. C. Vaughan, "Exponential sums with multiplicative coefficients", *Invent. Math.* **43**:1 (1977), 69–82. MR

[Montgomery and Vaughan 2007] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory, I: Classical theory*, Cambridge Stud. Adv. Math. **97**, Cambridge Univ. Press, 2007. MR

[Soundararajan 2008] K. Soundararajan, "The distribution of smooth numbers in arithmetic progressions", pp. 115–128 in *Anatomy of integers* (Montreal, 2006), edited by J.-M. De Koninck et al., CRM Proc. Lecture Notes **46**, Amer. Math. Soc., Providence, RI, 2008. MR

[Tao and Teräväinen 2023] T. Tao and J. Teräväinen, "Quantitative bounds for Gowers uniformity of the Möbius and von Mangoldt functions", *J. Eur. Math. Soc.* (online publication December 2023).

[Vaughan 1989] R. C. Vaughan, "A new iterative method in Waring's problem", *Acta Math.* **162**:1-2 (1989), 1–71. MR

[Wooley 1992] T. D. Wooley, "Large improvements in Waring's problem", *Ann. of Math.* (2) **135**:1 (1992), 131–164. MR

[Wooley 1995] T. D. Wooley, "New estimates for smooth Weyl sums", *J. Lond. Math. Soc.* (2) **51**:1 (1995), 1–13. MR

Communicated by Andrew Granville
Received 2022-12-10 Revised 2024-10-17 Accepted 2024-12-05

lilian.matthiesen@mathematik.uni-goettingen.de
Mathematisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany

mengdi.wang@epfl.ch
Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

ALGEBRA AND NUMBER THEORY 19:10 (2025)
https://doi.org/10.2140/ant.2025.19.1947

Stacked pseudo-convergent sequences and polynomial Dedekind domains

Giulio Peruginelli

To the memory of my mother

Let $p \in \mathbb{Z}$ be a prime, $\overline{\mathbb{Q}_p}$ a fixed algebraic closure of the field of p-adic numbers and $\overline{\mathbb{Z}_p}$ the absolute integral closure of the ring of p-adic integers. Given a residually algebraic torsion extension W of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$, by Kaplansky's characterization of immediate extensions of valued fields, there exists a pseudo-convergent sequence of transcendental type $E = \{s_n\}_{n \in \mathbb{N}} \subset \overline{\mathbb{Q}_p}$ such that

$$W = \mathbb{Z}_{(p),E} = \{\phi \in \mathbb{Q}(X) \mid \phi(s_n) \in \overline{\mathbb{Z}_p} \text{ for all sufficiently large } n \in \mathbb{N}\}.$$

We show here that we may assume that E is stacked, in the sense that, for each $n \in \mathbb{N}$, the residue field (resp. the value group) of $\overline{\mathbb{Z}_p} \cap \mathbb{Q}_p(s_n)$ is contained in the residue field (resp. the value group) of $\overline{\mathbb{Z}_p} \cap \mathbb{Q}_p(s_{n+1})$; this property of E allows us to describe the residue field and value group of W. In particular, if W is a DVR, then there exists α in the completion $\mathbb{C}_p$ of $\overline{\mathbb{Q}_p}$, α transcendental over $\mathbb{Q}$, such that $W = \mathbb{Z}_{(p),\alpha} = \{\phi \in \mathbb{Q}(X) \mid \phi(\alpha) \in \mathbb{O}_p\}$, where $\mathbb{O}_p$ is the unique local ring of $\mathbb{C}_p$; α belongs to $\overline{\mathbb{Q}_p}$ if and only if the residue field extension $W/M \supseteq \mathbb{Z}/p\mathbb{Z}$ is finite. As an application, we provide a full characterization of the Dedekind domains between $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$.

Introduction

The problem of characterizing the set of the extensions of a valuation domain V with quotient field K to the field of rational functions $K(X)$ has a long and rich tradition (for example, see [Alexandru and Popescu 1988; Alexandru et al. 1988; 1990a; 1990b; Kaplansky 1942; Matignon and Ohm 1988; Peruginelli 2017; Peruginelli and Spirito 2020; 2021]). One recent direction of research is to describe these extensions by means of pseudo-monotone sequences of K [Peruginelli and Spirito 2021] in the original spirit of Ostrowski [1935a; 1935b], who introduced the well-known notion of pseudo-convergent sequence, later expanded by Kaplansky [1942] to study immediate extensions of valued fields.

Here, given a prime $p \in \mathbb{Z}$ and the DVR $\mathbb{Z}_{(p)}$ of $\mathbb{Q}$, we are interested in describing residually algebraic torsion extensions of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$, that is, valuation domains W of $\mathbb{Q}(X)$ lying above $\mathbb{Z}_{(p)}$ such that the residue field extension $W/M \supseteq \mathbb{Z}/p\mathbb{Z}$ is algebraic and the value group Γ_w of the associated valuation w to W is contained in the divisible hull of the value group of $\mathbb{Z}_{(p)}$ (i.e., the rationals). These valuation

The author thanks the anonymous referee for helpful remarks.
MSC2020: 12J20, 13A18, 13F05, 13F20, 13F30.
Keywords: pseudo-convergent sequence, residually algebraic extension, distinguished pair, minimal pair, Dedekind domain, integer-valued polynomials.

domains arise naturally as overrings of rings of integer-valued polynomials and Dedekind domains between $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$ [Eakin and Heinzer 1973; Peruginelli 2023] and also in the description of closed subfields of $\mathbb{C}_p$ [Ioviţă and Zaharescu 1995], the completion of an algebraic closure $\overline{\mathbb{Q}_p}$ of the field of p-adic numbers $\mathbb{Q}_p$. In the case when W is a DVR and the residue field extension is finite, by [Peruginelli 2017, Theorem 2.5 & Proposition 2.2], there exists an element α in $\overline{\mathbb{Q}_p}$, transcendental over $\mathbb{Q}$, such that $W = \mathbb{Z}_{(p),\alpha} = \{\phi \in \mathbb{Q}(X) \mid \phi(\alpha) \in \overline{\mathbb{Z}_p}\}$, where $\overline{\mathbb{Z}_p}$ is the absolute integral closure of $\mathbb{Z}_p$ (i.e., the integral closure of $\mathbb{Z}_p$ in $\overline{\mathbb{Q}_p}$; note that $\overline{\mathbb{Z}_p}$ is the valuation domain of the unique extension of v_p to $\overline{\mathbb{Q}_p}$). In general, given a residually algebraic torsion extension W of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$, there exists a pseudo-convergent sequence $E = \{s_n\}_{n\in\mathbb{N}}$ in $\overline{\mathbb{Q}_p}$ such that

$$W = \mathbb{Z}_{(p),E} = \{\phi \in \mathbb{Q}(X) \mid \phi(s_n) \in \overline{\mathbb{Z}_p} \text{ for all sufficiently large } n \in \mathbb{N}\}$$

(Proposition 2.24). One of the main results of this paper is to show that we may assume that E is stacked (in a sense we make clear in Section 2; see Theorem 2.5). In particular, if W is a DVR of $\mathbb{Q}(X)$ extending $\mathbb{Z}_{(p)}$ such that the extension of the residue fields is infinite algebraic, then there exists α in $\mathbb{C}_p \setminus \overline{\mathbb{Q}_p}$ such that $W = \mathbb{Z}_{(p),\alpha} = \{\phi \in \mathbb{Q}(X) \mid \phi(\alpha) \in \mathbb{O}_p\}$, where $\mathbb{O}_p$ is the completion of $\overline{\mathbb{Z}_p}$ (equivalently, $\mathbb{O}_p$ is the valuation domain of the unique extension of v_p to $\mathbb{C}_p$; see Corollary 2.28). Necessarily, the (transcendental) extension $\mathbb{Q}_p(\alpha)/\mathbb{Q}_p$ has finite ramification.

It is worth recalling that in [Alexandru et al. 1990a, §5.1, & Theorem 5.1] a residually algebraic torsion extension W of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$ is realized as the limit of a sequence of residually transcendental extensions W_n of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$ (i.e., the residue field extension of W_n over $\mathbb{Z}_{(p)}$ is transcendental); moreover, for each $n \in \mathbb{N}$, W_n is defined by a minimal pair (s_n, δ_n) (as explained in [Alexandru et al. 1990a, p. 282]; for the definition of minimal pair see Section 1.2). Here, W is realized as the valuation domain $\mathbb{Z}_{(p),E}$, where, for each $n \in \mathbb{N}$, $(s_n, \delta_n = v_p(s_{n+1} - s_n))$ is a minimal pair.

The motivations behind these results are based on [Alexandru et al. 1998], in which the authors study closed subfields of $\mathbb{C}_p$ and show that any transcendental element of $\mathbb{C}_p$ is the limit of a particular kind of Cauchy sequence in $\overline{\mathbb{Q}_p}$ called distinguished [Alexandru et al. 1998, Proposition 2.2], which allows them to associate to such an element a set of invariants [Alexandru et al. 1998, Remark 2.4]. The notion of a stacked sequence that we introduce in this paper is a generalization of the notion of a distinguished sequence and falls into the well-known class of pseudo-convergent sequences. It allows us to describe the whole class of residually algebraic torsion extensions of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$, which strictly comprise the valuation domains $\mathbb{Z}_{(p),\alpha}$ arising from elements $\alpha \in \mathbb{C}_p \setminus \overline{\mathbb{Q}_p}$.

As an application of the above results, we are able to complete the classification of the family of Dedekind domains R between $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$ started in [Peruginelli 2023]. In that paper we described the Dedekind domains of this family whose residue fields of prime characteristic are finite fields [Peruginelli 2023, Theorem 2.17]; the description is obtained by means of the notion of rings of integer-valued polynomials over algebras. We also showed that, given a group G which is the direct sum of a countable family of finitely generated abelian groups, there exists a Dedekind domain R with finite residue fields of prime characteristic, $\mathbb{Z}[X] \subset R \subseteq \mathbb{Q}[X]$, with class group G [Peruginelli 2023, Theorem 3.1].

The paper is organized as follows. In Section 1 we recall the relevant notions we need in our paper: First, we review the definition of a pseudo-convergent sequence of a valued field K and the valuation domain of $K(X)$ associated to such a sequence in the spirit of Ostrowski [1935a; 1935b], as developed recently in [Peruginelli and Spirito 2020; 2021]. Then, we recall the notion of a distinguished pair introduced in [Popescu and Zaharescu 1995], which later was used in [Alexandru et al. 1998] to describe closed subfields of $\mathbb{C}_p$ in terms of a specific kind of pseudo-convergent Cauchy sequence called distinguished.

In Section 2, we introduce the notion of a stacked sequence $E = \{s_n\}_{n\in\mathbb{N}}$ in $\overline{\mathbb{Q}_p}$, which turns out to be a pseudo-convergent sequence of transcendental type such that, for each $n \in \mathbb{N}$, the value group (resp. the residue field) of $\overline{\mathbb{Z}_p} \cap \mathbb{Q}_p(s_n)$ is contained in the value group (resp. the residue field) of $\overline{\mathbb{Z}_p} \cap \mathbb{Q}_p(s_{n+1})$. By Theorem 2.5, every residually algebraic extension W of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$ can be realized by means of a stacked sequence $E \subset \overline{\mathbb{Q}_p}$, that is,

$$W = \mathbb{Z}_{p,E} = \{\phi \in \mathbb{Q}_p(X) \mid \phi(s_n) \in \overline{\mathbb{Z}_p} \text{ for all sufficiently large } n \in \mathbb{N}\}.$$

Moreover, the above specific property of stacked sequences is crucial for the description of the residue field and value group of W as the union of the ascending chain of residue fields and value groups of $\overline{\mathbb{Z}_p} \cap \mathbb{Q}_p(s_n)$, respectively (Proposition 2.7). We mentioned above that the elements $\alpha \in \mathbb{C}_p \setminus \overline{\mathbb{Q}_p}$ such that the extension $\mathbb{Q}_p(\alpha)/\mathbb{Q}_p$ has finite ramification give raise to DVRs of $\mathbb{Q}(X)$; we characterize such elements as the limits of sequences contained in the maximal unramified extension of a finite extension of $\mathbb{Q}_p$ (Proposition 2.20). We close this section by pointing out an incorrect statement in [Ioviţă and Zaharescu 1995], namely, that the completion of $\mathbb{Q}_p(X)$ with respect to a residually algebraic torsion extension W of $\mathbb{Z}_p$ is a subfield of $\mathbb{C}_p$; this is not true in general and it depends on whether the above sequence E is Cauchy or not. In Section 2.3, we use the result of Section 2.1 about residually algebraic torsion extensions of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$ to characterize the analogous extensions of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$ (Proposition 2.24). In Theorem 2.26, we show that, for any prescribed algebraic extension k of $\mathbb{F}_p$ and value group Γ, $\mathbb{Z} \subseteq \Gamma \subseteq \mathbb{Q}$, there exists $\alpha \in \mathbb{C}_p$, transcendental over $\mathbb{Q}$, such that $\mathbb{Z}_{(p),\alpha}$ has residue field k and value group Γ.

Finally, in Section 3 we provide the aforementioned classification of the Dedekind domains between $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$ by means of the notion of the ring of integer-valued polynomials over an algebra: given such a domain R, we show that, for each prime $p \in \mathbb{Z}$, there exists a finite set $E_p \subset \mathbb{C}_p$ of transcendental elements over $\mathbb{Q}$ such that $R = \{f \in \mathbb{Q}[X] \mid f(E_p) \subseteq \mathbb{O}_p, \forall p \in \mathbb{P}\}$ (Theorem 3.4).

1. Preliminaries

We refer to [Bourbaki 1985a; Engler and Prestel 2005; Ribenboim 1968; Zariski and Samuel 1960] for generalities about valuation theory. A valuation domain W of the field of rational functions $K(X)$ is an extension of a valuation domain V of K if $W \cap K = V$. We denote by w a valuation associated to W, by Γ_w the value group of w and by k_w the residue field of W. We recall that an extension W of V to $K(X)$ is called residually algebraic if the residue field extension is algebraic, and it is called torsion if Γ_w is contained in the divisible hull of the value group Γ_v of V; see [Alexandru et al. 1990a]. Given a valuation domain W with quotient field F, a subfield K of F and the valuation domain $V = W \cap K$, we say that W is an immediate extension of V (or simply immediate over V) if the value groups (resp. the residue fields) of V and W are the same. Given a field K with a valuation domain V, we denote by $\widehat{K}$ (resp. $\widehat{V}$) the completion of K (resp. V) with respect to V-adic topology.

1.1. ***Pseudo-convergent sequences.*** The following basic material about pseudo-convergent sequences can be found for example in [Kaplansky 1942; Peruginelli and Spirito 2020; 2021].

Given a valued field (K, v), a sequence $E = \{s_n\}_{n\in\mathbb{N}} \subset K$ is said to be *pseudo-convergent* if, for all $n < m < k$, we have

$$v(s_n - s_m) < v(s_m - s_k).$$

In particular, for all n and $m > n$, we have $v(s_n - s_m) = v(s_n - s_{n+1})$. For each $n \in \mathbb{N}$, we set $\delta_n = v(s_n - s_{n+1})$. The strictly increasing sequence $\{\delta_n\}_{n\in\mathbb{N}}$ of the value group Γ_v of v is called the *gauge* of E. The sequence E is a classical Cauchy sequence in K if and only if the gauge of E is cofinal in Γ_v. In this case, E converges to a unique limit $\alpha \in \widehat{K}$. In general, if $E = \{s_n\}_{n\in\mathbb{N}} \subset K$ is a pseudo-convergent sequence, we say that an element $\alpha \in K$ is a *pseudo-limit* of E if $v(s_n - \alpha)$ is a strictly increasing sequence. Equivalently, $v(s_n - \alpha) = \delta_n$ for each $n \in \mathbb{N}$. The set of pseudo-limits $\mathcal{L}_E$ in K of a pseudo-convergent sequence E is equal to $\mathcal{L}_E = \alpha + \mathrm{Br}(E)$ [Kaplansky 1942, Lemma 3], where

$$\mathrm{Br}(E) = \{x \in K \mid v(x) > \delta_n, \forall n \in \mathbb{N}\}$$

is a fractional ideal, called the *breadth ideal* of E. Clearly, E is a Cauchy sequence if and only if $\mathrm{Br}(E) = \{0\}$.

As in [Kaplansky 1942, Definitions, p. 306], a pseudo-convergent sequence $E = \{s_n\}_{n\in\mathbb{N}} \subset K$ is of *transcendental type* if, for all $f \in K[X]$, $v(f(s_n))$ is eventually constant. Otherwise, E is said to be of *algebraic type* if $v(f(s_n))$ is eventually strictly increasing for some $f \in K[X]$. The sequence E is of algebraic type if and only if, for some extension u of v to the algebraic closure $\overline{K}$ of K, there exists $\alpha \in \overline{K}$ which is a pseudo-limit of E with respect to u. If F is a subfield of K, then we say that E is of *transcendental type over* F if, for all $f \in F[X]$, $v(f(s_n))$ is eventually constant. Almost all the pseudo-convergent sequences considered in this paper in order to describe residually algebraic torsion extensions to the field of rational functions are of transcendental type.

Given a pseudo-convergent sequence $E = \{s_n\}_{n\in\mathbb{N}} \subset K$, the following is a valuation domain of $K(X)$ extending V associated to E [Peruginelli and Spirito 2020, Theorem 3.8]:

$$V_E = \{\phi \in K(X) \mid \phi(s_n) \in V \text{ for all sufficiently large } n \in \mathbb{N}\}.$$

Moreover, by the same theorem, X is a pseudo-limit of E with respect to the valuation v_E associated to V_E, so, in particular, $v_E(X - s_n) = \delta_n$ for every $n \in \mathbb{N}$. Also, if E is of transcendental type, then, for all $f \in K[X]$, we have $v_E(f) = v(f(s_n))$ for all n sufficiently large; see [Kaplansky 1942, Theorem 2] or [Peruginelli and Spirito 2020, Theorem 4.9, (a)].

In the case that E is a Cauchy sequence converging to $\alpha \in \widehat{K}$, we have

$$V_E = V_\alpha = \{\phi \in K(X) \mid \phi(\alpha) \in \widehat{V}\};$$

see [Peruginelli and Spirito 2020, Remark 3.10].

Given two pseudo-convergent sequences $E = \{s_n\}_{n\in\mathbb{N}}$, $E' = \{s'_n\}_{n\in\mathbb{N}} \subset K$, we say that E and E' are equivalent if $\mathrm{Br}(E) = \mathrm{Br}(E')$ and, for each $k \in \mathbb{N}$, there exist i_0, $j_0 \in \mathbb{N}$ such that $v(s_i - s'_j) > v(s_{k+1} - s_k)$ for each $i \geq i_0$ and $j \geq j_0$; see [Peruginelli and Spirito 2020, §5]. By Proposition 5.3 in that work, E and E' are equivalent if and only if $V_E = V_{E'}$.

1.2. *Distinguished pairs.* We suppose in this section that (K, v) is a complete valued field, where v is a rank-1 discrete valuation (so, in particular, (K, v) is Henselian). Let $\overline{K}$ be a fixed algebraic closure, and let v denote the unique extension of v to $\overline{K}$. Let also $\Gamma_{\bar{v}} = \Gamma_v \otimes \mathbb{Q}$ be the divisible hull of Γ_v. Given an element $a \in \overline{K}$, let O_a, k_a and Γ_a be the valuation domain of the restriction of v to $K(a)$, the residue field of O_a and the value group of O_a, respectively.

As in [Khanduja and Saha 1999], given $a \in \overline{K} \setminus K$, we set

$$\delta_K(a) = \sup\{v(a - c) \mid c \in \overline{K}, [K(c) : K] < [K(a) : K]\},$$
$$\omega_K(a) = \sup\{v(a - a') \mid a' \neq a \text{ runs over the } K\text{-conjugates of } a\}.$$

The following is the well-known Krasner's lemma. Essentially, given a separable element $a \in \overline{K}$, if another element $b \in \overline{K}$ is closer to a than to any of its other conjugates, then $K(a)$ is a subfield of $K(b)$.

Lemma 1.1 (Krasner). *If $a \in \overline{K}^{\mathrm{sep}}$ and $b \in \overline{K}$ are such that $v(a - b) > \omega_K(a)$, then $K(a) \subseteq K(b)$.*

In particular, for every $a \in \overline{K}^{\mathrm{sep}}$, we have $\delta_K(a) \leq \omega_K(a)$. Moreover, it follows also that $\delta_K(a)$ is a maximum, since v is supposed to be discrete. This is known (see, for example, [Popescu and Zaharescu 1995, p. 105]), but for the sake of the reader we give a short proof.

Lemma 1.2. *In the above setting*,

$$\delta_K(a) = \max\{v(a - c) \mid c \in \overline{K}, [K(c) : K] < [K(a) : K]\}.$$

Proof. By Krasner's lemma, for each of the relevant c we have $v(a-c) \leq \omega_K(a)$. Note that the ramification index of $K(a, c)$ over K is (strictly) bounded by $[K(a) : K]^2$. In particular, since the value $v(a - c)$ belongs to Γ_{a-c}, it follows that there exists $N \in \mathbb{N}$, independent from each of the above c, such that $Nv(a - c) \in \Gamma_v \cong \mathbb{Z}$. Hence the set

$$\{v(a - c) \mid c \in \overline{K}, [K(c) : K] < [K(a) : K]\}$$

(which is a subset of $\Gamma_{\bar{v}}$) is bounded from above and its elements have bounded torsion. It follows that this set has a maximum, which is equal to $\delta_K(a)$ by its very definition. □

Similar to Krasner's lemma, we have the following fundamental principle (see [Khanduja and Saha 1999, Theorem 1.1]), first discovered in [Popescu and Zaharescu 1995].

Theorem 1.3. *Suppose that $a, b \in \overline{K}$ are such that $v(a-b) > \delta_K(b)$. Then*:

(i) $\Gamma_b \subseteq \Gamma_a$.

(ii) $k_b \subseteq k_a$.

(iii) $[K(b):K] \mid [K(a):K]$.

Next, we recall the definition of a distinguished pair introduced in [Popescu and Zaharescu 1995, p. 105].

Definition 1.4. A pair of elements $(b, a) \in \overline{K}^2$ is said to be *distinguished* if the following hold:

(i) $[K(b):K] < [K(a):K]$.

(ii) For all $c \in \overline{K}$ such that $[K(c):K] < [K(a):K]$, we have $v(a-c) \leq v(a-b)$.

(iii) For all $c \in \overline{K}$ such that $[K(c):K] < [K(b):K]$, we have $v(a-c) < v(a-b)$.

Part of the definition of a distinguished pair is related to the notion of a minimal pair, which we now recall (see, for example, [Alexandru et al. 1988; 1990a; 1990b]).

Definition 1.5. Let $(a, \delta) \in \overline{K} \times \Gamma_{\bar{v}}$. We say that (a, δ) is a *minimal pair* if, for every $c \in \overline{K}$ such that $[K(c):K] < [K(a):K]$, we have $v(a-c) < \delta$.

In other words, (a, δ) is a minimal pair if, for every $b \in B(a, \delta) = \{x \in \overline{K} \mid v(a-x) \geq \delta\}$, we have $[K(b):K] \geq [K(a):K]$ (i.e., a is a "center" of the ball $B(a, \delta)$ of minimal degree). By Lemma 1.2, (a, δ) is a minimal pair if and only if $\delta > \delta_K(a)$. In particular, if $\delta > \omega_K(a)$, then (a, δ) is a minimal pair.

Remarks 1.6. Let (b, a) be a distinguished pair.

(1) Note that conditions (i) and (ii) above imply that $v(a-b) = \delta_K(a)$. In fact, by (i) and (ii), it immediately follows that the inequality "$\leq$" holds. Conversely, by (ii) we also have that $v(a-b) \geq v(a-c)$ for all c such that $[K(c):K] < [K(a):K]$; that is, $v(a-b) \geq \delta_K(a)$.

(2) Note that (iii) is equivalent to the following:

(iii′) For all $c \in \overline{K}$ such that $[K(c):K] < [K(b):K]$, we have $v(b-c) < v(a-b)$.

This precisely says that $(b, v(a-b))$ is a minimal pair with respect to K. In fact, if (iii) holds and $c \in \overline{K}$ is such that $[K(c):K] < [K(b):K]$, then $v(b-c) = v(b-a+a-c) = v(a-c) < v(a-b)$. Similarly, (iii′) implies (iii). Note also that (iii′) is equivalent to

$$v(a-b) > \delta_K(b).$$

In particular, by the above theorem, $\Gamma_b \subseteq \Gamma_a$, $k_b \subseteq k_a$ and $[K(b):K] \mid [K(a):K]$.

(3) Finally, note also that $\delta_K(b) < \delta_K(a)$.

2. Stacked pseudo-convergent sequences of $\overline{\mathbb{Q}_p}$

Let $\mathbb{P} \subset \mathbb{Z}$ be the set of prime numbers, and let $p \in \mathbb{P}$ be a fixed prime. We let $\mathbb{Z}_{(p)}$ be the localization of $\mathbb{Z}$ at the prime ideal $p\mathbb{Z}$, $\mathbb{Z}_p$ the ring of p-adic integers and $\mathbb{Q}_p$ its field of fractions, the field of p-adic numbers. If v_p denotes the usual p-adic valuation, then $\mathbb{Z}_p$ (resp. $\mathbb{Q}_p$) is the completion of $\mathbb{Z}$ (resp. $\mathbb{Q}$) with respect to the p-adic valuation. We denote by $\overline{\mathbb{Q}_p}$ a fixed algebraic closure of $\mathbb{Q}_p$ and still denote the unique extension of v_p to $\overline{\mathbb{Q}_p}$ by v_p. Note that $\overline{\mathbb{Q}_p}$ is a rank-1 nondiscrete valued field with valuation domain denoted by $\overline{\mathbb{Z}_p}$, the integral closure of $\mathbb{Z}_p$ in $\overline{\mathbb{Q}_p}$. We will use the well-known fact that $\mathbb{Q}_p$ has only finitely many extensions of a given degree; see, for example, [Narkiewicz 2004, Corollary 2, Chapter V, p. 202].

Finally, we let $\mathbb{C}_p$ be the completion of $\overline{\mathbb{Q}_p}$ with respect to the p-adic valuation, and we denote by $\mathbb{O}_p$ the completion of $\overline{\mathbb{Z}_p}$; v_p still denotes the unique extension of v_p to $\mathbb{C}_p$. For $\alpha \in \overline{\mathbb{Q}_p} \setminus \mathbb{Q}_p$, we write the abbreviations $\delta_{\mathbb{Q}_p}(\alpha) = \delta(\alpha)$ and $\omega_{\mathbb{Q}_p}(\alpha) = \omega(\alpha)$. For $\alpha \in \mathbb{C}_p$, we denote by e_α (resp. f_α) the ramification index (resp. the residue field degree) of $\mathbb{Q}_p(\alpha)$ over $\mathbb{Q}_p$. Clearly, if $\alpha \in \overline{\mathbb{Q}_p}$, then $e_\alpha \cdot f_\alpha < \infty$; we show that the converse holds in Remark 2.15. Note that each element of $\mathbb{C}_p \setminus \overline{\mathbb{Q}_p}$ is transcendental over $\mathbb{Q}_p$; we call such elements simply transcendental. For a transcendental element $\alpha \in \mathbb{C}_p$, even if $e_\alpha \cdot f_\alpha = \infty$, we will show in Theorem 2.21 that either one of e_α or f_α can be finite.

2.1. ***Residually algebraic torsion extensions of $\mathbb{Z}_p$.*** In this section we describe residually algebraic torsion extensions of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$ by means of a suitable class of pseudo-convergent sequences of transcendental type contained in $\overline{\mathbb{Q}_p}$, called a stacked sequence, which we now introduce. This definition is a generalization of [Alexandru et al. 1998, p. 135].[1]

Definition 2.1. Let $E = \{s_n\}_{n \geq 0} \subset \overline{\mathbb{Q}_p}$ be a sequence with $s_0 \in \mathbb{Q}_p$. For every $n \geq 0$, we consider the following properties:

(i) $[\mathbb{Q}_p(s_n) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_{n+1}) : \mathbb{Q}_p]$.

(ii) For every $c \in \overline{\mathbb{Q}_p}$ such that $[\mathbb{Q}_p(c) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_{n+1}) : \mathbb{Q}_p]$, we have $v(s_{n+1} - c) \leq v(s_{n+1} - s_n)$.

(iii) For every $c \in \overline{\mathbb{Q}_p}$ such that $[\mathbb{Q}_p(c) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$, we have $v(s_n - c) < v(s_{n+1} - s_n)$.

We say that E is *unbounded* if (i) holds for every n, *stacked* if (i) and (iii) hold for every n, and *strongly stacked* if (i), (ii), (iii) hold for every n. Equivalently, E is stacked if (i) holds and $(s_n, \delta_n = v(s_{n+1} - s_n))$ is a minimal pair for every $n \geq 0$, and E is strongly stacked if (s_n, s_{n+1}) is distinguished for every $n \geq 0$.

Remark 2.2. Let $E = \{s_n\}_{n \in \mathbb{N}} \subset \overline{\mathbb{Q}_p}$ be a stacked sequence. Note that the sequence $\{v(s_{n+1} - s_n) = \delta_n\}_{n \in \mathbb{N}}$ is strictly increasing since $[\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$ and (s_n, δ_n) is a minimal pair. In the original definition of a distinguished sequence E in [Alexandru et al. 1998], the sequence $\{\delta_n\}_{n \in \mathbb{N}}$ is unbounded; thus, in this case E is a Cauchy sequence. In our setting we are not imposing that restriction; we show in Lemma 2.3 below that a stacked sequence is a pseudo-convergent sequence of transcendental type of $\overline{\mathbb{Q}_p}$.

[1] The notion of a distinguished sequence was introduced in [Alexandru et al. 1998]. We cannot borrow that term here for our sequences for the following reason: by Lemma 2.3, a stacked sequence is pseudo-convergent, and distinguished pseudo-convergent sequences have already been defined by P. Ribenboim [1958, p. 474] to denote pseudo-convergent sequences of a valued field whose breadth ideal is a nonmaximal prime ideal.

The motivation for the terminology of these kind of sequences is due to the following fact. For each $n \in \mathbb{N}$, we abbreviate $\Gamma_n = \Gamma_{s_n}$ and $k_n = k_{s_n}$ (i.e., the value group and the residue field of the valuation domain O_{s_n} of $\mathbb{Q}_p(s_n)$, respectively). By Remarks 1.6, $v(s_{n+1} - s_n) > \delta(s_n)$. Hence, by Theorem 1.3, we have $\Gamma_n \subseteq \Gamma_{n+1}$ and $k_n \subseteq k_{n+1}$. For each $n \in \mathbb{N}$, we set $e_n = e(\mathbb{Q}_p(s_n) \,|\, \mathbb{Q}_p)$ and $f_n = f(\mathbb{Q}_p(s_n) \,|\, \mathbb{Q}_p)$, the ramification index and the residue field degree of O_{s_n} over $\mathbb{Z}_p$, respectively; we remark that $[\mathbb{Q}_p(s_n) : \mathbb{Q}_p] = e_n f_n = d_n$ for each $n \in \mathbb{N}$, and since $\{d_n\}_{n \in \mathbb{N}}$ is unbounded by assumption, either $\{e_n\}_{n \in \mathbb{N}}$ is unbounded or $\{f_n\}_{n \in \mathbb{N}}$ is unbounded. Since $e_n \,|\, e_{n+1}$ for each $n \in \mathbb{N}$, $\{e_n\}_{n \in \mathbb{N}}$ is bounded if and only if $e_n = e$ for all $n \in \mathbb{N}$ sufficiently large. Similarly for $\{f_n\}_{n \in \mathbb{N}}$.

By Remarks 1.6, condition (ii) is equivalent to $\delta_n = v(s_n - s_{n+1}) = \delta(s_{n+1})$ (note that in general the inequality $\delta_n \leq \delta(s_{n+1})$ holds). In other words, among all the elements $c \in \overline{\mathbb{Q}_p}$ such that

$$[\mathbb{Q}_p(s_n) : \mathbb{Q}_p] \leq [\mathbb{Q}_p(c) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_{n+1}) : \mathbb{Q}_p],$$

s_n is one of those which is closest to s_{n+1}.

Let $E = \{t_n\}_{n \in \mathbb{N}} \subset \overline{\mathbb{Q}_p}$ be a pseudo-convergent sequence. If $\{[\mathbb{Q}_p(t_n) : \mathbb{Q}_p] \mid n \in \mathbb{N}\}$ is bounded, then E is contained in a finite extension K of $\mathbb{Q}_p$, and hence E is Cauchy and therefore converges to an element $\alpha \in K$. In particular, if E is of transcendental type, then the set $\{[\mathbb{Q}_p(t_n) : \mathbb{Q}_p]\}_{n \in \mathbb{N}}$ is necessarily unbounded. Stacked sequences are of this kind, as the next lemma shows.

Lemma 2.3. *Let $E \subset \overline{\mathbb{Q}_p}$ be a stacked sequence. Then E is a pseudo-convergent sequence of transcendental type.*

Proof. Let $E = \{s_n\}_{n \in \mathbb{N}}$, and set $\delta_n = v(s_{n+1} - s_n)$ for each $n \in \mathbb{N}$. We have already observed in Remark 2.2 that $\{\delta_n\}_{n \in \mathbb{N}}$ is a strictly increasing sequence. Moreover, for every $m > n$, we have $v(s_n - s_m) > v(s_n - s_{n-1})$. In particular, $v(s_{n-1} - s_m) = v(s_{n-1} - s_n)$ for every $m \geq n$. Let now $n < m < k$. Then

$$v(s_n - s_m) = v(s_n - s_{n+1}) < v(s_m - s_{m+1}) = v(s_m - s_k),$$

which shows that E is a pseudo-convergent sequence.

We prove now that E is of transcendental type. Let $\alpha \in \overline{\mathbb{Q}_p}$. Then there exists $n \in \mathbb{N}$ such that

$$[\mathbb{Q}_p(\alpha) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_n) : \mathbb{Q}_p].$$

Since (s_n, δ_n) is a minimal pair, $v(s_n - \alpha) < \delta_n$, so, in particular, α cannot be a pseudo-limit of E. This shows that E has no pseudo-limits in $\overline{\mathbb{Q}_p}$, and thus E is of transcendental type. □

Let $E = \{s_n\}_{n \in \mathbb{N}} \subset \overline{\mathbb{Q}_p}$ be a stacked sequence. In particular, by Lemma 2.3, the sequence

$$\{\delta_n = v(s_{n+1} - s_n)\}_{n \in \mathbb{N}}$$

is the gauge of the pseudo-convergent sequence E. Moreover, by the same lemma, if E is Cauchy, then E converges to a transcendental element $\alpha \in \mathbb{C}_p$.

The next proposition shows that any residually algebraic torsion extension of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$ is obtained by means of a pseudo-convergent sequence of transcendental type of $\overline{\mathbb{Q}_p}$. We recall that if $E \subset \overline{\mathbb{Q}_p}$ is a pseudo-convergent sequence of transcendental type, then $\overline{\mathbb{Z}_p}_E$, the associated valuation domain of $\overline{\mathbb{Q}_p}(X)$, is an immediate extension of $\overline{\mathbb{Z}_p}$ and conversely every immediate extension of $\overline{\mathbb{Z}_p}$ to $\overline{\mathbb{Q}_p}(X)$ can be realized in this way; see, for example, [Kaplansky 1942; Peruginelli and Spirito 2021]. If $\mathbb{Z}_{p,E} = \overline{\mathbb{Z}_p}_E \cap \mathbb{Q}_p(X)$, then $\mathbb{Z}_{p,E}$ is a residually algebraic torsion extension of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$.

Proposition 2.4. *Let W be a residually algebraic torsion extension of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$. Then there exists a pseudo-convergent sequence $E \subset \overline{\mathbb{Q}_p}$ of transcendental type such that*

$$W = \mathbb{Z}_{p,E} = \{\phi \in \mathbb{Q}_p(X) \mid \phi(s_n) \in \overline{\mathbb{Z}_p} \text{ for all sufficiently large } n \in \mathbb{N}\}.$$

Proof. Let $\overline{W}$ be an extension of W to $\overline{\mathbb{Q}_p}(X)$. Then $\overline{W}$ is an immediate extension of $\overline{\mathbb{Z}_p}$ to $\overline{\mathbb{Q}_p}(X)$ (and, in particular, is a residually algebraic torsion extension of $\overline{\mathbb{Z}_p}$). By [Kaplansky 1942, Theorems 1 and 2] or [Peruginelli and Spirito 2021, Theorem 6.2 (a)], there exists a pseudo-convergent sequence $E \subset \overline{\mathbb{Q}_p}$ of transcendental type such that $\overline{W} = \overline{\mathbb{Z}_p}_E$. The claim follows by contracting down to $\mathbb{Q}_p(X)$. □

Clearly, not every pseudo-convergent sequence of transcendental type in $\overline{\mathbb{Q}_p}$ is stacked. However, the next theorem is the converse of Lemma 2.3: it shows that any pseudo-convergent sequence of transcendental type is equivalent to a strongly stacked sequence. In particular, every stacked sequence is equivalent to a strongly stacked sequence. Moreover, given a valuation domain $\mathbb{Z}_{p,E}$ of $\mathbb{Q}_p(X)$ associated to a pseudo-convergent sequence $E \subset \overline{\mathbb{Q}_p}$ of transcendental type, without loss of generality, we may also assume that E is strongly stacked.

By [Alexandru et al. 1998, Proposition 2.2], every transcendental element $t \in \mathbb{C}_p$ is the limit of a strongly stacked sequence E of $\overline{\mathbb{Q}_p}$. The next theorem is the analog of that result for residually algebraic extensions W of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$: for such a valuation W, there exists a strongly stacked sequence $E \subset \overline{\mathbb{Q}_p}$ such that $W = \mathbb{Z}_{p,E}$; it is not difficult to show that, for a transcendental element $t \in \mathbb{C}_p$, the valuation domain

$$\mathbb{Z}_{p,t} = \{\phi \in \mathbb{Q}_p(X) \mid \phi(t) \in \mathbb{O}_p\}$$

is a residually algebraic torsion extension of $\mathbb{Z}_p$.

Theorem 2.5. *Let $E \subset \overline{\mathbb{Q}_p}$ be a pseudo-convergent sequence of transcendental type. Then there exists a strongly stacked sequence $E' \subset \overline{\mathbb{Q}_p}$ which is equivalent to E. In particular, given a residually algebraic torsion extension W of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$, there exists a strongly stacked sequence $E' \subset \overline{\mathbb{Q}_p}$ such that $W = \mathbb{Z}_{p,E'}$.*

Proof. Let $E = \{t_n\}_{n\in\mathbb{N}}$, and let $\bar{v}_E$ be a valuation associated to $\overline{\mathbb{Z}_p}_E \subset \overline{\mathbb{Q}_p}(X)$.

First, we consider the following subset of $\Gamma_{v_E} \subseteq \mathbb{Q}$:

$$M_E(X, \mathbb{Q}_p) = \{v_E(X - s) \mid s \in \mathbb{Q}_p\}.$$

If $M_E(X, \mathbb{Q}_p)$ is not bounded, then there exists a sequence $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{Q}_p$ such that $v_E(X - s_n)$ tends to ∞. Necessarily, the sequence $\{s_n\}_{n\in\mathbb{N}}$ is Cauchy and so converges to an element s of $\mathbb{Q}_p$. Now, for every n, $v_E(X - s_n) = \bar{v}_E(X - s_n) = v(t_m - s_n)$ for all m sufficiently large since E is of transcendental type (see Section 1.1). Hence E would be a Cauchy sequence equivalent to $\{s_n\}_{n\in\mathbb{N}}$ and E would converge to s, too, which is not possible. Let then $\delta_0 = \sup M_E(X, \mathbb{Q}_p) \in \mathbb{R}$. We claim that $\delta_0 \in M_E(X, \mathbb{Q}_p)$; that is, δ_0 is a maximum. Suppose otherwise: there exists a sequence $\{r_k\}_{k\in\mathbb{N}} \subset \mathbb{Q}_p$ such that $v_E(X - r_k) \nearrow \delta_0$. Then $\{r_k\}_{k\in\mathbb{N}} \subset \mathbb{Q}_p$ would be a pseudo-convergent sequence which is not Cauchy, which is not possible, since $\mathbb{Q}_p$ is a complete valued field. Hence there exists $s_0 \in \mathbb{Q}_p$ such that $v_E(X - s_0) = \delta_0$.

For $n > 0$, we now choose $s_n \in \overline{\mathbb{Q}_p}$ so that (s_{n-1}, s_n) is distinguished. Let B_n be the subset of the α in $\overline{\mathbb{Q}_p}$ satisfying the following properties:

(i) $[\mathbb{Q}_p(\alpha) : \mathbb{Q}_p] > [\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p]$.

(ii) $\bar{v}_E(X - \alpha) > \bar{v}_E(X - s_{n-1})$.

(iii) The positive integer $[\mathbb{Q}_p(\alpha) : \mathbb{Q}_p] - [\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p]$ is minimal.

Note that since $\mathbb{N}$ is well-ordered, condition (iii) can be satisfied (that is, among the $\alpha \in \overline{\mathbb{Q}_p}$ satisfying (i) and (ii), we can find one which also satisfies (iii)). Since $\bar{v}_E(X - s_{n-1}) = v(t_m - s_{n-1})$ for all m sufficiently large, for all such m we also have $\bar{v}_E(X - s_{n-1}) < \bar{v}_E(X - t_m)$. Moreover, without loss of generality, we may also assume that $[\mathbb{Q}_p(t_m) : \mathbb{Q}_p] > [\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p]$ since $\{[\mathbb{Q}_p(t_m) : \mathbb{Q}_p]\}_{m\in\mathbb{N}}$ is unbounded. This shows that the set B_n is nonempty. Let

$$M_E(X, B_n) = \{\bar{v}_E(X - \alpha) \mid \alpha \in B_n\},$$

which is a subset of $\mathbb{Q}$. Let $\delta_n = \sup M_E(X, B_n)$. Since each element of B_n has the same degree over $\mathbb{Q}_p$, it follows that B_n is contained in a finite extension K of $\mathbb{Q}_p$. In particular, it follows as above that $M_E(X, B_n)$ is bounded above. Let $\delta_n = \sup M_E(X, B_n) \in \mathbb{R}$. Next, we show that $M_E(X, B_n)$ contains its upper bound (which is, in particular, a rational number). Suppose otherwise: then there exists a sequence $\{\alpha_k\}_{k\in\mathbb{N}} \subset B_n$ such that $\bar{v}_E(X - \alpha_k) \nearrow \delta_n$. In particular, $\{\alpha_k\}_{k\in\mathbb{N}}$ would be a pseudo-convergent sequence of a finite extension of $\mathbb{Q}_p$ which is not Cauchy, which is impossible. Let $s_n \in B_n$ be such that $\bar{v}_E(X - s_n) = \delta_n$. Note that

$$v_p(s_n - s_{n-1}) = \bar{v}_E(s_n - X + X - s_{n-1}) = \bar{v}_E(X - s_{n-1}) = \delta_{n-1}.$$

We now show that (s_{n-1}, s_n) is distinguished. Clearly, $[\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$.

Let $c \in \overline{\mathbb{Q}_p}$ be such that $[\mathbb{Q}_p(c) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$.

If $[\mathbb{Q}_p(c) : \mathbb{Q}_p] > [\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p]$, then, by the minimality of the degree of s_n, we have

$$\bar{v}_E(X - c) \leq \bar{v}_E(X - s_{n-1}) = \delta_{n-1},$$

so

$$v_p(s_n - c) = \bar{v}_E(s_n - X + X - c) = \bar{v}_E(X - c) \leq \delta_{n-1} = v_p(s_n - s_{n-1}).$$

Suppose now that $[\mathbb{Q}_p(c) : \mathbb{Q}_p] = [\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p]$.

If $\bar{v}_E(X - c) \leq \bar{v}_E(X - s_{n-2})$, then $\bar{v}_E(X - c) < \delta_{n-1}$.

If $\bar{v}_E(X - c) > \bar{v}_E(X - s_{n-2})$, then $c \in B_{n-1}$, so

$$\bar{v}_E(X - c) \leq \delta_{n-1} = \bar{v}_E(X - s_{n-1}).$$

In either case,

$$v_p(s_n - c) = \bar{v}_E(s_n - X + X - c) = \bar{v}_E(X - c) \leq \delta_{n-1} = v_p(s_n - s_{n-1}).$$

Note that, in particular, for $n = 1$, we have that (s_0, s_1) is distinguished since condition (iii) of Definition 2.1 is empty, since $s_0 \in \mathbb{Q}_p$.

Suppose now that $n \geq 2$, and assume by induction that (s_{n-2}, s_{n-1}) is distinguished. Let $c \in \overline{\mathbb{Q}_p}$ be such that $[\mathbb{Q}_p(c) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p]$. Since (s_{n-2}, s_{n-1}) is distinguished, we have

$$v_p(s_{n-1} - c) \leq v_p(s_{n-1} - s_{n-2}) = \delta_{n-2} < \delta_{n-1}.$$

Hence

$$v_p(s_n - c) = v_p(s_n - s_{n-1} + s_{n-1} - c) = v_p(s_{n-1} - c) < v_p(s_n - s_{n-1}).$$

We now show that $E' = \{s_n\}_{n\in\mathbb{N}}$ is equivalent to $E = \{t_n\}_{n\in\mathbb{N}}$. Let $\{\lambda_n\}_{n\in\mathbb{N}}$ and $\{\delta_n\}_{n\in\mathbb{N}}$ be the gauges of E and E', respectively. We need to show that, for each $k \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that $\lambda_k \leq \delta_n$. Since E' is unbounded, there exists $n \in \mathbb{N}$ such that $[\mathbb{Q}_p(t_k) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$. Since (s_n, δ_n) is a minimal pair, we have $v_p(s_n - t_k) < \delta_n$, so that

$$\lambda_k = \bar{v}_E(X - t_k) = \bar{v}_E(X - s_n + s_n - t_k) < \bar{v}_E(X - s_n) = \delta_n. \tag{2.6}$$

Conversely, let $n \in \mathbb{N}$. We need to show that there exists $k \in \mathbb{N}$ such that $\delta_n \leq \lambda_k$. For all m sufficiently large, we have

$$\bar{v}_E(X - s_n) = v_p(t_m - s_n) = \bar{v}_E(t_m - X + X - s_n),$$

and since n is fixed and $\bar{v}_E(t_m - X) = \lambda_m$ is strictly increasing, it follows that

$$\bar{v}_E(t_m - X) = \lambda_m > \bar{v}_E(X - s_n)$$

for all such m.

Hence $\mathrm{Br}(E) = \mathrm{Br}(E')$.

Finally, we need to show that, if $k \in \mathbb{N}$, then there exist $n_0, m_0 \in \mathbb{N}$ such that, for each $n \geq n_0$ and $m \geq m_0$, we have $v_p(t_n - s_m) > \lambda_k$. Let n_0 be the smallest integer such that $[\mathbb{Q}_p(t_k) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_{n_0}) : \mathbb{Q}_p]$. As in (2.6) above, $\lambda_k < v_E(X - s_{n_0}) = \delta_{n_0}$. Let now $m > k$ and $n \geq n_0$. Then,

$$v(t_m - s_n) = \bar{v}_E(t_m - X + X - s_n) > \lambda_k$$

since

$$\bar{v}_E(t_m - X) = \lambda_m > \lambda_k \quad \text{and} \quad \bar{v}_E(X - s_n) \geq \bar{v}_E(X - s_{n_0}) = \delta_{n_0} > \lambda_k.$$

Hence E and E' are equivalent.

By [Peruginelli and Spirito 2021, Proposition 5.3], $\overline{\mathbb{Z}_p}_E = \overline{\mathbb{Z}_p}_{E'}$, so, in particular, $\mathbb{Z}_{p,E} = \mathbb{Z}_{p,E'}$. The final claim follows by Proposition 2.4. □

The following proposition describes the value group and the residue field of a residually algebraic torsion extension W of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$. By Theorem 2.5, W is equal to $\mathbb{Z}_{p,E}$ for some strongly stacked sequence $E \subset \overline{\mathbb{Q}_p}$. We keep the notation of Remark 2.2.

Proposition 2.7. *Let $E = \{s_n\}_{n\in\mathbb{N}} \subset \overline{\mathbb{Q}_p}$ be a stacked sequence and $W = \mathbb{Z}_{p,E}$. Then we have*

$$\bigcup_{n\in\mathbb{N}} \Gamma_n = \Gamma_w, \quad \bigcup_{n\in\mathbb{N}} k_n = k_w.$$

Proof. Let $w = v_E$ be the valuation associated to $\mathbb{Z}_{p,E}$ and $\bar{v}_E$ the valuation associated to $\overline{\mathbb{Z}_p}_E$.

Since E is of transcendental type, for each $f \in \mathbb{Q}_p[X]$, we have $v_E(f) = v(f(s_n))$ for all n sufficiently large (see Section 1.1). It follows that, for each $\phi \in \mathbb{Q}_p(X)$ with $\phi = f/g$, for some $f, g \in \mathbb{Q}_p[X]$, we have that $v_E(\phi) = v_E(f) - v_E(g)$ is in Γ_n for all n sufficiently large. Hence $\Gamma_w \subseteq \bigcup_n \Gamma_n$. Conversely, let $n \in \mathbb{N}$ and $f \in \mathbb{Q}_p[X]$ be of degree smaller than $[\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$. Then, each root α_i of $f(X)$ in $\overline{\mathbb{Q}_p}$ has degree smaller than $[\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$ and so, since (s_n, δ_n) is a minimal pair, we have

$$v_p(s_n - \alpha_i) < \delta_n, \tag{2.8}$$

which implies that

$$\bar{v}_E(X - \alpha_i) = \bar{v}_E(X - s_n + s_n - \alpha_i) = v_p(s_n - \alpha_i), \tag{2.9}$$

and so

$$v_E(f(X)) = \sum_i \bar{v}_E(X - \alpha_i) = \sum_i v_p(s_n - \alpha_i) = v_p(f(s_n)), \tag{2.10}$$

which shows that $\Gamma_n \subseteq \Gamma_w$. Note that $\bar{v}_E(X - \alpha_i) = v(s_m - \alpha_i)$ for each $m \geq n$, and so $v_E(f(X)) = v(f(s_m))$ for each $m \geq n$.

Let now $n \in \mathbb{N}$ and $\bar{c} = \overline{f(s_n)} \in k_n$ for some $f(s_n) \in O_n^*$, where $f \in \mathbb{Q}_p[X]$ has degree strictly smaller than $[\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$. In particular, $\bar{c} \neq 0$. As in (2.10), $v_E(f(X)) = v(f(s_m)) = 0$ for each $m \geq n$. Let α_i be a root of $f(X)$ in $\overline{\mathbb{Q}_p}$. Then, by (2.9), $\bar{v}_E(X - \alpha_i) = v_p(s_n - \alpha_i) = v_p(d_i)$ for some $d_i \in \overline{\mathbb{Q}_p}$. Then

$$\bar{v}_E\left(\frac{(X - \alpha_i)/d_i}{(s_n - \alpha_i)/d_i} - 1\right) = \bar{v}_E\left(\frac{X - s_n}{s_n - \alpha_i}\right) = \delta_n - v_p(s_n - \alpha_i) > 0,$$

where the last inequality holds by (2.8). Therefore, $(X - \alpha_i)/d_i$ and $(s_n - \alpha_i)/d_i$ coincide over the residue field of W. In particular,

$$\frac{f(X)}{f(s_n)} = \prod_i \frac{(X - \alpha_i)}{(s_n - \alpha_i)} = \prod_i \frac{(X - \alpha_i)/d_i}{(s_n - \alpha_i)/d_i}, \tag{2.11}$$

and since each factor of the last product has residue $\bar{1}$ in W, it follows that $f(X)$ and $f(s_n)$ coincide over the residue field of W (which contains both $f(X)$ and $f(s_n)$). Since $f \in \mathbb{Z}_{p,E} = W$, this shows that k_n is contained in the residue field k_w of W.

Conversely, let $\phi = f/g \in W \subset \mathbb{Q}_p(X)$ for some $f, g \in \mathbb{Q}_p[X]$. Let α_i and β_j be the roots in $\overline{\mathbb{Q}_p}$ of f and g, respectively. There exists $n \in \mathbb{N}$ such that $[\mathbb{Q}_p(\alpha_i) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$ and $[\mathbb{Q}_p(\beta_j) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_n) : \mathbb{Q}_p]$ for all i and j. Hence, as in (2.9), we have

$$\bar{v}_E(X - \alpha_i) = v_p(s_n - \alpha_i), \quad \bar{v}_E(X - \beta_j) = v(s_n - \beta_j) \quad \text{for all } i, j,$$

which again, as in (2.10), shows that

$$v_E(\phi(X)) = v(\phi(s_n)).$$

Moreover, this last equation holds if we replace s_n by s_m for all $m \geq n$. If $v_E(\phi(X)) = 0$, then, as in (2.11), one can show that $\phi(X)$ and $\phi(s_n)$ coincide over the residue field of W, so that $k_w \subseteq k_n$. □

The following corollary gives a further characterization of the residue field and the value group of a residually algebraic torsion extension W of $\mathbb{Z}_p$: either the residue field of W is an infinite algebraic extension of $\mathbb{F}_p$, or the value group Γ_w is nondiscrete.

Corollary 2.12. *Let W be a residually algebraic torsion extension of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$, and let $e = e(W \mid \mathbb{Z}_p)$ and $f = f(W \mid \mathbb{Z}_p)$ be the ramification index and the residue field degree, respectively. Then $e \cdot f = \infty$.*

Proof. By Theorem 2.5, there exists a stacked sequence $E = \{s_n\}_{n\in\mathbb{N}} \subset \overline{\mathbb{Q}_p}$ such that $W = \mathbb{Z}_{p,E}$. By Proposition 2.7, $\Gamma_w = \bigcup_n \Gamma_n$ and $k_w = \bigcup_n k_n$. Remark 2.2 shows that either the sequence $\{e_n = [\Gamma_n : \mathbb{Z}]\}_{n\in\mathbb{N}}$ or $\{f_n = [k_n : \mathbb{F}_p]\}_{n\in\mathbb{N}}$ is unbounded; therefore, either $e = e(W \mid \mathbb{Z}_p)$ or $f = f(W \mid \mathbb{Z}_p)$ is infinite. □

The following proposition is analogous to [Alexandru et al. 1998, Proposition 2.3]. It shows that the sequence of ramification indexes, residue field degrees and gauges attached to a residually algebraic torsion extension W of $\mathbb{Z}_p$ do not depend on the strongly stacked sequence $E \subset \overline{\mathbb{Q}_p}$ such that $W = \mathbb{Z}_{p,E}$ (Theorem 2.5).

Proposition 2.13. *Let $W \subset \mathbb{Q}_p(X)$ be a residually algebraic torsion extension of $\mathbb{Z}_p$. Let $E = \{s_n\}_{n\in\mathbb{N}}$, $E' = \{t_n\}_{n\in\mathbb{N}} \subset \overline{\mathbb{Q}_p}$ be strongly stacked sequences with gauges $\{\delta_n\}_{n\in\mathbb{N}}$, $\{\delta'_n\}_{n\in\mathbb{N}}$, respectively, such that $W = \mathbb{Z}_{p,E} = \mathbb{Z}_{p,E'}$. Then, for each $n \in \mathbb{N}$, we have*

(i) *$[\mathbb{Q}_p(s_n) : \mathbb{Q}_p] = [\mathbb{Q}_p(t_n) : \mathbb{Q}_p]$ and $\delta_n = \delta'_n$,*

(ii) *$e_{s_n} = e_{t_n}$ and $f_{s_n} = f_{t_n}$.*

Proof. Without loss of generality, we may assume that in $\overline{\mathbb{Q}_p}(X)$ we have $\overline{\mathbb{Z}_p}_E = \overline{\mathbb{Z}_p}_{E'}$; we will let $\overline{W} = \overline{\mathbb{Z}_p}_E = \overline{\mathbb{Z}_p}_{E'}$ and denote by w a valuation associated to $\overline{W}$.

(i) We have $s_0, t_0 \in \mathbb{Q}_p$. There exists $n \in \mathbb{N}$, $n \geq 1$, such that

$$w(X - s_{n-1}) \leq w(X - t_0) < w(X - s_n),$$

otherwise t_0 would be a pseudo-limit of E, which is not possible. In particular,

$$v_p(s_n - t_0) = w(s_n - X + X - t_0) = w(X - t_0) \geq w(X - s_{n-1}) = \delta_{n-1}.$$

If $n > 1$, we have $[\mathbb{Q}_p(t_0) : \mathbb{Q}_p] < [\mathbb{Q}_p(s_{n-1}) : \mathbb{Q}_p]$, so by (iii) of Definition 2.1 we have that

$$v_p(s_n - t_0) = v_p(s_{n-1} - t_0) < v_p(s_n - s_{n-1}) = \delta_{n-1},$$

which is impossible. Hence $n = 1$, so $v_p(s_1 - t_0) = w(X - t_0) \geq w(X - s_0)$. Reversing the roles of s_0 and t_0, we get the other inequality, so $w(X - s_0) = w(X - t_0) = \delta_0 = \delta'_0$.

Let $n \in \mathbb{N}$, and suppose that, for each $m \leq n$, we have $[\mathbb{Q}_p(s_m) : \mathbb{Q}_p] = [\mathbb{Q}_p(t_m) : \mathbb{Q}_p]$ and $\delta_m = \delta'_m$.

Since $[\mathbb{Q}_p(s_n) : \mathbb{Q}_p] = [\mathbb{Q}_p(t_n) : \mathbb{Q}_p] < [\mathbb{Q}_p(t_{n+1}) : \mathbb{Q}_p]$, by (ii) of Definition 2.1 we have

$$v_p(t_{n+1} - s_n) \leq v_p(t_{n+1} - t_n) = \delta'_n = \delta_n.$$

Now,

$$v_p(t_{n+1} - s_n) = v_p(t_{n+1} - t_n + t_n - s_n) \geq \delta_n$$

since $v_p(t_n - s_n) = w(t_n - X + X - s_n) \geq \delta_n = \delta'_n$. This implies that $v_p(t_{n+1} - s_n) = \delta_n$, and so (s_n, t_{n+1}) is distinguished. Moreover, we have

$$v_p(t_{n+1} - s_{n+1}) = w(t_{n+1} - X + X - s_{n+1}) > \delta_n = \delta'_n = v_p(t_{n+1} - s_n).$$

Now, if $[\mathbb{Q}_p(s_{n+1}) : \mathbb{Q}_p] < [\mathbb{Q}_p(t_{n+1}) : \mathbb{Q}_p]$, then, since (s_n, t_{n+1}) is distinguished, we would have $v_p(s_{n+1} - t_{n+1}) \leq v_p(t_{n+1} - s_n)$, which is impossible. Hence $[\mathbb{Q}_p(s_{n+1}) : \mathbb{Q}_p] \geq [\mathbb{Q}_p(t_{n+1}) : \mathbb{Q}_p]$. The other inequality is proved in a symmetrical way, so $[\mathbb{Q}_p(s_{n+1}) : \mathbb{Q}_p] = [\mathbb{Q}_p(t_{n+1}) : \mathbb{Q}_p]$.

Suppose now that $w(X - s_{n+1}) < w(X - t_{n+1})$. Then

$$v_p(s_{n+2} - t_{n+1}) = w(s_{n+2} - X + X - t_{n+1}) > w(X - s_{n+1}) = v_p(s_{n+2} - s_{n+1}),$$

which is not possible since (s_{n+1}, s_{n+2}) is distinguished. Hence $w(X - s_{n+1}) \geq w(X - t_{n+1})$. The other inequality is proved similarly, so $\delta_{n+1} = \delta'_{n+1}$ as claimed.

(ii) For each $n \in \mathbb{N}$, let Γ_n and Γ'_n and k_n and k'_n be the value groups and residue fields, respectively, of $\mathbb{Q}_p(s_n)$ and $\mathbb{Q}_p(t_n)$. Let $e_n = e_{s_n}$, $e'_n = e_{t_n}$, $f_n = f_{s_n}$, $f'_n = f_{t_n}$.

Clearly, $e_0 = e'_0$ and $f_0 = f'_0$ since $s_0, t_0 \in \mathbb{Q}_p$.

Let $n \geq 1$. If $f \in \mathbb{Q}_p[X]$ has degree strictly smaller than $[\mathbb{Q}_p(s_n) : \mathbb{Q}_p] = [\mathbb{Q}_p(t_n) : \mathbb{Q}_p]$, then by (2.10) we have $w(f(X)) = v_p(f(s_n))$ and also $w(f(X)) = v_p(f(t_n))$, so $v_p(f(s_n)) = v_p(f(t_n))$. This proves that $\Gamma_n = \Gamma'_n$, and so $e_n = e'_n$.

Suppose now that $f \in \mathbb{Q}_p[X]$ of degree strictly smaller than $[\mathbb{Q}_p(s_n) : \mathbb{Q}_p] = [\mathbb{Q}_p(t_n) : \mathbb{Q}_p]$ is such that $v_p(f(s_n)) = v_p(f(t_n)) = 0$. In particular, $w(f(X)) = 0$ by (2.10). By (2.11) and the analogous equation where s_n is replaced by t_n, we get that $f(s_n)$ and $f(t_n)$ have the same residue as $f(X)$, so, in particular, $k_n = k'_n$. Therefore, $f_n = f'_n$. □

2.2. *Residually algebraic extensions of $\mathbb{Z}_p$ which are DVRs.* In this section we characterize DVRs of $\mathbb{Q}_p(X)$ extending $\mathbb{Z}_p$ such that the residue field extension is algebraic, necessarily of infinite degree by Corollary 2.12; this fact has already been noted in a different way in [Peruginelli 2017, p. 4217]. We will see in Section 2.3 that there is no such restriction on the residue field degree for DVRs of $\mathbb{Q}(X)$ which are residually algebraic extensions of $\mathbb{Z}_{(p)}$ (see Corollary 2.28).

Given $\alpha \in \mathbb{C}_p$, we denote by $\mathbb{O}_{p,\alpha}$ the unique valuation domain of $\mathbb{Q}_p(\alpha)$ lying over $\mathbb{Z}_p$ (i.e., $\mathbb{O}_{p,\alpha} = \mathbb{O}_p \cap \mathbb{Q}_p(\alpha)$). We also set

$$\mathbb{Z}_{p,\alpha} = \{\phi \in \mathbb{Q}_p(X) \mid \phi(\alpha) \in \mathbb{O}_p\},$$

which is a valuation domain of $\mathbb{Q}_p(X)$ and coincides with the previous definition if $\alpha \in \overline{\mathbb{Q}_p}$.

Proposition 2.14. *Let $\alpha \in \mathbb{C}_p$ be a transcendental element. Then there exists a Cauchy stacked sequence $E \subseteq \overline{\mathbb{Q}_p}$ converging to α. Moreover, the valued fields $(\mathbb{Q}_p(X), \mathbb{Z}_{p,\alpha})$ and $(\mathbb{Q}_p(\alpha), \mathbb{O}_{p,\alpha})$ are isomorphic. In particular, the ramification index $e(\mathbb{Z}_{p,\alpha} \mid \mathbb{Z}_p)$ is equal to e_α, the residue field degree $f(\mathbb{Z}_{p,\alpha} \mid \mathbb{Z}_p)$ is equal to f_α and $e_\alpha \cdot f_\alpha = \infty$.*

Note that the last condition implies that either e_α or f_α is infinite. It can happen that exactly one of these two quantities is finite (see Theorem 2.21).

Proof. The proof of the first claim follows also by [Alexandru et al. 1998, Proposition 2.2], but we give here a different proof based on the previous results.

By Theorem 2.5, there exists a stacked sequence $E \subset \overline{\mathbb{Q}_p}$ such that $\mathbb{Z}_{p,\alpha} = \mathbb{Z}_{p,E}$. Since the valuation domains $\overline{\mathbb{Z}_p}_E, \overline{\mathbb{Z}_p}_\alpha \subset \overline{\mathbb{Q}_p}(X)$ contract down to $\mathbb{Q}_p(X)$ to the same valuation domain, there exists $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$ such that $\sigma(\overline{\mathbb{Z}_p}_\alpha) = \overline{\mathbb{Z}_p}_{\sigma(\alpha)} = \overline{\mathbb{Z}_p}_E$. By [Peruginelli and Spirito 2021, Proposition 5.3], E is then a Cauchy sequence converging to $\sigma(\alpha)$. Since $\mathbb{Z}_{p,\alpha} = \mathbb{Z}_{p,\sigma(\alpha)}$, without loss of generality, we may assume that E converges to α.

Since α is transcendental over $\mathbb{Q}_p$, the evaluation homomorphism $\mathrm{ev}_\alpha : \mathbb{Q}_p(X) \to \mathbb{Q}_p(\alpha)$, $\phi(X) \mapsto \phi(\alpha)$, is an isomorphism. It is easy to see that $\mathrm{ev}_\alpha(\mathbb{Z}_{p,\alpha}) = \mathbb{O}_{p,\alpha}$. Hence $\mathbb{Z}_{p,\alpha}$ and $\mathbb{O}_{p,\alpha}$ have the same ramification indexes and residue field degrees over $\mathbb{Z}_p$.

Finally, the last claim follows by Corollary 2.12. □

Remark 2.15. By Proposition 2.14, we may conclude that, in general,

$$\text{for } \alpha \in \mathbb{C}_p, \text{ we have } e_\alpha \cdot f_\alpha < \infty \text{ if and only if } \alpha \in \overline{\mathbb{Q}_p}.$$

The next lemma may be well known, but lacking a reference we give a short proof.

Lemma 2.16. *Let $p \in \mathbb{Z}$ be a prime, K_1 and K_2 finite extensions of $\mathbb{Q}_p$ and $L = K_1K_2$ the compositum. Let e_1 be the ramification index of K_1 over $\mathbb{Q}_p$ and e the ramification index of L over K_2. Then $e \leq e_1$.*

Proof. If K_1 is a tame extension of $\mathbb{Q}_p$, then the ramification index of L over $\mathbb{Q}_p$ is equal to

$$\mathrm{lcm}\{e(K_1 \,|\, \mathbb{Q}_p), e(K_2 \,|\, \mathbb{Q}_p)\}$$

(see, for example, [Chabert and Halberstadt 2018]), so e divides e_1 and the claim is true.

We give a self-contained proof which works in general. Let L' be the normal closure of L over $\mathbb{Q}_p$ and I the inertia group of the maximal ideal $M_{L'}$ of $O_{L'}$ over $\mathbb{Z}_p$. Let G_i be the Galois group $\mathrm{Gal}(L' \,|\, K_i)$ for $i = 1, 2$ and G the Galois group $\mathrm{Gal}(L' \,|\, L)$. Since $L = K_1K_2$, we have $G = G_1 \cap G_2$. The inertia group of $M_{L'}$ over M_{K_1} is equal to $I \cap G_1$, and the inertia group of $M_{L'}$ over M_L is equal to $I \cap G$. We have

$$e = \frac{e(L' \,|\, K_2)}{e(L' \,|\, L)} = \frac{\#(I \cap G_2)}{\#(I \cap G)}, \quad e_1 = \frac{e(L' \,|\, \mathbb{Q}_p)}{e(L' \,|\, K_1)} = \frac{\#I}{\#(I \cap G_1)}.$$

Note that $I \cap G = (I \cap G_1) \cap (I \cap G_2)$. Therefore, the claim follows by the following general fact for finite groups: given a finite group G with two subgroups H_1 and H_2, we have

$$\frac{\#H_2}{\#(H_1 \cap H_2)} = [H_2 : H_1 \cap H_2] \leq \frac{\#G}{\#H_1} = [G : H_1],$$

which follows immediately since the map $h_2H_1 \cap H_2 \mapsto h_2H_1$ from the set $\{h_2(H_1 \cap H_2) \mid h_2 \in H_2\}$ of left cosets of $H_1 \cap H_2$ in H_2 to the set $\{gH_1 \mid g \in G\}$ of left cosets of H_1 in G is injective. □

The following result is analogous to [Peruginelli 2017, Theorem 2.5].

Theorem 2.17. *Let W be a DVR of $\mathbb{Q}_p(X)$ which is a residually algebraic extension of $\mathbb{Z}_p$. Then there exists a transcendental element $\alpha \in \mathbb{C}_p$ such that $W = \mathbb{Z}_{p,\alpha}$.*

Proof. Note that, by Corollary 2.12, the residue field of W is an infinite algebraic extension of $\mathbb{F}_p$.

By Theorem 2.5, there exists a stacked sequence $E=\{s_n\}_{n\in\mathbb{N}}\subset\overline{\mathbb{Q}_p}$ such that $W=\mathbb{Z}_{p,E}$. By assumption, the ramification index $e(W\,|\,\mathbb{Z}_p)=e$ is finite. By Remark 2.2 and Proposition 2.7, there exists $n_0\in\mathbb{N}$ such that $\Gamma_w=\Gamma_n=\Gamma_{n_0}$ for each $n\geq n_0$. Equivalently, $e_n=e_{n_0}=e$ for each $n\geq n_0$. Let $n\geq n_0$. Note that $\delta_n=v_p(s_{n+1}-s_n)\in\Gamma_{O_{K_n}}$, where $K_n=\mathbb{Q}_p(s_n,s_{n+1})$. Note that the ramification index of $\mathbb{Q}_p(s_i)$ over $\mathbb{Q}_p$ is equal to e for $i=n,n+1$. By Lemma 2.16, the ramification index of K_n over $\mathbb{Q}_p$ is bounded by e^2. If $d=\prod_{i=1}^{e^2}i$, then $d\delta_n\in\mathbb{Z}$ for each $n\geq n_0$. This shows that the gauge $\{\delta_n\}_{n\in\mathbb{N}}$ of E has bounded denominator, so $\delta_n\nearrow\infty$, and thus E is Cauchy and converges to a (unique) element α of $\mathbb{C}_p\setminus\overline{\mathbb{Q}_p}$ since E is of transcendental type by Lemma 2.3. In particular, $W=\mathbb{Z}_{p,\alpha}$. □

Remark 2.18. We say an element $\alpha\in\mathbb{C}_p$ has *bounded ramification* if the extension $\mathbb{Q}_p(\alpha)\supseteq\mathbb{Q}_p$ has finite ramification. We denote by $\mathbb{C}_p^{\mathrm{br}}$ the set of all elements of $\mathbb{C}_p$ of bounded ramification; clearly, $\overline{\mathbb{Q}_p}\subset\mathbb{C}_p^{\mathrm{br}}$. A transcendental element $\alpha\in\mathbb{C}_p$ has bounded ramification if and only if the set of ramification indexes $\{e_n\}_{n\in\mathbb{N}}$ attached to a stacked sequence $E\subset\overline{\mathbb{Q}_p}$ converging to α is bounded; in fact, by Theorem 2.17, the integer e such that $e=e_n$ for all n sufficiently large is equal to $e(\mathbb{Q}_p(\alpha)\,|\,\mathbb{Q}_p)$.

We remark that not all the transcendental elements $\alpha\in\mathbb{C}_p$ have bounded ramification. For example, according to [Iovită and Zaharescu 1995], there exist *generic* transcendental elements $t\in\mathbb{C}_p$ for $\mathbb{C}_p$; that is, the completion of $\mathbb{Q}_p(t)$ is equal to $\mathbb{C}_p$. In particular, the value group of the unique valuation of $\mathbb{O}_{p,t}$ is equal to $\mathbb{Q}$, so the corresponding ramification index is ∞. Hence, by Proposition 2.7, $\mathbb{Z}_{p,t}$ has value group equal to $\mathbb{Q}$ and therefore the set of ramification indexes $\{e_n\}_{n\in\mathbb{N}}$ is unbounded.

We show in Theorem 2.21 that given any algebraic extension k of $\mathbb{F}_p$ and group Γ such that $\mathbb{Z}\subseteq\Gamma\subseteq\mathbb{Q}$, there exists a transcendental element $\alpha\in\mathbb{C}_p$ such that $\mathbb{Z}_{p,\alpha}$ has residue field k and value group Γ, provided that either $[k:\mathbb{F}_p]$ is infinite or Γ is not discrete (this condition being necessary by Corollary 2.12).

Lemma 2.19. *Let l be an infinite algebraic extension of $\mathbb{Q}_p$ such that $e(l\,|\,\mathbb{Q}_p)$ is finite. Then l is contained in the maximal unramified extension K^{unr} of a finite extension K of $\mathbb{Q}_p$.*

Proof. For each $n\in\mathbb{N}$, let $\mathbb{Q}_p^{(n)}$ be the compositum of all the extensions of $\mathbb{Q}_p$ of degree bounded by n. Clearly, $\overline{\mathbb{Q}_p}=\bigcup_{n\in\mathbb{N}}\mathbb{Q}_p^{(n)}$ and $\mathbb{Q}_p^{(n)}\subset\mathbb{Q}_p^{(n+1)}$ for each $n\in\mathbb{N}$. Since $\mathbb{Q}_p$ has only finitely many extensions of bounded degree, $\mathbb{Q}_p^{(n)}=\mathbb{Q}_p(t_n)$ for some $t_n\in\overline{\mathbb{Q}_p}$. Now, for each $n\in\mathbb{N}$, we let $\mathbb{Q}_p(t_n)\cap l=\mathbb{Q}_p(s_n)$ for some $s_n\in l$. Clearly, $l=\bigcup_{n\in\mathbb{N}}\mathbb{Q}_p(s_n)$ and $\mathbb{Q}_p(s_n)\subset\mathbb{Q}_p(s_{n+1})$ for each $n\in\mathbb{N}$. Since $\Gamma_{s_n}\subseteq\Gamma_{s_{n+1}}\subseteq\Gamma_l$ for each $n\in\mathbb{N}$ and Γ_l is discrete by assumption, there exists $n_0\in\mathbb{N}$ such that $\Gamma_{s_n}=\Gamma_{s_{n_0}}$ for each $n\geq n_0$. Therefore, if $K=\mathbb{Q}_p(s_{n_0})$, then $s_n\in K^{\mathrm{unr}}$ for each $n\geq n_0$, so that $l\subseteq K^{\mathrm{unr}}$. □

The next proposition shows that a transcendental element t of $\mathbb{C}_p$ with bounded ramification arise as the limit of sequences contained in the maximal unramified extension K^{unr} of a finite extension K of $\mathbb{Q}_p$. We don't know whether there exists a stacked sequence in K^{unr} which converges to t.

Proposition 2.20. *Let $t\in\mathbb{C}_p^{\mathrm{br}}$. Then t is the limit of a sequence contained in the maximal unramified extension of a finite extension of $\mathbb{Q}_p$.*

Proof. By [Iovită and Zaharescu 1995, Theorem 1], the completion of $\widehat{\mathbb{Q}_p(t)} \cap \overline{\mathbb{Q}_p}$ is equal to $\widehat{\mathbb{Q}_p(t)}$. In particular, there exists a Cauchy sequence $E = \{s_n\}_{n\in\mathbb{N}} \subset \widehat{\mathbb{Q}_p(t)} \cap \overline{\mathbb{Q}_p}$ converging to t. Now, since $\mathbb{Q}_p(t) \subset \widehat{\mathbb{Q}_p(t)}$ and $\widehat{\mathbb{Q}_p(t)} \cap \overline{\mathbb{Q}_p} \subset \widehat{\mathbb{Q}_p(t)}$ are immediate extensions, it follows that $\widehat{\mathbb{Q}_p(t)} \cap \overline{\mathbb{Q}_p}$ has value group Γ_t and residue field k_t. By Lemma 2.19, $\widehat{\mathbb{Q}_p(t)} \cap \overline{\mathbb{Q}_p}$ is contained in the maximal unramified extension of a finite extension of $\mathbb{Q}_p$. The statement follows. □

The following result is not new; see for example [Lampert 1986, Lemma 2]. The present proof is different because it employs the notion of stacked sequence.

Theorem 2.21. *Let k be an algebraic extension of $\mathbb{F}_p$ and Γ a totally ordered group with $\mathbb{Z} \subseteq \Gamma \subseteq \mathbb{Q}$ such that either $[k : \mathbb{F}_p]$ or $[\Gamma : \mathbb{Z}]$ is infinite (the last condition is equivalent to Γ being not discrete). Then there exists a transcendental element $\alpha \in \mathbb{C}_p$ such that $k_\alpha = k$ and $\Gamma_\alpha = \Gamma$. In particular, $\mathbb{Z}_{p,\alpha}$ has residue field k and value group Γ.*

Note that, by Corollary 2.12, the last claim shows that $[k : \mathbb{F}_p] \cdot [\Gamma : \mathbb{Z}] = \infty$ is necessary.

Proof. Since $\overline{\mathbb{F}_p}$ is countable, we may suppose that $k = \bigcup_{n\in\mathbb{N}} k_n$, where k_n is a finite extension of $\mathbb{F}_p$, $k_n \subseteq k_{n+1}$ and $k_0 = \mathbb{F}_p$. Similarly, $\Gamma = \bigcup_{n\in\mathbb{N}} \Gamma_n$, where Γ_n is a discrete group, $\Gamma_n \subseteq \Gamma_{n+1}$ and $\Gamma_0 = \mathbb{Z}$. Let $f = [k : \mathbb{F}_p]$ and $e = [\Gamma : \mathbb{Z}]$; then, either e or f is infinite. Without loss of generality, we may assume that, for each n, $[k_{n+1} : k_n][\Gamma_{n+1} : \Gamma_n] > 1$.

For each $n \in \mathbb{N}$, there exists a local field $K_n = \mathbb{Q}_p(s_n)$ with residue field k_n and value group Γ_n. By induction, we may also assume that $K_n \subset K_{n+1}$. Let $\{\lambda_n\}_{n\in\mathbb{N}} \subset \mathbb{Q}$ be a strictly increasing sequence in $\mathbb{Q}$ which is unbounded and $\lambda_0 < \delta_0 = v(s_1 - s_0)$.

We define now a sequence $E = \{t_n\}_{n\in\mathbb{N}} \subset \overline{\mathbb{Q}_p}$ such that, for each $n \in \mathbb{N}$, $n \geq 1$, we have

(i) $\mathbb{Q}_p(t_n) = \mathbb{Q}_p(s_n)$,

(ii) $(t_{n-1}, \delta_{n-1} = v_p(t_n - t_{n-1}))$ is a minimal pair,

(iii) $\delta_{n-1} > \lambda_{n-1}$.

In particular, E is a stacked sequence by conditions (i) and (ii) and Cauchy by condition (iii) and the assumption on $\{\lambda_n\}_{n\in\mathbb{N}}$.

We set $t_0 = s_0 \in \mathbb{Q}_p$, $t_1 = s_1 \notin \mathbb{Q}_p$ and $\delta_0 = v_p(t_1 - t_0)$. Note that (t_0, δ_0) is a minimal pair. We proceed by induction on n. We assume that, for all $m < n$, we have chosen $t_m \in \overline{\mathbb{Q}_p}$ such that conditions (i), (ii) and (iii) above are satisfied.

We now show how to choose t_n. We choose $a_n \in \mathbb{Q}_p$, $a_n \neq 0$, such that

$$v_p(a_n) > \max\{\omega(t_{n-1}) - v_p(s_n), \lambda_{n-1} - v_p(s_n)\}.$$

We then set

$$t_n = a_n s_n + t_{n-1}.$$

Note that $\mathbb{Q}_p(t_n) \subseteq \mathbb{Q}_p(s_n)$ since by induction $\mathbb{Q}_p(t_{n-1}) = \mathbb{Q}_p(s_{n-1})$ and the last field is contained in $\mathbb{Q}_p(s_n)$. Now, since $\delta_{n-1} = v_p(t_n - t_{n-1}) > \omega(t_{n-1})$, it follows by Krasner's lemma that $\mathbb{Q}_p(t_{n-1}) \subseteq \mathbb{Q}_p(t_n)$. This containment and the fact that $s_n = (t_n - t_{n-1})/a_n$ show that s_n is in $\mathbb{Q}_p(t_n)$, so that $\mathbb{Q}_p(t_n) = \mathbb{Q}_p(s_n)$.

Moreover, note also that $\delta_{n-1} > \lambda_{n-1}$. Hence $E = \{t_n\}_{n\in\mathbb{N}}$ is a stacked sequence which is Cauchy, so E converges to a transcendental element α of $\mathbb{C}_p$. By Proposition 2.7, $\mathbb{Z}_{p,E} = \mathbb{Z}_{p,\alpha}$ has residue field k and value group Γ, as desired. By Proposition 2.14, $\mathbb{Z}_{p,\alpha}$ is isomorphic to $\mathbb{O}_{p,\alpha}$, so it follows that $\Gamma_\alpha = \Gamma$ and $k_\alpha = k$. □

Remark 2.22. We remark that, without condition (iii) above in the proof of Theorem 2.21, in general we may only conclude that there exists a stacked sequence $E \subset \overline{\mathbb{Q}_p}$ (which may not be Cauchy) such that the valuation domain $\mathbb{Z}_{p,E}$ has residue field k and value group Γ. If instead Γ is discrete by assumption, condition (iii) is not necessary: in fact, there exists $n_0 \in \mathbb{N}$ such that $\Gamma_n = \Gamma_{n_0} = \Gamma$ for all $n \geq n_0$; that is, $K_n = \mathbb{Q}_p(s_n)$ is an unramified extension of K_{n_0} for all $n > n_0$. Hence $E \subset \bigcup_{n\in\mathbb{N}} K_n$ is Cauchy, and so $\mathbb{Z}_{p,E} = \mathbb{Z}_{p,\alpha}$, where $\alpha \in \mathbb{C}_p^{\text{br}}$ is the transcendental limit of E.

We close this section showing that the statement of [Ioviţă and Zaharescu 1995, Proposition 1] is wrong, namely, in general the completion of $\mathbb{Q}_p(X)$ with respect to a residually algebraic torsion extension W of $\mathbb{Z}_p$ may not be a subfield of $\mathbb{C}_p$. The mistake is due to the fact that if $W = \mathbb{Z}_{p,E}$ for some pseudo-convergent sequence $E \subset \overline{\mathbb{Q}_p}$ of transcendental type, then X is a pseudo-limit of E with respect to w and may not be a limit (that is, E may not be Cauchy).

Proposition 2.23. *Let W be a residually algebraic torsion extension of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$. Then the completion $\widehat{\mathbb{Q}_p(X)}$ with respect to W is (isomorphic to) a subfield of $\mathbb{C}_p$ if and only if there exists a transcendental element α in $\mathbb{C}_p$ such that $W = \mathbb{Z}_{p,\alpha}$.*

Proof. By Theorem 2.5, there exists a pseudo-convergent sequence $E = \{s_n\}_{n\in\mathbb{N}} \subset \overline{\mathbb{Q}_p}$ of transcendental type such that $W = \mathbb{Z}_{p,E}$.

Suppose that $\widehat{\mathbb{Q}_p(X)} \subseteq \mathbb{C}_p$. In particular, $X \in \mathbb{C}_p$, so there exists a Cauchy sequence $F = \{t_n\}_{n\in\mathbb{N}} \subset \overline{\mathbb{Q}_p}$ which tends to X. Since $\mathbb{Q}_p(X) \subset \overline{\mathbb{Q}_p}(X)$ is an algebraic extension and $\mathbb{C}_p$ is algebraically closed, then also the completion of $\overline{\mathbb{Q}_p}(X)$ with respect to $\overline{W} = \overline{\mathbb{Z}_p}_E$ is contained in $\mathbb{C}_p$. Without loss of generality, we may suppose that the restriction of v_p to $\overline{\mathbb{Q}_p}(X)$ is equal to $\overline{w}$. In particular, $\overline{w}(X - t_n) = v_p(X - t_n) \nearrow \infty$. Since E is of transcendental type, for each n, there exists m_0 such that $\overline{w}(X - t_n) < \overline{w}(X - s_m)$ for each $m \geq m_0$. This shows that the gauge of E tends to infinity, and thus E is Cauchy; in particular, E converges to a transcendental element $\alpha \in \mathbb{C}_p$. Therefore, $W = \mathbb{Z}_{p,\alpha}$.

Conversely, let $W = \mathbb{Z}_{p,\alpha}$ for some transcendental element $\alpha \in \mathbb{C}_p$. Then, by Proposition 2.14, the completion $\widehat{\mathbb{Q}_p(X)}$ with respect to $\mathbb{Z}_{p,\alpha}$ is isomorphic to the completion of $\mathbb{Q}_p(\alpha)$ and therefore can be identified to a subfield of $\mathbb{C}_p$. □

In particular, if $W = \mathbb{Z}_{p,E}$ for some stacked non-Cauchy sequence $E \subset \overline{\mathbb{Q}_p}$, then $\widehat{\mathbb{Q}_p(X)} \not\subseteq \mathbb{C}_p$.

2.3. *Residually algebraic torsion extensions of* $\mathbb{Z}_{(p)}$. We now characterize residually algebraic torsion extensions of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$. We remark that such a valuation domain may have an extension to $\mathbb{Q}_p(X)$ which is a residually algebraic extension of $\mathbb{Z}_p$ but is not torsion. For example, let $\alpha \in \overline{\mathbb{Q}_p}$ be transcendental over $\mathbb{Q}$. Then $\mathbb{Z}_{(p),\alpha}$ is torsion but $\mathbb{Z}_{p,\alpha}$ is not (the one dimensional valuation overring of $\mathbb{Z}_{p,\alpha}$ is $\mathbb{Q}_p[X]_{(p_\alpha(X))}$, where $p_\alpha(X)$ is the minimal polynomial of α over $\mathbb{Q}_p$).

Given $\alpha \in \mathbb{C}_p$, we consider the following valuation domain of $\mathbb{Q}(X)$:

$$\mathbb{Z}_{(p),\alpha} = \{\phi \in \mathbb{Q}(X) \mid \phi(\alpha) \in \mathbb{O}_p\},$$

which is just the contraction to $\mathbb{Q}(X)$ of $\mathbb{Z}_{p,\alpha}$ considered in Section 2.1. Similarly, if $E = \{s_n\}_{n\in\mathbb{N}} \subset \overline{\mathbb{Q}_p}$ is a pseudo-convergent sequence of transcendental type, then we set

$$\mathbb{Z}_{(p),E} = \{\phi \in \mathbb{Q}(X) \mid \phi(s_n) \in \overline{\mathbb{Z}_p} \text{ for all sufficiently large } n \in \mathbb{N}\},$$

which is equal to $\mathbb{Z}_{p,E} \cap \mathbb{Q}(X)$.

The next proposition is analogous to Proposition 2.4 and characterizes residually algebraic torsion extensions of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$ in terms of pseudo-convergent sequences of $\overline{\mathbb{Q}_p}$ which are of transcendental type over $\mathbb{Q}$; clearly, every pseudo-convergent sequence of transcendental type of $\overline{\mathbb{Q}_p}$ belongs to this class. As a particular case, we find again part of the result of [Peruginelli 2017, Theorem 2.5].

Proposition 2.24. *Let $p \in \mathbb{P}$, and let W be a residually algebraic torsion extension of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$. Then there exists a pseudo-convergent sequence $E \subset \overline{\mathbb{Q}_p}$ of transcendental type over $\mathbb{Q}$ such that $W = \mathbb{Z}_{(p),E}$. More precisely, let e and f be the ramification index and residue field degree of W over $\mathbb{Z}_{(p)}$, respectively. Let $\widehat{\mathbb{Q}(X)}$ be the completion of $\mathbb{Q}(X)$ with respect to the W-adic topology. Then the following conditions are equivalent*:

(1) *$\widehat{\mathbb{Q}(X)}$ is a finite extension of $\mathbb{Q}_p$.*

(2) *X is algebraic over $\mathbb{Q}_p$.*

(3) *$W = \mathbb{Z}_{(p),\alpha}$ for some $\alpha \in \overline{\mathbb{Q}_p}$ transcendental over $\mathbb{Q}$.*

(4) *$ef < \infty$.*

If any one of these conditions holds, then the sequence E above is Cauchy and converges to α (and E is therefore of algebraic type over $\mathbb{Q}_p$). Moreover, we have $\Gamma_w = \Gamma_\alpha$ and $k_w = k_\alpha$.

If $ef = \infty$, then $E \subset \overline{\mathbb{Q}_p}$ is of transcendental type over $\mathbb{Q}_p$ and $\mathbb{Z}_{(p),E} \subset \mathbb{Z}_{p,E}$ is an immediate extension.

Proof. Note that, since W is a torsion extension of $\mathbb{Z}_{(p)}$, the p-adic completion $\mathbb{Q}_p$ of $\mathbb{Q}$ is contained in $\widehat{\mathbb{Q}(X)}$; see for example the arguments given in the proof of [Alexandru et al. 1988, Corollary 2.6].

If $\widehat{\mathbb{Q}(X)}$ is a finite extension of $\mathbb{Q}_p$, then clearly X is algebraic over $\mathbb{Q}_p$, so (1) implies (2). If X is algebraic over $\mathbb{Q}_p$, we may identify X with some $\alpha \in \overline{\mathbb{Q}_p}$; $\mathbb{Q}_p(\alpha)$ is a finite extension of $\mathbb{Q}_p$ and is hence complete. So, $\widehat{\mathbb{Q}(X)} = \mathbb{Q}_p(\alpha)$. As in the proof of [Peruginelli 2017, Theorem 2.5] it follows easily that $W = \mathbb{Z}_{(p),\alpha}$. Therefore, (2) implies (3).

If $W = \mathbb{Z}_{(p),\alpha}$ for some $\alpha \in \overline{\mathbb{Q}_p}$ transcendental over $\mathbb{Q}$, then, by [Peruginelli 2017, Proposition 2.2], $ef < \infty$, so (3) implies (4). Finally, (4) implies (1) by [Peruginelli 2017, Lemma 2.4] because $e(\widehat{W} \mid \mathbb{Z}_p) = e$ and $f(\widehat{W} \mid \mathbb{Z}_p) = f$.

Note that if $E \subset \overline{\mathbb{Q}_p}$ is a pseudo-convergent sequence such that $\mathbb{Z}_{(p),E} = \mathbb{Z}_{(p),\alpha}$, then by Lemma 2.27 below we have $\mathbb{Z}_{p,E} = \mathbb{Z}_{p,\alpha}$, so by [Peruginelli and Spirito 2021, Proposition 5.3] we have that E is Cauchy and converges to α.

The claims about the value group and residue field of $\mathbb{Z}_{(p),\alpha}$ follow by [Peruginelli 2017, Proposition 2.2].

If $ef = \infty$, then X is transcendental over $\mathbb{Q}_p$ by the previous part of the proof; in particular, the field of rational functions $\mathbb{Q}_p(X)$ is contained in the completion $\widehat{\mathbb{Q}(X)}$. If $\widetilde{W} = \widehat{W} \cap \mathbb{Q}_p(X)$, then $\widetilde{W}$ is a residually algebraic torsion extension of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$, so by Theorem 2.5 there exists a stacked sequence $E \subset \overline{\mathbb{Q}_p}$ such that $\widetilde{W} = \mathbb{Z}_{p,E}$ (by Lemma 2.3, E is a pseudo-convergent sequence of transcendental type, necessarily unbounded). Restricting down to $\mathbb{Q}(X)$, we get $W = \mathbb{Z}_{(p),E}$. Finally, since $W \subset \widehat{W}$ is an immediate extension, it follows that $\mathbb{Z}_{(p),E} \subset \mathbb{Z}_{p,E}$ is an immediate extension, too. Hence the value group and residue field of $\mathbb{Z}_{(p),E}$ are the same as those of $\mathbb{Z}_{p,E}$, respectively (see Proposition 2.7). □

The following statement is the analog of Proposition 2.23 for residually algebraic torsion extensions of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$.

Corollary 2.25. *Let W be a residually algebraic torsion extension of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$. Then the completion $\widehat{\mathbb{Q}(X)}$ with respect to W is (isomorphic to) a subfield of $\mathbb{C}_p$ if and only if there exists $\alpha \in \mathbb{C}_p$, transcendental over $\mathbb{Q}$, such that $W = \mathbb{Z}_{(p),\alpha}$.*

Proof. According to Proposition 2.24, when passing to the completion, either X is algebraic over $\mathbb{Q}_p$ or X is transcendental over $\mathbb{Q}_p$, and consequently either $\widehat{\mathbb{Q}(X)} \subset \overline{\mathbb{Q}_p} \subset \mathbb{C}_p$ or $\mathbb{Q}_p(X) \subset \widehat{\mathbb{Q}(X)}$, respectively. In the first case, $W = \mathbb{Z}_{(p),\alpha}$ for some $\alpha \in \overline{\mathbb{Q}_p} \subset \mathbb{C}_p$ transcendental over $\mathbb{Q}$. In the second case, $\widehat{\mathbb{Q}_p(X)} = \widehat{\mathbb{Q}(X)}$, where the completion of $\mathbb{Q}_p(X)$ is considered with respect to the valuation domain $\widetilde{W} = \widehat{W} \cap \mathbb{Q}_p(X)$. In particular, by Proposition 2.23, in this case we get that $\widehat{\mathbb{Q}(X)} \subseteq \mathbb{C}_p$ if and only if there exists a transcendental element $\alpha \in \mathbb{C}_p$ such that $W = \mathbb{Z}_{(p),\alpha}$. □

In particular, if $W = \mathbb{Z}_{(p),E}$ for some stacked non-Cauchy sequence $E \subset \overline{\mathbb{Q}_p}$, then $\widehat{\mathbb{Q}(X)}$ is not contained in $\mathbb{C}_p$.

The following result is the analog of Theorem 2.21 for building residually algebraic torsion extensions W of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$ with prescribed residue field k and value group Γ. Note that, contrary to that theorem, we are no longer assuming that $[k : \mathbb{F}_p] \cdot [\Gamma : \mathbb{Z}] = \infty$.

Theorem 2.26. *Let k be an algebraic extension of $\mathbb{F}_p$ and Γ a totally ordered group such that $\mathbb{Z} \subseteq \Gamma \subseteq \mathbb{Q}$. Then there exists $\alpha \in \mathbb{C}_p$, transcendental over $\mathbb{Q}$, such that $\mathbb{Z}_{(p),\alpha}$ has residue field k and value group Γ.*

Proof. Let $e = [\Gamma : \mathbb{Z}]$ and $f = [k : \mathbb{F}_p]$. If $ef < \infty$, then it is well known that there exists $\alpha \in \overline{\mathbb{Q}_p}$ transcendental over $\mathbb{Q}$ such that $\mathbb{O}_{p,\alpha}$ has residue field k and value group Γ. Hence, by [Peruginelli 2017, Proposition 2.2], $\mathbb{Z}_{(p),\alpha}$ is the desired extension of $\mathbb{Z}_{(p)}$.

If $ef = \infty$, then, by Theorem 2.21, there exists a transcendental element $\alpha \in \mathbb{C}_p$ such that $\mathbb{Z}_{p,\alpha}$ has residue field k and value group Γ. Clearly, $\mathbb{Z}_{p,\alpha} \cap \mathbb{Q}(X) = \mathbb{Z}_{(p),\alpha}$ is a residually algebraic torsion extension of $\mathbb{Z}_{(p)}$ to $\mathbb{Q}(X)$. Moreover, by Proposition 2.14, $\mathbb{Z}_{p,\alpha} = \mathbb{Z}_{p,E}$ for some stacked Cauchy sequence $E \subset \overline{\mathbb{Q}_p}$ which converges to α. In particular, $\mathbb{Z}_{(p),\alpha} = \mathbb{Z}_{(p),E}$. By the last part of Proposition 2.24, $\mathbb{Z}_{(p),E} \subset \mathbb{Z}_{p,E}$ is an immediate extension, so $\mathbb{Z}_{(p),\alpha}$ has residue field k and value group Γ. □

Now we are able to describe the DVRs of $\mathbb{Q}(X)$ which are residually algebraic extensions of $\mathbb{Z}_{(p)}$ for some $p \in \mathbb{P}$. We recall that every $\sigma \in G_{\mathbb{Q}_p} = \mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$ extends uniquely to a continuous automorphism of $\mathbb{C}_p$; see [Alexandru et al. 1998, §3]. Given $\alpha, \beta \in \mathbb{C}_p$, we say that α and β are conjugate (over $\mathbb{Q}_p$) if there exists $\sigma \in G_{\mathbb{Q}_p} = \mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$ such that $\sigma(\alpha) = \beta$; the orbit of an element $\alpha \in \mathbb{C}_p$ is finite if and only if $\alpha \in \overline{\mathbb{Q}_p}$; see [Alexandru et al. 1998, Remark 3.2].

We prove first the following lemma.

Lemma 2.27. *Let $p \in \mathbb{P}$ and W be a valuation domain of $\mathbb{Q}_p(X)$ such that $W \cap \mathbb{Q}(X) = \mathbb{Z}_{(p),\alpha}$ for some $\alpha \in \mathbb{C}_p$. Then $W = \mathbb{Z}_{p,\alpha}$.*

Proof. Let $n \geq 0$ be an integer such that $p^n \cdot \alpha = \alpha_0 \in \mathbb{O}_p$. The field isomorphism $X \mapsto X/p^n$ maps $\mathbb{Z}_{(p),\alpha}$ to $\mathbb{Z}_{(p),\alpha_0}$ and $\mathbb{Z}_{p,\alpha}$ to $\mathbb{Z}_{p,\alpha_0}$, respectively. Hence, in order to prove the statement, without loss of generality, we may assume that $\alpha \in \mathbb{O}_p$.

Let w be a valuation associated to W. We note first that, since $X \in \mathbb{Z}_{(p),\alpha}$, it follows that $w(X) \geq 0$. Let $f \in W \cap \mathbb{Q}_p[X]$, say $f(X) = \sum_{i=0}^{d} \alpha_i X^i$. Then, for $g(X) = \sum_{i=0}^{d} a_i X^i \in \mathbb{Q}[X]$, we have

$$w(f-g) \geq \min_{0 \leq i \leq d} \{v_p(\alpha_i - a_i) + i\,w(X)\}.$$

Therefore, if we choose $a_i \in \mathbb{Q}$ sufficiently v_p-adically close to α_i for each $i = 0, \ldots, d$, we have $w(f-g) \geq 0$. In particular, $g \in W \cap \mathbb{Q}(X) = \mathbb{Z}_{(p),\alpha}$. The polynomial $h = f - g$ is in $\mathbb{Z}_p[X]$; therefore $f(\alpha) = h(\alpha) + g(\alpha) \in \mathbb{O}_p$, so that $f \in \mathbb{Z}_{p,\alpha}$. Therefore $W \cap \mathbb{Q}_p[X] \subseteq \mathbb{Z}_{p,\alpha} \cap \mathbb{Q}_p[X]$. Similarly, one can easily show that the other containment holds, so $W \cap \mathbb{Q}_p[X] = \mathbb{Z}_{p,\alpha} \cap \mathbb{Q}_p[X]$. In the same way, $M_W \cap \mathbb{Q}_p[X] = M_{p,\alpha} \cap \mathbb{Q}_p[X]$, where M_W and $M_{p,\alpha}$ are the maximal ideals of W and $\mathbb{Z}_{p,\alpha}$, respectively.

Let now $\psi \in \mathbb{Z}_{p,\alpha}$; since $\mathbb{Z}_p[X] \subset \mathbb{Q}_p[X] \cap \mathbb{Z}_{p,\alpha}$, we may suppose that $\psi = f/g$, where $f, g \in \mathbb{Q}_p[X] \cap \mathbb{Z}_{p,\alpha}$. Clearly, $g(\alpha) \neq 0$; then, there exists $n \in \mathbb{N}$, $n \geq 1$, and $c \in \mathbb{Q}_p$, $c \neq 0$, such that $v_p(c) + v_p(g(\alpha)^n) = 0$. We consider then the rational function $\psi^n = cf^n/cg^n = f_1/g_1$, which still is in $\mathbb{Z}_{p,\alpha}$. Note that $f_1 \in \mathbb{Z}_{p,\alpha} \cap \mathbb{Q}_p[X] = W \cap \mathbb{Q}_p[X]$ and $g_1 \in \mathbb{Z}_{p,\alpha}^* \cap \mathbb{Q}_p[X] = W^* \cap \mathbb{Q}_p[X]$ since $v_{p,\alpha}(f_1) \geq v_{p,\alpha}(g_1) = 0$ (the * denotes the set of units of the valuation domains). In particular,

$$w(f_1) \geq 0 = w(g_1),$$

which proves that $\psi^n \in W$. Since W is integrally closed, it follows that $\psi \in W$. Hence $\mathbb{Z}_{p,\alpha} \subseteq W$. The equality follows because both rings are extensions of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$, and in the case that α is algebraic over $\mathbb{Q}_p$, the one-dimensional valuation overring of $\mathbb{Z}_{p,\alpha}$ is nonunitary (i.e., $\mathbb{Q}_p[X]_{(q)}$, where $q \in \mathbb{Q}_p[X]$ is the minimal polynomial of α). □

Corollary 2.28. *Let W be a DVR of $\mathbb{Q}(X)$ which is a residually algebraic extension of $\mathbb{Z}_{(p)}$ for some $p \in \mathbb{P}$. Then there exists $\alpha \in \mathbb{C}_p^{\mathrm{br}}$, transcendental over $\mathbb{Q}$, such that $W = \mathbb{Z}_{(p),\alpha}$. The element α belongs to $\overline{\mathbb{Q}_p}$ if and only if the residue field extension $\mathbb{Z}/p\mathbb{Z} \subseteq W/M$ is finite.*

Moreover, for $\alpha, \beta \in \mathbb{C}_p$, we have $\mathbb{Z}_{(p),\alpha} = \mathbb{Z}_{(p),\beta}$ if and only if there exists $\sigma \in G_{\mathbb{Q}_p}$ such that $\sigma(\alpha) = \beta$.

Proof. Let $f = [W/M : \mathbb{Z}/p\mathbb{Z}]$. If $f < \infty$, then the claim follows by [Peruginelli 2017, Theorem 2.5] and corresponds to the first case of Proposition 2.24: $W = \mathbb{Z}_{(p),\alpha}$ for some $\alpha \in \overline{\mathbb{Q}_p}$ which is transcendental over $\mathbb{Q}$. If $f=\infty$, then we are in the last case of Proposition 2.24, so $W=\mathbb{Z}_{(p),E}$ for some pseudo-convergent sequence in $\overline{\mathbb{Q}_p}$ of transcendental type. As in the proof of Proposition 2.24, we denote by $\widehat{W}$ the completion of W; since the ramification index $e(W\,|\,\mathbb{Z}_{(p)})$ is finite, $\widetilde{W}=\widehat{W}\cap\mathbb{Q}_p(X)$ is a residually algebraic torsion extension of $\mathbb{Z}_p$ to $\mathbb{Q}_p(X)$ which is a DVR, so by Theorem 2.17, $\widetilde{W}=\mathbb{Z}_{p,\alpha}$ for some $\alpha\in\mathbb{C}_p^{\mathrm{br}}\setminus\overline{\mathbb{Q}_p}$. Hence $W=\widetilde{W}\cap\mathbb{Q}(X)=\mathbb{Z}_{(p),\alpha}$. Note that α is transcendental over $\mathbb{Q}_p$ and hence also over $\mathbb{Q}$.

We prove now the final claim. Suppose there exists $\sigma\in G_{\mathbb{Q}_p}$ such that $\sigma(\alpha)=\beta$. If $\phi\in\mathbb{Z}_{(p),\alpha}$, then $\phi(\alpha)$ is defined and belongs to $\mathbb{O}_p$. In particular, $\sigma(\phi(\alpha))=\phi(\sigma(\alpha))=\phi(\beta)\in\overline{\mathbb{Z}_p}$. Hence $\mathbb{Z}_{(p),\alpha}\subseteq\mathbb{Z}_{(p),\beta}$, and the other containment is proved in a symmetrical way.

Conversely, suppose that $\mathbb{Z}_{(p),\alpha}=\mathbb{Z}_{(p),\beta}$. By Lemma 2.27, it follows that $\mathbb{Z}_{p,\alpha}=\mathbb{Z}_{p,\beta}$. Note that the last two valuation domains are the contraction to $\mathbb{Q}_p(X)$ of the valuation domains

$$\overline{\mathbb{Z}_p}_{\alpha}=\{\phi\in\overline{\mathbb{Q}_p}(X)\mid\phi(\alpha)\in\mathbb{O}_p\}\quad\text{and}\quad\overline{\mathbb{Z}_p}_{\beta}=\{\phi\in\overline{\mathbb{Q}_p}(X)\mid\phi(\alpha)\in\mathbb{O}_p\}$$

of $\overline{\mathbb{Q}_p}(X)$, respectively. By [Bourbaki 1985b, Chapter VI, §8, 6., Corollary 1], there exists a $\mathbb{Q}_p(X)$-automorphism σ of $\overline{\mathbb{Q}_p}(X)$ such that $\sigma(\overline{\mathbb{Z}_p}_{\alpha})=\overline{\mathbb{Z}_p}_{\beta}$. It is easy to check that $\sigma(\overline{\mathbb{Z}_p}_{\alpha})=\overline{\mathbb{Z}_p}_{\sigma(\alpha)}$. In particular, $\overline{\mathbb{Z}_p}_{\sigma(\alpha)}=\overline{\mathbb{Z}_p}_{\beta}$. If $\sigma(\alpha)-\beta\neq0$, let $c\in\overline{\mathbb{Z}_p}$ be such that $v_p(c)>v_p(\sigma(\alpha)-\beta)$. Let $a\in\overline{\mathbb{Q}_p}$ be such that $v_p(a-\sigma(\alpha))\geq v_p(c)$. Then the polynomial $(X-a)/c$ is in $\overline{\mathbb{Z}_p}_{\sigma(\alpha)}$ and not in $\overline{\mathbb{Z}_p}_{\beta}$, which is a contradiction. □

Note that, for a DVR W as in the statement of Corollary 2.28, there exists $\alpha\in\mathbb{O}_p\subset\mathbb{C}_p$ of bounded ramification such that $W=\mathbb{Z}_{(p),\alpha}$ if and only if $X\in W$. This last condition occurs for example if W is an overring of $\mathbb{Z}[X]$.

3. Polynomial Dedekind domains

In order to describe the family of Dedekind domains lying between $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$, we briefly recall the notion of integer-valued polynomials on algebras; see [Chabert and Peruginelli 2016; Peruginelli and Werner 2017], for example. Let D be an integral domain with quotient field K and A a torsion-free D algebra. We embed K and A into the extended K-algebra $B=A\otimes_D K$, and this allows us to evaluate polynomials over K at elements of A. If $f\in K[X]$ and $a\in A$ are such that $f(a)\in A$, then we say that f is integer-valued at a. In general, given a subset S of A, we denote by

$$\mathrm{Int}_K(S,A)=\{f\in K[X]\mid f(s)\in A,\forall s\in S\}$$

the ring of integer-valued polynomials over S. We omit the subscript K if $A=D$.

In our setting, let

$$\mathcal{O}=\prod_{p\in\mathbb{P}}\mathbb{O}_p\subset\prod_{p\in\mathbb{P}}\mathbb{C}_p.$$

Given $\alpha=(\alpha_p)\in\prod_{p\in\mathbb{P}}\mathbb{C}_p$ and $f\in\mathbb{Q}[X]$, we have $f(\alpha)=(f(\alpha_p))$, which is an element of $\prod_{p\in\mathbb{P}}\mathbb{C}_p$. If $\underline{E}=\prod_{p\in\mathbb{P}}E_p$ is a subset of $\prod_p\mathbb{C}_p$, then

$$\mathrm{Int}_{\mathbb{Q}}(\underline{E},\mathcal{O})=\{f\in\mathbb{Q}[X]\mid f(\alpha)\in\mathcal{O},\forall\alpha\in\underline{E}\};$$

that is, a polynomial f is in $\mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})$ if $f(\alpha_p) \in \mathbb{O}_p$ for each $\alpha_p \in E_p$ and $p \in \mathbb{P}$. By an argument similar to [Chabert and Peruginelli 2016, Remark 6.3], there is no loss in generality to suppose that a subset of $\prod_{p\in\mathbb{P}} \mathbb{C}_p$ is of the form $\prod_{p\in\mathbb{P}} E_p$ when dealing with such rings of integer-valued polynomials.

We remark that we have the following representation for the ring $\mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})$ as an intersection of valuation overrings (see [Peruginelli 2023, (2.2)], for example):

$$\mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O}) = \bigcap_{p\in\mathbb{P}} \bigcap_{\alpha_p\in E_p} \mathbb{Z}_{(p),\alpha_p} \cap \bigcap_{q\in\mathcal{P}^{\mathrm{irr}}} \mathbb{Q}[X]_{(q)}, \tag{3.1}$$

where $\mathcal{P}^{\mathrm{irr}}$ denotes the set of irreducible polynomials in $\mathbb{Q}[X]$. By [Peruginelli 2017, Proposition 2.2], the valuation domain $\mathbb{Z}_{(p),\alpha_p}$ of $\mathbb{Q}(X)$ has rank 1 if and only if α_p is transcendental over $\mathbb{Q}$ and has rank 2 otherwise (in the last case, note that necessarily $\alpha \in \overline{\mathbb{Q}_p}$).

A totally similar argument to [Peruginelli 2023, Lemma 2.5] shows that, for $p \in \mathbb{P}$, we have

$$(\mathbb{Z} \setminus p\mathbb{Z})^{-1}(\mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})) = \mathrm{Int}_{\mathbb{Q}}(E_p, \mathbb{O}_p) = \{f \in \mathbb{Q}[X] \mid f(E_p) \subseteq \mathbb{O}_p\}.$$

We also need to recall the following definition introduced in [Peruginelli 2023].

Definition 3.2. We say that a subset $\underline{E}$ of $\mathcal{O}$ is *polynomially factorizable* if, for each $g \in \mathbb{Z}[X]$ and $\alpha = (\alpha_p) \in \underline{E}$, there exist $n, d \in \mathbb{Z}$, $n, d \geq 1$, such that $g(\alpha)^n/d$ is a unit of $\mathcal{O}$; that is, $v_p(g(\alpha_p)^n/d) = 0$ for all $p \in \mathbb{P}$.

The next theorem characterizes which rings of integer-valued polynomials $\mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})$ are Dedekind domains. Given $p \in \mathbb{P}$ and a subset E_p of $\mathbb{C}_p$, we say that E_p has finitely many $G_{\mathbb{Q}_p} = \mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$-orbits if E_p contains finitely many equivalence classes under the relation of conjugacy over $\mathbb{Q}_p$ (we stress that E_p may not necessarily contain a full $G_{\mathbb{Q}_p}$-orbit). By Corollary 2.28, this condition holds if and only if the set $\{\mathbb{Z}_{(p),\alpha_p} \mid \alpha_p \in E_p\}$ is finite. Furthermore, if $E_p \subseteq \overline{\mathbb{Q}_p}$, then the number of $G_{\mathbb{Q}_p}$-orbits is finite if and only if E_p is a finite set.

Theorem 3.3. *Let* $\underline{E} = \prod_{p\in\mathbb{P}} E_p \subset \prod_p \mathbb{C}_p$. *Then* $\mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})$ *is a Dedekind domain if and only if, for each prime* p, E_p *is a subset of* $\mathbb{C}_p^{\mathrm{br}}$ *of transcendental elements over* $\mathbb{Q}$ *with finitely many* $G_{\mathbb{Q}_p}$*-orbits and* $\underline{E}$ *is polynomially factorizable.*

Moreover, if the above conditions hold, then the class group of $\mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})$ *is isomorphic to the direct sum of the class groups* $\mathrm{Int}_{\mathbb{Q}}(E_p, \mathbb{O}_p)$, $p \in \mathbb{P}$, *and, if* $E_p = \{\alpha_1, \ldots, \alpha_n\}$, *where the* α_i *are pairwise nonconjugate over* $\mathbb{Q}_p$, *then* $\mathrm{Cl}(\mathrm{Int}_{\mathbb{Q}}(E_p, \mathbb{O}_p)) = \mathbb{Z}/e\mathbb{Z} \oplus \mathbb{Z}^{n-1}$, *where e is the greatest common divisor of the ramifications indexes of* α_i *over* $\mathbb{Q}_p$.

In particular, assuming that E_p *is formed by pairwise nonconjugate elements over* $\mathbb{Q}_p$ *for each* $p \in \mathbb{P}$, $\mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})$ *is a PID if and only if* $\underline{E}$ *is polynomially factorizable and, for each* $p \in \mathbb{P}$, E_p *contains at most one element* $\alpha_p \in \mathbb{O}_p \cap \mathbb{C}_p^{\mathrm{br}}$ *such that* α_p *is transcendental over* $\mathbb{Q}$ *and unramified over* $\mathbb{Q}_p$.

Proof. Let $R = \mathrm{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})$.

Suppose that the above conditions on $\underline{E}$ are satisfied. By (3.1), R is equal to an intersection of DVRs. Moreover, R has finite character; that is, for every nonzero $f \in R$, f belongs to finitely many maximal ideals of the family of DVRs appearing in (3.1): in fact, if $f(X) = g(X)/n$ for some $g \in \mathbb{Z}[X]$ and $n \in \mathbb{Z}$, $n \neq 0$, then f is divisible only by finitely many $q \in \mathcal{P}^{\text{irr}}$; since $\underline{E}$ is polynomially factorizable, by [Peruginelli 2023, Lemma 2.12], the set $\{p \in \mathbb{P} \mid \exists \alpha_p \in E_p, v_p(g(\alpha_p)) > 0\}$ is finite, so that f belongs to finitely many maximal ideals of the family $\mathbb{Z}_{(p),\alpha_p}$, $\alpha_p \in E_p$, $p \in \mathbb{P}$. Hence R is a Krull domain.

Suppose that R is not a Dedekind domain. By [Heitmann 1974, Proposition 2.2], there exists a maximal ideal $M \subset R$ of height strictly greater than one. If $M \cap \mathbb{Z} = (0)$, then, since $\mathbb{Z}[X] \subseteq R \subseteq \mathbb{Q}[X]$, it follows that $R_{\mathbb{Z}\setminus\{0\}} = \mathbb{Q}[X]$ and $2 \leq ht M = ht(M_{\mathbb{Z}\setminus\{0\}}) \leq \dim(\mathbb{Q}[X]) = 1$, a contradiction. Hence $M \cap \mathbb{Z} = p\mathbb{Z}$ for some $p \in \mathbb{P}$. If we now localize at p, we have that $(\mathbb{Z} \setminus p\mathbb{Z})^{-1}R = R_p = \text{Int}_{\mathbb{Q}}(E_p, \mathbb{O}_p)$, which is a Dedekind domain by [Eakin and Heinzer 1973, Theorem]. So $(\mathbb{Z} \setminus p\mathbb{Z})^{-1}M \subset R_p$ cannot have dimension strictly greater than one, a contradiction.

Conversely, suppose that R is a Dedekind domain. In particular, for each $p \in \mathbb{P}$,

$$(\mathbb{Z} \setminus p\mathbb{Z})^{-1}R = R_p = \text{Int}_{\mathbb{Q}}(E_p, \mathbb{O}_p)$$

is a Dedekind domain, so $\{\mathbb{Z}_{(p),\alpha_p} \mid \alpha_p \in E_p\}$ is a finite set of DVRs (because p is contained in only finitely many maximal ideals of these valuation overrings) which implies that E_p is a subset of $\mathbb{C}_p^{\text{br}}$ of transcendental elements over $\mathbb{Q}$ and E_p has finitely many $G_{\mathbb{Q}_p}$-orbits. Since every polynomial of R is contained in only finitely many maximal ideals, it follows easily that $\underline{E}$ is polynomially factorizable.

Finally, suppose that R is a Dedekind domain. As in [Peruginelli 2023, Lemma 2.14], we have $\text{Cl}(R) = \bigoplus_{p\in\mathbb{P}} \text{Cl}(R_p)$, where $R_p = \text{Int}_{\mathbb{Q}}(E_p, \mathbb{O}_p)$ for $p \in \mathbb{P}$. The claim about the class group of R_p follows by [Peruginelli 2023, Proposition 2.10] or by [Eakin and Heinzer 1973, Theorem], since, for each $p \in \mathbb{P}$, we are assuming that $E_p = \{\alpha_1, \dots, \alpha_n\}$ is formed by pairwise nonconjugate elements over $\mathbb{Q}_p$.

The claim about when $\text{Int}_{\mathbb{Q}}(\underline{E}, \mathbb{O})$ is a PID is now straightforward. □

Let $\overline{\overline{\mathbb{Z}}} = \prod_{p\in\mathbb{P}} \overline{\mathbb{Z}_p}$. In [Peruginelli 2023, Theorem 2.17], we show that if R is a Dedekind domain between $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$ such that the residue fields of prime characteristic are finite fields, then $R = \text{Int}_{\mathbb{Q}}(\underline{E}, \overline{\overline{\mathbb{Z}}})$, for some $\underline{E} = \prod_p E_p \subset \overline{\overline{\mathbb{Z}}}$ such that $\underline{E}$ is polynomially factorizable and, for each $p \in \mathbb{P}$, E_p is a finite subset of $\overline{\mathbb{Z}_p}$ of transcendental elements over $\mathbb{Q}$. Now, we are able to complete the classification of the Dedekind domains R, $\mathbb{Z}[X] \subset R \subseteq \mathbb{Q}[X]$, without any restriction on the residue fields.

Theorem 3.4. *Let R be a Dedekind domain such that $\mathbb{Z}[X] \subset R \subseteq \mathbb{Q}[X]$. Then R is equal to $\text{Int}_{\mathbb{Q}}(\underline{E}, \mathcal{O})$ for some polynomially factorizable subset $\underline{E} = \prod_{p\in\mathbb{P}} E_p \subset \mathcal{O}$ such that, for each prime p, $E_p \subset \mathbb{O}_p \cap \mathbb{C}_p^{\text{br}}$ is a finite set of transcendental elements over $\mathbb{Q}$.*

Proof. Note first that, by [Peruginelli 2018, Theorem 3.14], no valuation overring of W of R can be a residually transcendental extension of $W \cap \mathbb{Q}$ since, for such a valuation domain W, the domain $W \cap \mathbb{Q}[X]$ is not Prüfer. Hence, for each prime ideal $P \subset R$ such that $P \cap \mathbb{Z} = p\mathbb{Z}$, $p \in \mathbb{P}$, R_P is a DVR of $\mathbb{Q}(X)$ which is a residually algebraic extension of $\mathbb{Z}_{(p)}$. By Corollary 2.28, there exists $\alpha \in \mathbb{O}_p \cap \mathbb{C}_p^{\text{br}}$ such that $R_p = \mathbb{Z}_{(p),\alpha}$. Let E_p be the subset of $\mathbb{C}_p^{\text{br}}$ formed by all such α_p. Note that, since p is contained in only finitely many maximal ideals P of R, it follows that E_p is a finite set; moreover, each element of E_p is transcendental over $\mathbb{Q}$ since R_P is a DVR. It now follows that

$$R=\bigcap_{p\in\mathbb{P}}\bigcap_{\substack{P\subset R\\ P\cap\mathbb{Z}=p\mathbb{Z}}} R_P\cap\mathbb{Q}[X]=\bigcap_{p\in\mathbb{P}}\mathrm{Int}_{\mathbb{Q}}(E_p,\mathbb{O}_p)=\mathrm{Int}_{\mathbb{Q}}(\underline{E},\mathcal{O}).$$

The rest of the statement follows by Theorem 3.3. □

Finally, the next corollary describes the PIDs among the family of Dedekind domains between $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$.

Corollary 3.5. *Let R be a PID such that $\mathbb{Z}[X]\subset R\subseteq\mathbb{Q}[X]$. Then R is equal to $\mathrm{Int}_{\mathbb{Q}}(\underline{E},\mathcal{O})$ for some $\underline{E}=\prod_{p\in\mathbb{P}}E_p\subset\mathcal{O}$ such that, for each prime p, E_p contains at most one element $\alpha_p\in\mathbb{O}_p\cap\mathbb{C}_p^{\mathrm{br}}$ such that α_p is transcendental over $\mathbb{Q}$ and unramified over $\mathbb{Q}_p$ and $\underline{E}=\{\alpha=(\alpha_p)\}$ is polynomially factorizable.*

Proof. By Theorem 3.4, the ring R is equal to $\mathrm{Int}_{\mathbb{Q}}(\underline{E},\mathcal{O})$ for some polynomially factorizable subset $\underline{E}=\prod_{p\in\mathbb{P}}E_p\subset\mathcal{O}$ such that, for each prime p, $E_p\subset\mathbb{O}_p\cap\mathbb{C}_p^{\mathrm{br}}$ is a set of transcendental elements over $\mathbb{Q}$ with finitely many $G_{\mathbb{Q}_p}$-orbits. Since by hypothesis the class group of R is trivial, it follows by Theorem 3.3 that, for each $p\in\mathbb{P}$, E_p contains at most one element, which is transcendental over $\mathbb{Q}$ and unramified over $\mathbb{Q}_p$. □

Remark 3.6. As we mentioned in the Introduction, given a group G which is the direct sum of a countable family of finitely generated abelian groups, there exists a Dedekind domain R between $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$ with class group G [Peruginelli 2023, Theorem 3.1]. The domain R of that construction is equal to $\mathrm{Int}_{\mathbb{Q}}(\underline{E},\mathcal{O})$ for some polynomially factorizable subset $\underline{E}=\prod_{p\in\mathbb{P}}E_p$, where E_p is a finite subset of $\overline{\mathbb{Q}_p}$ of transcendental elements over $\mathbb{Q}$. In particular, R has finite residue fields of prime characteristic [Peruginelli 2023, Theorem 2.17]; the reason is that the valuation overrings $\mathbb{Z}_{(p),\alpha_p}$ of R in (3.1) have finite residue fields precisely because α_p is chosen in $\overline{\mathbb{Q}_p}$ for each $p\in\mathbb{P}$ (Proposition 2.24).

Now, by means of Theorem 2.26, with the same method used in [Peruginelli 2023, Theorem 3.1], we can build a Dedekind domain R, $\mathbb{Z}[X]\subset R\subseteq\mathbb{Q}[X]$, with prescribed class group G as above and prescribed residue fields of prime characteristic, which can be finite or infinite algebraic extensions of the prime field $\mathbb{F}_p$ according to whether the above elements $\alpha_p\in E_p\subset\mathbb{C}_p^{\mathrm{br}}$ transcendental over $\mathbb{Q}$ are either algebraic or transcendental over $\mathbb{Q}_p$.

References

[Alexandru and Popescu 1988] V. Alexandru and N. Popescu, “Sur une classe de prolongements à $K(X)$ d’une valuation sur un corps K”, *Rev. Roumaine Math. Pures Appl.* **33**:5 (1988), 393–400. MR

[Alexandru et al. 1988] V. Alexandru, N. Popescu, and A. Zaharescu, “A theorem of characterization of residual transcendental extensions of a valuation”, *J. Math. Kyoto Univ.* **28**:4 (1988), 579–592. MR

[Alexandru et al. 1990a] V. Alexandru, N. Popescu, and A. Zaharescu, “All valuations on $K(X)$”, *J. Math. Kyoto Univ.* **30**:2 (1990), 281–296. MR

[Alexandru et al. 1990b] V. Alexandru, N. Popescu, and A. Zaharescu, “Minimal pairs of definition of a residual transcendental extension of a valuation”, *J. Math. Kyoto Univ.* **30**:2 (1990), 207–225. MR

[Alexandru et al. 1998] V. Alexandru, N. Popescu, and A. Zaharescu, “On the closed subfields of $\mathbb{C}_p$”, *J. Number Theory* **68**:2 (1998), 131–150. MR

[Bourbaki 1985a] N. Bourbaki, *Éléments de mathématique*: *Algèbre commutative*, *Chapitres* 1 *à* 4, Masson, Paris, 1985. MR

[Bourbaki 1985b] N. Bourbaki, *Éléments de mathématique*: *Algèbre commutative*, *Chapitres* 5 *à* 7, Masson, Paris, 1985. MR

[Chabert and Halberstadt 2018] J.-L. Chabert and E. Halberstadt, "On Abhyankar's lemma about ramification indices", preprint, 2018. arXiv 1805.08869

[Chabert and Peruginelli 2016] J.-L. Chabert and G. Peruginelli, "Polynomial overrings of Int($\mathbb{Z}$)", *J. Commut. Algebra* **8**:1 (2016), 1–28. MR

[Eakin and Heinzer 1973] P. Eakin and W. Heinzer, "More noneuclidian PID's and Dedekind domains with prescribed class group", *Proc. Amer. Math. Soc.* **40** (1973), 66–68. MR

[Engler and Prestel 2005] A. J. Engler and A. Prestel, *Valued fields*, Springer, 2005. MR

[Heitmann 1974] R. C. Heitmann, "PID's with specified residue fields", *Duke Math. J.* **41** (1974), 565–582. MR

[Ioviţă and Zaharescu 1995] A. Ioviţă and A. Zaharescu, "Completions of r.a.t.-valued fields of rational functions", *J. Number Theory* **50**:2 (1995), 202–205. MR

[Kaplansky 1942] I. Kaplansky, "Maximal fields with valuations", *Duke Math. J.* **9** (1942), 303–321. MR

[Khanduja and Saha 1999] S. K. Khanduja and J. Saha, "A generalized fundamental principle", *Mathematika* **46**:1 (1999), 83–92. MR

[Lampert 1986] D. Lampert, "Algebraic p-adic expansions", *J. Number Theory* **23**:3 (1986), 279–284. MR

[Matignon and Ohm 1988] M. Matignon and J. Ohm, "A structure theorem for simple transcendental extensions of valued fields", *Proc. Amer. Math. Soc.* **104**:2 (1988), 392–402. MR

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

[Ostrowski 1935a] A. Ostrowski, "Untersuchungen zur arthmetischen Theorie der Körper, I", *Math. Z.* **39**:1 (1935), 269–320. MR

[Ostrowski 1935b] A. Ostrowski, "Untersuchungen zur arithmetischen Theorie der Körper, II–III", *Math. Z.* **39**:1 (1935), 321–404. MR

[Peruginelli 2017] G. Peruginelli, "Transcendental extensions of a valuation domain of rank one", *Proc. Amer. Math. Soc.* **145**:10 (2017), 4211–4226. MR

[Peruginelli 2018] G. Peruginelli, "Prüfer intersection of valuation domains of a field of rational functions", *J. Algebra* **509** (2018), 240–262. MR

[Peruginelli 2023] G. Peruginelli, "Polynomial Dedekind domains with finite residue fields of prime characteristic", *Pacific J. Math.* **324**:2 (2023), 333–351. MR

[Peruginelli and Spirito 2020] G. Peruginelli and D. Spirito, "The Zariski–Riemann space of valuation domains associated to pseudo-convergent sequences", *Trans. Amer. Math. Soc.* **373**:11 (2020), 7959–7990. MR

[Peruginelli and Spirito 2021] G. Peruginelli and D. Spirito, "Extending valuations to the field of rational functions using pseudo-monotone sequences", *J. Algebra* **586** (2021), 756–786. MR

[Peruginelli and Werner 2017] G. Peruginelli and N. J. Werner, "Non-triviality conditions for integer-valued polynomial rings on algebras", *Monatsh. Math.* **183**:1 (2017), 177–189. MR

[Popescu and Zaharescu 1995] N. Popescu and A. Zaharescu, "On the structure of the irreducible polynomials over local fields", *J. Number Theory* **52**:1 (1995), 98–118. MR

[Ribenboim 1958] P. Ribenboim, "Corps maximaux et complets par des valuations de Krull", *Math. Z.* **69** (1958), 466–479. MR

[Ribenboim 1968] P. Ribenboim, *Théorie des valuations*, Sémin. Math. Sup. **9**, Presses Univ. Montréal, 1968. MR

[Zariski and Samuel 1960] O. Zariski and P. Samuel, *Commutative algebra, II*, Van Nostrand, Princeton, NJ, 1960. MR

Communicated by Anand Pillay
Received 2023-03-21 Revised 2024-06-10 Accepted 2024-10-21

gperugin@math.unipd.it *Dipartimento di Matematica "Tullio Levi-Civita", Università di Padova, Padova, Italy*

ALGEBRA AND NUMBER THEORY 19:10 (2025)
https://doi.org/10.2140/ant.2025.19.1973

Affine Deligne–Lusztig varieties via the double Bruhat graph, I: Semi-infinite orbits

Felix Schremmer

We introduce a new language to describe the geometry of affine Deligne–Lusztig varieties in affine flag varieties. This first part of a two-paper series develops the definition and fundamental properties of the double Bruhat graph by studying semi-infinite orbits. This double Bruhat graph was originally introduced by Naito and Watanabe to study periodic R-polynomials. We use it to describe the geometry of many affine Deligne–Lusztig varieties, overcoming a previously ubiquitous regularity condition.

1. Introduction

Shimura varieties play a central role in the Langlands program. By giving the Shimura variety an interpretation as a moduli space (e.g., of certain abelian varieties), one obtains an integral model whose generic fibre recovers the original Shimura variety [Rapoport 2005; Kisin and Pappas 2018; Pappas 2023]. The special fibre of such an integral model is then known as the mod p reduction of the Shimura variety. In the case of a parahoric level structure, the geometry of each special fibre is closely related to the geometry of corresponding affine Deligne–Lusztig varieties. Similar affine Deligne–Lusztig varieties occur in the special fibres of moduli spaces of local G-shtukas [Viehmann 2018].

We consider a reductive group G defined over a local field F, whose completion of the maximal unramified extension we denote by $\breve{F}$. Given a parahoric subgroup $K \subset G(\breve{F})$, we associate the affine Deligne–Lusztig variety $X_x^K(b)$ to any two elements x and b in $G(\breve{F})$ [Rapoport 2005, Definition 4.1]. It is defined as a locally closed subvariety of the partial flag variety associated with K. It has the structure of a finite-dimensional scheme or perfect scheme over the residue field k of $\breve{F}$, whose geometric points are given by

$$X_x(b) = X_x^K(b) = \{g \in G(\breve{F})/K \mid g^{-1}b\sigma(g) \in KxK\} \subset G(\breve{F})/K.$$

Here, σ denotes the Frobenius of $\breve{F}/F$. One notes that the affine Deligne–Lusztig variety depends, up to isomorphism, only on the double coset $KxK \subset G(\breve{F})$ and the σ-conjugacy class

$$[b] = \{g^{-1}b\sigma(g) \mid g \in G(\breve{F})\}.$$

MSC2020: 11G25, 20G25.
Keywords: affine Weyl group, affine Deligne–Lusztig variety, double Bruhat graph, semi-infinite orbit, Shimura variety, Langlands program.

The intersection of these two sets $KxK \cap [b]$ is known as Newton stratum and its geometry is closely related to that of the affine Deligne–Lusztig variety. The most important questions, in increasing order of difficulty, are the following:

(Q1) When is $X_x(b)$ empty? Equivalently, when is the Newton stratum empty?

(Q2) If $X_x(b) \neq \varnothing$, what is its dimension?

(Q3) How many irreducible components of any given dimension does $X_x(b)$ have?

The final question is especially interesting when the given dimension is equal to $\dim X_x(b)$, i.e., if one asks for the top-dimensional irreducible components. The number of such irreducible components will in general be infinite. However, the σ-centralizer of b,

$$J_b(F) = \{g \in G(\breve{F}) \mid g^{-1}b\sigma(g) = b\},$$

acts by left multiplication on $X_x(b)$. There are only finitely many orbits of irreducible components up to the $J_b(F)$-action, which is how (Q3) should be understood. Equivalently, one may ask for the number of top-dimensional irreducible components of the Newton stratum.

The first step towards answering these three questions is to give suitable parametrizations for the double cosets

$$K \setminus G(\breve{F})/K = \{KxK \mid x \in G(\breve{F})\}$$

and σ-conjugacy classes

$$B(G) = \{[b] \mid b \in G(\breve{F})\}.$$

We will assume that the group G is split and choose a split maximal torus T. For this introduction, this merely provides a slightly more convenient notation. More importantly, this restriction is essential for the remainder of the article due to the dependence on the earlier work [Görtz et al. 2006] with this assumption.

We first consider the case of a hyperspecial subgroup K. If G is already defined over the ring of integers $\mathcal{O}_{\breve{F}}$ of $\breve{F}$, then $K = G(\mathcal{O}_{\breve{F}})$ would be a typical example of this. For hyperspecial K, the double cosets $K\setminus G(\breve{F})/K$ are parametrized by the dominant elements of the cocharacter lattice $X_*(T)$. Explicitly, evaluation at a uniformizer assigns to each cocharacter $\mu \in X_*(T)$ a representative $\dot{\mu} \in T(\breve{F})$, and then each double coset KxK contains the representative of precisely one dominant cocharacter μ. We also write $K\mu K$ for $K\dot{\mu}K$. Then the *Cartan decomposition* is given by

$$G(\breve{F}) = \bigsqcup_{\substack{\mu \in X_*(T)\\ \text{dominant}}} K\mu K.$$

If $K = I$ is an Iwahori subgroup, the double cosets $I \setminus G(\breve{F})/I$ are parametrized by the extended affine Weyl group $\widetilde{W}$. This group can be defined as $N_G(T)(\breve{F})/T(\mathcal{O}_{\breve{F}})$ and it is isomorphic to the semidirect product of the Weyl group $W = N_G(T)(\breve{F})/T(\breve{F})$ of G with the cocharacter lattice $X_*(T)$. Here, we write $N_G(T)$ for the normalizer of T inside G. Choosing for each $x \in \widetilde{W}$ a representative $\dot{x} \in N_G(T)(\breve{F})$, the double coset $IxI = I\dot{x}I$ is independent of this choice. We obtain the *Iwahori–Bruhat decomposition*

$$G(\breve{F}) = \bigsqcup_{x \in \widetilde{W}} IxI.$$

For general parahoric levels, one may parametrize $K \setminus G(\breve{F})/K$ by suitable double cosets in $\widetilde{W}$, but we will not consider this case.

The σ-conjugacy class of an element $b \in G(\breve{F})$ is uniquely determined by two invariants; this is a celebrated result of Kottwitz [1985; 1997]. These invariants are known as the (dominant) Newton point $\nu(b) \in X_*(T) \otimes \mathbb{Q}$ and the Kottwitz point $\kappa(b) \in \pi_1(G)$. Here, $\pi_1(G) = X_*(T)/\mathbb{Z}\Phi^\vee$ is the Borovoi fundamental group and $\mathbb{Z}\Phi^\vee$ is the coroot lattice. Following [He 2014, Theorem 3.7], one may also parametrize the set $B(G)$ using σ-conjugacy classes in $\widetilde{W}$.

Since the Kottwitz point $\kappa : G(\breve{F}) \to \pi_1(G)$ parametrizes the connected components of the partial flag variety, we get that $\kappa(x) = \kappa(b)$ is a necessary condition for $X_x(b) \neq \varnothing$. Once this condition is imposed, we may focus on comparing the above parametrization for x with the Newton point $\nu(b)$.

In the case of hyperspecial level, our three initial questions have been mostly solved after concentrated effort by many researchers. For the split case under consideration, we can summarize the results as follows (while still providing references for the general case).

Theorem 1.1. *Assume that K is hyperspecial. Let $[b] \in B(G)$ and $\mu \in X_*(T)$ be a dominant coweight.*

(a) *The affine Deligne–Lusztig variety $X_\mu(b)$ is nonempty if and only if the **Mazur inequality** is satisfied, that is, $\kappa(b) = \kappa(\mu)$ and $\nu(b) \leq \mu$ in the dominance order of $X_*(T) \otimes \mathbb{Q}$. This was conjectured by Kottwitz and Rapoport, and proved in* [Rapoport and Richartz 1996; Gashi 2010; He 2014].

(b) *If $X_\mu(b) \neq \varnothing$, it is equidimensional of dimension*

$$\dim X_\mu(b) = \tfrac{1}{2}\big(\langle \mu - \nu(b), 2\rho\rangle - \operatorname{def}(b)\big).$$

Here, $\operatorname{def}(b)$ *denotes the **defect** of* b, *which is defined as* $\operatorname{rk}_F(G) - \operatorname{rk}_F(J_b)$; *see* [Chai 2000; Kottwitz 2006]. *This was conjectured by Rapoport, and proved in* [Görtz et al. 2006; Viehmann 2006; Hamacher 2015; Takaya 2025].

(c) *The $J_b(F)$-orbits of irreducible components of $X_\mu(b)$ are in bijection with a certain basis of the weight space $M_\mu(\lambda(b))$ of the irreducible quotient M_μ of the highest-weight Verma module V_μ. Here, $\lambda(b) \in X_*(T)$ is the largest cocharacter satisfying $\lambda(b) \leq \nu(b)$ and $\kappa(\lambda(b)) = \kappa(b)$. This was conjectured by Chen and Zhu, and proved in* [Zhou and Zhu 2020; Nie 2022].

We see that once $\kappa(\mu) = \kappa(b)$ is required, the difference $\mu - \nu(b)$, resp. $\mu - \lambda(b)$, determines most properties of $X_\mu(b)$, using, e.g., the fact that the dimension of the weight space $M_\mu(\lambda(b))$ can be approximated using the dimension of $V_\mu(\lambda(b))$, which is *Kostant's partition function* applied to the difference $\mu - \lambda(b)$. Under certain regularity conditions, the dimensions of the two weight spaces will be equal.

Let us now summarize the most important results known in the case of Iwahori level structure. Assume that $K = I$ is an Iwahori subgroup. Pick an element $x \in \widetilde{W} \cong W \ltimes X_*(T)$ and write it as $x = wt^\mu$, where $w \in W$ and $\mu \in X_*(T)$. The element $t \in F$ is the uniformizer, so the representative of t^μ in G is given by the image of t under the cocharacter μ. We set

$$B(G)_x = \{[b] \in B(G) \mid IxI \cap [b] \neq \varnothing\} = \{[b] \in B(G) \mid X_x(b) \neq \varnothing\}.$$

It should not be surprising that $B(G)_x$ contains a unique minimal and a unique maximal element, and both have been explicitly described [Viehmann 2014; 2021; Schremmer 2022].

For any $[b] \in B(G)_x$, we know that $\dim X_x(b) \leq d_x(b)$ [He 2016, Theorem 2.30], where $d_x(b)$ is the virtual dimension defined in [He 2014, Section 10]. It is defined as

$$d_x(b) = \tfrac{1}{2}\big(\ell(x) + \ell(\eta(x)) - \langle \nu(b), 2\rho\rangle - \operatorname{def}(b)\big).$$

The definition of $\eta(x) \in W$ is somewhat technical, so we will not recall it here. A striking feature of the virtual dimension is that it is a simple sum of four terms, the first two only depending on $x \in \widetilde{W}$ and the latter two only depending on $[b] \in B(G)$. The virtual dimension behaves best when the element x satisfies a certain regularity condition known as being in a shrunken Weyl chamber; see [Schremmer 2022, Example 2.8].

Theorem 1.2. *Let $x \in \widetilde{W}$, denote the largest element in $B(G)_x$ by $[b_x]$ and the smallest one by $[m_x]$.*

(a) *Suppose that* $\dim X_x(b_x) = d_x(b_x)$. *Then*

$$B(G)_x = \{[b] \in B(G) \mid [m_x] \leq [b] \leq [b_x]\}.$$

For each $[b] \in B(G)_x$, the variety $X_x(b)$ is equidimensional of dimension $\dim X_x(b) = d_x(b)$ [Milićević and Viehmann 2020, Theorem 1.1]. *The elements x satisfying this condition have been classified*; *see* [Schremmer 2022, Theorem 1.2].

(b) *Suppose that x lies in a shrunken Weyl chamber. Then* $\dim X_x(m_x) = d_x(m_x)$ [Viehmann 2021, Theorem 1.1(2)]. *For each $[m_x] \leq [b] \in B(G)$ with $\nu(b_x) - \nu(b)$ sufficiently large*, *we have $X_x(b) \neq \varnothing$ and* $\dim X_x(b) = d_x(b)$ [He 2021, Theorem 1.1].

While in a quantitative sense "most" elements of $\widetilde{W}$ lie in a shrunken Weyl chamber, the examples coming from Shimura varieties typically do not. In fact, the difference between virtual dimension and dimension for basic $[b]$ can be quite large for these examples.

It should not be a surprise that there are many examples where dimension and virtual dimension differ. For nonshrunken elements, the notion of virtual dimension behaves poorly. For example, it is not compatible with certain natural automorphisms of the reductive group G that preserve the Iwahori subgroup I (and hence induce isomorphisms of affine Deligne–Lusztig varieties). Even for shrunken elements, we expect to have $\dim X_x(b) = d_x(b)$ only for "small" elements $[b] \in B(G)_x$.

The case (a) in Theorem 1.2 is known as the *cordial* case. While (Q1) and (Q2) have "ideal" answers in this case, these descriptions are too good to be true in general. It is known that the set $B(G)_x$ will in general contain gaps and that affine Deligne–Lusztig varieties may fail to be equidimensional. If x is cordial, the answer to (Q3) does not seem to be known in general.

We may summarize that (Q1) and (Q2) are well understood if $[b] \in B(G)$ is small relative to x and x is in a shrunken Weyl chamber, or if x enjoys some exceptionally good properties.

Moreover, all three questions are perfectly understood in the case that $[b] = [b_x]$ is the largest σ-conjugacy class in $B(G)_x$, also known as the generic σ-conjugacy class of IxI. We have $\dim X_x(b_x) = \ell(x) - \langle \nu(b_x), 2\rho\rangle$ [He 2016, Theorem 2.23]. Up to the $J_b(F)$-action, there is only one irreducible component in $X_x(b_x)$ [Milićević and Viehmann 2020, Lemma 3.2]. In order to describe $[b_x]$ in terms of x, one may (and arguably should) use the quantum Bruhat graph [Milićević 2021; Schremmer 2022].

The goal of this paper and its sequel is to introduce a new concept, which generalizes the virtual dimension in the case of Theorem 1.2(b) and also generalizes the known theory for the generic σ-conjugacy class. We give answers to all three above questions in many cases that were previously intractable.

In this article, we follow one of the oldest approaches towards affine Deligne–Lusztig varieties in the affine flag variety, namely the one developed by Görtz, Haines, Kottwitz and Reuman [Görtz et al. 2006, Section 6]. They consider the case of a split group G, an equal characteristic field F and an integral σ-conjugacy class $[b] \in B(G)$; the element $x \in \widetilde{W}$ is allowed to be arbitrary. They compare the geometric properties of $X_x(b)$ (especially questions (Q1)–(Q3)) to similar geometric properties of intersections, in the affine flag variety, of IxI with certain *semi-infinite orbits*.

Given a Borel $B = TU$, we get another decomposition of $G(\breve{F})$, resp. the affine flag variety:

$$G(\breve{F}) = \bigsqcup_{y \in \widetilde{W}} U(L)yI.$$

The individual pieces $U(L)yI$ are called *semi-infinite orbits*. Each Borel containing our fixed torus T gives rise to a different decomposition. In our notation, we will fix B and then consider the semi-infinite orbit decompositions associated with the conjugates $uBu^{-1} = {}^uB$ for various $u \in W$.

In order to understand $X_x(b)$ following [Görtz et al. 2006, Theorem 6.3.1], we have to understand the intersections

$$IxI \cap {}^uU(\breve{F})yI \subset G(L)/I \tag{1.3}$$

for various $u \in W$ and $y \in \widetilde{W} \cap [b]$. One may ask questions (Q1)–(Q3) analogously for these intersections. Unfortunately, not many answers to these questions have been given in the previous literature, leaving basic geometric properties of (1.3) largely open. There is a decomposition of (1.3) into subvarieties parametrized by folded alcove walks [Parkinson et al. 2009, Theorem 7.1], which has been used to study affine Deligne–Lusztig varieties [Milićević et al. 2019], but these results have often been difficult to apply in practice.

One may always find an element $v \in W$ such that $IxI \subseteq {}^vU(L)xI$, and we will use this semi-infinite orbit to approximate IxI. Doing so (in the proof of Theorem 5.2 below), we can compare the intersection (1.3) to the intersection

$$({}^vU(\breve{F}) \cap {}^{uw_0}U(\breve{F}))xI \cap {}^uU(\breve{F})yI \tag{1.4}$$

for $v \in W$ such that $IxI \subseteq {}^vU(\breve{F})xI$. We write $w_0 \in W$ for the longest element, so that ${}^{uw_0}B$ is the Borel subgroup opposite to ${}^uB$. As an application of our findings, we will later see in Proposition 4.14 that (1.4) can equivalently be expressed as the intersection of three semi-infinite orbits, since

$$({}^vU(\breve{F}) \cap {}^{uw_0}U(\breve{F}))xI = ({}^vU(\breve{F})xI) \cap ({}^{uw_0}U(\breve{F})xI).$$

The first part of this paper studies intersections of the form (1.4). This is a question of independent interest, whose answer we want to later apply to affine Deligne–Lusztig varieties. A different motivation to study intersections as in (1.4) is the following: One may naturally ask about the intersections of arbitrary semi-infinite orbits

$$ {}^{u}U(\breve{F})xI \cap {}^{v}U(\breve{F})yI \subset G(\breve{F})/I. \tag{1.5}$$

Observe that the group ${}^{u}U(\breve{F}) \cap {}^{v}U(\breve{F})$ acts by left multiplication on (1.5), and the orbits of this action will be infinite-dimensional (unless $u = vw_0$). However, each such orbit will contain a point of (1.4), so we may see (1.4) as a finite-dimensional space of representatives of (1.5). Moreover, the intersection (1.5) is empty if and only if the intersection (1.4) is empty.

We study the intersection (1.4) for arbitrary x, y, u, v in Section 3. By comparing the valuation of root subgroups with the extended affine Weyl group, we get a decomposition of (1.4) into finitely many locally closed subvarieties, each of them irreducible and finite-dimensional.

It turns out that there is very convenient combinatorial tool to parametrize the subvarieties of this decomposition and to describe their dimensions. This is the double Bruhat graph, a combinatorial object introduced in [Naito and Watanabe 2017, Section 5.1] in order to study periodic R-polynomials. The double Bruhat graph is a finite graph associated with the finite Weyl group W, and it generalizes the aforementioned quantum Bruhat graph. We compare the double Bruhat graph with some foundational literature on the quantum Bruhat graph in Section 4.

Thus, our first main result expresses the intersections of semi-infinite orbits using the double Bruhat graph.

Theorem 1.6 (see Theorem 4.6). *Let $x, y \in \widetilde{W}$ and $u, v \in W$. Denote by $w_0 \in W$ the longest element. Then the intersection*

$$ {}^{u}U(\breve{F})xI \cap ({}^{uw_0}U(\breve{F}) \cap {}^{v}U(\breve{F}))yI \subset G(\breve{F})/I $$

has finite dimension (or is empty). We provide a decomposition into finitely many locally closed subsets of the affine flag variety, parametrized by certain paths in the double Bruhat graph. Each subset is irreducible, smooth and we calculate its dimension explicitly.

Finally, in Section 5, we apply these results on semi-infinite orbits to questions on affine Deligne–Lusztig varieties. We review the theory of [Görtz et al. 2006] and study the approximation of IxI by semi-infinite orbits. We introduce a new regularity condition on elements $x \in \widetilde{W}$ that we call *superparabolic*. While this is a fairly restricting assumption, it covers in a quantitative sense "most" elements in the extended affine Weyl group.

Theorem 1.7 (see Theorem 5.7). *Let $x \in \widetilde{W}$ and choose an integral element $[b] \in B(G)$. We give a necessary condition for $X_x(b) \neq \varnothing$ and an upper bound d for its dimension, both in terms of the double Bruhat graph. This improves previously known estimates such as Mazur's inequality or He's virtual dimension. We also give an upper bound for the number of $J_b(F)$-orbits of d-dimensional irreducible components of $X_x(b)$.*

If $x = wt^{\mu}$ is superparabolic, and $\mu^{\mathrm{dom}} - \nu(b)$ is small relative to the superparabolicity condition imposed, then the above "necessary condition" for $X_x(b) \neq \varnothing$ becomes sufficient, and the above upper *bound for the dimension is sharp, i.e.,* $\dim X_x(b) = d$*. If moreover the Newton point of* $[b]$ *is regular, then the above upper bound for the number of irreducible components is sharp.*

If x is in a shrunken Weyl chamber, the superparabolicity condition is simply a superregularity condition like the ones typically studied in the literature, e.g., [Milićević 2021; Milićević and Viehmann 2020; He and Yu 2021]. While some affine Deligne–Lusztig varieties associated with superregular elements x have been described in the past, this was only possible in the cases where $[b] \in B(G)$ is either the largest element $[b_x] \in B(G)_x$ or relatively small in $B(G)_x$ (in the sense of Theorem 1.2(b)). Our result "fills the gap", describing the geometry of $X_x(b)$ when $[b]$ is relatively large with respect to x.

Moreover, there are plenty of superparabolic elements which do not lie in any shrunken Weyl chamber. In fact, in a quantitative sense, "most" elements which do not lie in shrunken Weyl chambers are superparabolic. These cases have rarely been studied in the past, and the geometry of $X_x(b)$ has only been understood in very specific situations (such as x being cordial or $[b] = [b_x]$). Theorem 1.7 fully answers our main questions for superparabolic $x \in \widetilde{W}$ and many $[b] \in B(G)$.

Theorem 1.7 crucially assumes that F is of equal characteristic, the group G is split and the element $[b]$ is integral. The assumption on F can easily be removed from the theorem by using formal arguments comparing the equal characteristic case with the mixed characteristic case; see [He 2014, Section 6.1]. It is reasonable to expect that the assumption of G being split can be lifted if one finds an appropriate generalization of [Görtz et al. 2006] to nonsplit groups. It is unfortunately unclear how to lift the assumption of $[b]$ being integral for the method of this paper to work. The generalization of [Görtz et al. 2006] to nonintegral σ-conjugacy classes $[b]$ is given in [Görtz et al. 2010], but the connection between the latter paper and the double Bruhat graph remains unclear. In the second part of this two-paper series, we will consider a different approach towards the geometry of $X_x(b)$. This approach comes without any assumptions on G, F, $[b]$, but requires superregular elements $x \in \widetilde{W}$ instead of the more permissible notion of superparabolic elements considered here.

By introducing the double Bruhat graph, we can capture the delicate interplay between $x \in \widetilde{W}$ and $[b] \in B(G)$, which is not accounted for, e.g., by the notion of virtual dimension. Using this new language, we give new insights on the geometry of affine Deligne–Lusztig varieties, filling a conceptual vacuum of what $\dim X_x(b)$ "should be" when it cannot be virtual dimension. In this paper and its sequel, we hope to give a glimpse of what a generalization of Theorem 1.1 to the Iwahori level might look like.

2. Notation

Let $\mathbb{F}_q$ be a finite field and $F = \mathbb{F}_q((t))$ the field of formal Laurent series. We denote the usual t-adic valuation by ν_t. Then $\mathcal{O}_F = \mathbb{F}_q[[t]]$ is its ring of integers. Choose an algebraic closure $k = \overline{\mathbb{F}}_q$ and denote by $L = \breve{F} = k((t))$ the completion of the maximal unramified extension of F. We write $\mathcal{O}_L = k[[t]]$ for its ring of integers. Denote the Frobenius of L/F by σ, i.e.,

$$\sigma\Big(\sum_i a_i t^i\Big) = \sum_i (a_i)^q t^i.$$

We consider a split reductive group G defined over $\mathcal{O}_F$. We fix a split maximal torus and a Borel $T \subset B \subset G$ both defined over $\mathcal{O}_F$. As our Iwahori subgroup I, we choose the preimage of $B(k)$ under the projection $G(\mathcal{O}_L) \to G(k)$.

Denote the (co)character lattices of T by $X^*(T)$, resp. $X_*(T)$, and the (co)root systems by $\Phi \subset X^*(T)$, resp. $\Phi^\vee \subset X_*(T)$. The positive roots defined by B are denoted by Φ^+. We let $W = N_G(T)/T$ be the Weyl group of W and $\widetilde{W} = N_G(T)(L)/T(\mathcal{O}_L)$ the extended affine Weyl group. Under the isomorphism $\widetilde{W} \cong W \ltimes X_*(T)$, we write elements $x \in \widetilde{W}$ as $x = wt^\mu$ for $w \in W,\ \mu \in X_*(T)$. A representative of $\varepsilon^\mu \in \widetilde{W}$ is given by evaluating the cocharacter μ at the *inverse* of the uniformizer $t \in L$, i.e., by $\mu(t^{-1}) \in N_G(T)(L)$.

Denote by U the unipotent radical of B, so that $B = UT$. For each $\alpha \in \Phi$, we denote the corresponding root subgroup by $U_\alpha \subset G$. These come with an isomorphism to $U_\alpha \cong G_a$, the 1-dimensional additive group over F, from the construction of the Bruhat–Tits building.

The set of affine roots is $\Phi_{\mathrm{af}} = \Phi \times \mathbb{Z}$. For each affine root $a = (\alpha, n)$, we define the affine root subgroup $U_a \subset U_\alpha(L)$ to be the set of elements of the form $U_\alpha(rt^n)$ with $r \in k$. The natural action of $\widetilde{W}$ on Φ_{af} is given by

$$(wt^\mu)(\alpha, n) = (w\alpha, n - \langle \mu, \alpha \rangle).$$

We denote the positive affine roots by Φ_{af}^+; these are those $a \in \Phi_{\mathrm{af}}$ with $U_a \subset I$. By abuse of notation, we denote the indicator function of positive roots by Φ^+ as well. Then

$$a = (\alpha, n) \in \Phi_{\mathrm{af}}^+ \quad \Longleftrightarrow \quad n \geq \Phi^+(-\alpha) := \begin{cases} 1, & \alpha \in \Phi^-, \\ 0, & \alpha \in \Phi^+. \end{cases}$$

Denote the set of simple roots by $\Delta \subseteq \Phi^+$ and the set of simple affine roots by $\Delta_{\mathrm{af}} \subseteq \Phi_{\mathrm{af}}^+$. The latter are given by the roots of the form $(\alpha, 0)$ for $\alpha \in \Delta$ as well as $(-\theta, 1)$ whenever θ is the highest root of an irreducible component of Δ.

For $x = wt^\mu \in \widetilde{W}$, we denote by $\mathrm{LP}(x) \subseteq W$ the set of length positive elements as introduced by [Schremmer 2022, Section 2.2]. We remark that $\mathrm{LP}(x)$ is always nonempty, and it collapses to one single element if and only if x satisfies a mild regularity condition known as a *shrunken Weyl chamber* [Görtz et al. 2010, Definition 7.2.1]. This is equivalent to x lying in the lowest two-sided Kazhdan–Lusztig cell. If $\mathrm{LP}(x) = \{v\}$, then the element $\eta(x)$ occurring in the definition of virtual dimension above is given by $v^{-1}wv$.

3. Semi-infinite orbits

For any $u \in W$, the affine flag variety can be decomposed into semi-infinite orbits

$$G(L)/I = \bigsqcup_{x \in \widetilde{W}} {}^u U(L) x I.$$

Each element of the finite Weyl group W yields a different decomposition of $G(L)/I$, so one may naturally ask how these decompositions are related. Given $u, v \in W$ and $x, y \in \widetilde{W}$, we would like to understand

$${}^u U(L) x I \cap {}^v U(L) y I \subset G(L)/I.$$

Up to multiplying both sides by x^{-1} on the left and relabelling, it suffices to study intersections

$${}^{uw_0} U(L) y I \cap {}^v U(L) I \subset G(L)/I.$$

Here, we write $w_0 \in W$ for the longest element of the Weyl group such that ${}^{uw_0}B(L)$ is the Borel subgroup opposite to ${}^{u}B(L)$.

Let us enumerate the positive roots as $\Phi^+ = \{\beta_1, \dots, \beta_{\#\Phi^+}\}$. Then every element $g \in {}^{v}U(L)$ can be written in the form $g = U_{v\beta_1}(g_1) \cdots U_{v\beta_{\#\Phi^+}}(g_{\#\Phi^+})$, with $g_1, \dots, g_{\#\Phi^+} \in L$. For each such element g, there exists a uniquely determined $y \in \widetilde{W}$ with $gI \in {}^{uw_0}U(L)yI$, and we wish to compute that element y in terms of the $g_i \in L$.

In order to facilitate this computation, we make two simplifications. First, let us restrict the enumeration of positive roots $\Phi^+ = \{\beta_1, \dots, \beta_{\#\Phi^+}\}$ such that

$$\{\beta \in \Phi^+ \mid (uw_0)^{-1}v\beta \in \Phi^+\} = \{\beta_1, \dots, \beta_n\}$$

for some $n \in \{0, \dots, \#\Phi^+\}$. Then

$$U_{v\beta_1}(g_1) \cdots U_{v\beta_n}(g_n) \in {}^{uw_0}U(L)$$

by choice of the labelling of the positive roots. Hence we may replace g by

$$g' = U_{v\beta_{n+1}}(g_{n+1}) \cdots U_{v\beta_{\#\Phi^+}}(g_{\#\Phi^+}),$$

using that $gI \in {}^{uw_0}U(L)yI$ if and only if $g'I \in {}^{uw_0}U(L)yI$.

For now, we expressed $g' \in {}^{v}U(L) \cap {}^{u}U(L)$ using an arbitrary enumeration of the roots

$$\{\beta \in \Phi^+ \mid u^{-1}v\beta \in \Phi^+\} = \{\beta_{n+1}, \dots, \beta_{\#\Phi^+}\}.$$

Our second simplification is to use not just any such enumeration, but rather a specific one with extra structure, namely a *reflection order*.

Lemma 3.1 [Dyer 1993, Proposition 2.13; Papi 1994]. *Let $\prec$ be a total order on Φ^+. Then the following are equivalent*:

(a) *For all $\alpha, \beta \in \Phi^+$ with $\alpha + \beta \in \Phi^+$, we have*

$$\alpha \prec \alpha + \beta \prec \beta \quad \text{or} \quad \beta \prec \alpha + \beta \prec \alpha.$$

(b) *There exists a uniquely determined reduced word for the longest element $w_0 = s_{\alpha_1} \cdots s_{\alpha_{\#\Phi^+}}$ with corresponding simple roots $\alpha_1, \dots, \alpha_{\#\Phi^+} \in \Delta$ such that*

$$\alpha_1 \prec s_{\alpha_1}(\alpha_2) \prec \cdots \prec s_{\alpha_1} \cdots s_{\alpha_{\#\Phi^+-1}}(\alpha_{\#\Phi^+}). \qquad \square$$

A total order satisfying these equivalent conditions is called a *reflection order*. The following important facts on reflection orders will be used frequently.

Lemma 3.2. *Let $\prec$ be a reflection order and write $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$.*

(a) *For $1 \le a \le b \le \#\Phi^+$ and $u \in W$, the subsets*

$$U_{u\beta_a}(L)U_{u\beta_{a+1}}(L) \cdots U_{u\beta_b}(L) \quad \text{and} \quad U_{u\beta_{b+1}}(L) \cdots U_{u\beta_{\#\Phi^+}}(L)U_{-u\beta_1}(L)U_{-u\beta_2}(L) \cdots U_{-u\beta_{a-1}}(L)$$

are subgroups of $G(L)$.

(b) *For* $n = 0, \ldots, \#\Phi^+$ *and* $u = s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}} \in W$, *we have*

$$\{\beta \in \Phi^+ \mid u^{-1}\beta \in \Phi^-\} = \{\beta_{n+1}, \ldots, \beta_{\#\Phi^+}\}.$$

Any $u \in W$ *arises in this way for some reflection order* $\prec$ *and some index* $n \in \{0, \ldots, \#\Phi^+\}$.

Proof. (a) If $\alpha, \beta \in \{\beta_a, \ldots, \beta_b\}$, then any positive linear combination of α, β that lies in Φ^+ will also lie in this set. The fact that the first subset of $G(L)$ is a subgroup thus follows from the known theory of root subgroups [Springer 1998, Proposition 8.2.3].

Let us study the second subset. If both α, β lie in $\{\beta_{b+1}, \ldots, \beta_{\#\Phi^+}\}$, or both lie in $\{-\beta_1, \ldots, -\beta_{a-1}\}$, so will their sum (if it is in Φ). So suppose that $\alpha \in \{\beta_{b+1}, \ldots, \beta_{\#\Phi^+}\}$ and $\beta \in \{-\beta_1, \ldots, -\beta_{a-1}\}$ satisfy $\alpha + \beta \in \Phi$.

If $\alpha + \beta \in \Phi^+$, then $\alpha = (\alpha + \beta) + (-\beta)$ is expressed as the sum of two positive roots, which cannot both be $\prec \alpha$. Hence $\alpha + \beta \succ \alpha$; thus $\alpha + \beta \in \{\beta_{b+1}, \ldots, \beta_{\#\Phi^+}\}$ as well.

If $\alpha + \beta \in \Phi^-$, then $-\beta = \alpha - (\alpha + \beta)$ is expressed as the sum of two positive roots, which cannot both be $\succ -\beta$. Hence $-(\alpha + \beta) \prec -\beta$, so $\alpha + \beta \in \{-\beta_1, \ldots, -\beta_{a-1}\}$. The claim follows as above.

(b) Let $w_0 = s_{\alpha_1} \cdots s_{\alpha_{\#\Phi^+}}$ be the reduced word such that $\beta_i = s_{\alpha_1} \cdots s_{\alpha_{i-1}}(\alpha_i)$ for $i = 1, \ldots, \#\Phi^+$. Then

$$uw_0 = s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}} \cdots s_{\beta_{\#\Phi^+}} \cdots s_{\beta_1} = s_{\beta_n} \cdots s_{\beta_1} = s_{\alpha_1} \cdots s_{\alpha_n}.$$

Hence

$$\begin{aligned}\{\beta \in \Phi^+ \mid u^{-1}\beta \in \Phi^-\} &= \{\beta \in \Phi^+ \mid (uw_0)^{-1}\beta \in \Phi^+\} \\ &= \Phi^+ \setminus \{\beta \in \Phi^+ \mid (uw_0)^{-1}\beta \in \Phi^-\} \\ &= \Phi^+ \setminus \{\alpha_1, s_{\alpha_1}(\alpha_2), \ldots, s_{\alpha_1} \cdots s_{\alpha_{n-1}}(\alpha_n)\} \\ &= \Phi^+ \setminus \{\beta_1, \ldots, \beta_n\} = \{\beta_{n+1}, \ldots, \beta_{\#\Phi^+}\}.\end{aligned}$$

This shows the first claim. Now for any given $u \in W$, we can find some reduced word $uw_0 = s_{\alpha_1} \cdots s_{\alpha_n}$. Continue it to the right to a reduced word for w_0 to obtain the desired reflection order. □

So when studying intersections as above, i.e.,

$$^{uw_0}U(L)yI \cap ({}^uU(L) \cap {}^vU(L))I \subset G(L)/I, \tag{3.3}$$

we may write

$$^uU(L) \cap {}^vU(L) = U_{u\beta_1}(L) \cdots U_{u\beta_n}(L)$$

for a suitable reflection order $\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}$. With this notation, the fundamental method to evaluate intersections as in (3.3) is given by the following lemma.

Lemma 3.4. *Let* $\prec$ *be a reflection order and write* $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$. *Let* $x \in \widetilde{W}$, $u \in W$ *and* $1 \leq n \leq \#\Phi^+$. *Consider an element of the form*

$$g = U_{u\beta_1}(g_1) \cdots U_{u\beta_n}(g_n) \in {}^uU(L).$$

Let $m = \nu_L(g_n) \in \mathbb{Z}$, $b = (u\beta_n, m) \in \Phi_{\mathrm{af}}$ *and* $u' = us_{\beta_n}$.

(a) *If* $x^{-1}b \in \Phi_{\mathrm{af}}^+$, *then*

$$gxI = U_{u\beta_1}(g_1)\cdots U_{u\beta_{n-1}}(g_{n-1})xI/I \in G(L)/I.$$

(b) *If* $x^{-1}b \in \Phi_{\mathrm{af}}^-$, *then there are polynomials* $f_1, \dots, f_{n-1}$ *with*

$$f_i \in \mathbb{Z}[X_{i+1}, \dots, X_{n-1}, Y],$$

allowing us to write

$$gxI \in U_{-u\beta_n}(L)\cdots U_{-u\beta_{\#\Phi^+}}(L)U_{u\beta_1}(\tilde{g}_1)\cdots U_{u\beta_{n-1}}(\tilde{g}_{n-1})r_b xI/I \subset G(L)/I,$$

where

$$\tilde{g}_i = g_i + f_i(g_{i+1}, \dots, g_{n-1}, g_n^{-1}) \in L.$$

The polynomial f_i *is a sum of monomials*

$$\varphi X_{i+1}^{e_{i+1}}\cdots X_{n-1}^{e_{n-1}}Y^f$$

satisfying the conditions $\varphi \in \mathbb{Z}$ *and*

$$e_{i+1}\beta_{i+1} + \cdots + e_{n-1}\beta_{n-1} - f\beta_n = \beta_i.$$

It depends only on the datum of $G, B, T, \prec$, *but not on* g *nor* x. *We have*

$$U_{u\beta_1}(g_1)\cdots U_{u\beta_{n-1}}(g_{n-1})U_{-u\beta_n}(g_n^{-1}) \in U_{-u\beta_n}(g_n^{-1})U_{-u\beta_{n+1}}(L)\cdots U_{-u\beta_{\#\Phi^+}}(L)U_{u\beta_1}(\tilde{g}_1)\cdots U_{u\beta_{n-1}}(\tilde{g}_{n-1}).$$

Proof. The statement in (a) is immediately verified, since $x^{-1}(u\beta_n, \nu_L(g_n)) \in \Phi_{\mathrm{af}}^+$ is equivalent to ${}^{x^{-1}}U_{u\beta_n}(g_n) \in I$. So let us prove (b).

Using the fact $x^{-1}(-b) \in \Phi_{\mathrm{af}}^+$, we get

$$U_{u\beta_n}(g_n)xI = U_{u\beta_n}(g_n)U_{-u\beta_n}(-g_n^{-1})xI.$$

Following the usual combinatorics of root subgroups, e.g., [Springer 1998, Lemma 8.1.4] or [Parkinson et al. 2009, equation (7.6)], we rewrite this as

$$\begin{aligned}\cdots &= U_{-u\beta_n}(g_n^{-1})U_{-u\beta_n}(-g_n^{-1})U_{u\beta_n}(g_n)U_{-u\beta_n}(-g_n^{-1})xI\\ &= U_{-u\beta_n}(g_n^{-1})(-u\beta_n)^\vee(-g_n^{-1})n_{-u\beta_n}xI\\ &= U_{-u\beta_n}(g_n^{-1})r_b xI.\end{aligned}$$

Here, the cocharacter $(-u\beta_n)^\vee$ is understood as function $L \to T(L)$ and $n_{-u\beta_n} \in N_G(T)(L)$ is a representative of the reflection $s_{-u\beta_n} \in W$.

It remains to evaluate

$$g' := U_{u\beta_1}(g_1)\cdots U_{u\beta_{n-1}}(g_{n-1})U_{-u\beta_n}(g_n^{-1}) \in {}^{u'}U(L),$$

where we write $u' = us_{\beta_n}\cdots s_{\beta_{\#\Phi^+}} \in W$. By [Springer 1998, Proposition 8.2.3], we may write

$$U_{u\beta_{n-1}}(g_{n-1})U_{-u\beta_n}(g_n^{-1}) = U_{-u\beta_n}(g_n^{-1})\Big[\prod_{i,j} U_{\beta_{i,j}}(c_{i,j}g_n^{-i}g_{n-1}^j)\Big]U_{u\beta_{n-1}}(g_{n-1}),$$

where the product is taken over all indices $i, j \in \mathbb{Z}_{\geq 1}$ with $\beta_{i,j} := -iu\beta_n + ju\beta_{n-1} \in \Phi$. The product can be evaluated in any fixed order, up to changing the structure constants $c_{i,j}$. By the construction of the

Bruhat–Tits building, the structure constants are in $\mathbb{Z}$; see [Bruhat and Tits 1972, Example 6.1.3(b)] or [Springer 1998, Chapter 9].

We want to iterate this procedure. We claim for all $1 \le j \le n$ that we can write

$$U_{u\beta_j}(g_j)\cdots U_{u\beta_{n-1}}(g_{n-1})U_{-u\beta_n}(g_n^{-1}) = U_{-u\beta_n}(g_n^{(j)})\cdots U_{-u\beta_{\#\Phi^+}}(g_{\#\Phi^+}^{(j)})U_{u\beta_1}(g_1^{(j)})\cdots U_{u\beta_{n-1}}(g_{n-1}^{(j)}) \quad (*)$$

subject to the conditions

$$\begin{aligned} g_n^{(j)} &= g_n^{-1}, \\ g_i^{(j)} &= g_i + f_i^{(j)}(g_{i+1},\dots,g_{n-1},g_n^{-1}) && \text{for } j \le i < n, \\ g_i^{(j)} &= f_i^{(j)}(g_{j+1},\dots,g_{n-1},g_n^{-1}) && \text{for } 1 \le i < j \text{ or } n < i \le \#\Phi^+. \end{aligned}$$

Here, the polynomials $f_i^{(j)}$ are required to have the analogous properties as claimed in the lemma, i.e., the monomial $\varphi X_{j+1}^{e_{j+1}}\cdots X_{n-1}^{e_{n-1}}Y^f$ may only occur $f_i^{(j)}$ if

$$e_{j+1}\beta_{j+1}+\cdots+e_{n-1}\beta_{n-1}-f\beta_n = \begin{cases}\beta_i, & i<n,\\ -\beta_i, & i>n.\end{cases}$$

This long claim is trivially verified for $j = n$. In an inductive step, assume it has been proved for some $1 < j \le n$. We multiply the right-hand side of $(*)$ by $U_{u\beta_{j-1}}(g_{j-1})$ and apply [Springer 1998, Proposition 8.2.3] to sort the resulting product into our usual order. By Lemma 3.2, the result indeed lies in $U_{-u\beta_n}(L)\cdots U_{-u\beta_{\#\Phi^+}}(L)U_{u\beta_1}(L)\cdots U_{u\beta_{n-1}}(L)$, so this defines the elements $g_i^{(j-1)} \in L$ for $i = 1,\dots,\#\Phi^+$.

It is straightforward to see (but cumbersome to write down in full detail) that our required conditions for the $g_i^{(j-1)}$ are true precisely because they are true for the $g_i^{(j)}$. This finishes the induction. Specializing to $j = 1$ proves the lemma. □

We want to iterate this lemma. Doing so, we obtain the following result.

Proposition 3.5. *Let $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ and $u \in W$. Pick $gI \in {}^uU(L)I/I$. For each $n = 0,\dots,\#\Phi^+$, consider the Borel subgroup of G associated with the element $us_{\beta_{n+1}}\cdots s_{\beta_{\#\Phi^+}} \in W$ and the corresponding decomposition of the affine flag variety into semi-infinite orbits. This allows us to define $x_0,\dots,x_{\#\Phi^+} = 1 \in \widetilde{W}$ to be the uniquely determined elements such that*

$$gI \in ({}^{us_{\beta_{n+1}}\cdots s_{\beta_{\#\Phi^+}}}U(L))x_nI/I, \quad n = 0,\dots,\#\Phi^+.$$

Define

$$\{n_1 < \cdots < n_N\} := \bigl\{n \in \{1,\dots,\#\Phi^+\} \mid x_n \ne x_{n-1}\bigr\}.$$

Choose a representative of gI in ${}^uU(L)$ and write

$$gI = U_{u\beta_1}(g_1)\cdots U_{u\beta_{\#\Phi^+}}(g_{\#\Phi^+})I, \quad g_1,\dots,g_{\#\Phi^+} \in L.$$

Then, for $n = 0,\dots,\#\Phi^+$, we have the following:

(1) *We may write*

$$gI \in U_{-u\beta_{n+1}}(L)\cdots U_{-u\beta_{\#\Phi^+}}(L)U_{u\beta_1}(g_1^{(n)})\cdots U_{u\beta_n}(g_n^{(n)})x_n I$$

for elements $g_1^{(n)},\dots,g_n^{(n)} \in L$, *which are determined uniquely through polynomial identities*

$$g_i^{(n)} - g_i = f_i^{(n)}(g_{i+1}^{(i+1)},\dots,g_{\#\Phi^+}^{(\#\Phi^+)}),\qquad f_i^{(n)} \in \mathbb{Z}[X_{i+1}^{\pm 1},\dots,X_{\#\Phi^+}^{\pm 1}],$$

subject to the following condition: The polynomial $f_i^{(n)}$ *depends only on the datum of* $G, B, T, u, \prec$ *and the indices in* $\{n_1,\dots,n_N\}\cap\{n+1,\dots,\#\Phi^+\}$. *It is a sum of monomials*

$$\varphi X_{i+1}^{e_{i+1}}\cdots X_{\#\Phi^+}^{e_{\#\Phi^+}},\qquad \varphi, e_{i+1},\dots,e_{\#\Phi^+}\in\mathbb{Z},$$

subject to the conditions $\beta_i = e_{i+1}\beta_{i+1}+\cdots+e_{\#\Phi^+}\beta_{\#\Phi^+}$ *and*

$$\forall h \in \{i+1,\dots,\#\Phi^+\},\quad e_h < 0 \implies h \in \{n+1,\dots,\#\Phi^+\}\cap\{n_1,\dots,n_N\}.$$

(2) *Suppose that* $n \geq 1$. *If* $g_n^{(n)} = 0$, *then* $n \notin \{n_1,\dots,n_N\}$ *and* $x_{n-1} = x_n$ *Otherwise, define*

$$b_n := (u\beta_n, \nu_L(g_n^{(n)})) \in \Phi_{\mathrm{af}}.$$

If $x_n^{-1}(b_n) \in \Phi_{\mathrm{af}}^+$, *then* $n \notin \{n_1,\dots,n_N\}$ *and* $x_{n-1} = x_n$.
If $x_n^{-1}(b_n) \in \Phi_{\mathrm{af}}^-$, *then* $n \in \{n_1,\dots,n_N\}$ *and* $x_{n-1} = r_b x_n$.

(3) *The values of* $\nu_L(g_{n_h}^{(n_h)}) \in \mathbb{Z}$ *for* $h \in \{1,\dots,N\}$ *depend only on* $gI \in G(L)/I$, *and not on the chosen representative* $U_{u\beta_1}(g_1)\cdots U_{u\beta_{\#\Phi^+}}(g_{\#\Phi^+}) \in G(L)$.

Proof. By Lemma 3.2, we get

$$s_{\beta_{n+1}}\cdots s_{\beta_{\#\Phi^+}}\Phi^+ = \{-\beta_{n+1},\dots,-\beta_{\#\Phi^+},\beta_1,\dots,\beta_n\}.$$

So indeed

$${}^{us_{\beta_{n+1}}\cdots s_{\beta_{\#\Phi^+}}}U(L) = U_{-u\beta_{n+1}}(L)\cdots U_{-u\beta_{\#\Phi^+}}(L)U_{u\beta_1}(L)\cdots U_{u\beta_n}(L),$$

as claimed indirectly in (a).

We explain how to find the elements $g_i^{(n)}$ via induction on $\#\Phi^+ - n$, proving (b) along the way. For the inductive start, note that we have to choose $f_\bullet^{(\#\Phi^+)} \equiv 0$, so that $g_i^{(\#\Phi^+)} = g_i$.

For the inductive step, suppose now that we have constructed the elements $g_1^{(n)},\dots,g_n^{(n)}$ for some $n \in \{1,\dots,\#\Phi^+\}$. Define b_n as in (b). If $(x_n)^{-1}b_n \in \Phi_{\mathrm{af}}^+$, we may apply Lemma 3.4(a) to

$$gI \in U_{-u\beta_{n+1}}(L)\cdots U_{-u\beta_{\#\Phi^+}}(L)U_{u\beta_1}(g_1^{(n)})\cdots U_{u\beta_n}(g_n^{(n)})x_n I.$$

By the choice of x_{n-1}, we get $x_{n-1} = x_n$. We set $g_i^{(n-1)} = g_i^{(n)}$ for $i = 1,\dots,n-1$.

If $(x_n)^{-1}b_n \in \Phi_{\mathrm{af}}^-$, we may apply Lemma 3.4(b) to see

$$gI \in U_{-u\beta_n}(L)\cdots U_{-u\beta_{\#\Phi^+}}(L)U_{u\beta_1}(\tilde g_1^{(n)})\cdots U_{u\beta_{n-1}}(\tilde g_{n-1}^{(n)})r_b x_n I.$$

By choice of x_{n-1}, we get $x_{n-1} = r_b x_n$. In particular $n \in \{n_1,\dots,n_N\}$. The elements $\tilde g_\bullet^{(n)}$ are polynomials in the $g_\bullet^{(n)}$ as in Lemma 3.4(b). We set $g_i^{(n-1)} := \tilde g_i^{(n)}$.

Observe that in any case, the value of $g_i^{(n-1)} - g_i^{(n)}$ is a polynomial with integer coefficients in $g_{i+1}^{(n)}, \dots, g_{n-1}^{(n)}, (g_{n-1}^{(n)})^{-1}$ subject to the conditions of Lemma 3.4(b). Using a simple induction on $\#\Phi^+ - i$, one can now see that $g_i^{(n)} - g_i$ has the desired shape as claimed in (a), by composition of these polynomials. This finishes the proof of (a).

If $n \in \{n_1, \dots, n_N\}$, the value of b_n is uniquely determined by $x_n x_{n+1}^{-1}$, which in turn is determined by $gI \in G(L)/I$ alone. Hence (c) follows. □

Corollary 3.6. *In the setting of Proposition 3.5, we have $gI = I$ if and only if $U_{u\beta_n}(g_n) \in I$ for $n = 1, \dots, \#\Phi^+$.*

Proof. If each $U_{u\beta_n}(g_n)$ lies in I, then so does their product; hence $gI = I$.

If conversely $gI = I$, then all $x_n \in \widetilde{W}$ must be equal to 1. Now part (b) of Proposition 3.5 shows that each g_n must be zero or satisfy $x_n^{-1}(u\beta_n, \nu_L(g_n)) \in \Phi_{\mathrm{af}}^+$. Since $x_n = 1$, the latter condition is equivalent to $U_{u\beta_n}(g_n) \in I$. □

Definition 3.7. Let $\prec, x_0, u, gI$ be as in Proposition 3.5. We define the *semi-infinite type* of gI to be the set

$$\mathrm{type}(gI) = \{(n_h, \nu_L(g_{n_h}^{(n_h)})) \mid h = 1, \dots, N\} \subset \mathbb{Z} \times \mathbb{Z}.$$

Any subset of $\mathbb{Z} \times \mathbb{Z}$ of the above form is called an *admissible type* for $(x_0, u, \prec)$.

Lemma 3.8. *Let $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ be a reflection order and $u \in W$. Choose an arbitrary subset $\{n_1 < \cdots < n_N\} \subseteq \{1, \dots, \#\Phi^+\}$ and values $\nu_h \in \mathbb{Z}$ for $h = 1, \dots, N$. Define $b_h := (u\beta_{n_h}, \nu_h) \in \Phi_{\mathrm{af}}$.*

(1) *The set $\{(n_1, \nu_1), \dots, (n_N, \nu_N)\}$ defines an admissible type for $(r_{b_1} \cdots r_{b_N}, u, \prec)$ if and only if*

$$r_{b_N} \cdots r_{b_{h+1}}(b_h) \in \Phi_{\mathrm{af}}^-$$

for $h = 1, \dots, N$.

(2) *There is a locally closed and reduced k-sub-ind-scheme*

$$\mathcal{T} = \mathcal{T}_{u,\prec,(n_1,\nu_1),\dots,(n_N,\nu_N)}$$

of the affine flag variety whose k-valued points are given by precisely those elements

$$gI \in U_{u\beta_1}(L) \cdots U_{u\beta_{\#\Phi^+}}(L)I/I \subset G(L)/I$$

which satisfy $\mathrm{type}(gI) = \{(n_1, \nu_1), \dots, (n_N, \nu_N)\}$.

Proof. (a) The given condition for the b_h is certainly necessary by Proposition 3.5. Conversely, if this condition is satisfied, one may iteratively choose values for g_i in Proposition 3.5 to construct an element $gI \in G(L)/I$ of the desired type.

(b) The definition of $\mathrm{type}(gI)$ in terms of L-valuations allows us to write $gI \in T(k)$ in terms of vanishing or nonvanishing of certain polynomials over k. Hence we get a well-defined reduced subscheme with these geometric points. □

We call $\mathcal{T}_{u,\prec,(n_1,\nu_1),\dots,(n_N,\nu_N)}$ a *type variety*. These type varieties are analogues of the Gelfand–Goresky–MacPherson–Serganova strata in the affine Grassmannian; see [Kamnitzer 2010]. We need the following numerical datum to describe their dimensions. This can be seen as finite a replacement for the infinite dimension of ${}^uU(L)xI$.

Definition 3.9. Let $x = wt^\mu \in \widetilde{W},\ u \in W$ and $\alpha \in \Phi$.

(a) We define the *length functional*, following [Schremmer 2022, Definition 2.5], as

$$\ell(x,\alpha) = \langle \mu, \alpha\rangle + \Phi^+(\alpha) - \Phi^+(w\alpha).$$

(b) We define

$$\ell_u(x) := \sum_{\alpha\in\Phi^+} \ell(x^{-1}, u\alpha) = \langle -u^{-1}w\mu, 2\rho\rangle - \ell(u) + \ell(w^{-1}u).$$

The claimed identity can easily be seen along the lines of [loc. cit., Corollary 2.11]. This result moreover proves that $\ell_u(x) \le \ell(x)$, with equality holding if and only if $u \in \mathrm{LP}(x^{-1})$. From [loc. cit., Lemma 2.9], we see

$$\ell_u(x) = \dim\bigl((I \cap {}^uU(L))xI/I\bigr) - \dim\bigl((I \cap {}^{uw_0}U(L))xI/I\bigr).$$

Proposition 3.10. *Let $\tau = \{(n_1,\nu_1),\dots,(n_N,\nu_N)\}$ be an admissible type for $(x,u,\prec)$. Then $\mathcal{T} = \mathcal{T}_{u,\prec,(n_1,\nu_1),\dots,(n_N,\nu_N)}$ is a finite-dimensional irreducible smooth affine scheme over k. We have*

$$\dim \mathcal{T} = \tfrac{1}{2}(N - \ell_u(x)).$$

Proof. First consider the case $N=0$. Then evidently $\mathcal{T}$ is just a point over k, given by $\mathcal{T}(k) = \{I\} \subset G(L)/I$.

Suppose now $N \ge 1$. We prove the claim via induction on n_N (with the inductive start being the case n_N undefined, i.e., $N = 0$ above). For $m \in \mathbb{Z}$, we define the *truncation map* $\mathrm{tr}_{\le m} : L \to L$ as

$$\mathrm{tr}_{\le m}\Bigl(\sum_{i\in\mathbb{Z}} a_i t^i\Bigr) = \sum_{i\le m} a_i t^i.$$

Denote its image by $L_{\le m}$, which is easily equipped with the structure of an k-ind-scheme. We define the map of k-ind-schemes

$$f_1 : \mathcal{T} \to L_{\le -\Phi^+(u\beta_1)},\quad U_{u\beta_1}(g_1)\cdots U_{u\beta_{\#\Phi^+}}(g_{\#\Phi^+})I \in \mathcal{T}(k) \mapsto \mathrm{tr}_{\le -\Phi^+(u\beta_1)}(g_1).$$

In order to check that this is well-defined, suppose that

$$U_{u\beta_1}(g_1)\cdots U_{u\beta_{\#\Phi^+}}(g_{\#\Phi^+})I = U_{u\beta_1}(\tilde g_1)\cdots U_{u\beta_{\#\Phi^+}}(\tilde g_{\#\Phi^+})I \in \mathcal{T}(k)$$

for some $g_\bullet, \tilde g_\bullet \in L$. Then

$$[U_{u\beta_1}(g_1)\cdots U_{u\beta_{\#\Phi^+}}(g_{\#\Phi^+})]^{-1}[U_{u\beta_1}(\tilde g_1)\cdots U_{u\beta_{\#\Phi^+}}(\tilde g_{\#\Phi^+})] \in {}^uU(L)\cap I.$$

By Corollary 3.6 and the reflection order property, we conclude $U_{u\beta_1}(\tilde g_1 - g_1) \in I$. Hence $\mathrm{tr}_{\le -\Phi^+(u\beta_1)}(g_1) = \mathrm{tr}_{\le -\Phi^+(u\beta_1)}(\tilde g_1)$. This shows well-definedness of the map f_1.

Define the reflection order $\prec' = \prec^{\beta_1}$ as in [Björner and Brenti 2005, Proposition 5.2.3], so

$$s_{\beta_1}(\beta_2) \prec' s_{\beta_1}(\beta_3) \prec' \cdots \prec' s_{\beta_1}(\beta_{\#\Phi^+}) \prec' \beta_1.$$

Define moreover the type

$$\tau' = \{(n_i - 1, \nu_i) \mid i \in \{1, \dots, N\} \text{ and } n_i > 1\}.$$

Write $\mathcal{T}' = \mathcal{T}_{us_{\beta_1}, \prec', \tau'}$. Then the inductive assumption applies to $\mathcal{T}'$. For all $U_{u\beta_1}(g_1) \cdots U_{u\beta_{\#\Phi^+}}(g_{\#\Phi^+})I \in \mathcal{T}(k)$, one easily checks $U_{u\beta_2}(g_2) \cdots U_{u\beta_{\#\Phi^+}}(g_{\#\Phi^+})I \in \mathcal{T}'(k)$.

Observe that $U_{u\beta_1}(L)$ normalizes $U_{u\beta_2}(L) \cdots U_{u\beta_{\#\Phi^+}}(L)$. By the definition of the variety $\mathcal{T}'$, we see that $(U_{u\beta_1}(L) \cap I)\mathcal{T}'(k) = \mathcal{T}'(k)$. Hence we obtain a well-defined map of k-ind-schemes $f_2 : \mathcal{T} \times (U_{u\beta_1}(L) \cap I) \to \mathcal{T}'$ sending $gI \in \mathcal{T}(k)$ and $U_{u\beta_1}(h) \in I$ to

$$U_{u\beta_1}(h - f_1(gI))gI \in \mathcal{T}'(k).$$

Define

$$x' = w't^{\mu'} = \begin{cases} x, & n_1 > 1, \\ r_{b_1}x, & n_1 = 1, \end{cases}$$

such that τ' is admissible for $(x', us_{\beta_1}, \prec')$. Thus $\mathcal{T}'(k) \subset {}^{us_{\beta_1}w_0}U(L)x'I$, allowing us to write elements $g'I \in \mathcal{T}'(k)$ in the form

$$g'I = U_{-u\beta_2}(g'_1) \cdots U_{-u\beta_{\#\Phi^+}}(g'_{\#\Phi^+-1})U_{u\beta_1}(g'_{\#\Phi^+})x'I.$$

Here, we have

$$\begin{aligned} U_{u\beta_1}(g'_{\#\Phi^+}) \in {}^{x'}I \quad &\iff \quad (x')^{-1}(u\beta_1, \nu_L(g'_{\#\Phi^+})) \in \Phi^+_{\mathrm{af}} \\ &\iff \quad \nu_L(g'_{\#\Phi^+}) \geq \langle -w'\mu', u\beta_1\rangle + \Phi^+(-(w')^{-1}u\beta_1) =: m + 1 \in \mathbb{Z}. \end{aligned}$$

Thus we obtain a well-defined morphism of k-ind-schemes $\varphi_1 : \mathcal{T}' \to L_{\leq m}$ sending $g'I \in \mathcal{T}'(k)$ as represented above to $\mathrm{tr}_{\leq m}(g'_{\#\Phi^+})$ (check well-definedness using Corollary 3.6 as above).

Let $S \subset L$ be the k-sub-ind-scheme defined by the following condition for $z \in L$:

$$z \in S(k) \quad :\iff \quad \begin{cases} \nu_L(z) \geq m + 1 & \text{if } n_1 > 1, \\ \nu_L(z) = \nu_1 & \text{if } n_1 = 1. \end{cases}$$

We would like to define the map of k-ind-schemes $\varphi_2 : \mathcal{T}' \times S \to \mathcal{T}$ sending $g'I \in \mathcal{T}'(k)$ and $z \in S(k)$ to

$$\varphi_2(g'I, z) = U_{u\beta_1}(z - \varphi_1(g'I))g'I.$$

Let us check that φ_2 is well-defined, i.e., takes values in $\mathcal{T}$ as claimed. For $i = 2, \dots, \#\Phi^+$, we have

$$\begin{aligned} {}^{us_{\beta_i}\cdots s_{\beta_{\#\Phi^+}}}U(L)\varphi_2(g'I, z) &= {}^{us_{\beta_i}\cdots s_{\beta_{\#\Phi^+}}}U(L)g'I \\ &= {}^{(us_{\beta_1})s_{s_{\beta_1}(\beta_i)}\cdots s_{s_{\beta_1}(\beta_{\#\Phi^+})}s_{\beta_1}}U(L)g'I. \end{aligned}$$

Moreover, computing

$$\varphi_2(g'I, z) \in U_{-u\beta_2}(L) \cdots U_{-u\beta_{\#\Phi^+}}(L)U_{u\beta_1}(z)x'I,$$

we can apply Lemma 3.4 to get $\varphi_2(g'I, z) \in {}^{uw_0}U(L)xI$ by the condition $z \in S(k)$. Comparing the definitions of τ and τ', we get $\varphi_2(g'I, z) \in \mathcal{T}(k)$.

For a sufficiently large integer $M \gg 0$, one checks that we have an isomorphism of k-ind-schemes

$$\mathcal{T} \times (U_{u\beta_1}(L_{\leq M}) \cap I) \to \mathcal{T}' \times (S \cap L_{\leq M})$$

sending $gI \in \mathcal{T}(k)$ and $U_{u\beta_1}(h) \in U_{u\beta_1}(L_{\leq M}) \cap I$ to $f_2(gI, h) \in \mathcal{T}'(k)$ and

$$-h + \varphi_1(f_2(gI, h)) + f_1(gI) \in S(k) \cap L_{\leq M}.$$

Its inverse is the map sending $g'I \in \mathcal{T}'(k)$ and $z \in S(k) \cap L_{\leq M}$ to $\varphi_2(g'I, z) \in \mathcal{T}(k)$ and

$$U_{u\beta_1}\big(-z + f_1(\varphi_2(g'I, z)) + \varphi_1(g'I)\big) \in U_{u\beta_1}(L_{\leq M}) \cap I.$$

By the inductive assumption, $\mathcal{T}'$ is a finite-dimensional irreducible smooth affine scheme over k. The same conditions hold true for $S \cap L_{\leq M}$ (which is either an affine space over k or the product of a pointed affine line with an affine space). Hence the same conditions all hold true for $\mathcal{T} \times (U_{u\beta_1}(L_{\leq M}) \cap I)$. It follows that they must also hold true for $\mathcal{T}$ itself. Moreover, we have

$$\begin{aligned}\dim \mathcal{T}' - \dim \mathcal{T} &= \dim(U_{u\beta_1}(L_{\leq M}) \cap I) - \dim(S \cap L_{\leq M}) \\ &= \begin{cases} m + 1 - \Phi^+(-u\beta_1) & \text{if } n_1 > 1, \\ \nu_1 - \Phi^+(-u\beta_1) & \text{if } n_1 = 1. \end{cases}\end{aligned}$$

By induction, we know $2\dim \mathcal{T}' = N' - \ell_{us_{\beta_1}}(x')$, with $N' = N$ if $n_1 > 1$ and $N' = N - 1$ if $n_1 = 1$. We show $2\dim \mathcal{T} = N - \ell_u(x)$, using a case distinction depending on whether $n_1 > 1$ or not.

First consider the case $n_1 = 1$. From [Schremmer 2022, Lemma 2.12] or direct calculation, we get

$$\ell_{us_{\beta_1}}(x') = \ell_{us_{\beta_1}}(r_{b_1}x) = \ell_{us_{\beta_1}}(r_{b_1}) + \ell_u(x).$$

We calculate

$$\ell_{us_{\beta_1}}(r_{b_1}) = \ell_{us_{\beta_1}}(s_{u\beta_1}t^{\nu_1 u\beta_1^\vee}) = \langle -\nu_1\beta_1^\vee, 2\rho\rangle - \ell(us_{\beta_1}) + \ell(u).$$

Since β_1 is simple, the above expression simplifies to

$$\cdots = -2\nu_1 - 1 + 2\Phi^+(-u\beta_1).$$

We conclude

$$\begin{aligned}\dim \mathcal{T} &= \dim \mathcal{T}' - \nu_1 + \Phi^+(-u\beta_1) = \tfrac{1}{2}(N' - \ell_{us_{\beta_1}}(x')) - \nu_1 + \Phi^+(-u\beta_1) \\ &= \tfrac{1}{2}\big(N - \ell_u(x) + 2\nu_1 - 2\Phi^+(-u\beta_1)\big) - \nu_1 + \Phi^+(-u\beta_1) = \tfrac{1}{2}(N - \ell_u(x)).\end{aligned}$$

Let us now consider the case $n_1 > 1$. Then we calculate

$$\begin{aligned}\ell_{us_{\beta_1}}(x) &= \sum_{\alpha \in \Phi^+} \ell(x^{-1}, us_{\beta_1}\alpha) = \sum_{\alpha \in s_{\beta_1}\Phi^+} \ell(x^{-1}, u\alpha) \\ &= \ell_u(x) - \ell(x^{-1}, u\beta_1) + \ell(x^{-1}, -u\beta_1) = \ell_u(x) - 2\ell(x^{-1}, u\beta_1).\end{aligned}$$

We compute

$$\ell(x^{-1}, u\beta_1) = \ell(w^{-1}t^{-w\mu}, u\beta_1) = \langle -w\mu, u\beta_1\rangle + \Phi^+(u\beta_1) - \Phi^+(w^{-1}u\beta_1) = m + \Phi^+(u\beta_1).$$

The claimed dimension formula for $\mathcal{T}$ follows just as above:

$$\begin{aligned}\dim \mathcal{T} &= \dim \mathcal{T}' - m - \Phi^+(u\beta_1) = \tfrac{1}{2}(N - \ell_{us_{\beta_1}}(x)) - m - \Phi^+(u\beta_1)\\ &= \tfrac{1}{2}(N - \ell_u(x) + 2\ell(x^{-1}, u\beta_1)) - m - \Phi^+(u\beta_1) = \tfrac{1}{2}(N - \ell_u(x)).\end{aligned}$$

This finishes the induction and the proof. □

Lemma 3.11. *Let $(n_1, \nu_1), \dots, (n_N, \nu_N)$ be an admissible type for $(y, u, \prec)$. Define $y = x_1, \dots, x_{\#\Phi^+ +1} = 1 \in \widetilde{W}$ as in Proposition 3.5, i.e.,*

$$x_{n-1} = \begin{cases} x_n, & n \notin \{n_1, \dots, n_N\},\\ r_{(u\beta_n, \nu_h)} x_n, & n = n_h \in \{n_1, \dots, n_N\}.\end{cases}$$

Define integers $m_1, \dots, m_{\#\Phi^+} \in \mathbb{Z}$ as follows: If $n = n_h \in \{n_1, \dots, n_N\}$, we put $m_n = \nu_h$. Otherwise, we let $m_n \in \mathbb{Z}$ be the smallest value such that

$$x_n^{-1}(u\beta_n, m_n) \in \Phi^+_{\mathrm{af}}.$$

Let $M > 0$. There is a reduced and finite-dimensional k-subscheme $\widetilde{T}_M \subset L_{\le M}^{\#\Phi^+}$ whose k-valued points are

$$\begin{aligned}\widetilde{T}_M(k) = \big\{(g_1, \dots, g_{\#\Phi^+}) \in L_{\le M}^{\#\Phi^+} \text{ such that } &\nu_L(g_n + f_n^{(n+1)}(g_{i+1}, \dots, g_{\#\Phi^+})) \ge m_n \ \forall n \in \{1, \dots, \#\Phi^+\},\\ &\nu_L(g_n + f_n^{(n+1)}(g_{n+1}, \dots, g_{\#\Phi^+})) = m_n \ \forall n \in \{n_1, \dots, n_N\}\big\}.\end{aligned}$$

If $M > m_i$ for all i, then

$$\dim \widetilde{T}_M = \sum_{n=1}^{\#\Phi^+} (M - m_n + 1).$$

We get a surjective map

$$\begin{aligned} f : \widetilde{T}_M &\to \big(U_{u\beta_1}(L_{\le M}) \cdots U_{u\beta_{\#\Phi^+}}(L_{\le M}) x I\big) \cap T_{(n_1,\nu_1),\dots,(n_N,\nu_N)},\\ (g_1, \dots, g_{\#\Phi^+}) &\mapsto U_{u\beta_1}(g_1) \cdots U_{u\beta_{\#\Phi^+}}(g_{\#\Phi^+}) x I.\end{aligned}$$

Proof. For each point $(g_1, \dots, g_{\#\Phi^+})$ in the desired set for $\widetilde{T}_M(k)$, the t-valuations of the individual coordinates $g_1, \dots, g_{\#\Phi^+}$ can be bounded from below in terms of the datum $\{(n_1, \nu_1), \dots, (n_N, \nu_N)\}$, by the defining properties of the polynomials $f_\bullet^{(\bullet)}$. So indeed this set can be identified with the k-valued points of a finite-dimensional k-scheme, which we denote by $\widetilde{T}_M$.

The dimension of $\widetilde{T}_M$ is can be determined as follows. For $g_{\#\Phi^+}$, the conditions simply state $g_{\#\Phi^+} \in L_{\le M}$ and $\nu_L(g_{\#\Phi^+}) \ge m_{\#\Phi^+}$ or $= m_{\#\Phi^+}$. In any case, the allowed values for $g_{\#\Phi^+}$ form an irreducible $(M - m_{\#\Phi^+} + 1)$-dimensional scheme over k.

We may continue like this. For any given $g_{n+1}, \dots, g_{\#\Phi^+} \in L$, the space of allowed values for g_n has the form

$$g_n \in L_{\le M} \quad \text{such that} \quad \nu_L(g_n - c) \ge m_n \text{ or } = m_n,$$

where c is a constant independent of g_n. We find a unique $\tilde{c} \in L_{\le m_n}$ such that $\nu_L(\tilde{c} - c) \ge m_n + 1$. Now the allowed g_n are precisely the elements of the form $g_n = -\tilde{c} + \tilde{g}_n$, where $\tilde{g}_n \in L_{\le M}$ has t-valuation $\ge m_n$

or $= m_n$. So as long as $m_n < M$, we see that the space of allowed g_n is irreducible of dimension $M - m_i + 1$. The dimension formula follows.

Well-definedness and surjectivity of f are immediate from Proposition 3.5. □

We note the following useful facts.

Lemma 3.12. *Let $x \in \widetilde{W}$ and $u \in W$.*

(a) *Let $z = w_z t^{\mu_z} \in \widetilde{W}$ such that, for all $\alpha \in \Phi^+$,*

$$\langle w_z, u\alpha \rangle \leq -2\ell(x).$$

Then $\ell_u(x) = \ell(zx) - \ell(z)$.

(b) *If $a = (\alpha, m) \in \Phi_{\mathrm{af}}$ satisfies $u^{-1}\alpha \in \Phi^+$ and $x^{-1}a \in \Phi_{\mathrm{af}}^-$, then $\ell_u(r_a x) < \ell_u(x)$.*

(c) *Suppose that $y = x_1, \ldots, x_{\#\Phi^+ + 1} = 1 \in \widetilde{W}$ and $m_1, \ldots, m_{\#\Phi^+}$ are as in Lemma 3.11. Then*

$$\sum_{n=1}^{\#\Phi^+} (\Phi^+(-u\beta_n) - m_n) = 2(\ell_{uw_0}(y) + N).$$

Proof. (1) This follows easily from the theory of length functionals from [Schremmer 2022]. Write $x = wt^\mu$ and observe $\mathrm{LP}(z^{-1}) = \mathrm{LP}(x^{-1}z^{-1}) = \{w_z u\}$. Then

$$\begin{aligned}\ell(zx) = \ell(x^{-1}z^{-1}) &= \sum_{\alpha \in \Phi^+} \ell(x^{-1}z^{-1}, w_z u x) \\ &= \sum_{\alpha \in \Phi^+} \ell(x^{-1}, u\alpha) + \ell(z^{-1}, w_z u x) = \ell_u(x) + \ell(z),\end{aligned}$$

where we used [loc. cit., Corollary 2.10 and Lemma 2.12].

(2) Let z be as in (a). Then $u^{-1}\alpha \in \Phi^+$ is equivalent to $\ell(z, \alpha) \ll 0$, so $za \in \Phi_{\mathrm{af}}^+$. This shows $\ell(zr_a x) < \ell(zx)$. We conclude by (a).

(3) We show for all $q \in \{1, \ldots, \#\Phi^+ + 1\}$ that

$$2 \sum_{n=q}^{\#\Phi^+} (\Phi^+(-u\beta_n) - m_n) = \ell_{u_q}(x_q) + \#\{n \in \{n_1, \ldots, n_N\} \mid n > q\},$$

where $u_q = us_{\beta_q} \cdots s_{\beta_{\#\Phi^+}} \in W$.

For $q = 1$, this is the desired statement. We do induction on $\#\Phi^+ + 1 - q$. In the case $q = \#\Phi^+ + 1$, both sides of the equation are easily seen to vanish.

In the inductive step, let us write $x_q = w_q t^{\mu_q}$. Then

$$m_q = \langle -w_q \mu_q, u\beta_q \rangle + \Phi^+(w_q^{-1} u\beta_q) \quad \Longrightarrow \quad \Phi^+(-u\beta_q) - m_q = \ell(x_q^{-1}, -u\beta_q).$$

Note that

$$u_q = s_{u\beta_q} u_{q+1} = u_{q+1} s_{s_{\beta_{\#\Phi^+}} \cdots s_{\beta_{q+1}}(\beta_q)},$$

with $s_{\beta_{\#\Phi^+}} \cdots s_{\beta_{q+1}}(\beta_q)$ being a simple root by Lemma 3.1. We conclude that

$$u_q \Phi^+ = \{u\beta_q\} \cup (u_{q+1}\Phi^+) \setminus \{-u\beta_q\}.$$

Consider the case $q+1 \notin \{n_1, \dots, n_N\}$. Then $x_q = x_{q+1}$, and

$$\begin{aligned}\ell_{u_q}(x_q) &= \ell_{u_{q+1}}(x_q) + \ell(x_q^{-1}, u\beta_q) - \ell(x_q^{-1}, -u\beta_q)\\ &= \ell_{u_{q+1}}(x_{q+1}) + 2\ell(x_q^{-1}, u\beta_q).\end{aligned}$$

Hence

$$\begin{aligned}2\sum_{n=q}^{\#\Phi^+}(\Phi^+(-u\beta_n) - m_n) &= 2\ell(x_q^{-1}, -u\beta_q) + 2\sum_{n=q+1}^{\#\Phi^+}(\Phi^+(-u\beta_n) - m_n)\\ &\underset{\text{ind.}}{=} 2\ell(x_q^{-1}, -u\beta_q) + \ell_{u_{q+1}}(x_{q+1}) + \#\{n \in \{n_1, \dots, n_N\} \mid n > q+1\}\\ &= \ell_{u_q}(x_q) + \#\{n \in \{n_1, \dots, n_N\} \mid n > q\}.\end{aligned}$$

Consider now the case $q = n_h - 1$ for some $h \in \{1, \dots, N\}$. Then $x_q = r_{b_{q+1}} x_{q+1} = r_{(u\beta_q, \nu_h)} x_{q+1}$. By [Schremmer 2022, Lemma 2.12], we calculate

$$\begin{aligned}\ell_{u_q}(x_q) = \ell_{u_q}(r_{(u\beta_q,\nu_h)x_q} x_{q+1}) &= \ell_{u_q}(r_{(u\beta_q,\nu_h)}) + \ell_{s_{u\beta_q}u_q}(x_{q+1})\\ &= \ell_{u_q}(s_{u\beta_q} t^{\nu_h u\beta_q^\vee}) + \ell_{u_{q+1}}(x_{q+1}).\end{aligned}$$

Here,

$$\begin{aligned}\ell_{u_q}(s_{u\beta_q} t^{\nu_h u\beta_q^\vee}) &= \nu_h \langle -u_q^{-1} s_{u\beta_q} u\beta_q^\vee, 2\rho\rangle - \ell(u_q) + \ell(u_{q+1})\\ &= \nu_h \langle u_q^{-1} u\beta_q^\vee, 2\rho\rangle - \ell(u_q) + \ell(u_q(u_q^{-1}u\beta_q))\\ &= 2(\Phi^+(-u\beta_q) - \nu_h) - 1,\end{aligned}$$

since $u_q^{-1}u\beta_q = -s_{\beta_{\#\Phi^+}} \cdots s_{\beta_{q+1}}(\beta_q)$ is the negative of a simple root.

We conclude

$$\begin{aligned}2\sum_{n=q}^{\#\Phi^+}(\Phi^+(-u\beta_n) - m_n) &= 2(\Phi^+(-u\beta_q) - \nu_h) + 2\sum_{n=q+1}^{\#\Phi^+}(\Phi^+(-u\beta_n) - m_n)\\ &\underset{\text{ind.}}{=} 2(\Phi^+(-u\beta_q) - \nu_h) + \ell_{u_{q+1}}(x_{q+1}) + \#\{n \in \{n_1, \dots, n_N\} \mid n > q+1\}\\ &= \ell_{u_{q+1}}(x_{q+1}) + \ell_{u_q}(s_{u\beta_q} t^{\nu_h u\beta_q^\vee}) + 1 + \#\{n \in \{n_1, \dots, n_N\} \mid n > q+1\}\\ &= \ell_{u_q}(x_q) + \#\{n \in \{n_1, \dots, n_N\} \mid n > q\}.\end{aligned}$$

This finishes the induction and the proof. □

We reformulate this proposition to describe arbitrary intersections of semi-infinite orbits.

Theorem 3.13. *Let $u, v \in W$ and $x = w_x t^{\mu_x}$, $y \in \widetilde{W}$. Pick a reflection order $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ and an index $n \in \{0, \dots, \#\Phi^+\}$ such that*

$$u^{-1}v = s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}}.$$

Then we get a decomposition into locally closed subsets

$$\big((^{u}U(L) \cap {}^{v}U(L))xI\big) \cap (^{uw_0}U(L)yI) = \bigsqcup_{\tau} x\mathcal{T}_\tau \subset G(L)/I,$$

where τ runs through all $\tau = \{(n_1, \nu_1), \dots, (n_N, \nu_N)\}$ which are admissible types for $(x^{-1}y, w_x^{-1}u, \prec)$ and satisfy the additional constraint $N = 0$ or $n_N \le n$. Each piece $x\mathcal{T}_\tau = x\mathcal{T}_{w_x^{-1}u,\prec,\tau} \subset G(L)/I$ is a *locally closed subset of the affine flag variety, and an irreducible smooth affine k-scheme of dimension*

$$\dim x\mathcal{T}_\tau = \tfrac{1}{2}(\ell_u(x) - \ell_u(y) + \#\tau).$$

Proof. For all $\beta \in \Phi^+$, we have

$$U_{u\beta}(L) \subseteq {}^vU(L) \quad\Longleftrightarrow\quad v^{-1}u\beta \in \Phi^+ \underset{\text{Lem. 3.2}}{\Longleftrightarrow} \beta \in \{\beta_1, \dots, \beta_n\}.$$

Hence

$$\begin{aligned} {}^uU(L) \cap {}^vU(L) &= U_{u\beta_1}(L) \cdots U_{u\beta_n}(L), \\ x^{-1}({}^uU(L) \cap {}^vU(L))x &= U_{w_x^{-1}u\beta_1}(L) \cdots U_{w_x^{-1}u\beta_n}(L). \end{aligned}$$

By Proposition 3.5, we obtain a decomposition of the corresponding subset of the affine flag variety

$$U_{w_x^{-1}u\beta_1}(L) \cdots U_{w_x^{-1}u\beta_n}(L)I \subset G(L)/I$$

into types $\{(n_1, \nu_1), \dots, (n_N, \nu_N)\}$ with $n_N \le n$. Denote by $\mathcal{T} = \mathcal{T}_{w_x^{-1}u,\prec,\tau}$ the subset associated with such a type τ as in Lemma 3.8. We have

$$\mathcal{T} \subset x^{-1uw_0}U(L)yI = {}^{w_x^{-1}uw_0}U(L)x^{-1}yI$$

if and only if τ is admissible for $(x^{-1}y, w_x^{-1}u, \prec)$. It remains to compute the dimension of $\mathcal{T}$ using Proposition 3.10 and [Schremmer 2022, Lemma 2.12]:

$$\begin{aligned} 2\dim\mathcal{T} &= N - \ell_{w_x^{-1}u}(x^{-1}y) \\ &= N - \ell_{w_x^{-1}u}(x^{-1}) - \ell_u(y) = N + \ell_u(x) - \ell_u(y). \end{aligned}$$ □

Remark 3.14. (a) It seems reasonable to expect that each type variety $\mathcal{T}_{u,\prec,\tau}$ should be a product of affine lines and pointed affine lines over k, but the proof of such a statement would probably require undue analysis of the polynomials $f_\bullet^{(\bullet)}$ or some major progress towards Zariski's cancellation problem. We don't need such a precise description.

(b) Given u, v, there are in general several possible reflection orders satisfying $v^{-1}u = s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}}$ for $n = \ell(w_0v^{-1}u)$. While the geometry of the intersection

$$\big(({}^uU(L) \cap {}^vU(L))xI\big) \cap ({}^{uw_0}U(L)yI)$$

does not depend on the choice of reflection order, the decomposition into subsets indexed by types tends to do that; i.e., different reflection orders yield different subsets. It is not clear how these are related, aside from the simple observation that subsets of maximal dimension parametrize irreducible components of maximal dimension. We will prove that the *number* of such subsets $x\mathcal{T}_\tau$ of any given dimension is independent of the chosen reflection order.

(c) For each given $(x, u, \prec)$, there exist only finitely many admissible types. This is straightforward to prove directly, and will immediately follow from a later result (see Lemma 4.4). So Theorem 3.13 provides a decomposition into finitely many locally closed pieces.

Example 3.15. Consider the group $G = \mathrm{GL}_3$. Let T be the torus of diagonal matrices and B be the upper triangular matrices. Let $u = v = w_0$ and

$$x = w_0 t^{\rho^\vee} = \begin{pmatrix} 0 & 0 & t^{-1} \\ 0 & 1 & 0 \\ t & 0 & 0 \end{pmatrix}.$$

We denote our simple roots by $\Delta = \{\alpha, \beta\}$ corresponding to the diagonal matrices $\dot{\alpha} = \mathrm{diag}(t, t^{-1}, 0)$ and $\dot{\beta} = (0, t, t^{-1})$. We choose the reflection order $\alpha \prec \alpha + \beta \prec \beta$. Then the admissible types for $(x, u, \prec)$ are given by the cardinality-1 type $\{(-\alpha - \beta, -1)\}$ as well as the cardinality-3 types

$$\{(-\beta, 0), (-\alpha - \beta, -1), (-\alpha, 0)\}, \quad \{(-\alpha, -1), (-\alpha - \beta, 0), (-\beta, -1)\}.$$

We see that the intersection

$${}^{w_0}U(L)I \cap U(L)xI \subset G(L)/I$$

has dimension

$$\tfrac{1}{2}(3 - \ell_{w_0}(x)) = \tfrac{1}{2}\big(3 - \langle -\rho^\vee, 2\rho\rangle + \ell(w_0)\big) = \tfrac{1}{2}(3 + 4 + 3) = 5,$$

and the number of 5-dimensional irreducible components is two.

4. Double Bruhat graph

There is a more convenient and natural way to encode the datum of an admissible type $(n_1, \nu_1), \dots, (n_N, \nu_N)$. This construction is due to [Naito and Watanabe 2017, Section 5.1], used originally to describe periodic R-polynomials of affine Weyl groups.

Definition 4.1. Let $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ be a reflection order and $v, w \in W$.

(a) The *double Bruhat graph* $\mathrm{DBG}(W)$ is a finite directed graph. Its set of vertices is W. For each $w \in W$ and $\alpha \in \Phi^+$, there is an edge $w \xrightarrow{\alpha} ws_\alpha$.

(b) A *unlabelled path* $\bar{p}$ in $\mathrm{DBG}(W)$ is a sequence of adjacent edges

$$\bar{p} : v = u_1 \xrightarrow{\alpha_1} u_2 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_\ell} u_{\ell+1} = w.$$

We call $\bar{p}$ an unlabelled path from v to w of length $\ell(\bar{p}) = \ell$. We say $\bar{p}$ is *increasing* with respect to $\prec$ if $\alpha_1 \prec \cdots \prec \alpha_\ell$. We say that $\bar{p}$ *is bounded by* $n \in \mathbb{Z}$ if each occurring root α_i has the form $\alpha_i = \beta_j$ for $j \le n$.

(c) A *labelled path* or *path* p in $\mathrm{DBG}(W)$ consists of an unlabelled path

$$\bar{p} : v = u_1 \xrightarrow{\alpha_1} u_2 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_\ell} u_{\ell+1} = w$$

together with integers $m_1, \dots, m_\ell \in \mathbb{Z}$ subject to the condition

$$m_i \ge \Phi^+(-u_i\alpha_i) = \begin{cases} 0, & \ell(u_{i+1}) > \ell(u_i), \\ 1, & \ell(u_{i+1}) < \ell(u_i). \end{cases}$$

We write p as

$$p : v = u_1 \xrightarrow{(\alpha_1, m_1)} u_2 \xrightarrow{(\alpha_2, m_2)} \cdots \xrightarrow{(\alpha_\ell, m_\ell)} u_{\ell+1} = w.$$

The *weight* of p is

$$\mathrm{wt}(p) = m_1\alpha_1^\vee + \cdots + m_\ell\alpha_\ell^\vee \in \mathbb{Z}\Phi^\vee.$$

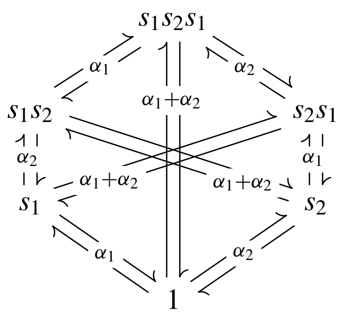
The *length* of p is $\ell(p) = \ell(\bar{p}) = \ell$. We say that p is *increasing* with respect to $\prec$ if $\bar{p}$ is. We say that p is bounded by $n \in \mathbb{Z}$ if $\bar{p}$ is.

(d) The set of all paths from v to w that are increasing with respect to $\prec$ and bounded by $n \in \mathbb{Z}$ is denoted by $\mathrm{paths}_{\leq n}^{\prec}(v \Rightarrow w)$. We also write

$$\mathrm{paths}^{\prec}(v \Rightarrow w) := \mathrm{paths}_{\leq \#\Phi^+}^{\prec}(v \Rightarrow w)$$

for the set of all increasing paths from v to w.

Example 4.2. This is the double Bruhat graph of type A_2, where we denote the simple roots by $\Delta = \{\alpha_1, \alpha_2\}$ and the corresponding simple reflections by $S = \{s_1, s_2\}$. For each root $\alpha \in \Phi^+ = \{\alpha_1, \alpha_2, \alpha_1+\alpha_2\}$ and each $w \in W$, there is an edge $w \to ws_\alpha$ with label α and the converse edge $ws_\alpha \to w$ with the same label, making each edge appear doubled (which explains the graph's name):



The two reflection orders are given by

$$\alpha_1 \prec \alpha_1 + \alpha_2 \prec \alpha_2,$$

$$\alpha_2 \prec \alpha_1 + \alpha_2 \prec \alpha_1.$$

Remark 4.3. One could similarly study paths in the double Bruhat graph bounded from below by a root β_m, or even paths such that $\alpha_1, \ldots, \alpha_{\ell(p)} \in \{\beta_m, \ldots, \beta_n\}$ for fixed indices $m, n \in \mathbb{Z}$. However, this extra generality would not yield any new information:

If $w_0 = s_{\alpha_1} \cdots s_{\alpha_{\#\Phi^+}}$ is the reduced expression corresponding to our reflection order and $0 \leq m \leq n \leq \#\Phi^+$, we could consider a different reflection order $\prec'$ coming from the reduced word

$$w_0 = s_{\alpha_{m+1}} \cdots s_{\alpha_{\#\Phi^+}} s_{-w_0\alpha_1} \cdots s_{-w_0\alpha_m}.$$

If we write $\Phi^+ = \{\beta_1' \prec' \cdots \prec' \beta_{\#\Phi^+}'\}$, then

$$\beta_i = s_{\alpha_1} \cdots s_{\alpha_m} \beta_{i-m}', \quad i = m+1, \ldots, n.$$

This would give a one-to-one correspondence of paths which are increasing for $\prec$ and where all roots lie *between* m and n and paths which are increasing for $\prec'$ and bounded by $n - m$.

The comparison between types and paths is given as follows.

Lemma 4.4. *Let $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ be a reflection order, $x = wt^\mu \in \widetilde{W}$ and $u \in W$. Let moreover $n \in \{0, \ldots, \#\Phi^+\}$.*

Let A be the set of all admissible types $\{(n_1, \nu_1), \ldots, (n_N, \nu_N)\}$ for $(x, u, \prec)$ satisfying the condition $n_1 < \cdots < n_N \leq n$. Define

$$P := \{p \in \mathrm{paths}^{\prec}_{\leq n}(w^{-1}u \Rightarrow u) \mid \mathrm{wt}(p) = u^{-1}w\mu\}.$$

Then we get a bijective map $A \xrightarrow{\sim} P$ sending a type $\tau = \{(n_1, \nu_1), \ldots, (n_N, \nu_N)\}$ of cardinality N to the length-N path

$$p : w^{-1}u \xrightarrow{(\beta_{n_1}, m_1')} \cdots \xrightarrow{(\beta_{n_N}, m_N')} u$$

in P, where $(\beta_h', m_h') := r_{b_N} \cdots r_{b_h}(b_h) \in \Phi_{\mathrm{af}}^+$ and $b_h = (u\beta_{n_h}, \nu_h) \in \Phi_{\mathrm{af}}$ for $h = 1, \ldots, N$.

Remark 4.5. We remark that if $\mu \notin \mathbb{Z}\Phi^\vee$ in the above lemma, then one trivially gets $P = \varnothing$ by the definition of the double Bruhat graph. Moreover, it is clear that such an x cannot be a product of affine reflections; hence trivially $A = \varnothing$ by the definition of admissible types. The statement is more interesting when $\mu \in \mathbb{Z}\Phi^\vee$, that is, when x lies in the nonextended affine Weyl group.

Proof of Lemma 4.4. For now, we fix arbitrary integers $1 \leq n_1 < \cdots < n_N \leq n$ and $\nu_1, \ldots, \nu_N \in \mathbb{Z}$. Define the set

$$\tau = \{(n_1, \nu_1), \ldots, (n_N, \nu_N)\} \subset \mathbb{Z} \times \mathbb{Z}.$$

For $h \in \{1, \ldots, N\}$, define $b_h := (u\beta_{n_h}, \nu_h) \in \Phi_{\mathrm{af}}$ and

$$b_h' = (\beta_h', m_h') := r_{b_N} \cdots r_{b_h}(b_h) \in \Phi_{\mathrm{af}}.$$

Finally, we write down something that may or may not be a path in $\mathrm{DBG}(W)$ as

$$p : us_{\beta_{n_N}} \cdots s_{\beta_{n_1}} \xrightarrow{(\beta_{n_1}, m_1')} \cdots \xrightarrow{(\beta_{n_N}, m_N')} u.$$

We want to show that $\tau \in A$ if and only if $p \in P$, since the desired bijection is immediate from this.

The key observation is that, for $h \in \{1, \ldots, N\}$,

$$\begin{aligned} r_{b_N} \cdots r_{b_{h+1}}(b_h) \in \Phi_{\mathrm{af}}^- \quad &\iff \quad r_{b_N} \cdots r_{b_h}(b_h) \in \Phi_{\mathrm{af}}^+ \\ &\iff \quad \beta_h' \in \Phi_{\mathrm{af}}^+ \\ &\iff \quad (us_{\beta_{n_N}} \cdots s_{\beta_{n_h}}\beta_{n_h}, m_h') \in \Phi_{\mathrm{af}}^+ \\ &\iff \quad m_h' \geq \Phi^+(us_{\beta_{n_N}} \cdots s_{\beta_{n_h}}(\beta_{n_h})). \end{aligned}$$

Thus τ is an admissible type for some $(\tilde{x}, u, \prec)$ if and only if p is a well-defined path in the double Bruhat graph. In this case, it is clear that p is increasing with respect to $\prec$ and bounded by n.

We calculate

$$r_{b_1} \cdots r_{b_N} = r_{b_N'} \cdots r_{b_1'} = s_{\beta_N'} \cdots s_{\beta_1'} t^{m_N' s_{\beta_1'} \cdots s_{\beta_{N-1}'}(\beta_N')^\vee + \cdots + m_1'(\beta_1')^\vee}.$$

In the case

$$s_{u\beta_{n_1}} \cdots s_{u\beta_{n_N}} \neq w,$$

we see that τ cannot be admissible for $(x, u, \prec)$ and p does not start in $w^{-1}u$. So assume from now on that

$$s_{u\beta_{n_1}} \cdots s_{u\beta_{n_N}} = w.$$

Then we can simplify

$$r_{b_1} \cdots r_{b_N} = \cdots = wt^{m'_N s_{\beta'_1} \cdots s_{\beta'_{N-1}} (\beta'_N)^\vee + \cdots + m'_1 (\beta'_1)^\vee}.$$

We compute for $h = 1, \dots, N$ that

$$m'_h s_{\beta'_1} \cdots s_{\beta'_{h-1}} (\beta'_h) = m'_h (s_{\beta'_1} \cdots s_{\beta'_N})(s_{\beta'_N} \cdots s_{\beta'_h}) \beta'^\vee_h = m'_h w^{-1} s_{\beta'_N} \cdots s_{\beta'_h} \beta'^\vee_h = m'_h w^{-1} u \beta^\vee_h.$$

We conclude

$$r_{b_1} \cdots r_{b_N} = wt^{w^{-1}u \operatorname{wt}(p)}.$$

Hence τ is admissible for $(x, u, \prec)$ if and only if $\operatorname{wt}(p) = u^{-1} w \mu$. □

We can state the double Bruhat version of Theorem 3.13 as follows. It can be seen as an analogue of [Parkinson et al. 2009, Theorem 7.1]. Their result studies slightly different intersections in the affine flag variety and uses folded alcove walks rather than paths in the double Bruhat graph.

Theorem 4.6. *Let $u, v \in W$ and $x = w_x t^{\mu_x}$, $y = w_y t^{\mu_y} \in \widetilde{W}$. Pick a reflection order $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ and an index $n \in \{0, \dots, \#\Phi^+\}$ such that*

$$u^{-1} v = s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}}.$$

Then we get a decomposition into locally closed subsets

$$\big(({}^u U(L) \cap {}^v U(L)) x I\big) \cap ({}^{uw_0} U(L) y I) = \bigsqcup_p x \mathcal{T}_p,$$

where p runs through all paths $p \in \operatorname{paths}^{\prec}_{\leq n}(w_y^{-1} u \Rightarrow w_x^{-1} u)$ such that

$$u \operatorname{wt}(p) = w_y \mu_y - w_x \mu_x.$$

Each variety $x\mathcal{T}_p \subseteq G(L)/I$ is an irreducible smooth affine k-scheme of dimension

$$\dim(x\mathcal{T}_p) = \dim \mathcal{T}_p = \tfrac{1}{2}(\ell_u(x) - \ell_u(y) + \ell(p)).$$ □

The aim of this section is to develop the basic properties of the double Bruhat graph. For the remainder of this section, fix a reflection order $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$. We first introduce a couple of immediate properties about paths in $\operatorname{paths}^{\prec}_{\leq \cdot}(\cdot \Rightarrow \cdot)$. Part (a) of the following lemma is an adaption of the duality antiautomorphism from [Lenart et al. 2015, Proposition 4.3].

Lemma 4.7. *Let $u, v \in W$ and $n \in \mathbb{Z}$.*

(a) *Denote by $\succ$ the reflection order obtained by reversing $\prec$. Then we have a bijection*

$$\operatorname{paths}^{\prec}_{\leq n}(u \Rightarrow v) \to \operatorname{paths}^{\succ}_{\geq n}(w_0 v \Rightarrow w_0 u),$$
$$\big(w_1 \xrightarrow{(\alpha_1, n_1)} \cdots \xrightarrow{(\alpha_\ell, n_\ell)} w_{\ell+1}\big) \mapsto \big(w_0 w_{\ell+1} \xrightarrow{(\alpha_\ell, n_\ell)} \cdots \xrightarrow{(\alpha_1, n_1)} w_0 w_1\big),$$

that preserves both the weight and the length of each path.

(b) *Let $\prec'$ be the reflection order defined by*

$$\alpha \prec' \beta \quad \iff \quad -w_0\alpha \prec -w_0\beta.$$

Then we have a bijection

$$\begin{gathered}\operatorname{paths}^{\prec}_{\preceq n}(u \Rightarrow v) \to \operatorname{paths}^{\prec'}_{\preceq' n}(w_0 u w_0 \Rightarrow w_0 v w_0),\\ \left(w_1 \xrightarrow{(\alpha_1,n_1)} \cdots \xrightarrow{(\alpha_\ell,n_\ell)} w_{\ell+1}\right) \mapsto \left(w_0 w_1 w_0 \xrightarrow{(-w_0\alpha_1,n_1)} \cdots \xrightarrow{(-w_0\alpha_\ell,n_\ell)} w_0 w_{\ell+1} w_0\right),\end{gathered}$$

that preserves the lengths of paths. It sends path of weight ω to a path of weight $-w_0\omega$.

(c) *Let $x \in \Omega$ be of length zero in $\widetilde{W}$ and write it as $x = wt^\mu$. Then we have a bijection*

$$\begin{gathered}\operatorname{paths}^{\prec}_{\preceq n}(u \Rightarrow v) \to \operatorname{paths}^{\prec}_{\preceq n}(wu \Rightarrow wv),\\ \left(w_1 \xrightarrow{(\alpha_1,n_1)} \cdots \xrightarrow{(\alpha_\ell,n_\ell)} w_{\ell+1}\right) \mapsto \left(ww_1 \xrightarrow{(\alpha_1,n_1-\langle\mu,w_1\alpha_1\rangle)} \cdots \xrightarrow{(\alpha_\ell,n_\ell-\langle\mu,w_\ell\alpha_\ell\rangle)} ww_{\ell+1}\right),\end{gathered}$$

that preserves the lengths of paths. It sends a path of weight ω to a path of weight $\omega + v^{-1}\mu - u^{-1}\mu$.

Proof. This proof is a straightforward verification. □

It is rather natural and very fruitful to compare the definition of the double Bruhat graph with the similar concept of the *quantum Bruhat graph*. The quantum Bruhat graph can be defined as a subgraph of the double Bruhat graph, containing all those paths

$$p : u_1 \xrightarrow{(\alpha_1,m_1)} \cdots \xrightarrow{(\alpha_\ell,m_\ell)} u_{\ell+1}$$

satisfying the additional constraint that for each $n \in \{1, \dots, m_\ell\}$ either

- $\ell(u_{n+1}) = \ell(u_n) + 1$ and $m_n = 0$ or
- $\ell(u_{n+1}) = \ell(u_n) + 1 - \langle\alpha_n^\vee, 2\rho\rangle$ and $m_n = 1$.

The quantum Bruhat graph was introduced by Brenti, Fomin and Postnikov [Brenti et al. 1999] in order to study certain solutions to the Yang–Baxter equations related to the quantum Chevalley–Monk formula. It has since occurred frequently in literature on quantum cohomology of flag varieties, e.g., in [Postnikov 2005]. Due to its relationship to the affine Bruhat order discovered by Lam and Shimozono [2010], it since has played a major role in the study of affine Bruhat order and affine Deligne–Lusztig varieties, e.g., in [Milićević 2021; Milićević and Viehmann 2020; He and Yu 2021; Schremmer 2024].

The initial article [Brenti et al. 1999] is of special interest to us, since we may identify its *main result* as a crucial statement on the double Bruhat graph. The authors derive the fundamental properties of the quantum Bruhat graph as an application of their main result, which we may recognize as the aforementioned embedding of the quantum Bruhat graph into the double Bruhat graph.

Let us recall the main result of [loc. cit.] in the authors' language. They construct a family of solutions to the Yang–Baxter equations for the finite Weyl group W as follows: Choose a field K of characteristic zero, and *multiplicative functions* $E_1, E_2 : \Phi^+ \to K$, i.e., functions satisfying

$$E_i(\alpha + \beta) = E_i(\alpha)E_i(\beta) \quad \text{whenever } \alpha, \beta, \alpha + \beta \in \Phi^+.$$

We have to assume that there is no root $\alpha \in \Phi^+$ such that $E_1(\alpha) = E_2(\alpha) = 0$. Choose moreover constants $\kappa_\alpha \in K$ depending only on the length of $\alpha \in \Phi^+$. Define for each $\alpha \in \Phi^+$ the following K-linear endomorphism of the group algebra $K[W]$:

$$R_\alpha : K[W] \to K[W], \quad w \mapsto \begin{cases} w + p_\alpha s_\alpha w, & \ell(s_\alpha w) > \ell(w), \\ w + q_\alpha s_\alpha w, & \ell(s_\alpha w) < \ell(w). \end{cases}$$

Here, we define the scalars

$$p_\alpha = \frac{\kappa_\alpha E_1(\alpha)}{E_1(\alpha) - E_2(\alpha)}, \quad q_\alpha = \frac{\kappa_\alpha E_2(\alpha)}{E_1(\alpha) - E_2(\alpha)} \in K.$$

Then the linear functions $\{R_\alpha\}_{\alpha \in \Phi^+}$ satisfy the *Yang–Baxter equations*. We do not wish to recall how these equations are defined, referring the reader to the original article [Brenti et al. 1999] for the details. We do want to note, however, the following consequence of the Yang–Baxter equations, which is proved completely analogously to [loc. cit., Proposition 2.5]:

Proposition 4.8. *Let* $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\} = \{\gamma_1 \prec' \cdots \prec' \gamma_{\#\Phi^+}\}$ *be two reflection orders on* Φ^+ *and* $0 \le n \le \#\Phi^+$ *such that*

$$s_{\beta_1} \cdots s_{\beta_n} = s_{\gamma_1} \cdots s_{\gamma_n}.$$

Then

$$R_{\beta_1} \cdots R_{\beta_n} = R_{\gamma_1} \cdots R_{\gamma_n}$$

as endomorphisms on $K[W]$. □

How is this related to the double Bruhat graph? We can make the following choices, which are essentially universal:[1]

Write $\Delta = \{\alpha_1, \dots, \alpha_{\#\Delta}\}$ and let K be the field of formal Laurent series over $\mathbb{Q}$ in $2 + \#\Delta$ formal variables, denoted by

$$K = \mathbb{Q}((\kappa_s, \kappa_l, e^{\alpha_1}, \dots, e^{\alpha_{\#\Delta}})).$$

To an element of the root lattice $\lambda = c_1\alpha_1 + \cdots + c_{\#\Delta}\alpha_{\#\Delta}$, we associate the element

$$e^\lambda = (e^{\alpha_1})^{c_1} \cdots (e^{\alpha_{\#\Delta}})^{c_{\#\Delta}} \in K.$$

Define $E_1(\alpha) = 1$ and $E_2(\alpha) = e^\alpha \in K$ for all $\alpha \in \Phi^+$. We put $\kappa_\alpha = \kappa_s$ if α is short and $\kappa_\alpha = \kappa_l$ if α is long. Then we observe

$$R_\alpha(w) = w + \frac{\kappa_\alpha e^{\Phi^+(-w^{-1}\alpha)\alpha}}{1 - e^\alpha} s_\alpha w = w + \sum_{i \ge \Phi^+(-w^{-1}\alpha)} \kappa_\alpha e^{i\alpha} s_\alpha w.$$

This basically describes all paths from w^{-1} to w^{-1} or $w^{-1}s_\alpha$ in the double Bruhat graph associated to the dual root system, with the restriction that the only occurring edge may be $w^{-1} \xrightarrow{\alpha^\vee} w^{-1}s_\alpha$. Composition

[1] We may focus on irreducible root systems, where only one of the two functions E_1, E_2 is allowed to vanish on the highest root. If this is without loss of generality E_1, replacing (E_1, E_2) by $(1, E_2/E_1)$ yields the same solutions $R_\bullet$. Now a universal solution would be given by choosing $E_1(\alpha) = e^\alpha$ and $\kappa_\alpha \in \{\kappa_s, \kappa_l\}$ for the subring R contained in our choice of K generated by these values and the inverses $1/(E_1(\alpha) - 1)$. However, R is not a field, and we would like to expand the geometric series.

of the linear operators $R_\bullet$ along a reflection order recovers precisely our notion of paths in the double Bruhat graph which are increasing with respect to that reflection order.

Corollary 4.9. *Let $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\} = \{\gamma_1 \prec' \cdots \prec' \gamma_{\#\Phi^+}\}$ be two reflection orders on Φ^+ and $0 \leq n \leq \#\Phi^+$ such that*

$$s_{\beta_1} \cdots s_{\beta_n} = s_{\gamma_1} \cdots s_{\gamma_n}.$$

Let $u, v \in W$, $\mu \in \mathbb{Z}\Phi^\vee$ and $\ell_s, \ell_l \in \mathbb{Z}_{\geq 0}$.

Let $p_\prec$ be the number of paths $p \in \mathrm{paths}^{\prec}_{\leq n}(u \Rightarrow v)$ such that $\mathrm{wt}(p) = \mu$, $\ell(p) = \ell_s + \ell_l$ and the number of short (resp. long) roots occurring as labels in p is equal to ℓ_s (resp. ℓ_l). Define $p_{\prec'}$ similarly. Then $p_\prec = p_{\prec'}$.

Proof. Construct the operators $R^\vee_\bullet$ as above for the dual root system $\Phi^\vee$, and consider the $\mathbb{Q}$-coefficient of $\kappa_l^{\ell_s}\kappa_s^{\ell_l} e^\mu v^{-1}$ in the expansion of

$$R^\vee_{\beta_n^\vee} \cdots R^\vee_{\beta_1^\vee} u^{-1} \in K[W].$$

By construction, this number is equal to $p_\prec$. By Proposition 4.8, it is also equal to $p_{\prec'}$. □

It is convenient to use the language of *multisets* when keeping track of the lengths and weights of paths. Recall that a multiset is a modification of the concept of a set, where elements are allowed to be contained multiple times in a multiset.

Formally, we may define a multiset M as a tuple $(|M|, m)$ where $|M|$ is any set and $m : |M| \to \mathbb{Z}_{\geq 1} \cup \{+\infty\}$ is a function (to be thought of as counting how often an element occurs in M). We write $x \in M$ meaning $x \in |M|$, and say that x has multiplicity $m(x)$ in M. If $x \notin M$, we say that x has multiplicity zero in M.

If f is a map from $|M|$ to some abelian group (e.g., the real numbers), we write

$$\sum_{x \in M} f(x) := \sum_{x \in |M|} m(x) f(x),$$

meaning that elements are summed with multiplicity (depending on the function and the abelian group, such a sum may or may not be well-defined). The cardinality of M can then be defined as

$$\#M := \sum_{x \in M} 1 \in \mathbb{Z}_{\geq 0} \cup \{\infty\}.$$

If M, M' are two multisets, we define their additive union $M \cup M'$ by declaring that x has multiplicity $m_1 + m_2$ in $M \cup M'$, where $m_1 \in \mathbb{Z}_{\geq 0} \cup \{+\infty\}$ is the multiplicity of x in M and m_2 the multiplicity of x in M'.

When explicitly writing down multisets via a list of elements (the number of occurrences expressing the multiplicity), we use the notation $\{\cdot\}_m$ to distinguish from the usual set notation $\{\cdot\}$.

Definition 4.10. Let $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ be a reflection order and $n \in \{0, \ldots, \#\Phi^+\}$.

(a) We write

$$\pi_{\succ n} = s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}} \in W.$$

(b) If $u, v \in W$, we define

$$\operatorname{wts}(u \Rightarrow v \dashrightarrow v\pi_{\succ n})$$

to be the multiset

$$\{(\operatorname{wt}(p), \ell(p)) \mid p \in \operatorname{paths}^{\prec}_{\preceq n}(u \Rightarrow v)\}_m,$$

i.e., the multiplicity of $(\omega, e) \in \operatorname{wts}(u \Rightarrow v \dashrightarrow v\pi_{\succ n})$ is equal to the number of paths in $\operatorname{paths}^{\prec}_{\preceq n}(u \Rightarrow v)$ of weight ω and length e.

We use the shorthand notation

$$\operatorname{wts}(u \Rightarrow v) := \operatorname{wts}(u \Rightarrow v \dashrightarrow v).$$

The reflection order does not occur any more in the $\operatorname{wts}(\cdots)$-notation, due to Corollary 4.9 (and the usual observation $w_0 = s_{\beta_1} \cdots s_{\beta_{\#\Phi^+}}$). From Lemma 3.2, we see that $\ell(\pi_{\succ n}) = \#\Phi^+ - n$ and that, for each $u \in W$, one may find a suitable reflection order $\prec$ with $u = \pi_{\succ \#\Phi^+ - \ell(u)}$. Hence the notation $\operatorname{wts}(u \Rightarrow v \dashrightarrow w)$ is well-defined for all $u, v, w \in W$.

We note the following immediate properties.

Lemma 4.11. *Let $u, v, v' \in W$. Then the multiset* $\operatorname{wts}(u \Rightarrow v \dashrightarrow v')$ *is nonempty if and only if the following inequality on the Bruhat order of W is satisfied*:

$$v^{-1}v' \leq u^{-1}v'.$$

In this case, *we have*

$$\max\{e \mid (\omega, e) \in \operatorname{wts}(u \Rightarrow v \dashrightarrow v')\} = \ell(u^{-1}v') - \ell(v^{-1}v').$$

Proof. Let us pick a reflection order $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ with $v' = v\pi_{\succ n}$. Write the corresponding reduced word as $w_0 = s_{\alpha_1} \cdots s_{\alpha_{\#\Phi^+}}$.

The multiset in question is nonempty if there is a sequence $1 \leq i_1 < \cdots < i_e \leq n$ of indices such that

$$\begin{aligned} v = us_{\beta_{i_1}} \cdots s_{\beta_{i_e}} \quad &\iff \quad vs_{\alpha_1} \cdots s_{\alpha_n} = us_{\alpha_1} \cdots \hat{s}_{\alpha_{i_1}} \cdots \hat{s}_{\alpha_{i_e}} \cdots s_{\alpha_n} \\ &\iff \quad u^{-1}vs_{\alpha_1} \cdots s_{\alpha_n} = s_{\alpha_1} \cdots \hat{s}_{\alpha_{i_1}} \cdots \hat{s}_{\alpha_{i_e}} \cdots s_{\alpha_n}. \end{aligned}$$

Using the subword criterion for Bruhat order, the existence of such indices $i_1, \dots, i_e$ is equivalent to the Bruhat order inequality

$$u^{-1}vs_{\alpha_1} \cdots s_{\alpha_n} \leq s_{\alpha_1} \cdots s_{\alpha_n},$$

and the maximal number e is equal to the difference in lengths of the two sides. Now we compute

$$s_{\alpha_1} \cdots s_{\alpha_n} = s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}} s_{\alpha_1} \cdots s_{\alpha_{\#\Phi^+}} = \pi_{\succ n} w_0.$$

Hence the above Bruhat order condition becomes $u^{-1}v'w_0 \leq v^{-1}v'w_0$. The claim follows using the Bruhat order antiautomorphism induced by multiplication by w_0. ☐

Remark 4.12. We can use these multisets to summarize Theorem 4.6 as follows: The number of pieces $\mathcal{T}_p$ of dimension d occurring in

$$\big(({}^{u}U(L) \cap {}^{v}U(L))xI\big) \cap ({}^{uw_0}U(L)yI)$$

is equal to the multiplicity of the tuple

$$(u^{-1}w_y\mu_y - u^{-1}w_x\mu_x, 2d - \ell_u(x) + \ell_u(y))$$

in the multiset

$$\mathrm{wts}(w_y^{-1}u \Rightarrow w_x^{-1}u \dashrightarrow w_x^{-1}v).$$

We finish this section by comparing the double Bruhat graph more directly to the quantum Bruhat graph.

Proposition 4.13. *Let $u, v \in W$. Denote by $d(u \Rightarrow v)$ the distance of a shortest path in the quantum Bruhat graph from u to v, and by $\mathrm{wt}(u \Rightarrow v)$ the weight of such a path. Let $(\omega, e) \in \mathrm{wts}(u \Rightarrow v)$.*

(1) *We have $\omega \geq \mathrm{wt}(u \Rightarrow v)$.*

(2) *We have*

$$e \leq \langle \omega, 2\rho\rangle + \ell(v) - \ell(u).$$

If the equality holds, then $\omega = \mathrm{wt}(u \Rightarrow v)$ and $e = d(u \Rightarrow v)$.

(3) *The multiplicity of*

$$(\mathrm{wt}(u \Rightarrow v), d(u \Rightarrow v))$$

in the multiset $\mathrm{wts}(u \Rightarrow v)$ is equal to 1.

Proof. Pick a reflection order $\prec$, and a path $p \in \mathrm{paths}^{\prec}(u \Rightarrow v)$ of weight ω and length e. Write it as

$$p : u = u_1 \xrightarrow{(\alpha_1, m_1)} \cdots \xrightarrow{(\alpha_e, m_e)} u_{e+1} = v.$$

(a) Using the triangle inequality for the quantum Bruhat graph [Postnikov 2005, Lemma 1], we get

$$\mathrm{wt}(u \Rightarrow v) \leq \mathrm{wt}(u_1 \Rightarrow u_2) + \cdots + \mathrm{wt}(u_e \Rightarrow u_{e+1}).$$

At each step, apply [Schremmer 2022, Lemma 4.7] to get

$$\begin{aligned}\cdots &\leq \Phi^+(-u_1\alpha_1)\alpha_1^\vee + \cdots + \Phi^+(-u_e\alpha_e)\alpha_e^\vee\\ &\leq m_1\alpha_1^\vee + \cdots + m_e\alpha_e^\vee = \omega.\end{aligned}$$

(b) For each edge $u_i \xrightarrow{(\alpha_i, m_i)} u_{i+1}$, we estimate

$$\langle m_i\alpha_i^\vee, 2\rho\rangle = m_i\langle \alpha_i^\vee, 2\rho\rangle \geq \Phi^+(-u_i\alpha_i)\langle \alpha_i^\vee, 2\rho\rangle.$$

If $u_i\alpha_i$ is a positive root, we evaluate $\cdots = 0 \geq \ell(u_i) - \ell(u_{i+1}) + 1$. If $u_i\alpha_i$ is a negative root, we apply [Brenti et al. 1999, Lemma 4.3] to see $\langle \alpha_i^\vee, 2\rho\rangle \geq \ell(u_i) - \ell(u_{i+1}) + 1$. In any case, we get

$$\langle m_i\alpha_i^\vee, 2\rho\rangle \geq \ell(u_i) - \ell(u_{i+1}) + 1,$$

with equality holding if and only if there is an edge $u_i \to u_{i+1}$ in the quantum Bruhat graph of weight $m_i\alpha_i^\vee$.

Iterating this over all edges of the path p, we get the desired inequality.

If equality holds, we get a path in the quantum Bruhat graph

$$p' : u_1 \to \cdots \to u_e$$

of length e and weight ω. Moreover, p' is increasing with respect to our reflection order. By [Brenti et al. 1999, Theorem 6.4], p' must be a shortest path, so $e = d(u \Rightarrow v)$ and $\omega = \mathrm{wt}(u \Rightarrow v)$.

(c) In view of the proof of (b), we have to see that there exists a unique shortest path in the quantum Bruhat graph which is increasing for $\prec$. This statement is found again in [loc. cit., Theorem 6.4]. □

As an application of our findings so far, we are able to prove the following identity.

Proposition 4.14. *Let $x \in \widetilde{W}$ and $u \in W$. Then*

$$({}^{u}U(L)xI) \cap ({}^{v}U(L)xI) = ({}^{u}U(L) \cap {}^{v}U(L))xI.$$

Proof. The group $H := {}^{u}U(L) \cap {}^{v}U(L)$ acts on both sides of this equation by left multiplication. Moreover, each H-orbit in ${}^{u}U(L)$ contains an element of ${}^{u}U(L) \cap {}^{vw_0}U(L)$. We see that each H-orbit of

$$({}^{u}U(L)xI) \cap ({}^{v}U(L)xI)$$

contains an element of the intersection

$$Y := ({}^{u}U(L) \cap {}^{vw_0}U(L))xI \cap ({}^{v}U(L)xI).$$

We claim that $Y = xI$. Indeed, $Y/I \subseteq G(L)/I$ is decomposed into pieces according to the multiset

$$\mathrm{wts}(w_x^{-1}v \Rightarrow w_x^{-1}v \dashrightarrow w_x^{-1}u).$$

By Lemma 4.11, we see that this multiset only contains tuples of the form $(\omega, 0)$, i.e., all elements in there correspond to length-zero paths from $w_x^{-1}v$ to $w_x^{-1}v$ in the double Bruhat graph. Since there is exactly one such path, we conclude that Y/I is irreducible of dimension zero, that is, a point. So the inclusion $xI \subseteq Y$ is an equality.

We summarize that

$$({}^{u}U(L)xI) \cap ({}^{v}U(L)xI) = HY = HxI = ({}^{u}U(L) \cap {}^{v}U(L))xI. \qquad \square$$

5. Affine Deligne–Lusztig varieties

In this section, we study affine Deligne–Lusztig varieties in the affine flag variety. So the parahoric subgroup is I and the variety $X_x(b)$ is parametrized by $x \in \widetilde{W}$ and $[b] \in B(G)$. In addition to the restriction on split groups over a local field of equal characteristic, we also fix a σ-conjugacy class $[b] \in B(G)$ whose Newton point $\nu(b) \in X_*(T) \otimes \mathbb{Q}$ is *integral*, i.e., contained in $X_*(T)$. Then $b = t^{\nu(b)} \in \widetilde{W}$ is our canonical representative of $[b] \in B(G)$. Our three main questions regarding the geometry of $X_x(b)$ are answered by Görtz, Haines, Kottwitz and Reuman in terms of intersections of semi-infinite orbits with affine Schubert cells. Due to different choices of Iwahori subgroups, we have a few signs different from the original source. We define the dimension of the empty variety to be $-\infty$.

Theorem 5.1 [Görtz et al. 2006, Theorem 6.3.1]. *Let $x, z \in \widetilde{W}$. Then*

$$\dim\bigl(X_x(b) \cap U(L)z^{-1}I/I\bigr) = \dim\bigl(IxI/I \cap ({}^zU(L))zbz^{-1}I/I\bigr).$$

The number of $(J_b(F) \cap U(L))$-orbits of top-dimensional irreducible components in $X_x(b) \cap U(L)z^{-1}I/I$ is equal to the number of top-dimensional irreducible components in

$$IxI/I \cap ({}^zU(L))zbz^{-1}I/I. \qquad \square$$

The statement on irreducible components is missing in the cited source, but follows since the proof method allows us to compare admissible subsets in $X_x(b) \cap U(L)z^{-1}I/I$ with admissible subsets in the other intersection. A generalization of Theorem 5.1 to nonintegral $[b]$ can be found in [Görtz et al. 2010, Theorem 11.3.1], but it is unclear how our methods can be applied to that generalized statement, or how a connection to the double Bruhat graph would be given in general.

If we write $z = ut^{\mu_z}$, then

$$({}^zU(L))zbz^{-1}I = {}^uU(L)t^{u\nu(b)}I.$$

Our main result describing the intersection of that set with IxI is the following.

Theorem 5.2. *Let $x = w_x t^{\mu_x}$, $y = w_y t^{\mu_y} \in \widetilde{W}$. Let $v_x \in \mathrm{LP}(x)$ and $u \in W$. Pick a reflection order $\prec$ and an index $n \in \{0, \dots, \#\Phi^+\}$ such that $\pi_{\succ n} = u^{-1}w_x v_x w_0$. Then*

$$(IxI/I) \cap ({}^{uw_0}U(L)yI/I) = \bigsqcup_{p \in P} \widetilde{\mathcal{T}}_p \subset G(L)/I,$$

where

$$P = \{p \in \mathrm{paths}^{\prec}_{\leq n}(w_y^{-1}u \Rightarrow w_x^{-1}u) \mid \mathrm{wt}(p) = u^{-1}w_y\mu_y - u^{-1}w_x\mu_x\}$$

and each $\widetilde{\mathcal{T}}_p \subset G(L)/I$ is a locally closed k-subscheme of finite dimension

$$\dim \widetilde{\mathcal{T}}_p = \tfrac{1}{2}(\ell(x) - \ell_u(y) + \ell(p)) - \mathrm{codim}(\mathcal{T}_p \cap (x^{-1}IxI) \subseteq \mathcal{T}_p).$$

Here, $\mathcal{T}_p \subset G(L)/I$ is the variety from Theorem 4.6. If $\mathcal{T}_p \cap x^{-1}IxI$ is dense in $\mathcal{T}_p$, then $\widetilde{\mathcal{T}}_p$ is irreducible.

Proof. We may write $IxI/I = I(x)xI/I \cong I(x)$, where $I(x)$ is the finite-dimensional k-group

$$I(x) = \prod_a U_a,$$

with the product taken over all positive affine roots $a \in \Phi^+_{\mathrm{af}}$ such that $x^{-1}a \in \Phi^-_{\mathrm{af}}$. If $a = (\alpha, m)$ is such a root, recall that the root subgroup U_a is defined as the subvariety $U_a = \{U_\alpha(ht^m) \mid h \in k\}$. The length positivity of v_x implies $(w_x v_x w_0)^{-1}\alpha \in \Phi^+$. Thus we can rewrite the condition $v_x \in \mathrm{LP}(x)$ as $I(x) \subset {}^{w_x v_x w_0}U(L)$, or

$$IxI = (I \cap {}^{w_x v_x w_0}U(L))xI.$$

Observe that we have an isomorphism of k-ind-schemes

$$\begin{aligned}({}^{w_x v_x w_0}U(L) \cap {}^uU(L)) \times ({}^{w_x v_x w_0}U(L) \cap {}^{uw_0}U(L)) &\to {}^{w_x v_x w_0}U(L),\\ (g_1, g_1) &\mapsto g_1 g_2.\end{aligned} \tag{5.3}$$

Such an isomorphism may, e.g., be obtained by decomposing $U(L)$ into a suitable product of root subgroups $U_\alpha(L)$ as discussed in the beginning of Section 3.

Using (5.3), we can write each element $g\in I(x)$ uniquely in the form $g=g_1g_2$, with $g_1\in I(x)\cap{}^{uw_0}U(L)$ and $g_2\in I(x)\cap{}^{u}U(L)$. Then $gxI\in{}^{uw_0}U(L)yI$ holds if and only if $g_2xI\in{}^{uw_0}U(L)yI$. Hence the decomposition (5.3) yields an isomorphism

$$\begin{aligned}(IxI\cap{}^{uw_0}U(L)yI)/I&=(I(x)xI\cap{}^{uw_0}U(L)yI)/I\\&\cong\big((I(x)\cap{}^{uw_0}U(L))xI/I\big)\times\big((I(x)\cap{}^{u}U(L))xI\cap{}^{uw_0}U(L)yI\big)/I.\end{aligned}\tag{5.4}$$

The first variety $(I(x)\cap{}^{uw_0}U(L))xI/I\cong I(x)\cap{}^{uw_0}U(L)$ is just an affine space over k whose dimension is given by the number of positive affine roots $a=(\alpha,m)$ with $x^{-1}a\in\Phi^-_{\text{af}}$ and $(uw_0)^{-1}\alpha\in\Phi^+$. By [Schremmer 2022, Lemma 2.9], we can express this quantity as

$$S_1:=\sum_{\alpha\in\Phi^-}\max(0,\ell(x^{-1},u\alpha)).$$

Let us moreover define

$$S_2:=\sum_{\alpha\in\Phi^-}\min(0,\ell(x^{-1},u\alpha)).$$

Then $-S_1-S_2$ is simply the sum over all $\ell(x^{-1},-u\alpha)$ for $\alpha\in\Phi^+$, which we denoted by $\ell_u(x)$. Conversely, S_1-S_2 is the sum over all $|\ell(x^{-1},u\alpha)|$, which equals $\ell(x)$ by [loc. cit., Corollary 2.10, Lemma 2.6]. We conclude

$$\dim I(x)\cap{}^{u}U(L)=S_1=\tfrac{1}{2}(\ell(x)-\ell_u(x)).$$

It remains to study the second factor in (5.4). Following Theorem 4.6, we may take the decomposition

$$[{}^{w_xv_xw_0}U(L)\cap{}^{u}U(L)]xI=\bigsqcup_p x\mathcal{T}_p,$$

with the union taken over all paths $p\in\text{paths}^{\preceq}_{\leq n}(u'\Rightarrow w_x^{-1}u)$ and all $u'\in W$. Using Theorem 4.6, we conclude that $x\mathcal{T}_p\cap{}^{uw_0}U(L)yI$ is empty if $p\notin P$ and equal to $x\mathcal{T}_p$ if $p\in P$. Hence the second factor in (5.4) can be decomposed as

$$\big((I\cap{}^{w_xv_xw_0}U(L)\cap{}^{u}U(L))xI\cap{}^{uw_0}U(L)yI\big)/I=\bigsqcup_{p\in P}(x\mathcal{T}_p\cap IxI/I).$$

We have

$$\dim(x\mathcal{T}_p\cap IxI/I)=\dim\mathcal{T}_p-\operatorname{codim}(\mathcal{T}_p\cap x^{-1}IxI/I\subseteq\mathcal{T}_p).$$

So the piece $\widetilde{\mathcal{T}}_p$ corresponding to $I(x)\cap{}^{uw_0}U(L)$ and $x\mathcal{T}_p\cap IxI$ under (5.4) has dimension

$$\tfrac{1}{2}(\ell(x)-\ell_u(x))+\tfrac{1}{2}(\ell_u(x)-\ell_u(y)+\ell(p))-\operatorname{codim}(\mathcal{T}_p\cap x^{-1}IxI/I\subseteq\mathcal{T}_p).$$

Cancelling the common term $\ell_u(x)$, we get the claimed formula. If $\mathcal{T}_p\cap x^{-1}IxI$ is dense in $\mathcal{T}_p$, then $\mathcal{T}_p\cap x^{-1}IxI$ is irreducible itself. Thus $\widetilde{\mathcal{T}}_p$ is isomorphic to the direct product of the affine space $I(x)\cap{}^{uw_0}U(L)$ and the irreducible variety $\mathcal{T}_p\cap x^{-1}IxI$, and hence is irreducible itself. □

We are especially interested in those situations where all $p \in P$ satisfy the property $\mathcal{T}_p \subseteq x^{-1}IxI$. This is not guaranteed at all, as one may choose p and x independently to obtain examples where the inclusion is far from being satisfied. Nonetheless, we can develop some regularity conditions imposed on (x, y) that guarantee this inclusion.

Lemma 5.5. *Let $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$ be a reflection order, $u, v \in W$ and $p \in \mathrm{paths}^{\prec}(u \Rightarrow v)$. Pick $gI \in \mathcal{T}_p$ and write it as*

$$gI = U_{v\beta_1}(g_1) \cdots U_{v\beta_{\#\Phi^+}}(g_{\#\Phi^+})I \in \mathcal{T}_p.$$

Then, for $m = 1, \ldots, \#\Phi^+$, we have

$$\nu_L(g_m) \geq -3\langle \rho^\vee, \beta_m\rangle \langle \mathrm{wt}(p), \rho\rangle,$$

where ρ is the half-sum of positive roots and $\rho^\vee$ the half-sum of positive coroots.

Proof. Write our path as

$$p : u = u_1 \xrightarrow{(\alpha_1, m_1)} \cdots \xrightarrow{(\alpha_{\ell(p)}, m_{\ell(p)})} u_{\ell(p)+1} = v.$$

Denote the type corresponding to p under Lemma 4.4 by $\tau = \{(n_1, \nu_1), \ldots, (n_{\ell(p)}, \nu_{\ell(p)})\}$. This means $\alpha_i = \beta_{n_i}$ and

$$(s_{\alpha_{\ell(p)}} \cdots s_{\alpha_{i+1}}(\alpha_i), m_i) = r_{(\alpha_{\ell(p)}, \nu_{\ell(p)})} \cdots r_{(\alpha_{i+1}, \nu_{i+1})}(\alpha_i, \nu_i) \in \Phi_{\mathrm{af}}.$$

We also write $\alpha' := s_{\alpha_{\ell(p)}} \cdots s_{\alpha_{i+1}}(\alpha_i) \in \Phi$.

Let $f_\bullet^{(\bullet)}$ be the rational polynomials used in the definition of $\mathcal{T}_p = \mathcal{T}_{u,\prec,\tau}$, i.e., the polynomials from Proposition 3.5. We use them to define the values $(g_i^{(m)})_{1 \leq i \leq m \leq \#\Phi^+}$ via the identities

$$g_i^{(m)} = g_i + f_i^{(m)}(g'_{i+1}, \ldots, g'_{\#\Phi^+}),$$
$$g'_m = g_m^{(m)}.$$

Let $m \in \{1, \ldots, \#\Phi^+\}$ and let $h = h(m) \in \{0, \ldots, \ell(p)\}$ be maximal such that $\alpha_{h'} \prec \beta_m$ for all $h' \leq h$. Then the condition $g \in \mathcal{T}_p$ implies, by definition, that $g'_m = 0$ or

$$r_{(u\alpha_{\ell(p)}, \nu_{\ell(p)})} \cdots r_{(u\alpha_{h+1}, \nu_{h+1})}(u\beta_m, \nu_L(g'_m)) \in \Phi^+_{\mathrm{af}}.$$

Observe that we can write

$$\begin{aligned} r_{(u\alpha_{\ell(p)}, \nu_{\ell(p)})} \cdots r_{(u\alpha_{h+1}, \nu_{h+1})}(u\beta_m, \nu_L(g'_m)) &= r_{(u\alpha'_{h+1}, m_{h+1})} \cdots r_{(u\alpha'_{\ell(p)}, m_{\ell(p)})}(u\beta_m, \nu_L(g'_m)) \\ &= (v\beta_m, \nu_L(g'_m)) + \sum_{i=h+1}^{\ell(p)} c_i(u\alpha'_i, m_i) \end{aligned}$$

for elements $c_i \in \{0, \pm 1, \pm 2, \pm 3\}$ describing the pairings between the occurring roots. Thus

$$\nu_L(g'_m) \geq -3 \sum_{i=h+1}^{\ell(p)} m_i \geq -3 \sum_{i=h+1}^{\ell(p)} m_i \langle \alpha_i^\vee, \rho\rangle \geq -3\langle \mathrm{wt}(p), \rho\rangle.$$

We claim for all $m \in \{1, \ldots, \#\Phi^+\}$ and $i \in \{1, \ldots, n\}$ that

$$\nu_L(g_i^{(m)} - g'_i) \geq -3\langle \rho^\vee, \beta_i\rangle \langle \mathrm{wt}(p), \rho\rangle.$$

Induction on $m-i$: For $m=i$, we have $g_i^{(m)}=g_i'$ by definition, so there is nothing to prove.

So let now $i<m$ and suppose the claim has been proved for all pairs of smaller difference.

If $\beta_m\notin\{\alpha_1,\dots,\alpha_{\ell(p)}\}$, then applying Lemma 3.4 (which is how the polynomials $f_\bullet^{(\bullet)}$ were constructed) yields $g_i^{(m)}=g_i^{(m-1)}$, so we are done by induction immediately.

So suppose now that $\beta_m\in\{\alpha_1,\dots,\alpha_{\ell(p)}\}$. Applying Lemma 3.4, we see that $g_i^{(m-1)}$ has the form

$$g_i^{(m-1)}=g_i^{(m)}+\sum c_{e_{i+1},\dots,e_m}(g_{i+1}^{(m)})^{e_{i+1}}\cdots(g_m^{(m)})^{e_m},$$

with the sum taken over all possible integers $e_{i+1},\dots,e_{m-1}\geq 0,\ e_m<0$ such that

$$\beta_i=e_{i+1}\beta_{i+1}+\cdots+e_m\beta_m$$

and structure constants $c_{e_{i+1},\dots,e_m}\in\mathbb{Z}$. By the inductive assumption and the above estimate on $\nu_L(g_\bullet')$, we see

$$\nu_L[(g_{i+1}^{(m)})^{e_{i+1}}\cdots(g_{n-1}^{(m)})^{e_{n-1}}]\geq -3\langle\mathrm{wt}(p),\rho\rangle\langle\rho^\vee,e_{i+1}\beta_{i+1}+\cdots+e_{i_{m-1}}\beta_{m-1}\rangle.$$

An entirely similar argument to the one presented above shows moreover $\nu_L(g_m')\leq 3\langle\mathrm{wt}(p),2\rho\rangle$, so that

$$\nu_L((g_m')^{e_m})\geq -3e_m\langle\mathrm{wt}(p),\rho\rangle\langle\rho^\vee,\beta_m\rangle.$$

We conclude

$$\begin{aligned}\nu_L[c_{e_{i+1},\dots,e_m}(g_{i+1}^{(m)})^{e_{i+1}}\cdots(g_m^{(m)})^{e_m}]&\geq -3\langle\mathrm{wt}(p),\rho\rangle\langle\rho^\vee,e_{i+1}\beta_{i+1}+\cdots+e_m\beta_m\rangle\\&=-3\langle\mathrm{wt}(p),\rho\rangle\langle\rho^\vee,\beta_i\rangle.\end{aligned}$$

Hence

$$\nu_L(g_i^{(m)}-g_i^{(m-1)})\geq -3\langle\rho^\vee,\beta_i\rangle\langle\mathrm{wt}(p),\rho\rangle.$$

This finishes the induction.

In particular, we see

$$\nu_L(g_m)\geq -3\langle\rho^\vee,\beta_m\rangle\langle\mathrm{wt}(p),\rho\rangle$$

for all m. □

We finally define the class of elements in $\widetilde{W}$ where Theorem 5.2 and Lemma 5.5 describe affine Deligne–Lusztig varieties fully.

Definition 5.6. Let $x=wt^\mu\in\widetilde{W},\ J\subseteq\Delta$ and $C\in\mathbb{R}_{>0}$. We say that x is (J,C)-*superparabolic* if there exists $v\in W$ such that

(a) all $\alpha\in\Phi_J$ satisfy $\ell(x,v\alpha)=0$ and

(b) all $\alpha\in\Phi^+\setminus\Phi_J$ and $v'\in vW_J$ satisfy

$$\langle\mu,v'\alpha\rangle>C\langle\rho^\vee,\alpha\rangle.$$

This is a generalization of the J-adjusted and J-superdominant elements from [Lenart et al. 2015]. If x is $(J,2)$-superparabolic and v as in the above definition, then one easily checks $\mathrm{LP}(x)=vW_J$. We can interpret condition (b) above as a regularity condition of the length functional, in particular

$$\begin{aligned}\big(\forall\alpha\in\Phi^+\setminus\Phi_J,\quad \ell(x,v\alpha)>1+C\langle\rho^\vee,2\rho\rangle\big)\quad&\Longrightarrow\quad\text{condition (b) of Definition 5.6}\\&\Longrightarrow\quad\big(\forall\alpha\in\Phi^+\setminus\Phi_J,\quad\ell(x,v\alpha)>C-1\big).\end{aligned}$$

Theorem 5.7. *Let $x = wt^{\mu} \in \widetilde{W}$ and $b = t^{\nu(b)}$ be an element with integral dominant Newton point $\nu(b) \in X_*(T)^{\mathrm{dom}}$. Define for each $v \in \mathrm{LP}(x)$ and $u \in W$ the multiset*

$$E(u, v) := \{e \mid (u^{-1}\mu - \nu(b), e) \in \mathrm{wts}(u \Rightarrow wu \dashrightarrow wv)\}_m.$$

Put $\max \varnothing := -\infty$ *and define*

$$e := \max_{u \in W} \min_{v \in \mathrm{LP}(x)} \max(E(u, v)) \in \mathbb{Z} \cup \{-\infty\},$$
$$d := \tfrac{1}{2}\big(\ell(x) + e - \langle \nu(b), 2\rho\rangle\big) \in \mathbb{Z} \cup \{-\infty\}.$$

(a) *If there exists for every $u \in W$ some $v \in \mathrm{LP}(x)$ with $E(u, v) = \varnothing$, i.e., if $e = d = -\infty$, then $X_x(b) = \varnothing$.*

(b) *If $X_x(b) \neq \varnothing$, then* $\dim X_x(b) \leq d$.

(c) *Write $C = 3\langle \mu^{\mathrm{dom}} - \nu(b), \rho\rangle$ and suppose that x is (J, C)-superparabolic for some $J \subseteq \Delta$. Define the multiset E as the additive union*

$$E = \bigcup_{v \in \mathrm{LP}(x)} E(vw_0(J), v).$$

Then $X_x(b) \neq \varnothing$ if and only if $E \neq \varnothing$. In this case, $e = \max(E)$ and $\dim X_x(b) = d$.

(d) *Assume $X_x(b) \neq \varnothing$ and let Σ_d be the set of d-dimensional irreducible components of $X_x(b)$. Then the number of $J_b(F)$-orbits in Σ_d is*

$$\#(\Sigma_d / J_b(F)) \leq \sum_{u \in W} \min_{v \in W}(\text{multiplicity of } e \text{ in } E(u, v)).$$

If we are in the situation of (c) and b is regular, i.e., $\langle \nu(b), \alpha\rangle \neq 0$ for all $\alpha \in \Phi$, then $\#(\Sigma_d / J_b(F))$ is equal to the multiplicity of e in E.

Proof. We use Theorem 5.1 to reduce questions on the affine Deligne–Lusztig variety to the situation of Theorem 5.2.

So let $z = ut^{\mu_z} \in \widetilde{W}$. Then $X_x(b) \cap U(L)z^{-1}I/I$ is closely related to the intersection

$$(IxI/I) \cap ({}^{u}U(L)t^{-u\nu(b)}I/I).$$

Pick $v \in \mathrm{LP}(x)$. By Theorem 5.2, the latter intersection can be decomposed into pieces $(\widetilde{\mathcal{T}}_p)_{p \in P}$ with

$$P = \{p \in \mathrm{paths}^{\prec}_{\preceq n}(uw_0 \Rightarrow w^{-1}uw_0) \mid \mathrm{wt}(p) = (uw_0)^{-1}(u\nu(b)) - (uw_0)^{-1}w\mu\}.$$

Here, $\prec$ is a reflection order chosen such that $\pi_{\succ n} = w_0u^{-1}wvw_0$. The number of paths in P having a given length $\ell \in \mathbb{Z}_{\geq 0}$ is equal to the multiplicity of

$$(w_0(\nu(b) - u^{-1}w\mu), \ell)$$

in the multiset

$$\mathrm{wts}(uw_0 \Rightarrow w^{-1}uw_0 \dashrightarrow vw_0).$$

By Lemma 4.7(a) and (b), this is also equal to the multiplicity of $(u^{-1}w\mu - \nu(b), \ell)$ in

$$\mathrm{wts}(w_0u \Rightarrow w_0w^{-1}u \dashrightarrow w_0v) = \mathrm{wts}(w^{-1}u \Rightarrow u \dashrightarrow wv).$$

By definition, this is the multiplicity of ℓ in the multiset $E(w^{-1}u, v)$.

We see that if $E(w^{-1}u, v) = \varnothing$ then $X_x(b) \cap U(L)z^{-1}I = \varnothing$. Otherwise,

$$\dim\bigl(X_x(b) \cap U(L)z^{-1}I/I\bigr) = \max_p \dim \widetilde{\mathcal{T}}_p \leq \tfrac{1}{2}\bigl(\ell(x) + \max(E(w^{-1}u, v)) - \ell_{uw_0}(t^{u\nu(b)})\bigr).$$

Observe $\ell_{uw_0}(t^{u\nu(b)}) = \langle \nu(b), 2\rho\rangle$. This shows (a) and (b).

In particular, we have $\dim \widetilde{\mathcal{T}}_p \leq d$ for all p. If equality holds, then $\widetilde{\mathcal{T}}_p$ must be irreducible by Theorem 5.2. Hence the number of d-dimensional irreducible components in $X_x(b) \cap U(L)z^{-1}I/I$ is equal to the number of pieces $\widetilde{\mathcal{T}}_p$ satisfying $\dim \widetilde{\mathcal{T}}_p = d$, which is at most the multiplicity of e in $E(w^{-1}u, v)$. Observe that the action of $T(F) \subseteq J_B(F)$ simply permutes the intersections $X_x(b) \cap U(L)z^{-1}I/I$ by changing the value of μ_z. Thus the number $\#(\Sigma_d / J_b(F))$ is at most equal to

$$\sum_{u \in W} (\text{number of } d\text{-dimensional irreducible components in } (X_x(b) \cap U(L)u^{-1}I)).$$

We get the desired estimate in (d).

Let us now assume the regularity condition from (c). Let $v_1 \in W$ be chosen such that $v_1^{-1}\mu$ is dominant. We have $\mathrm{LP}(x) = v_1 W_J$.

If $u \in W$ satisfies $w^{-1}u \notin \mathrm{LP}(x) = v_1 W_J$, we find a positive root $\alpha \in \Phi^+ \setminus \Phi_J$ with $v_1^{-1}w^{-1}u\alpha \in \Phi^-$. Hence

$$\langle \mu, -w^{-1}u\alpha\rangle \geq C = 3\langle v_1^{-1}\mu - \nu(b), \rho\rangle.$$

We conclude

$$w^{-1}u\mu \leq \mu^{\mathrm{dom}} - C\alpha^\vee \not\geq \nu(b).$$

Hence $E(u, v) = \varnothing$ for all $v \in \mathrm{LP}(x)$, proving

$$IxI \cap {}^uU(L)t^{-u\nu(b)}I = \varnothing.$$

Let us now consider the case $w^{-1}u \in \mathrm{LP}(x)$. Then also $v := w^{-1}uw_0(J) \in \mathrm{LP}(x)$. Consider the pieces $\widetilde{\mathcal{T}}_p$ as constructed above for this pair (u, v), i.e., for paths p from uw_0 to $w^{-1}uw_0$. We claim $\mathcal{T}_p \subseteq x^{-1}IxI$ for all occurring paths p, using Lemma 5.5: Indeed if

$$gI = U_{w^{-1}uw_0\beta_1}(g_1) \cdots U_{w^{-1}uw_0\beta_n}(g_n)I \in \mathcal{T}_p$$

and $w_0w_0(J)w_0 = \pi_{\succ\beta_n}$, then $-w_0\Phi_J^+ = \{\beta_{n+1}, \ldots, \beta_{\#\Phi^+}\}$. The condition that p is bounded above by n yields

$$w^{-1}uw_0\beta_1, \ldots, w^{-1}uw_0\beta_n \in w^{-1}u(\Phi^- \setminus \Phi_J).$$

Since $w^{-1}u \in \mathrm{LP}(x)$, the superparabolicity condition implies

$$\langle \mu, w^{-1}uw_0\beta_i\rangle < -C\langle \rho^\vee, \beta_i\rangle.$$

By Lemma 5.5, we obtain $xU_{w^{-1}uw_0\beta_i}(g_i)x^{-1} \in I$. This shows the claim $\mathcal{T}_p \subseteq x^{-1}IxI$.

By Theorem 5.2, we see that $\widetilde{\mathcal{T}}_p$ is irreducible of dimension

$$\dim \widetilde{\mathcal{T}}_p = \tfrac{1}{2}\bigl(\ell(x) + \ell(p) - \langle \nu(b), 2\rho\rangle\bigr).$$

This completely describes nonemptiness, dimension and top-dimensional irreducible components of $X_x(b)\cap U(L)z^{-1}I/I$. So $X_x(b)\neq\varnothing$ if and only if $E\neq\varnothing$, and in this case

$$\dim X_x(b)=\tfrac{1}{2}\big(\ell(x)+\max(E)-\langle \nu(b),2\rho\rangle\big).$$

We saw $E(u,v)=\varnothing$ whenever $u\notin \mathrm{LP}(x)$, so $e\leq \max(E)$ by definition of e. Conversely, we get $\max(E)\leq e$ from (b) and the above dimension calculation. Thus $\max(E)=e$. We conclude (c).

Assume now that $[b]$ is regular as in the final claim of (d). Then $J_b(F)=T(F)$, so the number of $J_b(F)$-orbits of d-dimensional irreducible components in $X_x(b)$ is equal to

$$\sum_{u\in W}(\text{number of } d\text{-dimensional irreducible components in } (X_x(b)\cap U(L)u^{-1}I)).$$

Observe that each summand is equal to the number of d-dimensional pieces $\widetilde{\mathcal{T}}_p$ corresponding to $u\in W$. By the above analysis using the superparabolicity assumption, this number is equal to the multiplicity of e in $E(w^{-1}u, w^{-1}uw_0(J))$ (thus zero if $w^{-1}u\notin \mathrm{LP}(x)$). The final claim of (d) follows. □

Remark 5.8. If x is not superparabolic, we do not expect that the converse of Theorem 5.7(a) holds in general. Even if $X_x(b)\neq\varnothing$, we do not expect that equality holds in (b) or (d) in general. It is easy to find counterexamples using a computer search.

Corollary 5.9. *Let $x=wt^\mu\in\widetilde W$ and $[b]\in B(G)$. Let $v\in W$ such that $v^{-1}\mu$ is dominant. Put $C=3\langle v^{-1}\mu-\nu(b),\rho\rangle$ and assume that*

$$\langle v^{-1}\mu,\alpha\rangle\geq C$$

for all simple roots α. Define the multiset

$$E=\{e\in\mathbb{Z}\mid (v^{-1}\mu-\nu(b),e)\in \mathrm{wts}(v\Rightarrow wv)\}_m.$$

Then $X_x(b)\neq\varnothing$ if and only if $E\neq\varnothing$. In this case, the dimension of $X_x(b)$ is

$$\dim X_x(b)=\tfrac{1}{2}\big(\ell(x)+\max(E)-\langle \nu(b),2\rho\rangle\big),$$

and the number of $J_b(F)$-orbits of top-dimensional irreducible components is equal to the multiplicity of $\max(E)$ *in E.*

Proof. The regularity condition on (x,b) implies that $\nu(b)$ must be regular. In particular, $[b]$ is integral. Now apply the previous theorem. □

Remark 5.10. We saw in Proposition 4.13 that the set $\{\omega\mid(\omega,e)\in \mathrm{wts}(v\Rightarrow wv)\}$ contains a unique minimum, which is given by the weight of a shortest path in the quantum Bruhat graph from v to wv. The above corollary shows under some strong regularity conditions that the set $B(G)_x$ contains a unique maximum $[b_x]$, being the element of Newton point $\nu(b_x)=v^{-1}\mu-\mathrm{wt}(v\Rightarrow wv)$. It moreover follows that this element $[b_x]$ satisfies

$$d(v\Rightarrow wv)=\ell(x)-\langle\nu(b_x),2\rho\rangle=\dim X_x(b_x)$$

and that $X_x(b_x)$ has, up to $J_{b_x}(F)$-action, only one top-dimensional irreducible component.

It is a well-known result of Viehmann [2014, Section 5] that $B(G)_x$ always contains a unique maximum for arbitrary G and x. She moreover provides a combinatorial description in terms of the Bruhat order on the extended affine Weyl group.

The description of the generic σ-conjugacy class $[b_x]$ in terms of the quantum Bruhat graph is known due to Milićević [2021]. Her proof is more combinatorial in nature, following Viehmann's description of $[b_x]$ using the Bruhat order and a comparison of the Bruhat order with the quantum Bruhat graph due to Lam and Shimozono [2010]. Her combinatorial methods have been refined since, so that a description of $[b_x]$ using the quantum Bruhat graph is known for arbitrary G and x [Sadhukhan 2023; He and Nie 2024; Schremmer 2022]. Our corollary recovers Milićević's original result using an entirely different proof method, which moreover reveals how to find the quantum Bruhat graph itself in the affine flag variety.

The aforementioned geometric properties of $X_x(b_x)$ are well known for arbitrary G and x, as described in the Introduction. We can interpret Proposition 4.13, i.e., essentially [Brenti et al. 1999, Theorem 6.4], as a combinatorial shadow of these geometric facts.

Remark 5.11. Let us compare Corollary 5.9 to the situation of affine Deligne–Lusztig varieties in the affine Grassmannian, i.e., where the parahoric subgroup $K = G(\mathcal{O}_L)$ is hyperspecial. Given $[b] \in B(G)$ and a dominant $\mu \in X_*(T)$, we can compare the affine Deligne–Lusztig variety $X_\mu(b) \subset G(L)/K$ with $X_{w_0 t^\mu}(b) \subset G(L)/I$ following [He 2014, Theorem 10.1]. If $x = w_0 t^\mu$ satisfies the regularity conditions from Corollary 5.9, this means we should study the multiset $\operatorname{wts}(1 \Rightarrow w_0)$.

Given any reflection order $\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$, there exists a unique unlabelled path of maximal length from 1 to w_0, given by

$$\bar{p} : 1 \xrightarrow{\beta_1} s_{\beta_1} \xrightarrow{\beta_2} \cdots \xrightarrow{\beta_{\#\Phi^+}} w_0.$$

Each arrow in this path is increasing the length in W. Thus, the labelled paths p in the double Bruhat graph with underlying unlabelled path $\bar{p}$ are precisely the paths of the form

$$p : 1 \xrightarrow{(\beta_1, m_1)} s_{\beta_1} \xrightarrow{(\beta_2, m_2)} \cdots \xrightarrow{(\beta_{\#\Phi^+}, m_{\#\Phi^+})} w_0$$

for integers $m_1, \ldots, m_{\#\Phi^+} \geq 0$. In the situation of Corollary 5.9, we see that $X_x(b) \neq \varnothing$ if and only if $\mu - \nu(b)$ is a sum of positive coroots, in which case we get

$$\dim X_x(b) = \tfrac{1}{2}\big(\ell(x) + \#\Phi^+ - \langle \nu(b), 2\rho \rangle\big).$$

The number of top-dimensional irreducible components of $X_x(b)$ is equal to the number of different ways to express $\mu - \nu(b)$ as a sum of positive coroots. This latter quantity is known as *Kostant's partition function*, which is also known to describe the dimension of the $\nu(b)$-weight space associated with the Verma module V_μ. Under the regularity assumption made, this is also equal to the dimension of the $\nu(b)$-weight space of the irreducible quotient M_μ by Kostant's multiplicity formula.

In view of [He 2014, Theorem 10.1], we recover Theorem 1.1 in the setting of Corollary 5.9. While this is a fairly restrictive setting, one may expect that statements similar to Corollary 5.9 hold true in much higher generality.

Acknowledgements

I was partially supported by the German Academic Scholarship Foundation, the Marianne-Plehn programme, the DFG Collaborative Research Centre 326 *GAUS* and the Chinese University of Hong Kong. I would like to thank Eva Viehmann, Xuhua He, Qingchao Yu and Arghya Sadhukhan for inspiring discussions, and Eva Viehmann again for her comments on a preliminary version of this article. I am very grateful for the careful reading and helpful comments of the anonymous referee.

References

[Björner and Brenti 2005] A. Björner and F. Brenti, *Combinatorics of Coxeter groups*, Graduate Texts in Mathematics **231**, Springer, 2005. MR

[Brenti et al. 1999] F. Brenti, S. Fomin, and A. Postnikov, "Mixed Bruhat operators and Yang–Baxter equations for Weyl groups", *Internat. Math. Res. Notices* 8 (1999), 419–441. MR

[Bruhat and Tits 1972] F. Bruhat and J. Tits, "Groupes réductifs sur un corps local", *Inst. Hautes Études Sci. Publ. Math.* 41 (1972), 5–251. MR Zbl

[Chai 2000] C.-L. Chai, "Newton polygons as lattice points", *Amer. J. Math.* **122**:5 (2000), 967–990. MR Zbl

[Dyer 1993] M. J. Dyer, "Hecke algebras and shellings of Bruhat intervals", *Compositio Math.* **89**:1 (1993), 91–115. MR Zbl

[Gashi 2010] Q. R. Gashi, "On a conjecture of Kottwitz and Rapoport", *Ann. Sci. Éc. Norm. Supér.* (4) **43**:6 (2010), 1017–1038. MR

[Görtz et al. 2006] U. Görtz, T. J. Haines, R. E. Kottwitz, and D. C. Reuman, "Dimensions of some affine Deligne–Lusztig varieties", *Ann. Sci. École Norm. Sup.* (4) **39**:3 (2006), 467–511. MR Zbl

[Görtz et al. 2010] U. Görtz, T. J. Haines, R. E. Kottwitz, and D. C. Reuman, "Affine Deligne–Lusztig varieties in affine flag varieties", *Compos. Math.* **146**:5 (2010), 1339–1382. MR Zbl

[Hamacher 2015] P. Hamacher, "The geometry of Newton strata in the reduction modulo p of Shimura varieties of PEL type", *Duke Math. J.* **164**:15 (2015), 2809–2895. MR Zbl

[He 2014] X. He, "Geometric and homological properties of affine Deligne–Lusztig varieties", *Ann. of Math.* (2) **179**:1 (2014), 367–404. MR Zbl

[He 2016] X. He, "Hecke algebras and p-adic groups", pp. 73–135 in *Current developments in mathematics* 2015, edited by D. Jerison et al., International Press, Somerville, MA, 2016. MR Zbl

[He 2021] X. He, "Cordial elements and dimensions of affine Deligne–Lusztig varieties", *Forum Math. Pi* **9** (2021), art. id. e9. MR Zbl

[He and Nie 2024] X. H. He and S. A. Nie, "仿射Weyl群上的Demazur乘积", *Acta Math. Sinica* (*Chinese Ser.*) **67**:2 (2024), 296–306. English version available as "Demazure products of the affine Weyl groups" at https://arxiv.org/abs/2112.06376v1. MR Zbl

[He and Yu 2021] X. He and Q. Yu, "Dimension formula for the affine Deligne–Lusztig variety $X(\mu, b)$", *Math. Ann.* **379**:3-4 (2021), 1747–1765. MR Zbl

[Kamnitzer 2010] J. Kamnitzer, "Mirković–Vilonen cycles and polytopes", *Ann. of Math.* (2) **171**:1 (2010), 245–294. MR Zbl

[Kisin and Pappas 2018] M. Kisin and G. Pappas, "Integral models of Shimura varieties with parahoric level structure", *Publ. Math. Inst. Hautes Études Sci.* **128** (2018), 121–218. MR Zbl

[Kottwitz 1985] R. E. Kottwitz, "Isocrystals with additional structure", *Compositio Math.* **56**:2 (1985), 201–220. MR Zbl

[Kottwitz 1997] R. E. Kottwitz, "Isocrystals with additional structure, II", *Compositio Math.* **109**:3 (1997), 255–339. MR Zbl

[Kottwitz 2006] R. E. Kottwitz, "Dimensions of Newton strata in the adjoint quotient of reductive groups", *Pure Appl. Math. Q.* **2**:3 (2006), 817–836. MR Zbl

[Lam and Shimozono 2010] T. Lam and M. Shimozono, "Quantum cohomology of G/P and homology of affine Grassmannian", *Acta Math.* **204**:1 (2010), 49–90. MR Zbl

[Lenart et al. 2015] C. Lenart, S. Naito, D. Sagaki, A. Schilling, and M. Shimozono, "A uniform model for Kirillov–Reshetikhin crystals, I: Lifting the parabolic quantum Bruhat graph", *Int. Math. Res. Not.* **2015**:7 (2015), 1848–1901. MR Zbl

[Milićević 2021] E. Milićević, “Maximal Newton points and the quantum Bruhat graph”, *Michigan Math. J.* **70**:3 (2021), 451–502. MR Zbl

[Milićević and Viehmann 2020] E. Milićević and E. Viehmann, “Generic Newton points and the Newton poset in Iwahori-double cosets”, *Forum Math. Sigma* **8** (2020), art. id. e50. MR Zbl

[Milićević et al. 2019] E. Milićević, P. Schwer, and A. Thomas, *Dimensions of affine Deligne–Lusztig varieties: a new approach via labeled folded alcove walks and root operators*, Mem. Amer. Math. Soc. **1260**, Amer. Math. Soc., Providence, RI, 2019. MR Zbl

[Naito and Watanabe 2017] S. Naito and H. Watanabe, “A combinatorial formula expressing periodic R-polynomials”, *J. Combin. Theory Ser. A* **148** (2017), 197–243. MR Zbl

[Nie 2022] S. Nie, “Irreducible components of affine Deligne–Lusztig varieties”, *Camb. J. Math.* **10**:2 (2022), 433–510. MR Zbl

[Papi 1994] P. Papi, “A characterization of a special ordering in a root system”, *Proc. Amer. Math. Soc.* **120**:3 (1994), 661–665. MR Zbl

[Pappas 2023] G. Pappas, “On integral models of Shimura varieties”, *Math. Ann.* **385**:3-4 (2023), 2037–2097. MR Zbl

[Parkinson et al. 2009] J. Parkinson, A. Ram, and C. Schwer, “Combinatorics in affine flag varieties”, *J. Algebra* **321**:11 (2009), 3469–3493. MR Zbl

[Postnikov 2005] A. Postnikov, “Quantum Bruhat graph and Schubert polynomials”, *Proc. Amer. Math. Soc.* **133**:3 (2005), 699–709. MR Zbl

[Rapoport 2005] M. Rapoport, “A guide to the reduction modulo p of Shimura varieties: automorphic forms, I”, pp. 271–318 in *Formes automorphes, I: Actes du semestre du centre*, edited by J. Tilouine et al., Astérisque **298**, Soc. Math. France, Paris, 2005. MR Zbl

[Rapoport and Richartz 1996] M. Rapoport and M. Richartz, “On the classification and specialization of F-isocrystals with additional structure”, *Compositio Math.* **103**:2 (1996), 153–181. MR Zbl

[Sadhukhan 2023] A. Sadhukhan, “Affine Deligne–Lusztig varieties and quantum Bruhat graph”, *Math. Z.* **303**:1 (2023), art. id. 21. MR Zbl

[Schremmer 2022] F. Schremmer, “Generic Newton points and cordial elements”, preprint, 2022. arXiv 2205.02039

[Schremmer 2024] F. Schremmer, “Affine Bruhat order and Demazure products”, *Forum Math. Sigma* **12** (2024), art. id. e53. MR Zbl

[Springer 1998] T. A. Springer, *Linear algebraic groups*, 2nd ed., Progress in Mathematics **9**, Birkhäuser, Boston, MA, 1998. MR Zbl

[Takaya 2025] Y. Takaya, “Equidimensionality of affine Deligne–Lusztig varieties in mixed characteristic”, *Adv. Math.* **465** (2025), art. id. 110153. MR Zbl

[Viehmann 2006] E. Viehmann, “The dimension of some affine Deligne–Lusztig varieties”, *Ann. Sci. École Norm. Sup.* (4) **39**:3 (2006), 513–526. MR Zbl

[Viehmann 2014] E. Viehmann, “Truncations of level 1 of elements in the loop group of a reductive group”, *Ann. of Math.* (2) **179**:3 (2014), 1009–1040. MR Zbl

[Viehmann 2018] E. Viehmann, “Moduli spaces of local **G**-shtukas”, pp. 1425–1445 in *Proceedings of the International Congress of Mathematicians — Rio de Janeiro* 2018, *Vol. II: Invited lectures*, edited by B. Sirakov et al., World Sci., Hackensack, NJ, 2018. MR Zbl

[Viehmann 2021] E. Viehmann, “Minimal Newton strata in Iwahori double cosets”, *Int. Math. Res. Not.* **2021**:7 (2021), 5349–5365. MR Zbl

[Zhou and Zhu 2020] R. Zhou and Y. Zhu, “Twisted orbital integrals and irreducible components of affine Deligne–Lusztig varieties”, *Camb. J. Math.* **8**:1 (2020), 149–241. MR Zbl

Communicated by Wee Teck Gan
Received 2023-06-12 Revised 2024-08-08 Accepted 2024-10-18

schr@hku.hk *Department of Mathematics and New Cornerstone Laboratory, The University of Hong Kong, Hong Kong*

ALGEBRA AND NUMBER THEORY 19:10 (2025)
https://doi.org/10.2140/ant.2025.19.2015

Affine Deligne–Lusztig varieties via the double Bruhat graph, II: Iwahori–Hecke algebra

Felix Schremmer

We introduce a new language to describe the geometry of affine Deligne–Lusztig varieties in affine flag varieties. This second part of a two-paper series uses this new language, i.e., the double Bruhat graph, to describe certain structure constants of the Iwahori–Hecke algebra. As an application, we describe nonemptiness and dimension of affine Deligne–Lusztig varieties for most elements of the affine Weyl group and arbitrary σ-conjugacy classes.

1. Introduction

In a seminal paper, Deligne and Lusztig [1976] introduced a class of varieties, which they use to describe many representations of finite groups of Lie type. An analogous construction yields the so-called affine Deligne–Lusztig varieties, which play an important role, e.g., in the reduction of Shimura varieties [Rapoport 2005; He 2018]. Continuing the treatment of [Schremmer 2025], we study affine Deligne–Lusztig varieties in affine flag varieties.

Let G be a reductive group defined over a local field F, whose completion of the maximal unramified extension we denote by $\breve{F}$. Denote the Frobenius of $\breve{F}/F$ by σ and pick a σ-stable Iwahori subgroup $I \subseteq G(\breve{F})$. The affine Deligne–Lusztig variety $X_x(b)$ associated to two elements $x, b \in G(\breve{F})$ is the reduced ind-subscheme of the affine flag variety $G(\breve{F})/I$ with geometric points

$$X_x(b) = \{g \in G(\breve{F})/I \mid g^{-1} b\sigma(g) \in IxI\}.$$

Observe that the isomorphism type of $X_x(b)$ only depends on the σ-conjugacy class

$$[b] = \{g^{-1} b\sigma(g) \mid g \in G(\breve{F})\}$$

and the Iwahori double coset $IxI \subseteq G(\breve{F})$. These Iwahori double cosets are naturally parametrized by the extended affine Weyl group $\widetilde{W}$ of G, and we get

$$G(\breve{F}) = \bigsqcup_{x \in \widetilde{W}} I\dot{x}I.$$

MSC2020: 11G25, 20C08, 20G25.
Keywords: affine Weyl group, Iwahori-Hecke algebra, double Bruhat graph, class polynomial, structure constant, affine Deligne-Lusztig variety, Shimura variety, Langlands program.

Many geometric properties of the double cosets $I\dot{x}I$ for various $x \in \widetilde{W}$ can be understood via the corresponding Iwahori–Hecke algebra $\mathcal{H} = \mathcal{H}(\widetilde{W})$. This algebra and its representation theory received tremendous interest since the discovery of the Satake isomorphism [1963]. There are a few different and mostly equivalent constructions of this algebra in use. For now, we summarize that this is an algebra over a suitable base field or ring with a basis given by formal variables T_x for $x \in \widetilde{W}$. The element $T_x \in \mathcal{H}$ can be thought of as the representation-theoretic analogue of the Iwahori double coset $IxI \subseteq G(\breve{F})$. For example, if $x, y \in \widetilde{W}$, we can write

$$IxI \cdot IyI = \bigcup_z IzI,$$

where the union is taken over all $z \in \widetilde{W}$ such that the T_z-coefficient of $T_x T_y \in \mathcal{H}$ is nonzero. For a general overview over the structure theory of Iwahori–Hecke algebras and its applications to the geometry of the affine flag variety, we refer to [He 2016].

The set of σ-conjugacy classes $B(G) = \{[b] \mid b \in G(\breve{F})\}$ is the second main object of interest in the definition of affine Deligne–Lusztig varieties. It is a celebrated result of Kottwitz [1985; 1997] that each σ-conjugacy class $[b]$ is uniquely determined by two invariants, known as its Newton point and its Kottwitz point. From [He 2014, Theorem 3.7], we get a parametrization of $B(G)$ using the extended affine Weyl group $\widetilde{W}$. For each $x \in \widetilde{W}$, consider its σ-conjugacy class in $\widetilde{W}$, denoted by

$$\mathcal{O} = \{y^{-1}x\sigma(y) \mid y \in \widetilde{W}\}.$$

Two elements that are σ-conjugate in $\widetilde{W}$ will also be σ-conjugate in $G(\breve{F})$, but the converse does not hold true in general. We obtain a surjective but not injective map

$$\{\sigma\text{-conjugacy classes } \mathcal{O} \subseteq \widetilde{W}\} \to B(G),$$

sending $\mathcal{O}$ to $[\dot{x}] \in B(G)$ for any $x \in \mathcal{O}$.

The analogous construction in the Iwahori–Hecke algebra is the formation of a σ-twisted cocenter, i.e., the quotient of $\mathcal{H}$ by the submodule $[\mathcal{H}, \mathcal{H}]_\sigma$ generated by

$$[h, h']_\sigma = hh' - h'\sigma(h), \quad h, h' \in \mathcal{H}.$$

An important result of He and Nie [2014, Theorem C] gives a full description of this cocenter. For each σ-conjugacy class $\mathcal{O} \subseteq \widetilde{W}$ and any two elements of minimal length $x_1, x_2 \in \mathcal{O}$, they prove that the images of T_{x_1} and T_{x_2} in the cocenter of $\mathcal{H}$ agree. Denoting the common image by $T_{\mathcal{O}}$, they prove moreover that these $T_{\mathcal{O}}$ form a basis of the cocenter, parametrized by all σ-conjugacy classes $\mathcal{O} \subseteq \widetilde{W}$.

With these preferred bases $\{T_x\}$ of $\mathcal{H}$ and $\{T_{\mathcal{O}}\}$ of the quotient, we obtain structure constants expressing the image of each T_x in the cocenter as a linear combination of the $T_{\mathcal{O}}$'s. These are known as class polynomials, so we write

$$T_x \equiv \sum_{\substack{\mathcal{O} \subseteq \widetilde{W} \\ \sigma\text{-conj. class}}} f_{x,\mathcal{O}} T_{\mathcal{O}} \pmod{[\mathcal{H}, \mathcal{H}]_\sigma}.$$

These representation-theoretic structure constants are often hard to determine. However, they are very useful for studying affine Deligne–Lusztig varieties, especially the following main three questions:

(Q1) When is $X_x(b)$ empty? Equivalently, when is the Newton stratum empty?

(Q2) If $X_x(b) \neq \varnothing$, what is its dimension?

(Q3) How many top-dimensional irreducible components, up to the action of the σ-centralizer of b, does $X_x(b)$ have?

It is an important result of He that these main questions can be fully answered in terms of the class polynomials; see [He 2014, Theorem 6.1; 2016, Theorem 2.19]. The class polynomials can moreover be used to count rational points of Newton strata; see [He et al. 2024, Proposition 3.7].

In the previous article [Schremmer 2025], we showed that the same main questions can also be answered, in some cases, using the combinatorial notion of a double Bruhat graph. This is an explicitly described finite graph, introduced in [Naito and Watanabe 2017, Section 5.1] in order to describe periodic R-polynomials. Following a result of Görtz, Haines, Kottwitz and Reumann [Görtz et al. 2006, Section 6] comparing affine Deligne–Lusztig varieties with certain intersections in the affine flag variety, we showed that the double Bruhat graph appears naturally as a way to encode certain subvarieties of the affine flag variety.

Write $x = wt^{\mu} \in \widetilde{W},\ v \in W$, and assume that a regularity condition of the form

$$\forall \alpha \in \Phi^+, \quad \langle v^{-1}\mu, \alpha\rangle \gg \langle \mu^{\mathrm{dom}} - \nu(b), 2\rho\rangle$$

is satisfied. Assume moreover that the group G is split over F. Then [Schremmer 2025, Corollary 5.9] shows that the questions of nonemptiness, dimension and top-dimensional irreducible components are determined by the set of paths from v to wv in the double Bruhat graph that are increasing with respect to some fixed reflection order $\prec$ and of weight $\mu^{\mathrm{dom}} - \nu(b)$. Our first main result states that this set of paths determines the full class polynomial, and that the assumption of a split group can be removed.

Theorem 1.1 (see Theorem 4.10). *Assume that the group G is quasisplit. Let $x = w\varepsilon^{\mu} \in \widetilde{W},\ v \in W$ and $\mathcal{O} \subseteq \widetilde{W}$ such that a regularity condition of the form*

$$\forall \alpha \in \Phi^+, \quad \langle v^{-1}\mu, \alpha\rangle \gg \langle \mu^{\mathrm{dom}} - \nu(\mathcal{O}), 2\rho\rangle$$

is satisfied. Then the class polynomial $f_{x,\mathcal{O}}$ can be expressed in terms of paths in the double Bruhat graph from v to $\sigma(wv)$ that are increasing with respect to some fixed reflection order. For a suitable parametrization of the Iwahori–Hecke algebra as an algebra over the polynomial ring $\mathbb{Z}[Q]$ (Definition 4.1), the class polynomial is explicitly given by

$$f_{x,\mathcal{O}} = \sum_{p} Q^{\ell(p)},$$

where the sum is taken over all paths p in the double Bruhat graph from v to $\sigma(wv)$ that are increasing with respect to some fixed reflection order and such that $\nu(\mathcal{O})$ is the σ-average of $v^{-1}\mu - \mathrm{wt}(p)$.

The assumption of a quasisplit group can be removed following [Görtz et al. 2015, Section 2], though it requires more cumbersome notation to write down statements in full generality; see [Schremmer 2022, Section 4.2].

We will prove Theorem 1.1 as a consequence of the following more fundamental result, computing the structure constants of the multiplication of our standard basis vectors in $\mathcal{H}$.

Theorem 1.2 (see Theorem 4.2). *Let $x = w_x\varepsilon^{\mu_x}$, $z = w_z\varepsilon^{\mu_z} \in \widetilde{W}$, and $v_z \in W$ satisfying a regularity condition of the form*

$$\forall\alpha \in \Phi^+, \quad \langle v_z^{-1}\mu_z, \alpha\rangle \gg \ell(x).$$

Define polynomials $\varphi_{x,z,y}$ via

$$T_x T_z = \sum_{y\in\widetilde{W}} \varphi_{x,z,y} T_y \in \mathcal{H}(\widetilde{W}).$$

Pick an element $y = w_y\varepsilon^{\mu_y} \in \widetilde{W}$ and $v_x \in W$ such that a regularity condition of the form

$$\forall\alpha \in \Phi^+, \quad \langle v_x^{-1}\mu_x, \alpha\rangle \gg \ell(x) + \ell(z) - \ell(y)$$

is satisfied. Then we can describe the structure constant $\varphi_{x,z,y}$ in terms of paths in the double Bruhat graph. Explicitly, we have $\varphi_{x,z,y} = 0$ unless $w_y = (w_x v_x)^{-1} v_y$. In this case, we have

$$\varphi_{x,z,y} = \sum_p Q^{\ell(p)},$$

where the sum is taken over all paths in the double Bruhat graph from v_x to $w_z v_z$ that are increasing with respect to some reflection order and of weight

$$\mathrm{wt}(p) = v_x^{-1}\mu_x + v_z^{-1}\mu_z - (w_z v_z)^{-1}\mu_y.$$

Theorem 4.2 below actually proves a stronger statement, requiring only a weaker regularity condition of the form

$$\forall\alpha \in \Phi^+, \quad \langle v^{-1}\mu_x, \alpha\rangle \gg \ell(x) - \ell(y^{-1}z).$$

The resulting description of $\varphi_{x,z,y}$ is more involved, however, replacing the single path p by pairs of bounded paths in the double Bruhat graph. Theorem 1.2 as stated here is sufficient to derive Theorem 1.1.

So under some very strong regularity conditions, the double Bruhat graph may also be used to understand multiplications of Iwahori double cosets $IxI \cdot IzI$ in $G(\breve{F})$. Theorems 1.1 and 1.2 give insight in the generic behaviour of class polynomials and products in the Iwahori–Hecke algebra, solving infinitely many previously intractable questions using a finite combinatorial object. From a practical point of view, this allows us to quickly derive many crucial properties of the weight multisets of the double Bruhat graph by referring to known properties of the Iwahori–Hecke algebra or affine Deligne–Lusztig varieties. Using some of the most powerful tools available to describe affine Deligne–Lusztig varieties and comparing them to the double Bruhat graph, we obtain the following result.

Theorem 1.3 (see Theorem 5.4). *Let $x = w\varepsilon^{\mu} \in \widetilde{W}$ and $v \in W$, satisfying the regularity condition*

$$\forall\alpha \in \Phi^+, \quad \langle v^{-1}\mu, \alpha\rangle \geq 2\,\mathrm{rk}(G) + 14,$$

where $\mathrm{rk}(G)$ is the rank of a maximal torus in the group G.

Pick an arbitrary σ-conjugacy class $[b] \in B(G)$. Let P be the set of all paths p in the double Bruhat graph from v to $\sigma(wv)$ that are increasing with respect to some fixed reflection order such that the

***λ-invariant** of* $[b]$ (*see* [Hamacher and Viehmann 2018, Section 2]) *satisfies*

$$\lambda(b) = v^{-1}\mu - \mathrm{wt}(p).$$

Then $P \neq \varnothing$ *if and only if* $X_x(b) \neq \varnothing$. *If* p *is a path of maximal length in* P, *then*

$$\dim X_x(b) = \tfrac{1}{2}\big(\ell(x) + \ell(p) - \langle \nu(b), 2\rho\rangle - \mathrm{def}(b)\big).$$

We give a similar description in terms of the dominant Newton points of $[b]$ *rather than the* λ*-invariant.*

Theorem 1.3 gives full answers to the questions (Q1) and (Q2) for arbitrary $[b] \in B(G)$ as long as the element $x \in \widetilde{W}$ satisfies a somewhat mild regularity condition (being linear in the rank of G).

The proofs given in this article are mostly combinatorial in nature, and largely independent of its predecessor article [Schremmer 2025]. We will rely only on some basic facts on the double Bruhat graph established in [Schremmer 2025, Section 4]. The best known ways to compute the structure constants of Theorem 1.2 and the class polynomials $f_{x,\mathcal{O}}$ are given by certain recursive relations involving simple affine reflections in the extended affine Weyl group. Similarly, the Deligne–Lusztig reduction method [1976] of Görtz and He [2010, Section 2.5] provides such a recursive method to describe many geometric properties of affine Deligne–Lusztig varieties, in particular the ones studied in this paper series. On the double Bruhat side, these are mirrored by the construction of certain bijections between paths due to Naito and Watanabe [2017, Section 3.3]. We recall these bijections and derive the corresponding properties of the weight multisets in Section 3. We study the consequences for the Iwahori–Hecke algebra in Section 4, and the resulting properties of affine Deligne–Lusztig varieties in Section 5.

In Section 6, we finish this series of two papers by listing a number of further-reaching conjectures, predicting a relationship between the geometry of affine Deligne–Lusztig varieties and paths in the double Bruhat graph in various cases. These conjectures are natural generalizations of our results, and withstand an extensive computer search for counterexamples.

Recall that our main goal is to find and prove a description of the geometry of affine Deligne–Lusztig varieties in the affine flag variety that is as concise and precise as the known analogous statements for the affine Grassmannian (as summarized in [Schremmer 2025, Theorem 1.1]). Our conjectures and partial results towards proving them suggest that the language of the double Bruhat graph is very useful for this task, and might even be the crucial missing piece towards a full description.

We would like to remark that once a conjecture is found that describes the geometry of $X_x(b)$ for arbitrary x, b in terms of the double Bruhat graph, a proof of such a conjecture might simply consist of a straightforward comparison of the Deligne–Lusztig reduction method [1976] due to Görtz and He [2010] with the analogous recursive relations of the double Bruhat graph that are discussed in this article.

2. Notation

We fix a nonarchimedean local field F whose completion of the maximal unramified extension will be denoted by $\breve{F}$. We write $\mathcal{O}_F$ and $\mathcal{O}_{\breve{F}}$ for the respective rings of integers. Let $\varepsilon \in F$ be a uniformizer. The Galois group $\Gamma = \mathrm{Gal}(\breve{F}/F)$ is generated by the Frobenius σ.

In the context of Shimura varieties, one would choose F to be a finite extension of the p-adic numbers. When studying moduli spaces of shutkas, F would be the field of Laurent series over a finite field.

In any case, we fix a reductive group G over F. Via [Görtz et al. 2015, Section 2], we may reduce questions regarding affine Deligne–Lusztig varieties of G to the case of a quasisplit group. In order to minimize the notational burden, we assume that the group G is quasisplit throughout this paper.

We construct its associated affine root system and affine Weyl group following [Haines and Rapoport 2008; Tits 1979].

Fix a maximal $\breve{F}$-split torus $T_{\breve{F}} \subseteq G_{\breve{F}}$ and write T for its centralizer in $G_{\breve{F}}$, so T is a maximal torus of $G_{\breve{F}}$. Write $\mathcal{A} = \mathcal{A}(G_{\breve{F}}, T_{\breve{F}})$ for the apartment of the Bruhat–Tits building of $G_{\breve{F}}$ associated with $T_{\breve{F}}$. We pick a σ-invariant alcove $\mathfrak{a}$ in $\mathcal{A}$. Its stabilizer is a σ-invariant Iwahori subgroup $I \subset G(\breve{F})$.

Denote the normalizer of T in G by $N_G(T)$. Then the quotient

$$\widetilde{W} = N_G(T)(\breve{F})/(T(\breve{F}) \cap I)$$

is called the *extended affine Weyl group*, and $W = N_G(T)(\breve{F})/T(\breve{F})$ is the (*finite*) *Weyl group*. The Weyl group W is naturally a quotient of $\widetilde{W}$. We denote the Frobenius action on W and $\widetilde{W}$ by σ as well.

The affine roots as constructed in [Tits 1979, Section 1.6] are denoted by Φ_{af}. Each of these roots $a \in \Phi_{\mathrm{af}}$ defines an affine function $a : \mathcal{A} \to \mathbb{R}$. The vector part of this function is denoted by $\mathrm{cl}(a) \in V^*$, where $V = X_*(S) \otimes \mathbb{R} = X_*(T)_{\Gamma_0} \otimes \mathbb{R}$. Here, $\Gamma_0 = \mathrm{Gal}(\overline{F}/\breve{F})$ is the absolute Galois group of $\breve{F}$, i.e., the inertia group of $\Gamma = \mathrm{Gal}(\overline{F}/F)$. The set of (*finite*) *roots* is[1] $\Phi := \mathrm{cl}(\Phi_{\mathrm{af}})$.

Each affine root in Φ_{af} divides the standard apartment into two half-spaces, one being the positive and one the negative side. Those affine roots where our fixed alcove $\mathfrak{a}$ is on the positive side are called *positive affine roots*. If moreover the alcove $\mathfrak{a}$ is adjacent to the root hyperplane, it is called a *simple affine root*. We denote the sets of simple, resp. positive, affine roots by $\Delta_{\mathrm{af}} \subseteq \Phi_{\mathrm{af}}^+ \subseteq \Phi_{\mathrm{af}}$.

Writing W_{af} for the extended affine Weyl group of G, we get a natural σ-equivariant short exact sequence (see [Haines and Rapoport 2008, Lemma 14])

$$1 \to W_{\mathrm{af}} \to \widetilde{W} \to \pi_1(G)_{\Gamma_0} \to 1.$$

Here, $\pi_1(G) := X_*(T)/\mathbb{Z}\Phi^\vee$ denotes the Borovoi fundamental group.

For each $x \in \widetilde{W}$, we denote by $\ell(x) \in \mathbb{Z}_{\geq 0}$ the length of a shortest alcove path from $\mathfrak{a}$ to $x\mathfrak{a}$. The elements of length zero are denoted by Ω. The above short exact sequence yields an isomorphism of Ω with $\pi_1(G)_{\Gamma_0}$, realizing $\widetilde{W}$ as semidirect product $\widetilde{W} = \Omega \ltimes W_{\mathrm{af}}$.

Each affine root $a \in \Phi_{\mathrm{af}}$ defines an affine reflection r_a on $\mathcal{A}$. The group generated by these reflections is naturally isomorphic to W_{af} (see [Haines and Rapoport 2008]), so by abuse of notation, we also write $r_a \in W_{\mathrm{af}}$ for the corresponding element. We define $S_{\mathrm{af}} := \{r_a \mid a \in \Delta_{\mathrm{af}}\}$, called the set of *simple affine reflections*. The pair $(W_{\mathrm{af}}, S_{\mathrm{af}})$ is a Coxeter group with length function ℓ as defined above.

[1] This is different from the root system that [Tits 1979] and [Haines and Rapoport 2008] denote by Φ; it coincides with the root system called Σ in [Haines and Rapoport 2008].

We pick a special vertex $\mathfrak{x} \in \mathcal{A}$ that is adjacent to $\mathfrak{a}$. Since we assumed G to be quasisplit, we may and do choose $\mathfrak{x}$ to be σ-invariant. We identify $\mathcal{A}$ with V via $\mathfrak{x} \mapsto 0$. This allows us to take the decomposition $\Phi_{\mathrm{af}} = \Phi \times \mathbb{Z}$, where $a = (\alpha, k)$ corresponds to the function

$$V \to \mathbb{R}, \quad v \mapsto \alpha(v) + k.$$

From [Haines and Rapoport 2008, Proposition 13], we moreover get decompositions $\widetilde{W} = W \ltimes X_*(T)_{\Gamma_0}$ and $W_{\mathrm{af}} = W \ltimes \mathbb{Z}\Phi^\vee$. Using this decomposition, we write elements $x \in \widetilde{W}$ as $x = w\varepsilon^\mu$, with $w \in W$ and $\mu \in X_*(T)_{\Gamma_0}$. For $a = (\alpha, k) \in \Phi_{\mathrm{af}}$, we have $r_a = s_\alpha \varepsilon^{k\alpha^\vee} \in W_{\mathrm{af}}$, where $s_\alpha \in W$ is the reflection associated with α. The natural action of $\widetilde{W}$ on Φ_{af} can be expressed as

$$(w\varepsilon^\mu)(\alpha, k) = (w\alpha, k - \langle \mu, \alpha \rangle).$$

We define the *dominant chamber* $C \subseteq V$ to be the Weyl chamber containing our fixed alcove $\mathfrak{a}$. This gives a Borel subgroup $B \subseteq G$, and corresponding sets of positive/negative/simple roots Φ^+, Φ^-, $\Delta \subseteq \Phi$.

By abuse of notation, we denote by Φ^+ also the indicator function of the set of positive roots, i.e.,

$$\forall \alpha \in \Phi, \quad \Phi^+(\alpha) = \begin{cases} 1, & \alpha \in \Phi^+, \\ 0, & \alpha \in \Phi^-. \end{cases}$$

The sets of positive and negative affine roots can be expressed as

$$\begin{aligned} \Phi_{\mathrm{af}}^+ &= (\Phi^+ \times \mathbb{Z}_{\geq 0}) \sqcup (\Phi^- \times \mathbb{Z}_{\geq 1}) = \{(\alpha, k) \in \Phi_{\mathrm{af}} \mid k \geq \Phi^+(-\alpha)\}, \\ \Phi_{\mathrm{af}}^- &= -\Phi_{\mathrm{af}}^+ = \Phi_{\mathrm{af}} \setminus \Phi_{\mathrm{af}}^+ = \{(\alpha, k) \in \Phi_{\mathrm{af}} \mid k < \Phi^+(-\alpha)\}. \end{aligned}$$

One checks that Φ_{af}^+ are precisely the affine roots that are sums of simple affine roots.

Decompose Φ as a direct sum of irreducible root systems, $\Phi = \Phi_1 \sqcup \cdots \sqcup \Phi_c$. Each irreducible factor contains a uniquely determined highest root $\theta_i \in \Phi_i^+$. Now the set of simple affine roots is

$$\Delta_{\mathrm{af}} = \{(\alpha, 0) \mid \alpha \in \Delta\} \cup \{(-\theta_i, 1) \mid i = 1, \ldots, c\} \subset \Phi_{\mathrm{af}}^+.$$

We call an element $\mu \in X_*(T)_{\Gamma_0} \otimes \mathbb{Q}$ *dominant* if $\langle \mu, \alpha \rangle \geq 0$ for all $\alpha \in \Phi^+$. Similarly, we call it C-regular for a real number C if

$$|\langle \mu, \alpha \rangle| \geq C$$

for each $\alpha \in \Phi^+$. If $\mu \in X_*(T)_{\Gamma_0}$ is dominant, then the Newton point of $\varepsilon^\mu \in \widetilde{W}$ is given by the σ-average of μ, defined as

$$\mathrm{avg}_\sigma(\mu) = \frac{1}{N} \sum_{i=1}^{N} \sigma^i(\mu),$$

where $N > 0$ is any integer such that the action of σ^N on $X_*(T)_{\Gamma_0}$ is trivial.

An element $x = w\varepsilon^\mu \in \widetilde{W}$ is called C-regular if μ is. We write $\mathrm{LP}(x) \subseteq W$ for the set of length positive elements as introduced in [Schremmer 2022, Section 2.2]. If x is 2-regular, then $\mathrm{LP}(x)$ consists only of one element, namely the uniquely determined $v \in W$ such that $v^{-1}\mu$ is dominant.

For elements μ, μ' in $X_*(T)_{\Gamma_0} \otimes \mathbb{Q}$ (resp. $X_*(T)_{\Gamma_0}$ or $X_*(T)_\Gamma$), we write $\mu \leq \mu'$ if the difference $\mu' - \mu$ is a $\mathbb{Q}_{\geq 0}$-linear combination of positive coroots.

3. Double Bruhat graph

We recall the definition of the double Bruhat graph following [Naito and Watanabe 2017, Section 5.1]. It turns out that the paths we studied in order to understand affine Deligne–Lusztig varieties are a certain subset of the paths studied by Naito–Watanabe in order to study Kazhdan–Lusztig theory, or more precisely periodic R-polynomials.

Definition 3.1. Let $\prec$ be a total order on Φ^+, and let moreover $v, w \in W$.

(a) The *double Bruhat graph* $\mathrm{DBG}(W)$ is a finite directed graph. Its set of vertices is W. For each $w \in W$ and $\alpha \in \Phi^+$, there is an edge $w \xrightarrow{\alpha} ws_\alpha$.

(b) A *nonlabelled path* $\bar{p}$ in $\mathrm{DBG}(W)$ is a sequence of adjacent edges

$$\bar{p} : v = u_1 \xrightarrow{\alpha_1} u_2 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_\ell} u_{\ell+1} = w.$$

We call $\bar{p}$ a nonlabelled path from v to w of length $\ell(\bar{p}) = \ell$. We say $\bar{p}$ is *increasing* with respect to $\prec$ if $\alpha_1 \prec \cdots \prec \alpha_\ell$. In this case, we moreover say that $\bar{p}$ *is bounded by* $n \in \mathbb{Z}$ if $\alpha_\ell = \beta_i$ for some $i \leq n$.

(c) A *labelled path* or *path* p in $\mathrm{DBG}(W)$ consists of an unlabelled path

$$\bar{p} : v = u_1 \xrightarrow{\alpha_1} u_2 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_\ell} u_{\ell+1} = w$$

together with integers $m_1, \ldots, m_\ell \in \mathbb{Z}$ subject to the condition

$$m_i \geq \Phi^+(-u_i\alpha_i) = \begin{cases} 0, & \ell(u_{i+1}) > \ell(u_i), \\ 1, & \ell(u_{i+1}) < \ell(u_i). \end{cases}$$

We write p as

$$p : v = u_1 \xrightarrow{(\alpha_1, m_1)} u_2 \xrightarrow{(\alpha_2, m_2)} \cdots \xrightarrow{(\alpha_\ell, m_\ell)} u_{\ell+1} = w.$$

The *weight* of p is

$$\mathrm{wt}(p) = m_1\alpha_1^\vee + \cdots + m_\ell\alpha_\ell^\vee \in \mathbb{Z}\Phi^\vee.$$

The *length* of p is $\ell(p) = \ell(\bar{p}) = \ell$. We say that p is *increasing* with respect to $\prec$ if $\bar{p}$ is. In this case, we say that p is bounded by $n \in \mathbb{Z}$ if $\bar{p}$ is.

(d) The set of all paths from v to w that are increasing with respect to $\prec$ and bounded by $n \in \mathbb{Z}$ is denoted by $\mathrm{paths}^{\prec}_{\leq n}(v \Rightarrow w)$. We also write

$$\mathrm{paths}^{\prec}(v \Rightarrow w) = \mathrm{paths}^{\prec}_{\leq \#\Phi^+}(v \Rightarrow w).$$

(e) The order $\prec$ is called a *reflection order* if, for all roots $\alpha, \beta \in \Phi^+$ with $\alpha + \beta \in \Phi^+$, we have

$$\alpha \prec \alpha + \beta \prec \beta \quad \text{or} \quad \beta \prec \alpha + \beta \prec \alpha.$$

We will frequently use the immediate properties of these paths as developed in [Schremmer 2025, Section 4]. For this section, our main result describes how these paths behave with respect to certain simple affine reflections. Fix a reflection order

$$\Phi^+ = \{\beta_1 \prec \cdots \prec \beta_{\#\Phi^+}\}$$

and write

$$\pi_{\succ n} = s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}} \in W$$

as in [Schremmer 2025, Definition 4.10].

Theorem 3.2. *Let $u, v \in W$ and $n \in \{0, \dots, \#\Phi^+\}$. Pick a simple affine root $a = (\alpha, k) \in \Delta_{\mathrm{af}}$ such that $(v\pi_{\succ n})^{-1}\alpha \in \Phi^-$.*

(a) *If $u^{-1}\alpha \in \Phi^-$, then there exists an explicitly described bijection of paths*

$$\psi : \mathrm{paths}^{\prec}_{\preceq n}(s_\alpha u \Rightarrow s_\alpha v) \to \mathrm{paths}^{\prec}_{\preceq n}(u \Rightarrow v)$$

satisfying for each $p \in \mathrm{paths}^{\prec}_{\preceq n}(s_\alpha u \Rightarrow s_\alpha v)$ the conditions

$$\ell(\psi(p)) = \ell(p), \quad \mathrm{wt}(\psi(p)) = \mathrm{wt}(p) + k(v^{-1}\alpha^\vee - u^{-1}\alpha^\vee).$$

(b) *If $u^{-1}\alpha \in \Phi^+$, then there exists an explicitly described bijection of paths*

$$\varphi : \mathrm{paths}^{\prec}_{\preceq n}(s_\alpha u \Rightarrow s_\alpha v) \sqcup \mathrm{paths}^{\prec}_{\preceq n}(s_\alpha u \Rightarrow v) \to \mathrm{paths}^{\prec}_{\preceq n}(u \Rightarrow v)$$

satisfying for each $p \in \mathrm{paths}^{\prec}_{\preceq n}(s_\alpha u \Rightarrow s_\alpha v)$ and $p' \in \mathrm{paths}^{\prec}_{\preceq n}(s_\alpha u \Rightarrow v)$ the conditions

$$\ell(\varphi(p)) = \ell(p), \qquad \mathrm{wt}(\varphi(p)) = \mathrm{wt}(p) + k(v^{-1}\alpha^\vee - u^{-1}\alpha^\vee),$$
$$\ell(\varphi(p')) = \ell(p') + 1, \quad \mathrm{wt}(\varphi(p')) = \mathrm{wt}(p') - ku^{-1}\alpha^\vee.$$

The proof of this theorem can essentially be found in Section 3.3 of [Naito and Watanabe 2017], which is a rather involved and technical construction. One may obtain a weaker version of Theorem 3.2 by comparing the action of simple affine reflections on semi-infinite orbits with [Schremmer 2025, Theorem 4.6]. While such a weaker result would be sufficient for our geometric applications, we do need the full strength of Theorem 3.2 for our conclusions on the Iwahori–Hecke algebra. Moreover, we would like to explain the connection between our paper and [Naito and Watanabe 2017]. Let us hence recall some of the notation used by Naito and Watanabe:

Definition 3.3. (a) By $\leq_{\infty/2}$, we denote the semi-infinite order on $\widetilde{W}$ as introduced in [Lusztig 1980]. It is generated by inequalities of the form

$$w\varepsilon^\mu <_{\infty/2} r_{(\alpha,k)} w\varepsilon^\mu,$$

where $(\alpha, k) \in \Phi^+_{\mathrm{af}}$, $w \in W$ and $\mu \in X_*(T)_{\Gamma_0}$ satisfy $w^{-1}\alpha \in \Phi^+$.

(b) For $w, y \in \widetilde{W}$, we denote by $P^{\prec}_r(y, w)$ the set of paths in $\widetilde{W}$ of the form

$$\Pi : y = y_1 \xrightarrow{(\beta_1, m_1)} y_2 \xrightarrow{(\beta_2, m_2)} \cdots \xrightarrow{(\beta_\ell, m_\ell)} y_{\ell+1} = w$$

such that the following two conditions are both satisfied:

- For each $i = 1, \dots, \ell$, we have $y_{i+1} >_{\infty/2} y_i$. Writing $y_i = w_i \varepsilon^{\mu_i}$, we have
$$y_{i+1} = w_i s_{\beta_i} \varepsilon^{\mu_i + m_i \beta_i^\vee}.$$
- The roots β_i are all positive and satisfy $\beta_1 \prec \cdots \prec \beta_\ell$.

We denote the number of edges in Π by $\ell(\Pi) := \ell$.

These paths $P_r^\prec(\,\cdot\,,\cdot\,)$ occur with exactly the same name in [Naito and Watanabe 2017] and are called translation-free paths. They also consider a larger set of paths, where so-called translation edges are allowed, which is however less relevant for our applications.

From the definition of the semi-infinite order, we easily obtain the following relation between the paths in $\widetilde{W}$ and the paths in the double Bruhat graph. This can be seen as a variant of [Naito and Watanabe 2017, Proposition 5.2.1].

Lemma 3.4. *Let $y = w_1 \varepsilon^{\mu_1}$, $w = w_2 \varepsilon^{\mu_2} \in \widetilde{W}$. Then the map*
$$\Psi : P_r^\prec(y, w) \to \{p \in \mathrm{paths}^\prec(w_1 \Rightarrow w_2) \mid \mathrm{wt}(p) = \mu_2 - \mu_1\},$$
$$\big(\Pi : y = y_0 \xrightarrow{(\beta_1, m_1)} y_1 \xrightarrow{(\beta_2, m_2)} \cdots \xrightarrow{(\beta_\ell, m_\ell)} y_{\ell+1} = w\big)$$
$$\mapsto \big(\Phi(\Pi) : w_1 = \mathrm{cl}(y_0) \xrightarrow{(\beta_1, m_1)} \mathrm{cl}(y_1) \xrightarrow{(\beta_2, m_2)} \cdots \xrightarrow{(\beta_\ell, m_\ell)} \mathrm{cl}(y_{\ell+1})\big),$$
is bijective and length-preserving (*i.e.*, $\ell(\Psi(\Pi)) = \ell(\Pi)$). □

The main results of [Naito and Watanabe 2017, Section 3.3] can be summarized as follows.

Theorem 3.5. *Let $y, w \in \widetilde{W}$ and pick a simple affine reflection $s \in S_{\mathrm{af}}$ such that $y <_{\infty/2} sy$ and $sw <_{\infty/2} w$.*

(a) [Naito and Watanabe 2017, Proposition 3.3.2]: *There is an explicitly described bijection*
$$\psi : P_r^\prec(y, sw) \to P_r^\prec(sy, w).$$
The map ψ preserves the lengths of paths. Its inverse map $\psi' = \psi^{-1}$ is also explicitly described.

(b) [Naito and Watanabe 2017, Proposition 3.3.1]: *There is an explicitly described bijection*
$$\varphi : P_r^\prec(sy, sw) \sqcup P_r^\prec(sy, w) \to P_r^\prec(y, w).$$
For $\Pi \in P_r^\prec(sy, sw)$, we have $\ell(\varphi(\Pi)) = \ell(\Pi)$. For $\Pi \in P_r^\prec(sy, w)$, we have $\ell(\varphi(\Pi)) = \ell(\Pi) + 1$. Its inverse map $\varphi' = \varphi^{-1}$ is also explicitly described. □

In view of Lemma 3.4, we immediately get the special case of Theorem 3.2 for the sets $\mathrm{paths}^\prec(u \Rightarrow v)$, i.e., if $n = \#\Phi^+$. By inspecting the proof and the explicit constructions involved in the proof of Theorem 3.5, we will obtain the full statement of Theorem 3.2. In order to facilitate this task, we introduce a technique that we call "path padding".

Definition 3.6. Let $u, v \in W$ and $0 \le n \le \#\Phi^+$. Fix positive integers m_i for $i = 1, \dots, \#\Phi^+$. Then we define the *padding map*
$$\mathrm{pad}_{(m_i)} : \mathrm{paths}^\prec_{\le n}(u \Rightarrow v) \to \mathrm{paths}^\prec(u \Rightarrow v\pi_{\succ n}),$$

sending a path $p \in \mathrm{paths}^{\prec}_{\leq n}(u \Rightarrow v)$ to the composite path

$$\mathrm{pad}_{(m_i)}(p) : u \overset{p}{\Rightarrow} v \xrightarrow{(\beta_{n+1}, m_{n+1})} v s_{\beta_{n+1}} \xrightarrow{(\beta_{n+2}, m_{n+2})} \cdots \xrightarrow{(\beta_{\#\Phi^+}, m_{\#\Phi^+})} v s_{\beta_{n+1}} \cdots s_{\beta_{\#\Phi^+}} = v\pi_{\succ n}.$$

Lemma 3.7. *Let $u, v \in W$ and $0 \leq n \leq \#\Phi^+$. Pick a simple affine root $a = (\alpha, k) \in \Delta_{\mathrm{af}}$ such that $(v\pi_{\succ n})^{-1}\alpha \in \Phi^-$.*

(a) *Suppose that $u^{-1}\alpha \in \Phi^-$. For each collection of integers $(m_i \geq 4)_{1 \leq i \leq \#\Phi^+}$, there is a unique map*

$$\tilde{\psi} : \mathrm{paths}^{\prec}_{\leq n}(s_\alpha u \Rightarrow s_\alpha v) \to \mathrm{paths}^{\prec}_{\leq n}(u \Rightarrow v)$$

and a collection of integers $(m_i' \geq m_i - 3)_{1 \leq i \leq \#\Phi^+}$ such that the following diagram commutes:

$$\begin{array}{ccccc}
\mathrm{paths}^{\prec}_{\leq n}(s_\alpha u \Rightarrow s_\alpha v) & \xhookrightarrow{\mathrm{pad}_{(m_i)}} & \mathrm{paths}^{\prec}(s_\alpha u \Rightarrow s_\alpha v\pi_{\succ n}) & \xrightarrow[\sim]{\Psi^{-1}} & \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(r_a u, r_a v\pi_{\succ n}\varepsilon^\mu) \\
\Big\downarrow{\scriptstyle \tilde{\psi}} & & & & \sim\Big\downarrow{\scriptstyle \psi} \\
\mathrm{paths}^{\prec}_{\leq n}(u \Rightarrow v) & \xhookrightarrow{\mathrm{pad}_{(m_i')}} & \mathrm{paths}^{\prec}(u \Rightarrow v\pi_{\succ n}) & \xrightarrow[\sim]{\Psi^{-1}} & \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(u, v\pi_{\succ n}\varepsilon^\mu)
\end{array}$$

The map ψ on the right comes from Theorem 3.5(a). The map $\tilde{\psi}$ has an explicit description independent of the integers (m_i). Moreover, $\tilde{\psi}$ satisfies the weight and length constraints as required in Theorem 3.2(a).

Similarly, there exist integers $(m_i'' \geq m_i - 3)_i$ and a uniquely determined and explicitly described map $\tilde{\psi}'$ making the following diagram commute:

$$\begin{array}{ccccc}
\mathrm{paths}^{\prec}_{\leq n}(u \Rightarrow v) & \xhookrightarrow{\mathrm{pad}_{(m_i)}} & \mathrm{paths}^{\prec}(u \Rightarrow v\pi_{\succ n}) & \xrightarrow[\sim]{\Psi^{-1}} & \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(u, v\pi_{\succ n}\varepsilon^\mu) \\
\Big\downarrow{\scriptstyle \tilde{\psi}'} & & & & \sim\Big\downarrow{\scriptstyle \psi'} \\
\mathrm{paths}^{\prec}_{\leq n}(s_\alpha u \Rightarrow s_\alpha v) & \xhookrightarrow{\mathrm{pad}_{(m_i'')}} & \mathrm{paths}^{\prec}(s_\alpha u \Rightarrow s_\alpha v\pi_{\succ n}) & \xrightarrow[\sim]{\Psi^{-1}} & \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(r_a u, r_a v\pi_{\succ n}\varepsilon^\mu)
\end{array}$$

(b) *Suppose that $u^{-1}\alpha \in \Phi^+$. For each collection of integers $(m_i \geq 4)_{1 \leq i \leq \#\Phi^+}$, the explicitly described maps*

$$\begin{aligned}
\varphi_1 &: \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(r_a u, r_a v\pi_{\succ n}\varepsilon^\mu) \to \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(u, v\pi_{\succ n}\varepsilon^\mu), \\
\varphi_2 &: \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(r_a u, v\pi_{\succ n}\varepsilon^\mu) \to \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(u, v\pi_{\succ n}\varepsilon^\mu), \\
\varphi' &: \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(u, v\pi_{\succ n}\varepsilon^\mu) \to \bigsqcup_{\mu \in \mathbb{Z}\Phi^\vee} P_r^{\prec}(r_a u, r_a v\pi_{\succ n}\varepsilon^\mu) \sqcup P_r^{\prec}(r_a u, v\pi_{\succ n}\varepsilon^\mu)
\end{aligned}$$

from Theorem 3.5(b) can be lifted, up to padding and Ψ^{-1} as in (a), to uniquely determined maps

$$\begin{aligned}
\tilde{\varphi}_1 &: \mathrm{paths}^{\prec}_{\leq n}(s_\alpha u \Rightarrow s_\alpha v) \to \mathrm{paths}^{\prec}_{\leq n}(u \Rightarrow v), \\
\tilde{\varphi}_2 &: \mathrm{paths}^{\prec}_{\leq n}(s_\alpha u \Rightarrow v) \to \mathrm{paths}^{\prec}_{\leq n}(u \Rightarrow v), \\
\tilde{\varphi}' &: \mathrm{paths}^{\prec}_{\leq n}(u \Rightarrow v) \to \mathrm{paths}^{\prec}_{\leq n}(s_\alpha u \Rightarrow s_\alpha v) \sqcup \mathrm{paths}^{\prec}_{\leq n}(s_\alpha u \Rightarrow s_\alpha v).
\end{aligned}$$

All three maps are explicitly described in a way that is independent of the integers (m_i). The maps φ_1 and φ_2 moreover satisfy the desired length and weight compatibility relations from Theorem 3.2(b).

Proof. We only explain how to obtain the map $\tilde{\psi}$ from the map ψ, as the other cases are analogous. So pick any path $p \in \mathrm{paths}^{\prec}_{\leq n}(s_\alpha u \Rightarrow s_\alpha v)$. Write it as

$$p : s_\alpha u = w_1 \xrightarrow{(\gamma_1, n_1)} w_2 \xrightarrow{(\gamma_2, n_2)} \cdots \xrightarrow{(\gamma_{\ell(p)}, n_{\ell(p)})} w_{\ell(p)+1} = s_\alpha v.$$

Then

$$\mathrm{pad}_{(m_i)}(p) : s_\alpha u = w_1 \xrightarrow{(\gamma_1, n_1)} \cdots \xrightarrow{(\gamma_{\ell(p)}, n_{\ell(p)})} s_\alpha w_{\ell(p)+1} = s_\alpha v \xrightarrow{(\beta_{n+1}, m_{n+1})} \cdots \xrightarrow{(\beta_{\#\Phi^+}, m_{\#\Phi^+})} s_\alpha v \pi_{\succ n}.$$

Define $\gamma_{\ell(p)+i} = \beta_{n+i}$ and $n_{\ell(p)+i} = m_{\ell(p)+i}$ for $i = 1, \dots, \#\Phi^+ - \ell(p)$. Then we can write

$$\mathrm{pad}_{(m_i)}(p) : s_\alpha u = w_1 \xrightarrow{(\gamma_1, n_1)} \cdots \xrightarrow{(\gamma_{\ell'}, n_{\ell'})} w_{\ell'+1} = v\pi_{\succ n}$$

such that $\ell' = \ell(p) + (\#\Phi^+ - n)$. Writing $\mu := \mathrm{wt}(\mathrm{pad}_{(m_i)}(p)) + k((v\pi_{\succ n})^{-1}\alpha^\vee - u^{-1}\alpha^\vee)$, we may express the path $\Pi := \Psi^{-1}(\mathrm{pad}_{(m_i)}(p)) \in P_r^{\prec}(r_a u, r_a v \pi_{\succ \beta_n} \varepsilon^\mu)$ as

$$\Pi : r_a u = w_1 \varepsilon^{-k w_1^{-1}\alpha^\vee} \xrightarrow{(\gamma_1, n_1)} w_2 \varepsilon^{n_1 \gamma_1^\vee - k w_1^{-1}\alpha^\vee} \xrightarrow{(\gamma_2, k_2)} \cdots \xrightarrow{(\gamma_{\ell'}, n_{\ell'})} w_{\ell'+1}\varepsilon^{\mathrm{wt}(\mathrm{pad}_{(m_i)}(p)) - k w_1^{-1}\alpha^\vee} = r_a v \pi_{\succ n}\varepsilon^\mu.$$

We now apply the map ψ as defined in [Naito and Watanabe 2017, Section 3.3]. For this, we need to determine the set

$$D_{r_a}(\Pi) = \{d \in \{1, \dots, \ell'\} \mid (\alpha, k) = (w_d^{-1}\gamma_d, n_d)\}.$$

Since $m_i \geq 4$ for all i, we get

$$D_{r_a}(\Pi) = \bigl\{d \mid d \in \{1, \dots, \ell(p)\} \text{ and } (\alpha, k) = (w_d^{-1}\gamma_d, n_d)\bigr\} \subseteq [1, \ell(p)].$$

In particular, the set $D_{r_a}(\Pi)$ depends only on p and not the integers (m_i).

Naito–Watanabe construct the path $\psi(\Pi)$ as follows: Write $D_{r_a}(\Pi) = \{d_1 < \cdots < d_m\}$, which we allow to be the empty set.

For each index $q \in \{1, \dots, m\}$, we define $r_q \in \{d_q + 2, \dots, d_{q+1}\}$ (where $d_{m+1} = \ell' + 1$) to be the smallest index such that

$$w_{r_q}^{-1}\alpha \in \Phi^+ \quad \text{and} \quad \gamma_{r_q - 1} \prec w_{r_q}^{-1}\alpha \prec \gamma_{r_q}.$$

The existence of such an index r_q is proved in [Naito and Watanabe 2017, Lemma 2.3.2]. For $i = 1, \dots, \#\Phi^+ - n$, note that there is no positive root β satisfying $\gamma_i \prec \beta \prec \gamma_{i+1}$ (resp. $\gamma_{\ell'} \prec \beta$ if $i = \#\Phi^+ - n \geq 1$). Hence $r_1, \dots, r_m \leq n$ and they only depend on the path p, not the integers (m_i).

We introduce the shorthand notation

$$x_h := w_h \varepsilon^{n_1\gamma_1^\vee + \cdots + n_{h-1}\gamma_{h-1}^\vee - k w_1^{-1}\alpha^\vee}$$

such that Π is of the form $x_1 \to \cdots \to x_{\ell'+1}$. Then $\psi(\Pi)$ is defined as the composition of $\Pi_0', \dots, \Pi_m'$, given by

$$\Pi_0' : u = r_a x_1 \xrightarrow{(\gamma_1, n_1')} r_a x_2 \xrightarrow{(\gamma_2, n_2')} \cdots \xrightarrow{(\gamma_{d_1 - 1}, n'_{d_1 - 1})} r_a x_{d_1},$$

$$\Pi_q' : r_a x_{d_q} = x_{d_q + 1} \xrightarrow{(\gamma_{d_q+1}, n_{d_q+1})} \cdots \xrightarrow{(\gamma_{r_q - 1}, n_{r_q - 1})} x_{r_q} \xrightarrow{(w_{r_q}^{-1}\alpha, k)} r_a x_{r_q} \xrightarrow{(\gamma_{r_q}, n'_{r_q})} \cdots \xrightarrow{(\gamma_{d_{q+1}-1}, n'_{d_{q+1}-1})} r_a x_{d_{q+1}},$$

where we write

$$n_i' := n_i + k\langle\alpha^\vee, w_i\gamma_i\rangle,\quad i = 1,\dots,\ell'.$$

Since $r_1,\dots,r_m \le n$, we may write $\psi(\Pi) = \Psi^{-1}(\mathrm{pad}_{(m_i')}(p'))$, with

$$m_i' = m_i - k\langle\alpha^\vee, v s_{\beta_{n+1}}\cdots s_{\beta_{i-1}}(\beta_i)\rangle,\quad i > n.$$

The path p' is the composition of the paths $p_0',\dots,p_m'$ defined as

$$p_0' : u = s_\alpha w_1 \xrightarrow{(\gamma_1, n_1')} \cdots \xrightarrow{(\gamma_{d_1-1}, n_{d_1-1}')} s_\alpha w_{d_1-1},$$

$$p_q' : s_\alpha w_{d_q} = w_{d_q+1} \xrightarrow{(\gamma_{d_q+1}, n_{d_q+1})} \cdots \xrightarrow{(\gamma_{r_q-1}, n_{r_q-1})} w_{r_q} \xrightarrow{(w_{r_q}^{-1}\alpha, k)} s_\alpha w_{r_q} \xrightarrow{(\gamma_{r_q}, n_{r_q}')} \cdots \xrightarrow{(\gamma_{d_{q+1}-1}, n_{d_{q+1}-1}')} s_\alpha w_{d_{q+1}}.$$

We see that p' as defined above is explicitly described only in terms of p and independently of the (m_i).

To summarize, we chose integers (m_i') only depending on (m_i), u, v, n, $\prec$, a with the following property: for each path $p \in \mathrm{paths}^{\prec}_{\le n}(s_\alpha u \Rightarrow s_\alpha v)$, we may write

$$\psi(\Psi^{-1}\,\mathrm{pad}_{(m_i)}(p)) = \Psi^{-1}(\mathrm{pad}_{(m_i')}(p'))\quad\text{for some path } p' \in \mathrm{paths}^{\prec}_{\le n}(u \Rightarrow v).$$

It follows that the function $\tilde\psi$ as claimed exists. It is uniquely determined since Ψ^{-1} and $\mathrm{pad}_{(m_i')}$ are injective. Moreover, we saw that $p' := \tilde\psi(p)$ can be explicitly described depending only on p and not the integers (m_i).

The function $\tilde\psi$ preserves lengths of paths by construction. Using the explicit description, it is possible to verify that it also satisfies the weight constraint stated in Theorem 3.2(a). The interested reader is invited to verify that the constructions of ψ', φ_1, φ_2, φ' of Naito and Watanabe carry through in similar ways. □

With the main lemma proved, we can conclude Theorem 3.2 immediately. Indeed, it remains to show that the functions $\tilde\psi$ and $\tilde\varphi := (\tilde\varphi_1, \tilde\varphi_2)$ from Lemma 3.7 are bijective. Since ψ is bijective with ψ' being its inverse, it follows from the categorical definition and a bit of diagram chasing that $\tilde\psi$ is bijective with $\tilde\psi'$ its inverse. Similarly, one concludes that $\tilde\varphi$ is bijective with $\tilde\varphi'$ its inverse. The main result of this section is proved.

Remark 3.8. (a) Theorem 3.2 can be conveniently restated using the language of weight multisets from [Schremmer 2025, Definition 4.10]. For $u, v \in W$ and $0 \le n \le \#\Phi^+$, we write $\mathrm{wts}(u \Rightarrow v \dashrightarrow v\pi_{\succ n})$ for the multiset

$$\{(\mathrm{wt}(p), \ell(p)) \mid p \in \mathrm{paths}^{\prec}_{\le n}(u \Rightarrow v)\}_m.$$

We proved that this yields a well-defined multiset $\mathrm{wts}(u \Rightarrow v \dashrightarrow v')$ for all $u, v, v' \in W$.

If $a = (\alpha, k) \in \Delta_{\mathrm{af}}$ is a simple affine root with $(v')^{-1}\alpha \in \Phi^-$ and $u^{-1}\alpha \in \Phi^-$, then

$$\mathrm{wts}(u \Rightarrow v \dashrightarrow v') = \{(\omega + k(v^{-1}\alpha^\vee - u^{-1}\alpha^\vee), e) \mid (\omega, e) \in \mathrm{wts}(s_\alpha u \Rightarrow s_\alpha v \dashrightarrow s_\alpha v')\}_m.$$

If $(v')^{-1}\alpha \in \Phi^-$ and $u^{-1}\alpha \in \Phi^+$, then $\mathrm{wts}(u \Rightarrow v \dashrightarrow v')$ is the additive union of the two multisets

$$\{(\omega + k(v^{-1}\alpha^\vee - u^{-1}\alpha^\vee), e) \mid (\omega, e) \in \mathrm{wts}(s_\alpha u \Rightarrow s_\alpha v \dashrightarrow s_\alpha v')\}_m \cup \{(\omega - ku^{-1}\alpha^\vee, e) \mid (\omega, e) \in \mathrm{wts}(s_\alpha u \Rightarrow v \dashrightarrow v')\}_m.$$

(b) The double Bruhat graph can be seen as a generalization of the quantum Bruhat graph; see [Schremmer 2025, Proposition 4.13]. It is very helpful to compare results about the double Bruhat graph with the much better developed theory of the quantum Bruhat graph.

Under this point of view, one obtains a version of Theorem 3.2 for the quantum Bruhat graph. This is a well-known recursive description of weights in the quantum Bruhat graph; see [Lenart et al. 2015, Lemma 7.7].

(c) The remainder of this paper will mostly study consequences of recursive relations from Theorem 3.2. By studying the proof of Theorem 4.2 below, one may see that the weight multiset is already uniquely determined by these recursive relations together with a few additional facts to fix a recursive start. This can be seen as an alternative proof that the weight multiset is independent of the chosen reflection order; see [Schremmer 2025, Corollary 4.9].

4. Iwahori–Hecke algebra

Let us briefly motivate the definition of the Iwahori–Hecke algebra associated with an affine Weyl group.

Under suitable assumptions on our group and our fields, the *Hecke algebra* $\mathcal{H}(G, I)$ is classically defined to be the complex vector space of all compactly supported functions $f : G(F) \to \mathbb{C}$ satisfying $f(i_1 g i_2) = f(g)$ for all $g \in G(F),\ i_1, i_2 \in I \cap G(F)$. It becomes an algebra where multiplication is defined via convolution of functions. In this form, it occurs in the classical formulation of the Satake isomorphism [1963].

It is proved by Iwahori and Matsumoto [1965, Section 3] for split G that $\mathcal{H}(G, I)$ has a basis given by $\{S_x \mid x \in \widetilde{W}\}$ over $\mathbb{C}$ where the multiplication is uniquely determined by the conditions

$$\begin{aligned} S_x S_y &= S_{xy}, && x, y \in \widetilde{W} \text{ and } \ell(xy) = \ell(x) + \ell(y),\\ S_{r_a} S_x &= q S_{r_a x} + (q-1) S_x, && x \in \widetilde{W}, a \in \Delta_{\mathrm{af}} \text{ and } \ell(r_a x) < \ell(x). \end{aligned}$$

Here, $q := \#(\mathcal{O}_F / \mathfrak{m}_{\mathcal{O}_F})$ is the cardinality of the residue field of F. The basis element S_x corresponds to the indicator function of the coset $IxI \subseteq G(\breve{F})$.

With the convenient change of variables $T_x := q^{-\ell(x)/2} S_x \in \mathcal{H}(G, I)$, the above relations get the equally popular form

$$\begin{aligned} T_x T_y &= T_{xy}, && x, y \in \widetilde{W} \text{ and } \ell(xy) = \ell(x) + \ell(y),\\ T_{r_a} T_x &= T_{r_a x} + (q^{1/2} - q^{-1/2}) T_x, && x \in \widetilde{W}, a \in \Delta_{\mathrm{af}} \text{ and } \ell(r_a x) < \ell(x). \end{aligned}$$

Since the number q is independent of the choice of affine root system, we define the *Iwahori–Hecke algebra* of $\widetilde{W}$ as follows.

Definition 4.1. The *Iwahori–Hecke algebra* $\mathcal{H}(\widetilde{W})$ of $\widetilde{W}$ is the algebra over $\mathbb{Z}[Q]$ defined by the generators

$$T_x, \quad x \in \widetilde{W}$$

and the relations

$$\begin{aligned} T_x T_y &= T_{xy}, && x, y \in \widetilde{W} \text{ and } \ell(xy) = \ell(x) + \ell(y),\\ T_{r_a} T_x &= T_{r_a x} + Q T_x, && x \in \widetilde{W}, a \in \Delta_{\mathrm{af}} \text{ and } \ell(r_a x) < \ell(x). \end{aligned}$$

One easily sees that $\mathcal{H}(\widetilde{W})$ is a free $\mathbb{Z}[Q]$-module with basis $\{T_x \mid x \in \widetilde{W}\}$, and that each T_x is invertible, because

$$T_{r_a}(T_{r_a} - Q) = 1, \quad a \in \Delta_{\mathrm{af}}.$$

All results presented in this article can be immediately generalized to most other conventions for the Iwahori–Hecke algebra, e.g., by substituting $Q = q^{1/2} - q^{-1/2}$.

4.1. ***Products via the double Bruhat graph.*** We are interested in the question of how to express arbitrary products of the form $T_x T_y$ with $x, y \in \widetilde{W}$ in terms of this basis. This is related to understanding the structure of the subset $IxI \cdot IyI \subseteq G(\breve{F})$. While it might be too much to ask for a general formula, we can understand these products (and thus the Iwahori–Hecke algebra) better by relating it to the double Bruhat graph. Our main result of this section is the following:

Theorem 4.2. *Let $C_1 > 0$ be a constant and define $C_2 := (8\#\Phi^+ + 4)C_1$.*

Let $x = w_x\varepsilon^{\mu_x}$, $z = w_z\varepsilon^{\mu_z} \in \widetilde{W}$ such that x is C_2-regular and z is $2\ell(x)$-regular. Define polynomials $\varphi_{x,z,yz} \in \mathbb{Z}[Q]$ via

$$T_x T_z = \sum_{y \in \widetilde{W}} \varphi_{x,z,yz} T_{yz} \in \mathcal{H}(\widetilde{W}).$$

Pick an element $y = w_y\varepsilon^{\mu_y} \in \widetilde{W}$ such that $\ell(x) - \ell(y) < C_1$. Let

$$\mathrm{LP}(x) = \{v_x\}, \quad \mathrm{LP}(y) = \{v_y\}, \quad \mathrm{LP}(z) = \{v_z\}$$

and define the multiset

$$M := \big\{\ell_1 + \ell_2 \mid (\omega_1, \ell_1) \in \mathrm{wts}(v_x \Rightarrow v_y \dashrightarrow w_z v_z),\ (\omega_2, \ell_2) \in \mathrm{wts}(w_x v_x w_0 \Rightarrow w_y v_y w_0 \dashrightarrow w_y w_z v_z) \text{ such that } v_y^{-1}\mu_y = v_x^{-1}\mu_x - \omega_1 + w_0\omega_2\big\}_m.$$

Here, $w_0 \in W$ denotes the longest element. Then

$$\varphi_{x,z,yz} = \sum_{e \in M} Q^e.$$

Remark 4.3. (a) In principle, we have the following recursive relations to calculate $T_x T_z$ as long as all occurring elements are in shrunken Weyl chambers, e.g., 2-regular: Pick a simple affine root $a = (\alpha, k) \in \Delta_{\mathrm{af}}$. If $xr_a < x$ (i.e., $v_x^{-1}\alpha \in \Phi^+$), then

$$T_x T_z = T_{xr_a} T_{r_a} T_z = \begin{cases} T_{xr_a} T_{r_a z}, & r_a z > z \text{ (i.e., } (w_z v_z)^{-1}\alpha \in \Phi^+), \\ T_{xr_a} T_{r_a z} + Q T_{xr_a} T_z, & r_a z < z \text{ (i.e., } (w_z v_z)^{-1}\alpha \in \Phi^-). \end{cases}$$

This kind of recursive relation is analogous to the recursive behaviour of the multiset $\mathrm{wts}(v_x \Rightarrow v_y \dashrightarrow w_z v_z)$; see Theorem 3.2.

Similarly, if $r_a x < x$ (i.e., $(w_x v_x)^{-1}\alpha \in \Phi^-$), we get

$$\begin{aligned} T_x T_z &= T_{r_a} T_{r_a x} T_z = \sum_{y \in \widetilde{W}} \varphi_{r_a x, z, yz} T_{r_a} T_{yz} \\ &= \sum_{y \in \widetilde{W}} \varphi_{r_a x, z, yz} \cdot \begin{cases} T_{r_a yz}, & r_a yz > yz \text{ (i.e., } (w_y w_z v_z)^{-1}\alpha \in \Phi^+), \\ T_{r_a yz} + Q T_{yz}, & r_a yz < yz \text{ (i.e., } (w_y w_z v_z)^{-1}\alpha \in \Phi^-). \end{cases} \end{aligned}$$

This kind of recursive relation is analogous to the recursive behaviour of the multiset $\mathrm{wts}(w_x v_x w_0 \Rightarrow w_y v_y w_0 \dashrightarrow w_y w_z v_z)$; see Theorem 3.2.

For the proof of Theorem 4.2, we have to apply these recursive relations iteratively while keeping track of the length and regularity conditions to ensure everything happens inside the shrunken Weyl chambers.

(b) Let us compare Theorem 4.2 to the quantum Bruhat graph. In view of [Schremmer 2025, Proposition 4.13], it follows that $\varphi_{x,z,yz} = 0$ unless

$$v_y^{-1}\mu_y \leq v_x^{-1}\mu_x - \mathrm{wt}_{\mathrm{QB}(W)}(v_x \Rightarrow v_y) - \mathrm{wt}_{\mathrm{QB}(W)}(w_y v_y \Rightarrow w_x v_x).$$

By [Schremmer 2024, Theorem 4.2], this latter inequality is equivalent to the Bruhat order condition $y \leq x$, which is (by the definition of the Iwahori–Hecke algebra) always a necessary condition for $\varphi_{x,z,yz}$ to be nonzero.

(c) If the condition $\ell(x) - \ell(y) < C_1$ gets strengthened to $\ell(x) + \ell(z) - \ell(yz) < C_1$, it follows that the product yz must be length-additive, so $v_y = w_z v_z$ [Schremmer 2022, Lemma 2.13]. One of the simple facts on the double Bruhat graph [Schremmer 2025, Lemma 4.11] yields

$$\mathrm{wts}(w_x v_x w_0 \Rightarrow w_y v_y w_0 \dashrightarrow w_y w_z v_z) = \begin{cases} \varnothing, & w_y v_y \neq w_x v_x, \\ \{(0,0)\}_m, & w_y v_y = w_x v_x. \end{cases}$$

So the multiset M as defined in Theorem 4.2 is empty unless $w_y v_y = w_x v_x$, in which case it will be equal to

$$M = \{\ell \mid (\omega, \ell) \in \mathrm{wts}(v_x \Rightarrow v_y) \text{ such that } v_y^{-1}\mu_y = v_x^{-1}\mu_x - \omega\}_m.$$

This recovers Theorem 1.2.

The unique smallest element of $\mathrm{wts}(v_x \Rightarrow v_y)$ from [Schremmer 2025, Proposition 4.13] corresponds to the uniquely determined largest element in $\widetilde{W}$ having nonzero coefficient in $T_x T_z$. This element is known as the *Demazure product* of x and z in $\widetilde{W}$. We recover the formula for the Demazure product of x and z in terms of the quantum Bruhat graph from [He and Nie 2024, Proposition 3.3] in the situation of Theorem 4.2.

Definition 4.4. (a) For $x \in \widetilde{W}$ and $w \in W$, we define the multiset $Y(x, w)$ as follows: the underlying set $|Y(x, w)|$ is a subset of $\widetilde{W} \times \mathbb{Z}$, and the multiplicity of the pair $(y, e) \in \widetilde{W} \times \mathbb{Z}$ in $Y(x, w)$ is defined via

$$T_x T_{w\varepsilon^{2\rho^\vee \ell(x)}} = \sum_{(y,e)\in Y(x,w)} Q^e T_{yw\varepsilon^{2\rho^\vee \ell(x)}}.$$

(b) We define the usual product group structure on $\widetilde{W} \times \mathbb{Z}$, i.e.,

$$(y_1, e_1) \cdot (y_2, e_2) := (y_1 y_2, e_1 + e_2)$$

for $y_1, y_2 \in \widetilde{W}$ and $e_1, e_2 \in \mathbb{Z}$. If M is a multiset with $|M| \subseteq \widetilde{W} \times \mathbb{Z}$, we write $M \cdot (y, e)$ for the multiset obtained by the right action of $(y, e) \in \widetilde{W} \times \mathbb{Z}$.

Lemma 4.5. *Let $x, z \in \widetilde{W}$ such that z is $2\ell(x)$-regular.*

(a) *Write $z = w_z \varepsilon^{\mu_z}$ and $\mathrm{LP}(z) = \{v_z\}$. Then*

$$T_x T_z = \sum_{(y,e)\in Y(x,w_z v_z)} Q^e T_{yz}.$$

(b) *Let* $a=(\alpha,k)\in\Delta_{\mathrm{af}}$ *with* $xr_a<x$ *and* $w\in W$. *If* $w^{-1}\alpha\in\Phi^+$, *we have*

$$Y(x,w)=Y(xr_a,s_\alpha w)\cdot(r_a,0).$$

If $w^{-1}\alpha\in\Phi^-$, *we express* $Y(x,w)$ *as the additive union of multisets*

$$Y(x,w)=\big(Y(xr_a,s_\alpha w)\cdot(r_a,0)\big)\cup\big(Y(xr_a,w)\cdot(1,1)\big).$$

(c) *For* $y=w_y\varepsilon^{\mu_y}\in\widetilde{W}$ *and* $e\in\mathbb{Z}$, *the multiplicity of* $(y,e)\in Y(x,w)$ *agrees with the multiplicity of* (y^{-1},e) *in* $Y(x^{-1},\mathrm{cl}(y)w)$, *where* $\mathrm{cl}(y)\in W$ *is the classical part of* $y\in W\ltimes X_*(T)_{\Gamma_0}$.

Proof. (a) The regularity condition allows us to write z as the length-additive product

$$z=z_1\cdot z_2,\quad z_1=w_zv_z\varepsilon^{2\rho^\vee\ell(x)},\quad z_2=v_z^{-1}\varepsilon^{\mu_z-v_z2\rho^\vee\ell(x)}.$$

Then we get

$$T_xT_z=T_xT_{z_1}T_{z_2}=\sum_{(y,e)\in Y(x,w_zv_z)}T_{yz_1}T_{z_2}.$$

By the regularity of z_1, it follows that $\mathrm{LP}(yz_1)=\mathrm{LP}(z_1)=\{1\}$ for each $y\le x$ in the Bruhat order. Thus $T_{yz_1}T_{z_2}=T_{yz_1z_2}=T_{yz}$ for each $(y,e)\in Y(x,w_zv_z)$.

(b) Let $z=w\varepsilon^\mu$ with μ superregular and dominant, as in (a). Use the fact

$$T_xT_z=T_{xr_a}T_{r_a}T_z$$

and evaluate $T_{r_a}T_z$ depending on whether $w^{-1}\alpha$ is positive or negative.

(c) We consider the symmetrizing form of $\mathcal{H}(\widetilde{W})$ given by

$$\tau:\mathcal{H}(\widetilde{W})\to\mathbb{Z}[Q],\quad \sum_{x\in\widetilde{W}}a_xT_x\mapsto a_1.$$

One checks that $\tau(T_xT_{x^{-1}})=1$ and $\tau(T_xT_y)=0$ for $x,y\in\widetilde{W}$ with $xy\neq 1$; see [Bonnafé 2017, Section 4.1D]. It follows from this that $\tau(hh')=\tau(h'h)$ for all $h,h'\in\mathcal{H}(\widetilde{W})$, and that $\tau(T_{x^{-1}}h)$ is the T_x-coefficient of h for $x\in\widetilde{W}$.

Moreover, note that $T_x\mapsto T_{x^{-1}}$ defines an antiautomorphism of the $\mathbb{Z}[Q]$-algebra $\mathcal{H}(\widetilde{W})$, and that τ is invariant under this map.

Fix $y\in\widetilde{W}$ and assume that both z and yz are $2\ell(x)$-regular. We calculate

$$\begin{aligned}\sum_{e\in\mathbb{Z}}(\text{multiplicity of }(y,e)\text{ in }Y(x,w_zv_z))Q^e&=(\text{coefficient of }T_{yz}\text{ in }T_xT_z)\\&=\tau(T_{(yz)^{-1}}T_xT_z)\\&=\tau(T_{z^{-1}}T_{x^{-1}}T_{yz})\\&=(\text{coefficient of }T_z\text{ in }T_x^{-1}T_{yz})\\&=\sum_{e\in\mathbb{Z}}(\text{multiplicity of }(y^{-1},e)\text{ in }Y(x^{-1},w_yw_zv_z))Q^e.\end{aligned}$$

Comparing coefficients of Q^e in $\mathbb{Z}[Q]$, the claim follows. □

Remark 4.6. The connection to our previous article [Schremmer 2025] is given as follows: For x, z as in Lemma 4.5, the regularity condition on z basically ensures that zIz^{-1} behaves like ${}^{w_z v_z}U(L)$, so we can approximate IzI by the semi-infinite orbit $Iz\,{}^{v_z}U(L) = I\,{}^{w_z v_z}U(L)z$. Then $IxI \cdot IzI$ is very close to

$$IxI \cdot {}^{w_z v_z}U(L)z = \bigcup_{(y,e)\in Y(x,w_z v_z)} Iy\,{}^{w_z v_z}U(L)z \subseteq G(\breve{F}).$$

Now observe for any $y \in \widetilde{W}$ that

$$IxI \cap Iy\,{}^{w_z v_z}U(L) \neq \varnothing \quad \iff \quad y \in IxI \cdot {}^{w_z v_z}U(L).$$

So the multiset $Y(x, w)$ is the representation-theoretic correspondent of the main object of interest in [Schremmer 2025, Theorem 5.2].

Lemma 4.7. *Let $x = w_x \varepsilon^{\mu_x} \in \widetilde{W}$ and pick elements $u_1, u_2 \in W$, as well as $v_x \in \mathrm{LP}(x)$.*

(a) *The multiset* $\mathrm{wts}(v_x \Rightarrow u_1 \dashrightarrow u_2)$ *is equal to the additive union of multisets*

$$\bigcup_{(w_y\varepsilon^{\mu_y},e)\in Y(x,u_2)} \big\{(v_x^{-1}\mu_x - u_1^{-1}\mu_y + \omega, e+\ell) \mid (\omega,\ell) \in \mathrm{wts}(w_x v_x \Rightarrow w_y u_1 \dashrightarrow w_y u_2)\big\}_m.$$

(b) *The multiset* $\mathrm{wts}(w_x v_x w_0 \Rightarrow u_2 w_0 \dashrightarrow u_1)$ *is equal to the additive union of multisets*

$$\bigcup_{\substack{u_3\in W\\ (w_y\varepsilon^{\mu_y},e)\in Y(x,u_3)\\ \text{s.t. } w_y u_3 = u_1}} \big\{(w_0 u_2^{-1} w_y \mu_y - w_0 v_x^{-1}\mu_x + \omega, e+\ell) \mid (\omega,\ell) \in \mathrm{wts}(v_x w_0 \Rightarrow w_y^{-1} u_2 w_0 \dashrightarrow u_3)\big\}_m.$$

Proof. (a) Induction on $\ell(x)$. In the case $\ell(x) = 0$, we get $Y(x, u_2) = \{(x, 0)\}_m$. From [Schremmer 2025, Lemma 4.7(c)], we indeed get that $\mathrm{wts}(v_x \Rightarrow u_1 \dashrightarrow u_2)$ is equal to

$$\{(v_x^{-1}\mu_x - u_1^{-1}\mu_x + \omega, \ell) \mid (\omega,\ell) \in \mathrm{wts}(w_x v_x \Rightarrow w_x u_1 \dashrightarrow w_x u_2)\}_m.$$

Now in the inductive step, pick a simple affine root $a = (\alpha, k)$ with $xr_a < x$. This means $v_x^{-1}\alpha \in \Phi^+$ and $v_{x'} := s_\alpha v_x \in \mathrm{LP}(x')$, where

$$x' := w_{x'}\varepsilon^{\mu_{x'}} := xr_a = w_x s_\alpha \varepsilon^{s_\alpha(\mu_x)+k\alpha^\vee}.$$

Let us first consider the case $u_2^{-1}\alpha \in \Phi^+$. Then $Y(x, u_2) = Y(x', s_\alpha u_2) \cdot (r_a, 0)$ by Lemma 4.5(b). We get

$$\begin{aligned}
&\bigcup_{(w_y\varepsilon^{\mu_y},e)\in Y(x,u_2)} \big\{(v_x^{-1}\mu_x - u_1^{-1}\mu_y + \omega, e+\ell) \mid (\omega,\ell) \in \mathrm{wts}(w_x v_x \Rightarrow w_y u_1 \dashrightarrow w_y u_2)\big\}_m\\
&= \bigcup_{(w_{y'}\varepsilon^{\mu_{y'}},e)\in Y(x',s_\alpha u_2)} \big\{(v_{x'}^{-1}\mu_x' + k v_x^{-1}\alpha^\vee - (s_\alpha u_1)^{-1}\mu_{y'} - k u_1^{-1}\alpha^\vee + \omega, e+\ell) \mid\\
&\qquad\qquad (\omega,\ell) \in \mathrm{wts}(w_{x'}v_{x'} \Rightarrow w_{y'}(s_\alpha u_1) \dashrightarrow w_{y'}(s_\alpha u_2))\big\}_m.
\end{aligned}$$

By the inductive assumption, this is equal to

$$\{(\omega + k(v_x^{-1}\alpha^\vee - u_1^{-1}\alpha^\vee), \ell) \mid (\omega,\ell) \in \mathrm{wts}(s_\alpha v_x \Rightarrow s_\alpha u_1 \dashrightarrow s_\alpha u_2)\}_m.$$

By Theorem 3.2(a), this is equal to $\mathrm{wts}(v_x \Rightarrow u_1 \dashrightarrow u_2)$, using the assumption $u_2^{-1}\alpha \in \Phi^+$ again.

In the converse case where $u_2^{-1}\alpha \in \Phi^-$, we argue entirely similarly. Use Lemma 4.5 to write

$$Y(x, u_2) = \big(Y(x', s_\alpha u_2)\cdot(r_a, 0)\big) \cup \big(Y(x', u_2)\cdot(1, 1)\big).$$

Considering Theorem 3.2(b), the inductive claim follows.

(b) One may argue similarly to (a), tracing through somewhat more complicated expressions to reduce to Theorem 3.2 again. Instead, we show that (a) and (b) are equivalent. Recall that $w_x v_x w_0 \in \mathrm{LP}(x^{-1})$ [Schremmer 2022, Lemma 2.12]. By (a), we see that $\mathrm{wts}(w_x v_x w_0 \Rightarrow u_2 w_0 \dashrightarrow u_1)$ is equal to

$$\bigcup_{(w_y\varepsilon^{\mu_y}, e)\in Y(x^{-1}, u_1)} \big\{((w_x v_x w_0)^{-1}(-w_x\mu_x) - (u_2 w_0)^{-1}\mu_y + \omega, e+\ell) \mid (\omega, \ell) \in \mathrm{wts}(v_x w_0 \Rightarrow w_y u_2 w_0 \dashrightarrow w_y u_1)\big\}.$$

In view of Lemma 4.5(c), we recover the claim in (b). □

Lemma 4.8. *Let $C_1, e \geq 0$ be two nonnegative integers. Define $C_2 := (8e+4)C_1$.*

Let $x, y \in \widetilde{W}$ such that x is C_2-regular and $\ell(x) - \ell(y) < C_1$. Let $u \in W$. Write

$$x = w_x\varepsilon^{\mu_x}, \qquad y = w_y\varepsilon^{\mu_y},$$
$$\mathrm{LP}(x) = \{v_x\}, \qquad \mathrm{LP}(y) = \{v_y\}.$$

Define the multiset

$$M := \big\{\ell_1 + \ell_2 \mid (\omega_1, \ell_1) \in \mathrm{wts}(v_x \Rightarrow v_y \dashrightarrow u),\ (\omega_2, \ell_2) \in \mathrm{wts}(w_x v_x w_0 \Rightarrow w_y v_y w_0 \dashrightarrow w_y u) \text{ such that } v_y^{-1}\mu_y = v_x^{-1}\mu_x - \omega_1 + w_0\omega_2\big\}_m.$$

Then the multiplicity of (y, e) in $Y(x, u)$ agrees with the multiplicity of e in M.

Proof. Induction on e. Consider the inductive start $e = 0$. If $0 \in M$, then $\ell_1 = \ell_2 = 0$ and $v_x = v_y$ by definition of M. Hence $x = y$, and indeed $0 \in M$ has multiplicity 1. Similarly, $(y, 0)$ also has multiplicity 1 in $Y(x, u)$.

If $0 \notin M$, we see $x \neq y$ and indeed $(y, 0) \notin Y(x, u)$ for $x \neq y$. This settles the inductive start.

In the inductive step, let us write x as a length-additive product $x = x_1x_2x_3$, where

$$x_1 = \varepsilon^{4C_1 w_x v_x \rho^\vee}, \quad x_2 = w_x\varepsilon^{\mu_x - 8C_1 v_x\rho^\vee}, \quad x_3 = \varepsilon^{4C_1 v_x\rho^\vee}.$$

Note that the inductive assumptions are satisfied for $C_1, e-1, x_2$ and any element $y' \in \widetilde{W}$ such that $\ell(x_2) - \ell(y') < C_1$.

The length-additivity of $x = x_1x_2x_3$ implies

$$Y(x, u) = \big\{(y_1y_2y_3, e_1+e_2+e_3) \mid (y_3, e_3) \in Y(x_3, u), (y_2, e_2) \in Y(x_2, \mathrm{cl}(y_3)u),\ (y_1, e_1) \in Y(x_1, \mathrm{cl}(y_2)\,\mathrm{cl}(y_3)u)\big\}_m.$$

Pick elements

$$(y_3, e_3) \in Y(x_3, u), \quad (y_2, e_2) \in Y(x_2, \mathrm{cl}(y_3)u), \quad (y_1, e_1) \in Y(x_1, \mathrm{cl}(y_2)\,\mathrm{cl}(y_3)u)$$

such that $\ell(y_1y_2y_3) > \ell(x) - C_1$ and $e_1 + e_2 + e_3 = e$.

In this case, we certainly get $\ell(y_i) > \ell(x_i) - C_1$ for $i = 1, 2, 3$. Since x_1, x_2, x_3 are $4C_1$-regular by construction, it follows that each y_i is $2C_1$-regular by $y_i \le x_i$ and $\ell(y_i) > \ell(x_i) - C_1$ (studying how regularity behaves in a sequence of Bruhat covers from y_i to x_i). We claim that

$$\ell(y_1 y_2 y_3) = \ell(y_1) + \ell(y_2) + \ell(y_3).$$

We can study the question of length-additivity of such products using [Schremmer 2022, Lemma 2.13]. This lemma expresses the condition $\ell(xy) = \ell(x) + \ell(y)$ in terms of the *length functionals* $\ell(x, \cdot)$ and $\ell(y, \cdot)$ as defined in [loc. cit., Definition 2.5]. Using the aforementioned lemma, it suffices to see that $\ell(y_1 y_2) = \ell(y_1) + \ell(y_2)$ and $\ell(y_2 y_3) = \ell(y_2) + \ell(y_3)$ (using regularity). If $y_1 y_2$ is not a length-additive product, we use [loc. cit., Lemma 2.13] to find a root $\alpha \in \Phi$ with $\ell(y_1, \mathrm{cl}(y_2)\alpha) > 0$ and $\ell(y_2, \alpha) < 0$. By regularity, this means $\ell(y_1, \mathrm{cl}(y_2)\alpha) > C_1$ and $\ell(y_2, \alpha) < -C_1$. Using [loc. cit., Corollary 2.10 and Lemma 2.12], we get

$$\begin{aligned}\ell(y_1 y_2) &= \sum_{\beta\in\Phi} \tfrac{1}{2}|\ell(y_1, \mathrm{cl}(y_2)\beta) + \ell(y_2, \beta)| \\ &\le -C_1 + \sum_{\beta\in\Phi} \tfrac{1}{2}(|\ell(y_1, \mathrm{cl}(y_2)\beta)| + |\ell(y_2, \beta)|) = \ell(y_1) + \ell(y_2) - C_1.\end{aligned}$$

This contradicts the above assumption $\ell(y_1 y_2 y_3) > \ell(x) - C_1 \ge \ell(y_1) + \ell(y_2) + \ell(y_3) - C_1$. The proof that $y_2 y_3$ is length-additive is completely analogous.

Let us consider the special case $e_1 = e_3 = 0$ separately. Then $y_1 = x_1$ and $y_3 = x_3$. The length-additivity of the product $x_1 y_2 x_3$ implies that $\mathrm{LP}(y_2) = \{v_x\}$ and $\mathrm{cl}(y_2) = w_x$. Using Lemma 4.7(a), we can express $\{(0, 0)\}_m = \mathrm{wts}(v_x \Rightarrow v_x \dashrightarrow u)$ in the form

$$\bigcup_{(w_y \varepsilon^{\mu_y}, e')\in Y(x_2, u)} \{(\dots, e' + \ell) \mid (\omega, \ell) \in \mathrm{wts}(w_x v_x \Rightarrow w_y v_x \dashrightarrow w_y u)\}_m.$$

From this and [Schremmer 2025, Lemma 4.11], it follows that $Y(x_2, u)$ contains only one element (y', e') with $\mathrm{cl}(y') = w_x$, and that this element must be equal to $(x_2, 0)$.

We see that, if $e_1 = e_3 = 0$, we must also have $e_2 = 0$. This case has been settled before.

We hence assume that $e_1 + e_3 > 0$. In particular, we may apply the inductive assumption to x_2, y_2, e_2. Recall that the multiplicity of (y, e) in $Y(x, u)$ is equal to the number of tuples (with multiplicity)

$$(y_3, e_3) \in Y(x_3, u), \quad (y_2, e_2) \in Y(x_2, \mathrm{cl}(y_3)u), \quad (y_1, e_1) \in Y(x_1, \mathrm{cl}(y_2)\,\mathrm{cl}(y_3)u)$$

such that $e_1 + e_2 + e_3 = e$ and $y = y_1 y_2 y_3$ (necessarily length-additive). Hence $\mathrm{LP}(y_2) = \{\mathrm{cl}(y_3)v_y\}$ and $w_y = \mathrm{cl}(y_1)\,\mathrm{cl}(y_2)\,\mathrm{cl}(y_3)$. By induction, the multiplicity of (y, e) in $Y(x, u)$ is also equal to the number of tuples (with multiplicity)

$$\begin{aligned}&(y_3, e_3) \in Y(x_3, u),\\ &(\omega_1, \ell_1) \in \mathrm{wts}(v_x \Rightarrow \mathrm{cl}(y_3)v_y \dashrightarrow \mathrm{cl}(y_3)u),\\ &(\omega_2, \ell_2) \in \mathrm{wts}(w_x v_x w_0 \Rightarrow \mathrm{cl}(y_1)^{-1} w_y v_y w_0 \dashrightarrow \mathrm{cl}(y_1)^{-1} w_y u),\\ &(y_1, e_1) \in Y(x_1, \mathrm{cl}(y_1)^{-1} w_y u),\end{aligned}$$

satisfying $e = e_1 + \ell_1 + \ell_2 + e_3$ and

$$y_1^{-1} y y_3^{-1} = \mathrm{cl}(y_1)^{-1} w_y \,\mathrm{cl}(y_3)^{-1} \varepsilon^{(\mathrm{cl}(y_3) v_y)(v_x^{-1}\mu_{x_2} - \omega_1 + w_0\omega_2)}.$$

The latter identity can be rewritten, if we write $y_3 = w_3\varepsilon^{\mu_3}$ and $y_1 = w_1\varepsilon^{\mu_1}$, as

$$v_y^{-1}\mu_y = v_x^{-1}\mu_{x_2} - \omega_1 + w_0\omega_2 + v_y^{-1}\mu_3 + (w_y v_y)^{-1} w_1\mu_1.$$

We see that we may study the contributions of $(y_3, e_3, \omega_1, \ell_1)$ and $(y_1, e_1, \omega_2, \ell_2)$ separately.

We may combine the above data for $(y_3, e_3, \omega_1, \ell_1)$, noticing that we are only interested in the multiset

$$\bigl\{(-v_y^{-1}\mu_3 + \omega_1 + v_x^{-1}\mu_{x_3}, e_3 + \ell_1) \mid (w_3\varepsilon^{\mu_3}, e_3) \in Y(x_3, u),\ (\omega_1, \ell_1) \in \mathrm{wts}(v_x \Rightarrow w_3 v_y \dashrightarrow w_3 u)\bigr\}_m.$$

By Lemma 4.7(a), the above multiset agrees with $\mathrm{wts}(v_x \Rightarrow v_y \dashrightarrow u)$.

Similarly, we may combine the data for $(y_1, e_1, \omega_2, \ell_2)$, noticing that we are only interested in the multiset

$$\begin{aligned}\bigl\{(w_0(w_y v_y)^{-1} w_1\mu_1 + \omega_2 - w_0(w_x v_x)^{-1}\mu_{x_1}, e_1 + \ell_2) \mid \\ u' \in W,\ (w_1\varepsilon^{\mu_1}, e_1) \in Y(x_1, u') \text{ such that } w_1 u' = w_y u, \\ (\omega_2, \ell_2) \in \mathrm{wts}(w_x v_x w_0 \Rightarrow w_1^{-1} w_y v_y w_0 \dashrightarrow w_1^{-1} w_y u)\bigr\}_m.\end{aligned}$$

By Lemma 4.7(b), the above multiset agrees with $\mathrm{wts}(w_x v_x w_0 \Rightarrow w_y v_y w_0 \dashrightarrow w_y u)$.

We summarize that the multiplicity of (y, e) in $Y(x, u)$, i.e., the number of tuples

$$(y_3, e_3, \omega_1, \ell_1, \omega_2, \ell_2, y_1, e_1)$$

with multiplicity as above, is equal to the number of tuples

$$\begin{aligned}&(\lambda_1, f_1) \in \mathrm{wts}(v_x \Rightarrow v_y \dashrightarrow u),\\ &(\lambda_2, f_2) \in \mathrm{wts}(w_x v_x w_0 \Rightarrow w_y v_y w_0 \dashrightarrow w_y u)\end{aligned}$$

satisfying $e = f_1 + f_2$ and

$$v_y^{-1}\mu_y = v_x^{-1}\mu_{x_2} - \lambda_1 + v_x^{-1}\mu_{x_3} + w_0\lambda_2 + (w_x v_x)^{-1}\mu_{x_1}.$$

Up to evaluating the product $x = x_1x_2x_3 \in W \ltimes X_*(T)_{\Gamma_0}$, this finishes the induction and the proof. □

Corollary 4.9. *Let $x = w_x\varepsilon^{\mu_x}$, $z = w_z\varepsilon^{\mu_z} \in \widetilde{W}$. Write*

$$T_x T_z = \sum_{y\in\widetilde{W}} \sum_{e\geq 0} n_{y,e} Q^e T_{yz}, \quad n_{y,e} \in \mathbb{Z}_{\geq 0}.$$

Pick elements $v_x \in \mathrm{LP}(x_x)$, $v_z \in \mathrm{LP}(x_z)$, $e \in \mathbb{Z}_{\geq 0}$ and $y = w_y\varepsilon^{\mu_y} \in \widetilde{W}$. Then $n_{y,e}$ is at most equal to the multiplicity of the element

$$\bigl(v_x^{-1}(\mu_x - w_x^{-1} w_y\mu_y), e\bigr)$$

in the multiset

$$\mathrm{wts}(v_x \Rightarrow w_y^{-1} w_x v_x \dashrightarrow w_z v_z).$$

Proof. Let us write $\mathcal{H}(\widetilde{W})_{\geq 0}$ for the subset of those elements of $\mathcal{H}(\widetilde{W})$ which are nonnegative linear combinations of elements of the form $Q^e T_x$ for $e \in \mathbb{Z}_{\geq 0}$ and $x \in \widetilde{W}$. For dominant coweights $\lambda_1, \lambda_2 \in X_*(T)_{\Gamma_0}$, we obtain

$$\begin{aligned} T_{\varepsilon^{w_x v_x \lambda_1} x} T_{z \varepsilon^{v_z \lambda_2}} &= T_{\varepsilon^{w_x v_x \lambda_1}} T_x T_z T_{\varepsilon^{v_z \lambda_2}} \\ &= \sum_{y \in \widetilde{W}} \sum_{e \geq 0} n_{y,e} Q^e T_{\varepsilon^{w_x v_x \lambda_1}} T_{yz} T_{\varepsilon^{v_z \lambda_2}} \in \sum_{y \in \widetilde{W}} \sum_{e \geq 0} n_{y,e} Q^e T_{\varepsilon^{w_x v_x \lambda_1} yz \varepsilon^{v_z \lambda_2}} + \mathcal{H}(\widetilde{W})_{\geq 0}. \end{aligned}$$

So the quantity $n_{y,e}$ can only increase if we replace (x, y, z) by $(\varepsilon^{w_x v_x \lambda_1} x, \varepsilon^{w_x v_x \lambda_1} y, z \varepsilon^{w_z v_z \lambda_2})$. Choosing our dominant coweights λ_1, λ_2 appropriately regular, the claim follows from Lemma 4.8. ☐

Proof of Theorem 4.2. In view of Corollary 4.9 and the definition of paths in the double Bruhat graph, it follows easily that, for all $x, y, z \in \widetilde{W}$, the degree of the polynomial $\varphi_{x,y,z}$ in $\mathbb{Z}[Q]$ is bounded from above by $\#\Phi^+$ (reproving this well-known fact). Thus the theorem follows by assuming $e \leq \#\Phi^+$ in Lemma 4.8 (noticing that also the multiset M cannot contain elements $> \#\Phi^+$ using the definition of paths in the double Bruhat graph). ☐

4.2. *Class polynomial.* Choose for each σ-conjugacy class $\mathcal{O} \subseteq \widetilde{W}$ a minimal-length element $x_{\mathcal{O}} \in \mathcal{O}$. Then the class polynomials associated with each $x \in \widetilde{W}$ are the uniquely determined polynomials $f_{x,\mathcal{O}} \in \mathbb{Z}[Q]$ satisfying

$$T_x \equiv \sum_{\mathcal{O}} f_{x,\mathcal{O}} T_{x_{\mathcal{O}}} \pmod{[\mathcal{H}, \mathcal{H}]_\sigma},$$

where $[\mathcal{H}, \mathcal{H}]_\sigma$ is the $\mathbb{Z}[Q]$-submodule of $\mathcal{H}$ generated by the elements of the form

$$[h, h']_\sigma = hh' - h'\sigma(h) \in \mathcal{H}.$$

These polynomials $f_{x,\mathcal{O}} \in \mathbb{Z}[Q]$ are independent of the choice of minimal-length representatives $x_{\mathcal{O}} \in \mathcal{O}$, and there is an explicit algorithm to compute them; see [He and Nie 2014]. Using this algorithm, one easily sees the following boundedness property: Whenever $\ell(x) < \ell(x_{\mathcal{O}})$, we must have $f_{x,\mathcal{O}} = 0$. The main result of this section is the following.

Theorem 4.10. *Let $B > 0$ be any real number. There exists an explicitly described constant $B' > 0$, depending only on B and the root system Φ, such that the following holds true*:

Let $x = w\varepsilon^\mu \in \widetilde{W}$ be B'-regular and write $\mathrm{LP}(x) = \{v\}$. *For each σ-conjugacy class $\mathcal{O} \subseteq \widetilde{W}$ with $\langle v^{-1}\mu - \nu(\mathcal{O}), 2\rho \rangle \leq B$ and $\kappa(\mathcal{O}) = \kappa(x)$, we have*

$$f_{x,\mathcal{O}} = \sum_{\substack{(\omega, e) \in \mathrm{wts}(v \Rightarrow \sigma(wv)) \text{ s.t.} \\ \nu(\mathcal{O}) = \mathrm{avg}_\sigma(v^{-1}\mu - \omega)}} Q^e \in \mathbb{Z}[Q].$$

Remark 4.11. (a) Our proof reduces Theorem 4.10 to Theorem 4.2. This yields a short and instructive proof, but results in a very large value of B'. One may alternatively compare the aforementioned algorithm of He and Nie directly with Theorem 3.2 to obtain a significantly smaller value of B'.

(b) Explicit formulas for the full class polynomials, rather than just degree and sometimes leading coefficients, have been very rare in the past. One exception to this is the elements with finite Coxeter part as studied in [He et al. 2024]. In the setting of Theorem 4.10, this means that $v^{-1}\sigma(wv) \in W$ has a reduced expression in W where every occurring simple reflection lies in a different σ-orbit in S. Then the class polynomial from [loc. cit., Theorem 7.1] is, translating to our notation as above, given by $Q^{\ell(v^{-1}\sigma(wv))}$.

Write $v^{-1}\sigma(wv) = s_{\alpha_1}\cdots s_{\alpha_n}$ for such a reduced expression as above, and choose a reflection order $\prec$ with $\alpha_1 \prec \cdots \prec \alpha_n$. Then one sees that there is only one unlabelled $\prec$-increasing path from v to $\sigma(wv)$ in the double Bruhat graph, given by

$$v \to v s_{\alpha_1} \to \cdots \to v s_{\alpha_1}\cdots s_{\alpha_n} = \sigma(wv).$$

This path has length n. Since the simple coroots $\alpha_1, \ldots, \alpha_n$ lie in pairwise distinct σ-orbits, it follows for any coroot $\omega \in \mathbb{Z}\Phi^\vee$ that there is at most one choice of integers $m_1, \ldots, m_n \in \mathbb{Z}$ with

$$m_1\alpha_1^\vee + \cdots + m_n\alpha_n^\vee \equiv \omega \in X_*(T)_\Gamma.$$

With a bit of bookkeeping, one may explicitly describe $\mathrm{wts}(v \Rightarrow \sigma(wv))$ as a multiset of pairs (ω, n), each with multiplicity 1, for exactly those coweights ω which are nonnegative linear combinations of the simple coroots $\alpha_1^\vee, \ldots, \alpha_n^\vee$. This easy double Bruhat theoretic calculation recovers [He et al. 2024, Theorem 7.1] in the setting of Theorem 4.10.

(c) Let $J \subseteq \Delta$ be the support of $v^{-1}\sigma(wv)$ in W. Let $v^J \in W^J$ be the unique minimal-length element in vJ. Write $v = v^J v_1$ and $\sigma(wv) = v^J v_2$ so that $v_1, v_2 \in W_J$. Choosing a suitable reflection order, we get a one to one correspondence between paths in the double Bruhat graph of W from v to $\sigma(wv)$ and paths in the double Bruhat graph of W_J from v_1 to v_2. The resulting statement on class polynomials recovers [He and Nie 2015, Theorem C] in the setting of Theorem 4.10.

Proof of Theorem 4.10. Define $C_1 := B + 1$, and let $C_2 > 0$ be as in Theorem 4.2.

By choosing B' appropriately, we may assume that we can write x as a length-additive product

$$x = x_1x_2, \quad x_1 = wv\varepsilon^{\mu_1}, \quad x_2 = v^{-1}\varepsilon^{\mu_2}$$

such that x_1 is $2\ell(x_2)$-regular and x_2 is C_2-regular. Observe that $\mathrm{LP}(x_2) = \{v\}$ and $\mathrm{LP}(x_1) = \{1\}$. Then

$$T_x = T_{x_1}T_{x_2} \equiv T_{x_2}\sigma(T_{x_1}) \pmod{[\mathcal{H}, \mathcal{H}]_\sigma}.$$

Write $\mathcal{H}_{\leq \ell(x)-B-1}$ for the $\mathbb{Z}[Q]$-submodule of $\mathcal{H}$ generated by all elements T_z satisfying $\ell(z) < \ell(x) - B$.

Using Theorem 1.2, we may write

$$T_{x_2}T_{\sigma(x_1)} \equiv \sum_{(\omega,e)\in \mathrm{wts}(v\Rightarrow\sigma(wv))} Q^e T_{\varepsilon^{v^{-1}\mu-\omega}} \pmod{\mathcal{H}_{\leq \ell(x)-B-1}}.$$

So if $\mathcal{O}$ satisfies $\langle v^{-1}\mu - \nu(\mathcal{O}), 2\rho\rangle \leq B$, we see that

$$f_{x,\mathcal{O}} = \sum_{(\omega,e)\in \mathrm{wts}(v\Rightarrow\sigma(wv))} Q^e f_{\varepsilon^{v^{-1}\mu-\omega},\mathcal{O}}.$$

Here, we used the above observation that $f_{y,\mathcal{O}} = 0$ if $\ell(y) < \langle \nu(\mathcal{O}), 2\rho\rangle$. By regularity of $v^{-1}\mu$ with respect to ω, we see that $v^{-1}\mu - \omega$ is always dominant and 1-regular in the above sum. Hence

$$f_{\varepsilon^{v^{-1}\mu-\omega},\mathcal{O}} = \begin{cases} 1 & \text{if } \nu(\mathcal{O}) = \operatorname{avg}_\sigma(v^{-1}\mu - \omega), \\ 0 & \text{otherwise.} \end{cases}$$

The claim follows. □

5. Affine Deligne–Lusztig varieties

One crucial feature of the class polynomials $f_{x,\mathcal{O}}$ is that they encode important information on the geometry of affine Deligne–Lusztig varieties.

Theorem 5.1 [He 2016, Theorem 2.19]. *Let $x \in \widetilde{W}$ and $[b] \in B(G)$. Define*

$$f_{x,[b]} := \sum_{\mathcal{O}} Q^{\ell(\mathcal{O})} f_{x,\mathcal{O}} \in \mathbb{Z}[Q],$$

where the sum is taken over all σ-conjugacy classes $\mathcal{O} \subset \widetilde{W}$ whose image in $B(G)$ is $[b]$. For each such σ-conjugacy class $\mathcal{O}$, we write

$$\ell(\mathcal{O}) = \min\{\ell(y) \mid y \in \mathcal{O}\}.$$

Then $X_x(b) \neq \varnothing$ if and only if $f_{x,[b]} \neq 0$. In this case,

$$\dim X_x(b) = \tfrac{1}{2}(\ell(x) + \deg(f_{x,[b]})) - \langle \nu(b), 2\rho\rangle$$

and the number of $J_b(F)$-orbits of top-dimensional irreducible components in $X_x(b)$ is equal to the leading coefficient of $f_{x,[b]}$. □

Combining with the explicit description of class polynomials from Theorem 4.10, we conclude the following.

Proposition 5.2. *Let $B > 0$ be any real number. There exists an explicitly described constant $B' > 0$, depending only on B and the root system Φ, such that the following holds true*:

Let $x = w\varepsilon^\mu \in \widetilde{W}$ be B'-regular and write $\mathrm{LP}(x) = \{v\}$. Let $[b] \in B(G)$ such that $\langle v^{-1}\mu - \nu(b), 2\rho\rangle < B$ and $\kappa(b) = \kappa(x)$. Let E denote the multiset

$$E = \{e \mid (\omega, e) \in \mathrm{wts}(v \Rightarrow \sigma(wv)) \text{ such that } \nu(b) = \operatorname{avg}_\sigma(v^{-1}\mu - \omega)\}_m.$$

Then $X_x(b) \neq \varnothing$ if and only if $E \neq \varnothing$. In this case, set $e := \max(E)$. Then

$$\dim X_x(b) = \tfrac{1}{2}(\ell(x) + e - \langle \nu(b), 2\rho\rangle),$$

and the number of $J_b(F)$-orbits of top-dimensional irreducible components of $X_x(b)$ is equal to the multiplicity of e in E.

Proof. Let $\mathcal{O} \subseteq \widetilde{W}$ be the unique σ-conjugacy class whose image in $B(G)$ is $[b]$ (unique by regularity). Then $\ell(\mathcal{O}) = \langle \nu(b), 2\rho \rangle$ and $f_{x,[b]} = Q^{\ell(\mathcal{O})} f_{x,\mathcal{O}}$. Expressing

$$f_{x,\mathcal{O}} = \sum_{e \in E} Q^e$$

using Theorem 4.10, the statements follow immediately using Theorem 5.1. □

For split groups, this recovers [Schremmer 2025, Corollary 5.9] up to possibly different regularity constraints. In practice, one may use Proposition 5.2 to deduce statements on the double Bruhat graph from the well-studied theory of affine Deligne–Lusztig varieties.

Corollary 5.3. *Let $u, v \in W$ and let $J = \operatorname{supp}(u^{-1}v) \subseteq \Delta$ be the support of $u^{-1}v$ in W, and $\omega \in \mathbb{Z}\Phi^\vee$.*

(a) *Suppose that $\ell(u^{-1}v)$ is equal to $d_{\mathrm{QB}(W)}(u \Rightarrow v)$, the length of a shortest path from u to v in the quantum Bruhat graph. Then $(\omega, \ell(u^{-1}v)) \in \operatorname{wts}(u \Rightarrow v)$ whenever $\omega \geq \operatorname{wt}_{\mathrm{QB}(W)}(u \Rightarrow v)$ and $\omega \in \mathbb{Z}\Phi_J^\vee$.*

(b) *If $\omega \in \mathbb{Z}\Phi_J^\vee$ with $\omega \geq 2\rho_J^\vee$, which denotes the sum of positive coroots in $\Phi_J^\vee$, we have*

$$(\omega, \ell(u^{-1}v)) \in \operatorname{wts}(u \Rightarrow v).$$

Proof. Assume without loss of generality that the group G is split. Reducing to the double Bruhat graph of W_J as in Remark 4.11(d), we may and do assume that $J = \Delta$.

Let $B = \langle \omega, 2\rho \rangle + 1$ and $B' > 0$ as in Proposition 5.2. Choose $x = w\varepsilon^\mu \in \widetilde{W}$ to be B'-superregular such that $\mathrm{LP}(x) = \{u\}$ and $v = wu$. Let $[b] \in B(G)$ be the σ-conjugacy class containing $\varepsilon^{u^{-1}\mu - \omega}$, so that $\nu(b) = u^{-1}\mu - \omega$.

(a) By [Milićević and Viehmann 2020, Proposition 4.2], the element x is cordial. By [Milićević and Viehmann 2020, Theorem 1.1] and [Görtz et al. 2015, Theorem B], we get $X_x(b) \neq \varnothing$ and

$$\dim X_x(b) = \tfrac{1}{2}\big(\ell(x) + \ell(u^{-1}v) - \langle \nu(b), 2\rho \rangle\big).$$

The claim follows.

(b) Similar to (a), using [He 2021, Theorem 1.1]. This celebrated result of He shows that if $\omega \geq 2\rho^\vee$ and $\operatorname{supp}(u^{-1}v) = \Delta$, then $X_x(b) \neq \varnothing$ and

$$\dim X_x(b) = \tfrac{1}{2}\big(\ell(x) + \ell(u^{-1}v) - \langle \nu(b), 2\rho \rangle\big).$$

The claim follows again. □

The reader who wishes to familiarize themselves more with the combinatorics of double Bruhat graphs may take the challenge and prove the above corollary directly.

We now want to state the main result of this section, describing the nonemptiness pattern and dimensions of affine Deligne–Lusztig varieties associated with sufficiently regular elements $x \in \widetilde{W}$ and arbitrary $[b] \in B(G)$. We let $\lambda(b) \in X_*(T)_\Gamma$ be the λ-invariant as introduced in [Hamacher and Viehmann 2018, Section 2]. By $\operatorname{conv} : X_*(T)_\Gamma \to X_*(T)_{\Gamma_0} \otimes \mathbb{Q}$, we denote the convex hull map from [Schremmer 2022, Section 3.1], so that $\nu(b) = \operatorname{conv}(\lambda(b))$.

Our regularity condition is given as follows: Decompose the (finite) Dynkin diagram of Φ into its connected components, so we have $\Phi = \Phi_1 \sqcup \cdots \sqcup \Phi_c$. Denote by $\theta_i \in \Phi_i^+$ the uniquely determined highest root, and write it as linear combination of simple roots

$$\theta_i = \sum_{\alpha\in\Delta} c_{i,\alpha}\alpha.$$

Define the regularity constant C to be

$$C = 1 + \max_{i=1,\dots,c} \sum_{\alpha\in\Delta} c_{i,\alpha} \in \mathbb{Z}.$$

With that, we can state our main result as follows.

Theorem 5.4. *Let $x = w\varepsilon^\mu \in \widetilde{W}$ be C-regular and $[b] \in B(G)$ such that $\kappa(b) = \kappa(x)$. Write $\mathrm{LP}(x) = \{v\}$ and define E to be either of the following two sets E_1 or E_2:*

$$E_1 := \{e \mid (\omega, e) \in \mathrm{wts}(v \Rightarrow \sigma(wv)) \text{ such that } \lambda(b) \equiv v^{-1}\mu - \omega \in X_*(T)_\Gamma\},$$

$$E_2 := \{e \mid (\omega, e) \in \mathrm{wts}(v \Rightarrow \sigma(wv)) \text{ such that } \nu(b) = \mathrm{conv}(v^{-1}\mu - \omega)\}.$$

Then $X_x(b) \neq \varnothing$ if and only if $E \neq \varnothing$. In this case,

$$\dim X_x(b) = \tfrac{1}{2}\big(\ell(x) + \max(E) - \langle \nu(b), 2\rho\rangle - \mathrm{def}(b)\big).$$

Remark 5.5. (a) Since $\mathrm{conv}(\lambda(b)) = \nu(b)$, we have $E_1 \subseteq E_2$. The inclusion may be strict, and it is a nontrivial consequence of Theorem 5.4 that the two sets have the same maxima.

(b) If Φ is irreducible, the regularity constant C is equal to the *Coxeter number* of Φ and explicitly given as follows:

Cartan type	A_n	B_n	C_n	D_n	E_6	E_7	E_8	F_4	G_2
$C =$	$n+1$	$2n$	$2n$	$2n-2$	12	18	30	12	6

(c) Unlike in Proposition 5.2, we get no information on the number of top-dimensional irreducible components. The main advantage of Theorem 5.4 over Proposition 5.2 comes from the different regularity conditions, making Theorem 5.4 more applicable.

(d) The unique minimum in $\mathrm{wts}(v \Rightarrow \sigma(wv))$ from [Schremmer 2025, Proposition 4.13] corresponds to the unique maximum in $B(G)_x$. This recovers the formula for the generic Newton point from [He and Nie 2024, Proposition 3.1] in the setting of Theorem 5.4.

(e) If the difference between $v^{-1}\mu$ and $\nu(b)$ becomes sufficiently large, the maximum $\max(E)$ can be expected to be is $\ell(v^{-1}\sigma(wv))$ (see [Schremmer 2025, Lemma 4.11] or Corollary 5.3(b) above) and we recover the notion of virtual dimension from [He 2014, Section 10]. In fact, one may use Corollary 5.3(b) to recover [He 2021, Theorem 1.1] in the situation of Theorem 5.4. This line of argumentation is ultimately cyclic, since a special case of [He 2021, Theorem 1.1] was used in the proof of Corollary 5.3(b). We may however summarize that Corollary 5.3(b) is the double Bruhat theoretic correspondent of [He 2021, Theorem 1.1]. Similarly, most known results on affine Deligne–Lusztig varieties correspond to theorems on the double Bruhat graph and vice versa.

(f) The proof method for Theorem 5.4 is similar to the proof of [He 2014, Proposition 11.5] or equivalently the proof of [He 2021, Theorem 1.1].

Proof of Theorem 5.4. We assume without loss of generality that the group G is of adjoint type, following [Görtz et al. 2015, Section 2]. This allows us to find a coweight $\mu_v \in X_*(T)_{\Gamma_0}$ satisfying for each simple root $\alpha \in \Delta$ the condition

$$\langle \mu_v, \alpha\rangle = \Phi^+(-v\alpha) = \begin{cases} 1, & v\alpha \in \Phi^-,\\ 0, & v\alpha\in\Phi^+.\end{cases}$$

It follows that $\langle \mu_v, \beta\rangle \geq \Phi^+(-v\beta)$ for all $\beta\in\Phi^+$. Define

$$x_1 := wv\varepsilon^{v^{-1}\mu-\mu_v},\quad x_2 = v^{-1}\varepsilon^{v\mu_v}\in\widetilde{W}.$$

By the choice of μ_v, we see that $v^{-1}\mu-\mu_v$ is dominant and $(C-1)$-regular. The above estimate $\langle\mu_v,\beta\rangle\geq\Phi^+(-v\beta)$ implies $v\in \mathrm{LP}(x_2)$. Hence $x = x_1x_2$ is a length-additive product. We obtain

$$T_x = T_{x_1}T_{x_2}\equiv T_{\sigma^{-1}(x_2)}T_{x_1} \pmod{[\mathcal{H},\mathcal{H}]_\sigma}.$$

Define the multiset Y via

$$T_{\sigma^{-1}(x_2)}T_{x_1} = \sum_{(y,e)\in Y} Q^e T_{yx_1}\in\mathcal{H}. \tag{5.6}$$

Then each $(y,e)\in Y$ satisfies $y\leq\sigma^{-1}(x_2)$ in the Bruhat order. Writing $y = w_y\varepsilon^{\mu_y}$, we get $\mu_y^{\mathrm{dom}}\leq\sigma^{-1}(\mu_v)$ in $X_*(T)_{\Gamma_0}$. We estimate

$$\max_{\beta\in\Phi}|\langle\mu_y,\beta\rangle| = \max_{\beta\in\Phi^+}\langle\mu_y^{\mathrm{dom}},\beta\rangle = \max_i\langle\mu_y^{\mathrm{dom}},\theta_i\rangle\leq\max_i\langle\mu_v,\theta_i\rangle\leq C-1,$$

by the choice of C. It follows that

$$yx_1 = w_ywv\varepsilon^{v^{-1}\mu-\mu_v+(wv)^{-1}\mu_y},$$

with $v^{-1}\mu-\mu_v+(wv)^{-1}\mu_y$ being dominant. For any dominant coweight $\lambda\in X_*(T)_{\Gamma_0}$, we can multiply (5.6) by T_{ε^λ} to obtain

$$T_{\sigma^{-1}(x_2)}T_{x_1\varepsilon^\lambda} = T_{\sigma^{-1}(x_2)}T_{x_1}T_{\varepsilon^\lambda} = \sum_{(y,e)\in Y} T_{yx_1}T_{\varepsilon^\lambda} = \sum_{(y,e)\in Y} T_{yx_1\varepsilon^\lambda}.$$

In light of Lemma 4.5, we see that the multiset Y is equal to the multiset $Y(\sigma^{-1}(x_2), wv)$ defined earlier.

For each $(y,e)\in Y$, write $yx_1 = \tilde{w}_y\varepsilon^{\tilde{\mu}_y}$ to define the sets

$$E_1(yx_1) := \{e \mid (\omega, e)\in \mathrm{wts}(1\Rightarrow\sigma(\tilde{w}_y)) \text{ such that } \lambda(b) = \tilde{\mu}_y-\omega\in X_*(T)_\Gamma\},$$
$$E_2(yx_1) := \{e \mid (\omega, e)\in \mathrm{wts}(1\Rightarrow\sigma(\tilde{w}_y)) \text{ such that } \nu(b) = \mathrm{conv}(\tilde{\mu}_y-\omega)\}.$$

Define $E(yx_1)$ to be $E_1(yx_1)$ or $E_2(yx_1)$ depending on whether E was chosen as E_1 or E_2. By Lemma 4.7(a), we may write $\mathrm{wts}(\sigma^{-1}(v)\Rightarrow wv)$ as the additive union of multisets

$$\begin{aligned}\mathrm{wts}(\sigma^{-1}(v)\Rightarrow wv) &= \bigcup_{(w_y\varepsilon^{\mu_y},e)\in Y(\sigma^{-1}(x_2),wv)}\{(\mu_v-(wv)^{-1}\mu_y+\omega, e+\ell)\mid(\omega,\ell)\in\mathrm{wts}(1\Rightarrow w_ywv)\}_m\\ &= \bigcup_{(y,e)\in Y}\{(v^{-1}\mu-\tilde{\mu}_y+\omega, e+\ell)\mid(\omega,\ell)\in\mathrm{wts}(1\Rightarrow\tilde{w}_y)\}_m.\end{aligned}\tag{5.7}$$

Note that the definition of the sets E_1, E_2, $E_1(yx_1)$, $E_2(yx_1)$ does not change if we apply σ^{-1} to the occurring weights ω. Hence (5.7) implies

$$E = \bigcup_{(y,e)\in Y} \{e+\ell \mid \ell \in E(yx_1)\}.$$

By definition of the multiset Y, the class polynomials of $f_{x,\mathcal{O}}$ for arbitrary σ-conjugacy classes $\mathcal{O} \subset \widetilde{W}$ are given by

$$f_{x,\mathcal{O}} = \sum_{(y,e)\in Y} Q^e f_{yx_1,\mathcal{O}}.$$

By Theorem 5.1, we see that $X_x(b) \neq \varnothing$ if and only if $X_{yx_1}(b) \neq \varnothing$ for some $(y,e) \in Y$. In this case, the dimension of $X_x(b)$ is the maximum of

$$\dim X_{yx_1}(b) + \tfrac{1}{2}(\ell(x) - \ell(yx_1) + e),$$

where (y,e) runs through all elements of Y satisfying $X_{yx_1}(b) \neq \varnothing$.

We see that it suffices to prove the following claim for all $(y,e) \in Y$:

$$X_{yx_1}(b) \neq \varnothing \text{ if and only if } E(yx_1) \neq \varnothing \text{ and in this case, we have} \\ \dim X_{yx_1}(b) = \tfrac{1}{2}\big(\ell(yx_1) + \max(E(yx_1)) - \langle \nu(b), 2\rho\rangle - \operatorname{def}(b)\big). \tag{$*$}$$

Writing $yx_1 = \tilde{w}\varepsilon^{\tilde{\mu}}$, we saw above that $\tilde{\mu}$ is dominant. Applying [Milićević and Viehmann 2020, Theorem 1.2] to the inverse of yx_1, or equivalently [He 2021, Theorem 4.2] directly to yx_1, we see that the element yx_1 is *cordial* in the sense of [Milićević and Viehmann 2020]. This gives a convenient criterion to check $X_{yx_1}(b) \neq \varnothing$ and to calculate its dimension. We saw in Corollary 5.3(a) that the multiset $\mathrm{wts}(1 \Rightarrow \sigma(\tilde{w}_y))$ must satisfy the analogous conditions. Let us recall these results.

The uniquely determined largest Newton point in $B(G)_{yx_1} = B(G)_{\tilde{w}\varepsilon^{\tilde{\mu}}}$ is $\mathrm{avg}_\sigma(\tilde{\mu})$; see [He 2021, Theorem 4.2].

Let $J' = \mathrm{supp}(\tilde{w}) \subseteq \Delta$ be the support of $\tilde{w}$ and $J = \bigcup_i \sigma^i(J') = \mathrm{supp}_\sigma(\tilde{w})$ its σ-support. Let $\pi_J : X_*(T)_{\Gamma_0} \to X_*(T)_{\Gamma_0} \otimes \mathbb{Q}$ be the corresponding function from [Chai 2000, Definition 3.2] or equivalently [Schremmer 2022, Section 3.1]. Then $\pi_J(\tilde{\mu})$ is the unique smallest Newton point occurring in $B(G)_{yx_1}$; see [Viehmann 2021, Theorem 1.1].

The condition of cordiality [Milićević and Viehmann 2020, Theorem 1.1] implies that $B(G)_{yx_1}$ contains all those $[b] \in B(G)$ with the correct Kottwitz point $\kappa(b) = \kappa(yx_1) = \kappa(x)$ and Newton point

$$\pi_J(\tilde{\mu}) \leq \nu(b) \leq \mathrm{avg}_\sigma(\tilde{\mu}).$$

In this case, we know moreover from [Milićević and Viehmann 2020, Theorem 1.1] that $X_{yx_1}(b)$ is equidimensional of dimension

$$\dim X_{yx_1}(b) = \tfrac{1}{2}\big(\ell(yx_1) + \ell(\tilde{w}) - \langle \nu(b), 2\rho\rangle - \operatorname{def}(b)\big).$$

This condition on Newton points is equivalent to $\mathrm{avg}_\sigma(\tilde{\mu}) - \nu(b)$ being a nonnegative $\mathbb{Q}$-linear combination of simple coroots of J, or equivalently $\tilde{\mu} - \lambda(b)$ being a nonnegative $\mathbb{Z}$-linear combination of these coroots.

On the double Bruhat side, note that $(\omega, e) \in \mathrm{wts}(1 \Rightarrow \tilde{w})$ implies $\omega \in \mathbb{Z}\Phi_{J'}^{\vee}$ and $e \leq \ell(\tilde{w})$. This can either be seen directly, similar to the proof of [Schremmer 2025, Lemma 4.11], or as in Corollary 5.3, reducing to [Viehmann 2021, Theorem 1.1]. From Corollary 5.3, we know conversely that any $\omega \geq 0$ with $\omega \in \mathbb{Z}\Phi_{J'}^{\vee}$ satisfies $(\omega, \ell(\tilde{w})) \in \mathrm{wts}(1 \Rightarrow \tilde{w})$.

Comparing these explicit descriptions of $\dim X_{yx_1}(b)$ and $\max(E(yx_1))$, we conclude the claim $(*)$. □

6. Outlook

We saw that the weight multiset of the double Bruhat graph can be used to describe the geometry of affine Deligne–Lusztig varieties in many cases. This includes the case of superparabolic elements x together with sufficiently large integral $[b] \in B(G)$ in split groups [Schremmer 2025, Theorem 5.7], as well as the case of sufficiently regular elements x together with arbitrary $[b] \in B(G)$ (Theorem 5.4). One may ask how much the involved regularity constants can be improved, and whether a unified theorem simultaneously generalizing [Schremmer 2025, Theorem 5.7] and Theorem 5.4 can be found. Towards this end, we propose a number of conjectures that would generalize our theorems in a straightforward manner.

Let $x = w\varepsilon^{\mu} \in \widetilde{W}$ and $[b] \in B(G)$. If $X_x(b) \neq \varnothing$, define the integer $D \in \mathbb{Z}_{\geq 0}$ such that

$$\dim X_x(b) = \tfrac{1}{2}\big(\ell(x) + D - \langle \nu(b), 2\rho\rangle - \mathrm{def}(b)\big),$$

and denote the number of $J_b(F)$-orbits of top-dimensional irreducible components in $X_x(b)$ by $C \in \mathbb{Z}_{\geq 1}$. We would like to state the following conjectures. The first conjecture makes a full prediction of the nonemptiness pattern and the dimension for elements x in the shrunken Weyl chamber and arbitrary $[b] \in B(G)$.

Conjecture 6.1. *Suppose that x lies in a shrunken Weyl chamber, i.e.,* $\mathrm{LP}(x) = \{v\}$ *for a uniquely determined $v \in W$. Define E to be either of the multisets*

$$E_1 := \{e \mid (\omega, e) \in \mathrm{wts}(v \Rightarrow \sigma(wv)) \text{ such that } \lambda(b) \equiv v^{-1}\mu - \omega \in X_*(T)_\Gamma\}_m,$$
$$E_2 := \{e \mid (\omega, e) \in \mathrm{wts}(v \Rightarrow \sigma(wv)) \text{ such that } \nu(b) = \mathrm{conv}(v^{-1}\mu - \omega)\}_m.$$

We make the following predictions.

(a) *$X_x(b) \neq \varnothing$ if and only if $E \neq \varnothing$ and $\kappa(x) = \kappa(b) \in \pi_1(G)_\Gamma$ (the latter condition on Kottwitz points is automatically satisfied if $E = E_1$).*

(b) *If $X_x(b) \neq \varnothing$, then* $\max(E) = D$.

(c) *If $X_x(b) \neq \varnothing$, then C is at most the multiplicity of D in E (which may be $+\infty$ for E_2).*

The multiset E_1 is always contained in E_2, since $\nu(b) = \mathrm{conv}(\lambda(b))$. The inclusion may be strict. So in fact we are suggesting two different dimension formulas for shrunken x, and claim that both yield the same answer, which moreover agrees with the dimension.

For sufficiently regular x, Theorem 5.4 shows (a) and (b). Under some strong superregularity conditions, Proposition 5.2 shows (c) with equality. While both proofs can certainly be optimized with regards to the involved regularity constants, proving Conjecture 6.1 as stated will likely require further methods. It is unclear how to show the conjecture, e.g., for the particular element $x = w_0\varepsilon^{-2\rho^\vee}$, since the proof method for Theorem 5.4 fails.

It is easy to see that Conjecture 6.1 is compatible with many known results on affine Deligne–Lusztig varieties, such as the ones recalled in the introduction of the previous article [Schremmer 2025, Theorem 1.2]. By Corollary 5.3, we see that parts (a) and (b) of Conjecture 6.1 hold true for cordial elements x. If x is of the special form $x = w_0\varepsilon^\mu$ with μ dominant, then x is in a shrunken Weyl chamber and we know that (c) holds; see [Schremmer 2025, Remark 6.11].

Our second conjecture suggests how the double Bruhat graph can be used for elements x which are not necessarily in shrunken Weyl chambers.

Conjecture 6.2. *Suppose that $[b]$ is **integral**, i.e., of defect zero. Define for each $v \in \mathrm{LP}(x)$ and $u \in W$ the multiset*

$$E(u, v) := \{e \mid (\omega, e) \in \mathrm{wts}(u \Rightarrow \sigma(wu) \dashrightarrow \sigma(wv)) \text{ such that } u^{-1}\mu - \omega = \lambda(b) \in X_*(T)_\Gamma\}_m.$$

Set $\max\varnothing := -\infty$ *and define*

$$d := \max_{u\in W}\min_{v\in \mathrm{LP}(x)} \max(E(u, v)) \in \mathbb{Z}_{\geq 0} \cup \{-\infty\},$$
$$c := \sum_{u\in W}\min_{v\in \mathrm{LP}(x)} (\text{multiplicity of } d \text{ in } E(u, v)) \in \mathbb{Z}_{\geq 0}.$$

We make the following predictions:

(a) *If there exists for every $u \in W$ some $v \in \mathrm{LP}(x)$ with $E(u, v) = \varnothing$, i.e., if $d = -\infty$, then $X_x(b) = \varnothing$.*

(b) *If $X_x(b) \neq \varnothing$, then $D \leq d$.*

(c) *If $X_x(b) \neq \varnothing$ and $D = d$, then $C \leq c$.*

If the group is split, then [Schremmer 2025, Theorem 5.7] proves (a), (b) and (c). Moreover, under some strong superparabolicity assumptions, we get the full conjecture including equality results for (b) and sometimes (c). We expect that a similar superparabolicity statement holds true for nonsplit groups, but it is unclear what the involved regularity constants should be, which is why we did not formulate a precise, falsifiable conjecture.

If the element $x \in \widetilde{W}$ is in a shrunken Weyl chamber with $\mathrm{LP}(x) = \{v\}$, then the multiset E_1 from Conjecture 6.1 is equal to the multiset $E(v, v)$ from Conjecture 6.2. If we moreover assume that Conjecture 6.1 holds true, then we get parts (a), (b) and (c) of Conjecture 6.2.

Compatibility of Conjecture 6.2 with previously known results is a lot harder to verify. We expect that one does not have to account for all pairs (u, v) as in Conjecture 6.2 to accurately describe nonemptiness and dimension of $X_x(b)$, similar to [Schremmer 2025, Theorem 5.7(c)] or Conjecture 6.1. However, we cannot make a precise prediction how such a refinement of Conjecture 6.2 should look in general.

Nonetheless, extensive computer searches did not yield a single counterexample to either conjecture. Most straightforward generalizations of these conjectures, however, can be disproved quickly using such a computer search [SageMath 2020; Sage-Combinat 2008].

Example 6.3. For both conjectures, the estimate on the number of irreducible components is only an upper bound. Indeed, it suffices to consider elements of the form $x = w_0\varepsilon^\mu$ for dominant cocharacters μ. Then, as discussed in [Schremmer 2025, Remark 5.11], the number C is equal to the dimension of the $\lambda(b)$-weight space of the *irreducible Weyl module* M_μ. The element x lies in a shrunken Weyl chamber, and the multiplicity of $d = D$ in $E_1 = E(v, v)$ is equal to the dimension of the $\lambda(b)$-weight space in the Verma module V_μ. These numbers are not equal in general.

Example 6.4. One may ask whether it is possible to find for each nonshrunken x an element $v \in \mathrm{LP}(x)$ such that the analogous statement of Conjecture 6.1 holds true. While this is certainly possible, say, for cordial elements x, such a statement cannot be expected to hold true in general. We may choose $G = \mathrm{GL}_4$ and $x = s_3s_2s_1\varepsilon^\mu$, where the pairing of μ with the simple roots $\alpha_1, \alpha_2, \alpha_3$ is given by $1, -1, 1$ respectively. Then $\mathrm{LP}(x) = \{s_2, s_2s_3\}$. For $[b]$ basic, we have $D = 3$, yet the analogous statements of Conjecture 6.1 for both possible choices of v in $\mathrm{LP}(x)$ would predict $D = 5$.

Example 6.5. Conjecture 6.2 should not be expected to hold for nonintegral $[b]$. Indeed, it suffices to choose $G = \mathrm{GL}_3$ and $x = w\varepsilon^\mu$ to be of length zero such that the action of x on the affine Dynkin diagram is nontrivial. Let $[b] = [x]$, so that $B(G)_x = \{[b]\}$. Define

$$E(u, v) := \{e \mid (\omega, e) \in \mathrm{wts}(u \Rightarrow wu \dashrightarrow wv) \text{ such that } u^{-1}\mu - \omega = \lambda(b) \in X_*(T)_\Gamma\}_m$$

for $u, v \in W = \mathrm{LP}(x)$. Since $w \neq 1$, we have $E(u, v) = \varnothing$ whenever $v = uw_0$. A statement analogous to Conjecture 6.2(a) would thus predict that $X_x(b) = \varnothing$, which is absurd.

Example 6.6. The converse of Conjecture 6.2(a) should not be expected to hold, even for $[b]$ basic. The construction in Conjecture 6.2 can fail to detect (J, w, δ)-alcove elements, and hence falsely predict a nonempty basic locus. For a concrete example, one may choose $G = \mathrm{GL}_3$ and x to be the shrunken element $x = s_2\varepsilon^{\rho^\vee}$, with $\langle \rho^\vee, \alpha\rangle = 1$ for all simple roots α. Then $\mathrm{LP}(x) = \{1\}$. For $u = s_1s_2$ and $[b] = [1]$ basic, we have $E(u, 1) \neq \varnothing$.

Acknowledgements

I was partially supported by the German Academic Scholarship Foundation, the Marianne-Plehn programme, the DFG Collaborative Research Centre 326 *GAUS* and the Chinese University of Hong Kong. I thank Eva Viehmann, Xuhua He and Qingchao Yu for inspiring discussions, and Eva Viehmann again for her comments on a preliminary version of this article. I am very grateful for the careful reading and helpful comments by the anonymous referee.

References

[Bonnafé 2017] C. Bonnafé, *Kazhdan–Lusztig cells with unequal parameters*, Algebra Appl. **24**, Springer, 2017. MR Zbl

[Chai 2000] C.-L. Chai, "Newton polygons as lattice points", *Amer. J. Math.* **122**:5 (2000), 967–990. MR Zbl

[Deligne and Lusztig 1976] P. Deligne and G. Lusztig, "Representations of reductive groups over finite fields", *Ann. of Math.* (2) **103**:1 (1976), 103–161. MR Zbl

[Görtz and He 2010] U. Görtz and X. He, "Dimensions of affine Deligne–Lusztig varieties in affine flag varieties", *Doc. Math.* **15** (2010), 1009–1028. MR Zbl

[Görtz et al. 2006] U. Görtz, T. J. Haines, R. E. Kottwitz, and D. C. Reuman, "Dimensions of some affine Deligne–Lusztig varieties", *Ann. Sci. École Norm. Sup.* (4) **39**:3 (2006), 467–511. MR Zbl

[Görtz et al. 2015] U. Görtz, X. He, and S. Nie, "$\mathbb{P}$-alcoves and nonemptiness of affine Deligne–Lusztig varieties", *Ann. Sci. École Norm. Sup.* (4) **48**:3 (2015), 647–665. MR Zbl

[Haines and Rapoport 2008] T. Haines and M. Rapoport, "On parahoric subgroups", 2008. Appendix to G. Pappas and M. Rapoport, "Twisted loop groups and their affine flag varieties", *Adv. Math.* **219**:1 (2008), 118–198. Zbl

[Hamacher and Viehmann 2018] P. Hamacher and E. Viehmann, "Irreducible components of minuscule affine Deligne–Lusztig varieties", *Algebra Number Theory* **12**:7 (2018), 1611–1634. MR Zbl

[He 2014] X. He, "Geometric and homological properties of affine Deligne–Lusztig varieties", *Ann. of Math.* (2) **179**:1 (2014), 367–404. MR Zbl

[He 2016] X. He, "Hecke algebras and p-adic groups", pp. 73–135 in *Current developments in mathematics* (Cambridge, MA, 2015), Int. Press, Somerville, MA, 2016. MR Zbl

[He 2018] X. He, "Some results on affine Deligne–Lusztig varieties", pp. 1345–1365 in *Proceedings of the International Congress of Mathematicians*, *II* (Rio de Janiero, 2018), World Sci., Hackensack, NJ, 2018. MR Zbl

[He 2021] X. He, "Cordial elements and dimensions of affine Deligne–Lusztig varieties", *Forum Math. Pi* **9** (2021), art. id. e9. MR Zbl

[He and Nie 2014] X. He and S. Nie, "Minimal length elements of extended affine Weyl groups", *Compos. Math.* **150**:11 (2014), 1903–1927. MR Zbl

[He and Nie 2015] X. He and S. Nie, "P-alcoves, parabolic subalgebras and cocenters of affine Hecke algebras", *Selecta Math.* (*N.S.*) **21**:3 (2015), 995–1019. MR Zbl

[He and Nie 2024] X. H. He and S. A. Nie, "Demazure product of the affine Weyl groups", *Acta Math. Sinica* (*Chinese Ser.*) **67**:2 (2024), 296–306. In Chinese. MR Zbl

[He et al. 2024] X. He, S. Nie, and Q. Yu, "Affine Deligne–Lusztig varieties with finite Coxeter parts", *Algebra Number Theory* **18**:9 (2024), 1681–1714. MR Zbl

[Iwahori and Matsumoto 1965] N. Iwahori and H. Matsumoto, "On some Bruhat decomposition and the structure of the Hecke rings of $\mathfrak{p}$-adic Chevalley groups", *Inst. Hautes Études Sci. Publ. Math.* **1965**:25 (1965), 5–48. MR Zbl

[Kottwitz 1985] R. E. Kottwitz, "Isocrystals with additional structure", *Compos. Math.* **56**:2 (1985), 201–220. MR Zbl

[Kottwitz 1997] R. E. Kottwitz, "Isocrystals with additional structure, II", *Compos. Math.* **109**:3 (1997), 255–339. MR Zbl

[Lenart et al. 2015] C. Lenart, S. Naito, D. Sagaki, A. Schilling, and M. Shimozono, "A uniform model for Kirillov–Reshetikhin crystals, I: Lifting the parabolic quantum Bruhat graph", *Int. Math. Res. Not.* **2015**:7 (2015), 1848–1901. MR Zbl

[Lusztig 1980] G. Lusztig, "Hecke algebras and Jantzen's generic decomposition patterns", *Adv. Math.* **37**:2 (1980), 121–164. MR Zbl

[Milićević and Viehmann 2020] E. Milićević and E. Viehmann, "Generic Newton points and the Newton poset in Iwahori-double cosets", *Forum Math. Sigma* **8** (2020), art. id. e50. MR Zbl

[Naito and Watanabe 2017] S. Naito and H. Watanabe, "A combinatorial formula expressing periodic R-polynomials", *J. Combin. Theory Ser. A* **148** (2017), 197–243. MR Zbl

[Rapoport 2005] M. Rapoport, "A guide to the reduction modulo p of Shimura varieties", pp. 271–318 in *Formes automorphes*, *I* (Paris, 2000), edited by J. Tilouine et al., Astérisque **298**, Soc. Math. France, Paris, 2005. MR Zbl

[Sage-Combinat 2008] "Sage-Combinat: enhancing Sage as a toolbox for computer exploration in algebraic combinatorics", software project, 2008, available at http://combinat.sagemath.org.

[SageMath 2020] *SageMath, version 9.2*, 2020, available at https://www.sagemath.org.

[Satake 1963] I. Satake, "Theory of spherical functions on reductive algebraic groups over $\mathfrak{p}$-adic fields", *Inst. Hautes Études Sci. Publ. Math.* **1963**:18 (1963), 5–69. MR Zbl

[Schremmer 2022] F. Schremmer, "Generic Newton points and cordial elements", preprint, 2022. arXiv 2205.02039

[Schremmer 2024] F. Schremmer, "Affine Bruhat order and Demazure products", *Forum Math. Sigma* **12** (2024), art. id. e53. MR Zbl

[Schremmer 2025] F. Schremmer, "Affine Deligne–Lusztig varieties via the double Bruhat graph, I: Semi-infinite orbits", *Algebra Number Theory* **19**:10 (2025), 1973–2014.

[Tits 1979] J. Tits, "Reductive groups over local fields", pp. 29–69 in *Automorphic forms, representations and L-functions*, *I* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979. MR Zbl

[Viehmann 2021] E. Viehmann, "Minimal Newton strata in Iwahori double cosets", *Int. Math. Res. Not.* **2021**:7 (2021), 5349–5365. MR Zbl

Communicated by Wee Teck Gan
Received 2023-06-12 Revised 2024-08-08 Accepted 2024-10-18

schr@hku.hk *Department of Mathematics and New Cornerstone Laboratory, The University of Hong Kong, Hong Kong*

ALGEBRA AND NUMBER THEORY 19:10 (2025)
https://doi.org/10.2140/ant.2025.19.2049

Paucity of rational points on fibrations with multiple fibres

Tim Browning, Julian Lyczak and Arne Smeets

Given a family of varieties over the projective line, we study the density of fibres that are everywhere locally soluble in the case that components of higher multiplicity are allowed. We use log geometry to formulate a new sparsity criterion for the existence of everywhere locally soluble fibres and formulate new conjectures that generalise previous work of Loughran and Smeets. These conjectures involve geometric invariants of the associated multiplicity orbifolds on the base of the fibration in the spirit of Campana. We give evidence for the conjectures by providing an assortment of bounds using Chebotarev's theorem and sieve methods, with most of the evidence involving upper bounds.

1. Introduction

Let X be a smooth, proper, geometrically irreducible variety over $\mathbb{Q}$, which is equipped with a dominant morphism $\pi : X \to \mathbb{P}^1$ with geometrically integral generic fibre. We shall refer to such fibrations as *standard*. The focus of this article is on situations where multiple fibres are present. Work of Colliot-Thélène, Skorobogatov and Swinnerton-Dyer [Colliot-Thélène et al. 1997] shows that the set $X(\mathbb{Q})$ of $\mathbb{Q}$-rational points on X is not Zariski dense when there are at least 5 geometric double fibres. Our goal is to put this kind of result on a quantitative footing by analysing the simpler question of solubility over the ring of adèles $\boldsymbol{A}_{\mathbb{Q}}$. Let

$$N_{\text{loc}}(\pi, H, B) = \#\{x \in \mathbb{P}^1(\mathbb{Q}) \cap \pi(X(\boldsymbol{A}_{\mathbb{Q}})) : H(x) \leqslant B\}, \tag{1-1}$$

MSC2020: primary 14G05; secondary 11G35, 11N36, 14A21, 14D10.
Keywords: local solubility, fibration, sieve methods, multiple fibres, Halphen surface.

where H is a height function on $\mathbb{P}^1(\mathbb{Q})$. In general, we will need to allow the height H to be any adelic height on a line bundle $\mathcal{O}(d)$. However, most of the time we shall use an $\mathcal{O}(1)$-height. In this case we will simply write $N_{\mathrm{loc}}(\pi, H, B) = N_{\mathrm{loc}}(\pi, B)$. Usually we will take the naive height $H(x) = \max\{|x_0|, |x_1|\}$ if $x \in \mathbb{P}^1(\mathbb{Q})$ is represented by a vector $\boldsymbol{x} = (x_0, x_1) \in \mathbb{Z}^2_{\mathrm{prim}}$, in which case it is easy to prove that

$$\#\{x \in \mathbb{P}^1(\mathbb{Q}) : H(x) \leqslant B\} \sim \frac{2}{\zeta(2)} B^2 \quad \text{as } B \to \infty.$$

Loughran and Smeets [2016] have shown that

$$N_{\mathrm{loc}}(\pi, B) \ll \frac{B^2}{(\log B)^{\Delta(\pi)}} \tag{1-2}$$

for a certain exponent $\Delta(\pi) \geqslant 0$ that is defined in terms of the data of the fibration. (Here, as throughout our work, all implied constants are allowed to depend on the fibration π.) Roughly speaking, the size of $\Delta(\pi)$ is determined by the number of nonsplit fibres of π, thereby lending credence to a philosophy put forward by Serre [1990] and further developed by Loughran [2018]. In [Loughran and Smeets 2016, Conjecture 1.6] it is conjectured that the upper bound (1-2) is sharp provided that the fibre of π over every closed point of $\mathbb{P}^1$ has an irreducible component of multiplicity 1. (In fact, [Loughran and Smeets 2016] works over arbitrary number fields k and concerns fibrations $X \to \mathbb{P}^n$ over projective space of arbitrary dimension, but we shall restrict to $k = \mathbb{Q}$ and $n = 1$ in our work.) Our goal is to explore what happens to $N_{\mathrm{loc}}(\pi, B)$ when the assumption about components of multiplicity 1 is violated.

There are relatively few examples in the number theory literature that feature standard fibrations with multiple fibres. When the generic fibre of π is rationally connected, it follows from work of Graber, Harris and Starr [Graber et al. 2003] that every fibre contains a geometrically integral component of multiplicity 1. In particular, when $\dim X = 2$, we must look to fibrations over $\mathbb{P}^1$ into curves of positive genus to find examples with multiple fibres. Let $c, d, f \in \mathbb{Q}[t]$ be nonzero polynomials such that f is square-free of even degree and such that f and $c - d$ are coprime. Let $\pi : X \to \mathbb{P}^1$ be a smooth, proper model of the affine variety cut out by the pair of equations

$$x^2 - c(t) = f(t)y^2, \quad x^2 - d(t) = f(t)z^2. \tag{1-3}$$

Then it follows from [Colliot-Thélène et al. 1997, Proposition 4.1] that all the fibres of π over the zeros of f are double fibres and that the generic fibre is a geometrically integral curve whose projective model is isomorphic to a curve of genus 1. When $\deg(f) \geqslant 6$, as pointed out in [Loughran and Matthiesen 2024, Theorem 1.4], the argument of [Colliot-Thélène et al. 1997, Corollary 2.2] implies that $N_{\mathrm{loc}}(\pi, B) = O(1)$. Further examples involving genus-2 fibrations over $\mathbb{P}^1$ have been worked out in [Stoppino 2011].

In the spirit of [Campana 2005], our approach to this problem comes from relating the arithmetic of $\pi : X \to \mathbb{P}^1$ to the arithmetic of the *orbifold base* $(\mathbb{P}^1, \partial_\pi)$ for a certain $\mathbb{Q}$-divisor ∂_π, in the sense of Definition 4.6. For each closed point $D \in (\mathbb{P}^1)^{(1)}$, we let $m_D \geqslant 1$ denote the minimum multiplicity of the irreducible components of $\pi^{-1}(D)$. We will call the fibre over D *multiple* if $m_D > 1$. We emphasise that we have not defined the multiplicity of a fibre as the greatest common divisor of the multiplicities of its components, as is common in some applications, but rather as the minimum. Then we may define

$$\partial_\pi = \sum_{D\in(\mathbb{P}^1)^{(1)}} \Big(1-\frac{1}{m_D}\Big)[D]. \tag{1-4}$$

With this notation, we make the following conjecture.

Conjecture 1.1. *Let $\pi : X \to \mathbb{P}^1$ be a standard fibration such that the $\mathbb{Q}$-divisor $-(K_{\mathbb{P}^1}+\partial_\pi)$ is ample. Then*

$$N_{\mathrm{loc}}(\pi, B) = O_\varepsilon(B^{2-\deg\partial_\pi+\varepsilon})$$

for any $\varepsilon > 0$.

Note that $-\deg(K_{\mathbb{P}^1}+\partial_\pi) = 2-\deg\partial_\pi$. Hence $-(K_{\mathbb{P}^1}+\partial_\pi)$ is ample if and only if $\deg\partial_\pi < 2$. The main feature of Conjecture 1.1 is that we expect $N_{\mathrm{loc}}(\pi, B)$ to be much smaller in the presence of multiple fibres. Our remaining results give evidence towards this, as well as a proposal about the replacement of B^ε by an explicit *nonpositive* power of $\log B$.

In the case that $\deg\partial_\pi > 2$, the Mordell orbifold conjecture shows that the rational points of X can only lie in finitely many fibres of π. This conjecture follows from the *abc*-conjecture, as shown by Smeets [2017]. Examples where the conclusion can be proven unconditionally are found in [Colliot-Thélène et al. 1997]. It remains unclear what can be said generally about the number of everywhere locally soluble fibres in this situation. In the intermediate case $\deg\partial_\pi = 2$, very little is known about the number of soluble or everywhere locally soluble fibres.

1.1. ***Upper bounds.*** For each closed point $D \in (\mathbb{P}^1)^{(1)}$, let S_D be the set of geometrically irreducible components of $\pi^{-1}(D)$ of multiplicity m_D, and let $\kappa(D)$ be the residue field. For any number field $N/\mathbb{Q}$, we write

$$\delta_{D,N}(\pi) = \frac{\#\{\sigma\in\Gamma_{D,N} : \sigma \text{ acts with a fixed point on } S_D\}}{\#\Gamma_{D,N}}, \tag{1-5}$$

where $\Gamma_{D,N}$ is a finite group through which the action of $\mathrm{Gal}(\overline{N}/N)$ on S_D factors. (We take $\delta_{D,N}(\pi)=0$ when no such components exist.) Note that

$$0 \leqslant \delta_{D,N}(\pi) \leqslant 1. \tag{1-6}$$

Moreover, we shall write $\delta_D(\pi) = \delta_{D,\kappa(D)}(\pi)$. When $\pi^{-1}(D)$ has components of multiplicity 1, this agrees with the definition given in [Loughran and Smeets 2016, (1.4)]. A natural analogue of the exponent appearing in [Loughran and Smeets 2016, Theorem 1.2] is then

$$\Delta(\pi) = \sum_{D\in(\mathbb{P}^1)^{(1)}} (1-\delta_D(\pi)), \tag{1-7}$$

which agrees with the exponent appearing in (1-2) whenever $\pi^{-1}(D)$ contains a multiplicity-1 component for every $D \in (\mathbb{P}^1)^{(1)}$.

The following upper bound treats the case of one multiple fibre above a degree-1 point of $\mathbb{P}^1$ and is consistent with Conjecture 1.1.

Theorem 1.2. *Let $\pi : X \to \mathbb{P}^1$ be a standard fibration with a unique multiple fibre at* 0*. Then*

$$N_{\mathrm{loc}}(\pi, B) \ll \frac{B^{2-\deg \partial_\pi}}{(\log B)^{\Delta(\pi)}},$$

where $\Delta(\pi)$ is given by (1-7).

It is tempting to suppose that the same estimate continues to hold when there is more than one closed point of $\mathbb{P}^1$ above which multiple fibres exist. However, in Theorem 7.1, we shall illustrate that a smaller exponent than $\Delta(\pi)$ is sometimes necessary.

Let $\pi : X \to \mathbb{P}^1$ be a standard fibration, and let $D \in (\mathbb{P}^1)^{(1)}$, which we suppose is defined by an irreducible binary form $g \in \mathbb{Q}[x, y]$. Assume first that $g(1, 0) \neq 0$. Then the residue field is $\kappa(D) = \mathbb{Q}[x]/(g(x, 1))$. Moreover, for any $d \in \mathbb{N}$ and any $v \in \mathbb{Q}$, let $h(x) = h^{(1)}(x)^{e_1} \cdots h^{(s_D)}(x)^{e_s}$ be the factorisation of $h(x) = g(x^d, v)$ into distinct irreducible polynomial $h^{(i)}(x)$. We define $N^{(i)}_{D,d,v} = \mathbb{Q}[x]/(h^{(i)}(x))$ and

$$N_{D,d,v} = N^{(1)}_{D,d,v} \times \cdots \times N^{(s_D)}_{D,d,v}. \tag{1-8}$$

For typical v this forms a number field of degree $\deg(g) + d$, but in general an étale algebra of possibly lower degree is formed since h need not be irreducible nor separable. It still remains to deal with the case $g(1, 0) = 0$. But then $D = \infty$ and we apply the same construction to the polynomial $g(1, vy^d) \in \mathbb{Q}[y]$.

We may now define

$$\Theta_v(\pi) = \sum_{D \in (\mathbb{P}^1)^{(1)}} \sum_{k=1}^{s_D} (1 - \delta_{D, N^{(k)}_{D,d,v}}(\pi)) \tag{1-9}$$

in the notation of (1-5). Our main upper bound is as follows.

Theorem 1.3. *Let $\pi : X \to \mathbb{P}^1$ be a standard fibration with multiple fibres at* 0 *and* ∞ *and nowhere else. Let $d = \gcd(m_0, m_\infty)$. Then*

$$N_{\mathrm{loc}}(\pi, B) \ll \frac{B^{2-\deg \partial_\pi}}{(\log B)^{\min_{v \in \mathbb{Q}^\times/\mathbb{Q}^{\times,d}} \Theta_v(\pi)}}.$$

It will be convenient to put

$$\Theta(\pi) = \min_{v \in \mathbb{Q}^\times/\mathbb{Q}^{\times,d}} \Theta_v(\pi). \tag{1-10}$$

Let us first note that $\Theta(\pi) \geqslant 0$ by (1-6). Secondly, $\Delta(\pi)$ and $\Theta(\pi)$ can be different; in Theorem 7.1 we will see an example with $\Theta(\pi) = 0$ but $\Delta(\pi) = 1$. However, we will see that

$$\Theta(\pi) = \Delta(\pi) \quad \text{if } \gcd(m_0, m_\infty) = 1. \tag{1-11}$$

The following result shows that there are only finitely many values that $\Theta_v(\pi)$ can take.

Proposition 1.4. *Let $\pi : X \to \mathbb{P}^1$ be a standard fibration, and let $D \in (\mathbb{P}^1)^{(1)}$. Let E be the field of definition of the elements of S_D, and let $N/\mathbb{Q}$ be a number field. Then $\delta_{D,N}(\pi) = \delta_{D,N \cap E^{\mathrm{normal}}}(\pi)$, where E^{normal} is the normal closure of E.*

As we have seen, our understanding of $N_{\text{loc}}(\pi, B)$ is inexorably linked to the arithmetic of the orbifold base $(\mathbb{P}^1, \partial_\pi)$. The study of rational points on orbifolds is the focus of work by Pieropan, Smeets, Tanimoto and Várilly-Alvarado [Pieropan et al. 2021], which offers a far-reaching conjectural asymptotic formula for any orbifold (Y, ∂) with $\mathbb{Q}$-ample divisor $-(K_Y + \partial)$. Pieropan and Schindler [2024] have verified many cases of the conjecture when Y is a split toric variety over $\mathbb{Q}$. Their work covers the orbifolds that arise in the proof of Theorems 1.2 and 1.3 and would yield the upper bound $N_{\text{loc}}(\pi, B) = O(B^{2-\deg \partial_\pi})$. In order to achieve the desired nonpositive powers of $\log B$, we need to incorporate extra Chebotarev-type conditions that arise when counting locally soluble fibres.

The proofs of Theorems 1.2 and 1.3 are based on the large sieve and will be carried out in Section 6. A crucial ingredient will be a *sparsity criterion*, which gives explicit control over which fibres are everywhere locally soluble. This criterion will be proved in Section 5 using log geometry and may be of independent interest.

The condition $\deg \partial_\pi < 2$ restricts us to only considering fibrations over $\mathbb{P}^1$ with at most three multiple fibres, and the multiplicities of these fibres cannot be too large. Extending Theorem 1.3 to three multiple fibres represents a formidable challenge. The easiest such case corresponds to the $\mathbb{Q}$-divisor

$$\partial_\pi = \tfrac{1}{2}[0] + \tfrac{1}{2}[1] + \tfrac{1}{2}[\infty].$$

Conjecture 1.1 would predict that $N_{\text{loc}}(\pi, B) = O_\varepsilon(B^{1/2+\varepsilon})$ for any $\varepsilon > 0$. However, the best upper bound we have is due to [Browning and Van Valckenborgh 2012], which only yields the exponent $\frac{3}{5} + \varepsilon$.

1.2. ***A new conjecture.*** We are now ready to reveal a new conjecture for the density of locally soluble fibres for standard fibrations in which multiple fibres are allowed. Let $\pi : X \to \mathbb{P}^1$ be a standard fibration, and let $\theta : \mathbb{P}^1 \to (\mathbb{P}^1, \partial_\pi)$ be a finite étale orbifold morphism, as defined in Definition 4.2.

We assume that $(\mathbb{P}^1, \partial_\pi)$ does not admit a finite étale orbifold morphism which factors through θ, and θ is a G-torsor under a finite étale group scheme G. Let $\theta_v : C_v \to \mathbb{P}^1$ denote the twist of θ by any $v \in \mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), G)$, which is a torsor under the inner twist G_v of G [Skorobogatov 2001, p. 20]. Finally, let $\pi_v : X_v \to C_v$ denote the normalisation of the pullback of π along θ_v. We will only consider v for which $C_v(\mathbb{Q}) \neq \varnothing$, in which case we identify $C_v \cong \mathbb{P}^1$.

Conjecture 1.5. *Let $\pi : X \to \mathbb{P}^1$ be a standard fibration such that the $\mathbb{Q}$-divisor $-(K_{\mathbb{P}^1} + \partial_\pi)$ is ample and $X(\boldsymbol{A}_\mathbb{Q}) \neq \varnothing$. Then there exists a constant $c_\pi > 0$ such that*

$$N_{\text{loc}}(\pi, B) \sim c_\pi \frac{B^{2-\deg \partial_\pi}}{(\log B)^{\min_{v \in \mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), G)} \Delta(\pi_v)}},$$

where $\Delta(\pi_v)$ is given by (1-7).

Note that it follows from Proposition 1.4 that $\Delta(\pi_v)$ takes only finitely many values. In the special case that the orbifold base is simply connected as an orbifold, which in the setting of Theorem 1.3 covers the case $\gcd(m_0, m_\infty) = 1$, the exponent will simply equal $\Delta(\pi)$. Thus Conjecture 1.5 implies that

$$N_{\text{loc}}(\pi, B) \sim c_\pi \frac{B^{2-\deg \partial_\pi}}{(\log B)^{\Delta(\pi)}}$$

in this case, which is consistent with the upper bound in Theorem 1.2. In Corollary 4.9 we take $G = \mu_d$ and prove that $\Theta_v(\pi) = \Delta(\pi_v)$ in (1-9). Hence the upper bound in Theorem 7.1 is also consistent with Conjecture 1.5. In Section 7 we provide further evidence for the conjecture by establishing a range of estimates for the variant $N_{\mathrm{loc},S}(\pi, B)$ of $N_{\mathrm{loc}}(\pi, B)$, in which local solubility is only required away from a finite set S of primes. In Theorem 7.2, for example, we establish a precise lower bound for $N_{\mathrm{loc},S}(\pi, B)$ in the case that $\pi : X \to \mathbb{P}^1$ is a standard fibration for which the only nonsplit fibres lie over 0 and ∞.

One further source of examples that can be used to illustrate our conjectures is the class of *Halphen surfaces*. These were introduced by Halphen [1882] and correspond to standard fibrations admitting a unique multiple fibre. In Theorems 7.3–7.8 we provide several estimates for $N_{\mathrm{loc},S}(\pi, B)$ that are consistent with Conjecture 1.5 for appropriate surfaces of Halphen type. In the proof of Theorem 7.8 we are led to a concrete problem in analytic number theory that was solved by Friedlander and Iwaniec [2010, Theorem 11.31]. Indeed, we need matching upper and lower bounds for the number of positive integers a, b satisfying $a^6 + b^2 \leqslant x$, as $x \to \infty$, such that the only prime divisors of $a^6 + b^2$ are those that split in a given cubic Galois extension $K/\mathbb{Q}$. It would be useful to have a similar result for non-Galois extensions, but this appears to be difficult.

Remark 1.6. Returning to the example (1-3), we see that the associated $\mathbb{Q}$-divisor ∂_π has degree $\frac{1}{2}\deg(f)$. Since f is assumed to have even degree, it follows that $-(K_{\mathbb{P}^1} + \partial_\pi)$ is ample only when $\deg(f) = 2$. When f is a quadratic polynomial, Conjecture 1.1 implies that $N_{\mathrm{loc}}(\pi, B) = O_\varepsilon(B^{1+\varepsilon})$ for any $\varepsilon > 0$. The orbifold base $(\mathbb{P}^1, \partial_\pi)$ admits μ_2-covers, and it is possible to apply Conjecture 1.5 to predict an explicit power of $\log B$. The outcome will depend on the Galois action on the geometric components of the fibres.

1.3. *Further questions.* We expect similar conjectures to hold when looking at fibrations $\pi : X \to Y$ over other bases for which $-(K_Y + \partial_\pi)$ is $\mathbb{Q}$-ample. However, when $\dim(Y) > 1$ the sparsity criterion we work out in Section 5 will be significantly more complicated. Moreover, care also needs to be taken around the effect of thin subsets of $Y(\mathbb{Q})$ on the counting problem. A counter-example to the most naive expectation has recently been provided [Browning et al. 2023] in the case that Y is a split quadric in $\mathbb{P}^3$.

In a different direction, when $Y = \mathbb{P}^1$, we can extend the definition (1-1) by defining $N_{\mathrm{loc}}(\pi, B; Z)$ to be the number of $x \in (\mathbb{P}^1(\mathbb{Q}) \setminus Z) \cap \pi(X(\mathbf{A}_\mathbb{Q}))$ for which $H(x) \leqslant B$ for any *thin subset* $Z \subseteq \mathbb{P}^1(\mathbb{Q})$. It is then very natural to ask whether or not we should expect a bound of the shape

$$N_{\mathrm{loc}}(\pi, B; Z) \ll \frac{B^{1/m_0 + 1/m_\infty}}{(\log B)^{\Delta(\pi)}},$$

where $\Delta(\pi)$ is given by (1-7), if we have the freedom to remove any thin set Z. Of course, as pointed out by the anonymous referee, it is not completely clear whether anything is left if we are able to remove arbitrary thin sets from $\mathbb{P}^1(\mathbb{Q})$.

1.4. *Summary of the paper.* The main sparsity criterion for locally soluble fibres is Theorem 5.5. It is proved using log geometry in Section 5 and leads to Chebotarev-type conditions about the splitting behaviour of primes. In Section 2 we shall collect together some basic group-theoretic results that allow us to interpret the output from Chebotarev's theorem. Section 3 uses recent work of Arango-Piñeros, Keliher and Keyes [Arango-Piñeros et al. 2022] to count pairs of power-full integers which lie in the multiplicative span of Frobenian sets of primes. In Section 4 we shall introduce the necessary background on orbifolds that is required to interpret the exponent of $\log B$ in Conjecture 1.5. Section 6 contains the proof of Theorems 1.2 and 1.3 and is based on an application of the large sieve. Finally, Section 7 builds on the work in Section 3 and contains new evidence for Conjecture 1.5, including estimates for $N_{\mathrm{loc},S}(\pi, B)$ in the case of Halphen surfaces and other families admitting multiple fibres.

2. Group-theoretic results

We will need some preliminary results on the density of primes with a prescribed splitting behaviour. Using Chebotarev's theorem, we will be able to translate it into statements about groups and group actions. We begin by proving some results in elementary group theory.

2.1. *Group theory lemmas.* Let G be a finite group, and let $H \subseteq G$ be a subgroup. For an element $g \in G$, we will write $\mathrm{Fix}_g(G/H)$ for the set of fixed points of g under the natural action of G on G/H.

Lemma 2.1. *Let $C \subseteq G$ be a conjugacy class. Then we have*

$$\sum_{g\in C} \#\,\mathrm{Fix}_g(G/H) = \frac{\#G}{\#H}\#(C \cap H).$$

Proof. First note that, for conjugate elements $g, y \in C$, there is an element $u \in G$ such that $u^{-1}yu = g$. Hence

$$\{x \in G : x^{-1}gx = y\} = \{x \in G : (ux)^{-1}y(ux) = y\} = u^{-1}\,\mathrm{Stab}_y,$$

whose cardinality is $\#G/\#C$ by the orbit-stabiliser theorem since C is the orbit of y under conjugation. We now see that

$$\begin{aligned}\sum_{g\in C} \#\,\mathrm{Fix}_g(G/H) &= \#\{(g, xH) \in C \times G/H : gxH = xH\}\\ &= \#\{(g, xH) \in C \times G/H : x^{-1}gx \in H\}.\end{aligned}$$

Hence

$$\begin{aligned}\sum_{g\in C} \#\,\mathrm{Fix}_g(G/H) &= \frac{1}{\#H}\#\{(g, x) \in C \times G : x^{-1}gx \in H \cap C\}\\ &= \frac{1}{\#H}\#\{(g, x, y) \in C \times G \times (H \cap C) : x^{-1}gx = y\}\\ &= \frac{1}{\#H}\#C \cdot \frac{\#G}{\#C} \cdot \#(C \cap H),\end{aligned}$$

which proves the lemma. □

Lemma 2.2. *Let S and T be subgroups of G. Then*

$$\#S\#T = \#(S \cap T)\#(ST).$$

Proof. Consider the action $S \times T$ on G by $(s,t)g = sgt^{-1}$. The stabiliser of e_G equals the image of diagonal map $S \cap T \hookrightarrow S \times T$, and the set ST is the orbit of e_G. The result now follows from the orbit-stabiliser formula. □

2.2. ***Density of primes.*** Let $F/\mathbb{Q}$ be a number field with ring of integers $\mathscr{O}_F$. Define $\mathscr{P}_{F,m}$ to be the set of rational primes p unramified in F which are divisible by exactly m primes $\mathfrak{p}_i \subseteq \mathscr{O}_F$ of degree 1. Let

$$\mathscr{P}_F = \bigcup_{m \geqslant 1} \mathscr{P}_{F,m}.$$

We define

$$\delta(\mathscr{E}, K) = 1 - \sum_{m=1}^{d} m \operatorname{dens}\biggl(\mathscr{P}_{K,m} \cap \bigcup_{E \in \mathscr{E}} \mathscr{P}_E\biggr)$$

for a finite set $\mathscr{E}$ of number fields and any number field $K \subseteq \overline{\mathbb{Q}}$ with $d = [K : \mathbb{Q}]$. If $\mathscr{E} = \{E\}$ consists of a single number field $E \subseteq \overline{\mathbb{Q}}$, we will write $\delta(E, K) = \delta(\mathscr{E}, K)$. The main result of this section is the following result.

Theorem 2.3. *Let $\mathscr{E}$ be a finite set of number fields and $K \subseteq \overline{\mathbb{Q}}$ a number field with $d = [K : \mathbb{Q}]$. Define*

$$\delta(\mathscr{E}, K) = 1 - \sum_{m=1}^{d} m \operatorname{dens}\biggl(\mathscr{P}_{K,m} \cap \bigcup_{E \in \mathscr{E}} \mathscr{P}_E\biggr).$$

Let $L \subseteq \overline{\mathbb{Q}}$ be a Galois extension of $\mathbb{Q}$ which contains both K and all $E \in \mathscr{E}$. Then

$$\delta(\mathscr{E}, K) = 1 - \frac{\#\{\sigma \in \mathrm{Gal}(L/K) : \sigma \text{ fixes a conjugate of some } E \in \mathscr{E}\}}{\#\mathrm{Gal}(L/K)}.$$

The quantity $\delta(\mathscr{E}, K)$ generalises a quantity that is implicit in [Loughran and Smeets 2016, (1.4)]. Let $\pi : X \to \mathbb{P}^1$ be a standard fibration, and let D be a closed point of $\mathbb{P}^1$ with residue field $\kappa(D)$. Let $I_D(\pi)$ be the set of geometrically irreducible components of $\pi^{-1}(D)$ of multiplicity 1, and let $\mathscr{E}_D$ be the set of number fields obtained from taking the algebraic closure of $\mathbb{Q}$ in the function field of each irreducible component of $\pi^{-1}(D)$, i.e., the minimal finite extensions of $\mathbb{Q}$ over which the irreducible components of $\pi^{-1}(D)$ split into their geometrically irreducible components. Then

$$\delta_D(\pi) = 1 - \delta(\mathscr{E}_D, \kappa(D))$$

in [Loughran and Smeets 2016, (1.4)]. Moreover, if we take S_D to be the set of geometrically irreducible components of $\pi^{-1}(D)$ of minimal multiplicity m_D and we let $\mathscr{E}_D$ be the set of fields of definition of the elements of S_D, then we also have

$$\delta_{D,N}(\pi) = 1 - \delta(\mathscr{E}_D, N) \tag{2-1}$$

in (1-5) for any number field $N/\mathbb{Q}$.

Proof of Theorem 2.3. Write $G = \mathrm{Gal}(L/\mathbb{Q})$, and let K be the fixed field of the subgroup H_1 of G. Similarly, let $E \in \mathscr{E}$ be the fixed fields of the subgroups $H \in \mathscr{H}$ of G. Let $\mathscr{P}'_{K,m}$ denote the set of primes in $\mathscr{P}_{K,m}$ that are unramified in L, and similarly for each $\mathscr{P}'_E$. Then we have

$$\mathscr{P}'_{K,m} = \{\text{primes } p \in \mathbb{Z} \text{ unramified in } L \text{ for which } \#\operatorname{Fix}_{\operatorname{Frob}_p}(G/H_1) = m\}$$

and

$$\mathscr{P}'_E = \{\text{primes } p \in \mathbb{Z} \text{ unramified in } L \text{ for which } \#\operatorname{Fix}_{\operatorname{Frob}_p}(G/H) \geqslant 1\}.$$

Note that

$$C_m = \left\{ g \in G : \#\operatorname{Fix}_g(G/H_1) = m, \sum_{H \in \mathscr{H}} \#\operatorname{Fix}_g(G/H) \geqslant 1 \right\}$$

is closed under conjugation, since conjugate elements have the same number of fixed points. By Chebotarev's theorem, in the form presented in [Serre 2012, Theorem 3.4], for example, we therefore obtain

$$\operatorname{dens}\left(\mathscr{P}_{K,m} \cap \bigcup_{E \in \mathscr{E}} \mathscr{P}_E\right) = \operatorname{dens}\left(\mathscr{P}'_{K,m} \cap \bigcup_{E \in \mathscr{E}} \mathscr{P}'_E\right) = \frac{\#C_m}{\#G}.$$

Let $T = \bigcup_{t \in G, H \in \mathscr{H}} tHt^{-1}$, which we note is closed under conjugation. Since $g \in G$ has at least a fixed point on one of the G/H if and only if $g \in T$, we arrive at

$$\sum_{m=1}^{d} m \operatorname{dens}\left(\mathscr{P}_{K,m} \cap \bigcup_{E \in \mathscr{E}} \mathscr{P}_E\right) = \frac{1}{\#G} \sum_{m=1}^{d} m \#C_m = \frac{1}{\#G} \sum_{g \in T} \#\operatorname{Fix}_g(G/H_1).$$

We may now conclude from Lemma 2.1 that

$$\sum_{m=1}^{d} m \operatorname{dens}\left(\mathscr{P}_{K,m} \cap \bigcup_{E \in \mathscr{E}} \mathscr{P}_E\right) = \frac{\#(T \cap H_1)}{\#H_1}. \tag{2-2}$$

The statement of the theorem follows on noting that $H_1 = \operatorname{Gal}(L/K)$ and

$$T = \{\sigma \in G : \sigma \text{ fixes a conjugate of some } E \in \mathscr{E}\}. \qquad \square$$

Note that we could not have applied the Chebotarev theorem to $\#(T \cap H_1)$, since $T \cap H_1$ is not necessarily fixed under conjugation in G. It is however closed under conjugation in H_1.

2.3. ***Computation of δ in specific cases.*** Theorem 2.3 allows us to compute the density $\delta(\mathscr{E}, K)$ in the common Galois closure L of both K and each $E \in \mathscr{E}$. The following theorem says that this can be reduced to a computation in a Galois closure of the fields $E \in \mathscr{E}$.

Proposition 2.4. *Let E^{normal} be the normal closure of the compositum of the $E \in \mathscr{E}$ in $\overline{\mathbb{Q}}$. Then*

$$\delta(\mathscr{E}, K) = \delta(\mathscr{E}, E^{\text{normal}} \cap K).$$

Proof. We adopt the notation from the proof of Theorem 2.3. Let $A^{\{j\}}$ be the subgroups of G indexed by a set J, which are of the form tHt^{-1} for $t \in G$ and $H \in \mathscr{H}$. For a set $I \subseteq J$, we write $A^I = \bigcap_{i \in I} A^{\{i\}}$.

The field $E^{\mathrm{normal}} \cap K$ corresponds to the subgroup $\langle H_1, A^J \rangle \subseteq G$ generated by H_1 and A^J. (Since A^J is normal, one can actually show that $\langle H_1, A^J \rangle = H_1 A^J$.) It follows from Lemma 2.2 that

$$\frac{\#(S \cap H_1)}{\# H_1} = \frac{\#(S \cap \langle H_1, A^J \rangle)}{\# \langle H_1, A^J \rangle}$$

when S is equal to A^I for any $I \subseteq J$. Since both sides are additive in S, the statement extends to $S = T = \bigcup_{j \in J} A^{\{j\}}$ by the principle of inclusion and exclusion. □

Proof of Proposition 1.4. Combine Proposition 2.4 with (2-1). □

Our remaining results summarise some special situations in which we can use Theorem 2.3 and Proposition 2.4 to calculate the densities $\delta(\mathscr{E}, K)$ easily.

Lemma 2.5. *If $E \subseteq K$ for some $E \in \mathscr{E}$, then $\delta(\mathscr{E}, K) = 0$.*

Proof. Since K, E are the fixed fields of the subgroups $H_1, H' \subseteq \mathrm{Gal}(L/\mathbb{Q})$, we have $E \subseteq K$ if and only if $H' \supseteq H_1$. But then $H_1 \subseteq H' \subseteq T = \bigcup_{t \in G, H \in \mathscr{H}} tHt^{-1}$, whence $\#(T \cap H_1)/\#H_1 = 1$ in (2-2). □

Let us now consider some cases in which $\mathscr{E}$ contains a single element.

Lemma 2.6. *If $E/\mathbb{Q}$ is Galois, then $\delta(E, K) = 1 - \deg(E \cap K)/\deg E$.*

Proof. Since $E/\mathbb{Q}$ is Galois, E is also Galois over $E^{\mathrm{normal}} \cap K = E \cap K$. Thus we conclude $\delta(E, K) = \delta(E, E \cap K) = 1 - 1/[E : E \cap K]$. □

Lemma 2.7. *If $K/\mathbb{Q}$ is Galois and $KE = E^{\mathrm{normal}}$, then $\delta(E, K) = 1 - \deg(E \cap K)/\deg E$.*

Proof. Since $KE = E^{\mathrm{normal}}$ and $K/\mathbb{Q}$ is Galois, we have $H_1 \cap A^{\{j\}} = A^J$ for all $j \in J$. Thus

$$\frac{\#(T \cap H_1)}{\# H_1} = \frac{\# A^J}{\# H_1} = \frac{\deg K}{\deg E^{\mathrm{normal}}} = \frac{\deg K}{\deg KE}$$

in (2-2). Since K is Galois, we have $[KE : K] = [E : E \cap K]$, from which the lemma follows. □

3. Pairs of integers with Frobenian conditions

We say that a set $\mathscr{P}$ of rational primes is *Frobenian* if there is a finite Galois extension $K/\mathbb{Q}$ and a union of conjugacy classes H in $\mathrm{Gal}(K/\mathbb{Q})$ such that $\mathscr{P}$ is equal to the set of primes p that are unramified in K and for which the Frobenius conjugacy class of p in $\mathrm{Gal}(K/\mathbb{Q})$ lies in H. In this section we produce an asymptotic formula for the density of coprime integers a_0, a_1 which are both power-full and lie in the multiplicative span of a Frobenian set of primes.

It will be convenient to introduce the notation

$$c_S(\alpha) = \prod_{p \in S} \left(1 - \frac{1}{p^\alpha}\right) \tag{3-1}$$

for any $\alpha > 0$ and any finite set of primes S. We shall prove the following result.

Proposition 3.1. *For $i \in \{0,1\}$, let $m_i \in \mathbb{N}$ and let $\mathscr{P}_i$ be a Frobenian set of rational primes of density $\eth_i$. Then, for any finite set of primes S, we have*

$$\#\{(a_0,a_1) \in \mathbb{Z}^2_{\mathrm{prim}} : |a_i| \leqslant B,\ p \notin S \Rightarrow [m_i \mid v_p(a_i) \text{ and } (p \mid a_i \Rightarrow p \in \mathscr{P}_i)]\} \sim c_{m_i,\mathscr{P}_i,S} \frac{B^{1/m_0+1/m_1}}{(\log B)^{2-\eth_0-\eth_1}}$$

as $B \to \infty$, where

$$c_{m_i,\mathscr{P}_i,S} = \frac{4m_0^{1-\eth_0} m_1^{1-\eth_1}}{\Gamma(\eth_0)\Gamma(\eth_1)} \cdot \frac{c_S\big(\frac{1}{m_0}+\frac{1}{m_1}\big)}{c_S\big(\frac{1}{m_0}\big)c_S\big(\frac{1}{m_1}\big)} \prod_{\substack{p \in \mathscr{P}_0 \cap \mathscr{P}_1 \\ p \notin S}} \Big(1-\frac{1}{p^2}\Big)$$
$$\times \prod_{p \in \mathscr{P}_0 \cap S} \Big(1-\frac{1}{p}\Big) \prod_{p \in \mathscr{P}_0} \Big(1-\frac{1}{p}\Big)^{-1+\eth_0} \prod_{p \notin \mathscr{P}_0} \Big(1-\frac{1}{p}\Big)^{\eth_0}$$
$$\times \prod_{p \in \mathscr{P}_1 \cap S} \Big(1-\frac{1}{p}\Big) \prod_{p \in \mathscr{P}_1} \Big(1-\frac{1}{p}\Big)^{-1+\eth_1} \prod_{p \notin \mathscr{P}_1} \Big(1-\frac{1}{p}\Big)^{\eth_1}.$$

There are only $O(1)$ elements with $a_0a_1 = 0$ that contribute to the counting function. Let $M(B) = M(m_i, \mathscr{P}_i, B, S)$ denote the overall contribution with $a_0a_1 \neq 0$. Hence, on accounting for signs, we have

$$M(B) = 4\#\left\{(a_0,a_1) \in \mathbb{N}^2 : \begin{array}{l} a_0, a_1 \leqslant B,\ \gcd(a_0,a_1)=1, \\ \qquad p \notin S \Rightarrow [m_i \mid v_p(a_i) \text{ and } (p \mid a_i \Rightarrow p \in \mathscr{P}_i)] \end{array}\right\}.$$

For (a_0,a_1) appearing in the counting function, we may clearly write

$$a_0 = b_0 u_0^{m_0} \quad \text{and} \quad a_1 = b_1 u_1^{m_1},$$

where $p \mid b_0b_1 \Rightarrow p \in S$, $\gcd\big(u_0u_1, \prod_{p\in S} p\big) = 1$, and $p \mid u_i \Rightarrow p \in \mathscr{P}_i$. Moreover, we have $\gcd(b_0,b_1) = \gcd(u_0,u_1) = 1$. Let $\mathscr{Q} = \mathscr{P}_0 \cap \mathscr{P}_1$.

We proceed by introducing the counting functions

$$M_i(x) = \#\{v \leqslant x : p \mid v \Rightarrow p \in \mathscr{P}_{i,S}\}$$

for $i = 0, 1$, where $\mathscr{P}_{i,S} = \mathscr{P}_i \setminus (S \cap \mathscr{P}_i)$. On using the Möbius function to detect the condition $\gcd(u_0,u_1) = 1$, we may now write

$$M(B) = 4 \sum_{\substack{b_0,b_1 \in \mathbb{N} \\ \gcd(b_0,b_1)=1 \\ p \mid b_0b_1 \Rightarrow p \in S}} \sum_{\substack{k \in \mathbb{N} \\ p \mid k \Rightarrow p \in \mathscr{Q}_S}} \mu(k) M_0(k^{-1}(B/b_0)^{1/m_0}) M_1(k^{-1}(B/b_1)^{1/m_1}),$$

where $\mathscr{Q}_S = \mathscr{Q} \setminus (S \cap \mathscr{Q})$. The treatment of $M_i(x)$ is handled by the following result.

Lemma 3.2. *Let $i \in \{0,1\}$. Then*

$$M_i(x) \sim \frac{\kappa_{i,S}}{\Gamma(\eth_i)} \frac{x}{(\log x)^{1-\eth_i}}$$

as $x \to \infty$, where

$$\kappa_{i,S} = \prod_{p \in \mathscr{P}_i \cap S} \Big(1-\frac{1}{p}\Big) \prod_{p \in \mathscr{P}_i} \Big(1-\frac{1}{p}\Big)^{-1+\eth_i} \prod_{p \notin \mathscr{P}_i} \Big(1-\frac{1}{p}\Big)^{\eth_i}. \tag{3-2}$$

Proof. Let $i \in \{0, 1\}$. There are several approaches to estimating $M_i(x)$, but the one we shall adopt is via a general result of [Wirsing 1967] on mean values of multiplicative arithmetic functions $g : \mathbb{N} \to [0, 1]$. (In fact, this result applies to general nonnegative multiplicative arithmetic functions under further assumptions on the behaviour of g at prime powers.) Suppose that

$$\sum_{p \leqslant x} g(p) \log p \sim \tau x$$

for some $\tau > 0$. Then it follows that

$$\sum_{n \leqslant x} g(n) \sim \frac{e^{-\gamma\tau}}{\Gamma(\tau)} \frac{x}{\log x} \prod_{p \leqslant x} \left(1 + \frac{g(p)}{p} + \frac{g(p^2)}{p^2} + \cdots\right),$$

where γ is Euler's constant.

In our case we take

$$g(n) = \begin{cases} 1 & \text{if } p \mid n \Rightarrow p \in \mathscr{P}_{i,S}, \\ 0 & \text{otherwise.} \end{cases}$$

Then, since $\mathscr{P}_i$ is a Frobenian set of primes of density $\eth_i$, it follows from the Chebotarev density theorem that

$$\sum_{p \leqslant x} g(p) \log p = \sum_{\substack{p \leqslant x \\ p \in \mathscr{P}_{i,S}}} \log p \sim \eth_i \log x$$

as $x \to \infty$. Hence $\tau = \eth_i$ and we obtain

$$M(x) \sim \frac{e^{-\gamma \eth_i}}{\Gamma(\eth_i)} \frac{x}{\log x} \prod_{\substack{p \leqslant x \\ p \in \mathscr{P}_{i,S}}} \left(1 - \frac{1}{p}\right)^{-1}$$

as $x \to \infty$. It remains to study

$$\prod_{\substack{p \leqslant x \\ p \in \mathscr{P}_{i,S}}} \left(1 - \frac{1}{p}\right)^{-1} = \prod_{p \in \mathscr{P}_i \cap S} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \leqslant x \\ p \in \mathscr{P}_i}} \left(1 - \frac{1}{p}\right)^{-1}.$$

However, on appealing to [Arango-Piñeros et al. 2022, Theorem A], we quickly arrive at the expression

$$\prod_{\substack{p \leqslant x \\ p \in \mathscr{P}_i}} \left(1 - \frac{1}{p}\right)^{-1} \sim \left(\frac{\log x}{e^{-\gamma_K}}\right)^{\eth_i}$$

as $x \to \infty$, where

$$e^{-\gamma_K} = e^{-\gamma} \prod_{p \in \mathscr{P}_i} \left(1 - \frac{1}{p}\right)^{\eth_i^{-1} - 1} \prod_{p \notin \mathscr{P}_i} \left(1 - \frac{1}{p}\right)^{-1}.$$

It now follows that

$$\prod_{\substack{p \leqslant x \\ p \in \mathscr{P}_{i,S}}} \left(1 - \frac{1}{p}\right)^{-1} \sim \kappa_{i,S} (\log x)^{\eth_i} e^{\gamma \eth_i}$$

in the notation of lemma. Inserting this into our previous asymptotic formula for $M_i(x)$, we finally arrive at the statement of the lemma. □

We clearly have

$$(\log(k^{-1}(B/b_i)^{1/m_i}))^{-(1-\partial_i)} = m_i^{1-\partial_i}(\log B)^{-(1-\partial_i)}\left(1+O\left(\frac{\log kb_i}{\log B}\right)\right)$$

for $i = 0, 1$. Hence, on substituting Lemma 3.2 into our previous expression for $M(B)$, we thereby obtain

$$M(B) = 4 \sum_{\substack{b_0,b_1\in\mathbb{N} \\ \gcd(b_0,b_1)=1 \\ p\mid b_0b_1\Rightarrow p\in S}} \sum_{\substack{k\in\mathbb{N} \\ p\mid k\Rightarrow p\in\mathscr{Q}_S}} A_{b_0,b_1,k}(B) + o\left(\frac{B^{1/m_0+1/m_1}}{(\log B)^{2-\partial_0-\partial_1}}\right),$$

with

$$\begin{aligned} A_{b_0,b_1,k}(B) &= \frac{\kappa_{0,S}\kappa_{1,S}}{\Gamma(\partial_0)\Gamma(\partial_1)} \cdot \frac{\mu(k)m_0^{1-\partial_0}m_1^{1-\partial_1}(k^{-1}(B/b_0)^{1/m_0})(k^{-1}(B/b_1)^{1/m_1})}{(\log B)^{2-\partial_0-\partial_1}} \\ &= \frac{\kappa_{0,S}\kappa_{1,S}}{\Gamma(\partial_0)\Gamma(\partial_1)} \cdot m_0^{1-\partial_0}m_1^{1-\partial_1} \cdot \frac{B^{1/m_0+1/m_1}}{(\log B)^{2-\partial_0-\partial_1}} \cdot \frac{\mu(k)}{k^2} \cdot \frac{1}{b_0^{1/m_0}b_1^{1/m_1}} \end{aligned}$$

and where $\kappa_{0,S}$, $\kappa_{1,S}$ are given by (3-2)

Next, on recalling the notation of (3-1), a simple calculation furnishes the identities

$$\sum_{\substack{b_0,b_1\in\mathbb{N} \\ \gcd(b_0,b_1)=1 \\ p\mid b_0b_1\Rightarrow p\in S}} \frac{1}{b_0^{1/m_0}b_1^{1/m_1}} = \frac{c_S\left(\frac{1}{m_0}+\frac{1}{m_1}\right)}{c_S\left(\frac{1}{m_0}\right)c_S\left(\frac{1}{m_1}\right)} \quad \text{and} \quad \sum_{\substack{k\in\mathbb{N} \\ p\mid k\Rightarrow p\in\mathscr{Q}_S}} \frac{\mu(k)}{k^2} = \prod_{\substack{p\in\mathscr{P}_0\cap\mathscr{P}_1 \\ p\notin S}} \left(1-\frac{1}{p^2}\right).$$

Hence it follows that the asymptotic formula in Proposition 3.1 holds with the leading constant

$$c_{m_i,\mathscr{P}_i,S} = 4 \cdot \frac{\kappa_{0,S}\kappa_{1,S}}{\Gamma(\partial_0)\Gamma(\partial_1)} \cdot m_0^{1-\partial_0}m_1^{1-\partial_1} \cdot \frac{c_S\left(\frac{1}{m_0}+\frac{1}{m_1}\right)}{c_S\left(\frac{1}{m_0}\right)c_S\left(\frac{1}{m_1}\right)} \cdot \prod_{\substack{p\in\mathscr{P}_0\cap\mathscr{P}_1 \\ p\notin S}} \left(1-\frac{1}{p^2}\right),$$

where $\kappa_{0,S}$, $\kappa_{1,S}$ are given by (3-2). This therefore completes the proof of Proposition 3.1.

4. Orbifolds and étale orbifold morphisms

Campana [2004] related the study of fibrations $\pi : X \to Y$ of varieties over a fixed field k to orbifolds on the base. He studied *multiplicity orbifolds*, but since these are the only orbifolds in this paper we will simply call them *orbifolds*. In this section we summarise the construction of the most important invariant of orbifolds.

4.1. *Orbifold pairs.*

Throughout this section, let k be an arbitrary field of characteristic 0.

Definition 4.1. An *orbifold* is a pair (B, Δ), where B is a normal, proper k-scheme and Δ is a $\mathbb{Q}$-divisor

$$\Delta = \sum_D \left(1-\frac{1}{m_D}\right)[D]$$

for positive integers m_D associated to prime divisors D on B. We call m_D the *multiplicity* of the orbifold over D.

Definition 4.2. Let (B, Δ) be an orbifold on a normal and proper k-variety B. A *finite étale* (*orbifold*) *morphism* is a morphism $\theta : C \to B$, with C normal, which is

(i) finite,

(ii) étale away from Δ,

(iii) has the property $e(D'/D) \mid m_D$ for any prime divisor $D' \mid D$ (meaning any prime divisor $D' \subseteq C$ above $D \subseteq B$), where $e(D'/D)$ is the ramification index.

Let us explain the use of the word étale. Consider a finite dominant morphism $\theta : C \to B$ between integral, normal, proper k-varieties. Then we can always endow B with an orbifold structure such that θ becomes a finite étale orbifold morphism by assigning $m_D = \operatorname{lcm}\{e(D'/D) : D' \mid D\}$. If B has an orbifold divisor Δ under which θ is a finite étale orbifold morphism, then we can endow C with the $\mathbb{Q}$-divisor

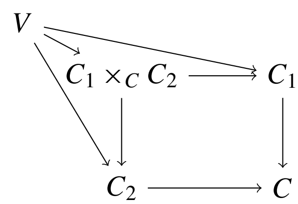
$$\Delta_C = \sum_{D'} \Bigl(1 - \frac{1}{m_{D'}}\Bigr)[D'], \quad \text{where } m_{D'} = \frac{m_D}{e(D'/D)}.$$

This is the unique orbifold structure on C such that the orbifold morphism $(C, \Delta_C) \to (B, \Delta)$ is étale in codimension 1, in the sense of [Campana 2011, Definition 2.21]. In the latter case, the Riemann–Hurwitz formula yields

$$K_{C,\Delta_C} = \theta^* K_{B,\Delta},$$

where $K_{B,\Delta} = K_B + \Delta$ is the *canonical divisor class* on an orbifold (B, Δ). (This statement can be proven along similar lines to the proof of Proposition 4.7 (c).)

Proposition 4.3. *Let $C_1, C_2 \to C$ be morphisms of normal k-varieties. Let $V = \widetilde{C_1 \times_C C_2}$ be the normalisation of the product $C_1 \times_C C_2$:*



Let $D_V \subseteq V$ be a prime divisor lying above prime divisors $D_i \subseteq C_i$ and $D \subseteq C$. Then

$$e(D_V/D_1) = \frac{e_2}{\gcd(e_1, e_2)},$$

where $e_i = e(D_i/D)$ for $i = 1, 2$.

Proof. Replacing the prime divisors with their generic points we can compute the normalisation étale locally over D. Hence we assume k is algebraically closed and consider the normalisation of the tensor product of the two homomorphisms $\varrho_i : k[\![t]\!] \to k[\![t_i]\!]$ given by $t \mapsto t_i^{e_i}$. The tensor product is

$R = k[\![t_1, t_2]\!]/(t_1^{e_1} - t_2^{e_2})$ generated by the images of the t_i. Let us define $d = \gcd(e_1, e_2)$, and let $\zeta \in k$ be a primitive d-th root of unity. We can write R as the product

$$R = \prod_i k[\![t_1, t_2]\!]/(t_1^{e_1/d} - \zeta^i t_2^{e_2/d}).$$

All factors are principal ideal domains, since polynomials $X^a - \lambda Y^b$ with $\lambda \in k^\times$ and $\gcd(a, b) = 1$ are irreducible over an algebraically closed field. We will compute the integral closure of each component separately. Let us write $\alpha_1 e_1 + \alpha_2 e_2 = d$ for $\alpha_i \in \mathbb{Z}$. Then $T = t_1^{\alpha_2} t_2^{\alpha_1}$ is integral in each factor since $T^{e_1/d} = t_2 \cdot (t_1^{e_1/d}/t_2^{e_2/d})^{\alpha_2}$ and $T^{e_2/d} = t_1 \cdot (t_2^{e_2/d}/t_1^{e_1/d})^{\alpha_1}$. It follows that

$$k[\![t_1, t_2]\!]/(t_1^{e_1/d} - \zeta^i t_2^{e_2/d}) \hookrightarrow k[\![T]\!]$$

is the integral closure. Finally, to compute $e(D_V/D_1)$, we look at the image of t_1 under the map

$$k[\![t_1]\!] \to k[\![T]\!],$$

which has valuation e_2/d. □

Remark 4.4. Campana [2011, Definition 11.1] defines the orbifold fundamental group $\pi_1(X|\Delta)$ for a complex orbifold $(X|\Delta)$ and relates it to *covers unramified away from* Δ. Likewise, we can define the (algebraic) orbifold fundamental group and relate it to the structure of all finite étale orbifold morphisms over a fixed base (B, Δ) of dimension 1. (Note that we could do this in arbitrary dimension if we allow finite étale morphisms to be defined away from a codimension-2 locus.) Consider the category $\mathbf{FEt}_{(B,\Delta)}$ of all finite étale orbifold morphisms to (B, Δ), where the morphisms are given by B-morphisms. Given a point $\bar{x} \in B(\bar{k}) \setminus \operatorname{supp}(\Delta)$, we have the fibre functor

$$F : \mathbf{FEt}_{(B,\Delta)} \to \mathbf{Sets}$$

given by $C \mapsto C_{\bar{x}}$, and one can show that $(\mathbf{FEt}_{(B,\Delta)}, F)$ is a Galois category. The only nontrivial part is to show that $\mathbf{FEt}_{(B,\Delta)}$ has products, but this follows from Proposition 4.3. In particular, this implies that, for any two finite étale covers of (C, ∂), there is another cover mapping to both. We define the (*algebraic*) *orbifold fundamental group* $\pi_1^{\mathrm{orb}}(B, \Delta)$ to be the automorphism group of the fibre functor F. Many relations between the topological and algebraic fundamental group can be directly translated to fundamental groups of orbifolds. For example, if $k \subseteq \mathbb{C}$ then

$$\pi_1^{\mathrm{orb}}(B, \Delta) = \widehat{\pi_1(B(\mathbb{C})|\Delta)}.$$

Campana [2011, Sections 11 and 12] studied the complex orbifold fundamental group and provided several results and conjectures about their structure.

For our application we will need the following definition.

Definition 4.5. Let G/k be a finite étale group and (B, Δ) an orbifold. Let $\theta : C \to B$ be a finite étale orbifold morphism endowed with a G-action on C, which is compatible with θ. We say that θ is a G-*torsor* (*of orbifolds*) if the restriction of θ away from the support of Δ is a G-torsor.

Since we are dealing with curves, it makes sense to talk about torsors. The natural morphism $G \times C \to C \times_B C$ is not necessarily an isomorphism over B, but it is so over $B \setminus \Delta$ by definition. Since $G \times C$ is a smooth curve over k, this morphism factors through the normalisation $G \times C \to \widetilde{C \times_B C} \to C \times_B C$. Now $G \times C \to \widetilde{C \times_B C}$ is a morphism between normal curves, which is an isomorphism on a dense open subset. Note that this agrees with the observation that $\widetilde{C \times_B C} \to C$ is unramified by Proposition 4.3; $\widetilde{C \times_B C}$ is just a union of copies of C.

4.2. ***Orbifold base of a fibration.*** As we saw in Section 1, we can associate a natural orbifold to any fibration. In this section we discuss this further before passing to our reasoning behind Conjecture 1.5.

Definition 4.6. Consider a fibration $\pi : X \to Y$, which we assume is a morphism between integral, normal, proper k-schemes such that the generic fibre is geometrically irreducible. For a prime divisor $D \subseteq Y$ with generic point η_D, we define m_D as the minimum multiplicity of the components of X_{η_D} as a divisor on X. The *orbifold base* of π is (Y, ∂_π), where

$$\partial_\pi = \sum_D \Big(1 - \frac{1}{m_D}\Big)[D].$$

Possibly up to thin sets, we expect the geometry of the base orbifold (Y, ∂_π) to govern the arithmetic properties of the fibration. We henceforth focus our attention on standard fibrations $\pi : X \to \mathbb{P}^1$ defined over $\mathbb{Q}$, with the aim of interpreting the growth of the counting function $N_{\mathrm{loc}}(\pi, B)$ that was defined in (1-1). Occasionally we will write $N^\circ_{\mathrm{loc}}(\pi, B)$ for the same counting function but excluding the finitely many points in the orbifold divisors ∂_π.

Let us begin by discussing the conjectured power of B in Conjecture 1.5, which is equal to

$$2 - \deg \partial_\pi = -\deg(K_{\mathbb{P}^1, \partial_\pi}), \tag{4-1}$$

where $K_{\mathbb{P}^1,\partial_\pi} = K_{\mathbb{P}^1} + \partial_\pi$. The following result relates the geometry of π to the geometry of a normalisation of the fibre product of π with a finite cover.

Proposition 4.7. *Let $\pi : X \to \mathbb{P}^1$ be a standard fibration, and let*

$$\theta : \mathbb{P}^1 \to \mathbb{P}^1$$

be a (possibly ramified) finite cover of degree d. We define $\pi_\theta : X_\theta \to \mathbb{P}^1$ to be the normalisation of the fibre product of θ and π. Then we have the following properties.

(a) *$\pi_\theta : X_\theta \to \mathbb{P}^1$ is a standard fibration.*

(b) *The orbifold multiplicities $m_{P'}$ for π_θ satisfy*

$$m_{P'} \geqslant \frac{m_P}{e(P'/P)}$$

for any prime divisor P' of $\mathbb{P}^1$, where $P = \theta(P')$. We have equality precisely when condition (iii) *in Definition 4.2 is satisfied at P'.*

(c) *We have*

$$\deg(K_{\mathbb{P}^1,\partial_{\pi_\theta}}) \geqslant d\,\deg(K_{\mathbb{P}^1,\partial_\pi}),$$

with equality precisely when θ is a finite étale orbifold morphism.

Proof. (a) This is clear from the definition.

(b) Consider a component Z' of the fibre of π_θ over a prime divisor P' of $\mathbb{P}^1$. Suppose that Z' lies over $Z \subseteq X$ and P' lies over P. Let $m_{P'}(Z')$ and $m_P(Z)$ denote the multiplicities of these components in their respective fibres. We wish to apply Proposition 4.3 with $C_1 \to C$ being the morphism $\theta : \mathbb{P}^1 \to \mathbb{P}^1$ and $C_2 \to C$ being the morphism $\pi : X \to \mathbb{P}^1$. Then $V \to C_1$ is the morphism $\pi_\theta : X_\theta \to \mathbb{P}^1$. It follows that

$$e(Z'/P') = \frac{e(Z/P)}{\gcd(e(Z/P), e(P'/P))}.$$

Hence, since the ramification indices over a codimension-1 point are precisely the multiplicities of the different components of the fibre, we obtain

$$m_{P'}(Z') = \frac{m_P(Z)}{\gcd(m_P(Z), e(P'/P))}.$$

Since $m_P(Z) \geqslant m_P$ and $\gcd(m_P(Z), e(P'/P)) \leqslant e(P'/P)$, we conclude

$$m_{P'}(Z') \geqslant m_P/e(P'/P)$$

for all components Z' in the fibre over P'.

Clearly, if $m_{P'} = m_P/e(P'/P)$, we have $e(P'/P) \mid m_P$. Now suppose that $e(P'/P) \mid m_P$. To prove the statement we must show that there is a component Z' over P' with $m_P(Z') = m_P/e(P'/P)$. By the definition of m_P, there exists a component Z over P with $m_P = m_P(Z)$. Now let Z' be any component over P' which lies over P. Then

$$m_{P'}(Z') = \frac{m_P(Z)}{\gcd(m_P(Z), e(P'/P))} = \frac{m_P}{\gcd(m_P, e(P'/P))} = \frac{m_P}{e(P'/P)}.$$

This concludes the proof of part (b).

(c) We will prove the result for orbifolds equipped with a degree-d morphism $(C', \partial') \to (C, \partial)$ for general smooth curves C and C', in order to distinguish between the two copies of $\mathbb{P}^1$. The statement is invariant under base change, so we can assume we are working over an algebraically closed field $k = \bar{k}$. We begin by noting that

$$\deg K_{C,\partial} = 2g(C) - 2 + \sum_{P \in C^{(1)}} \Big(1 - \frac{1}{m_P}\Big)$$

and

$$\deg K_{C',\partial'} = 2g(C') - 2 + \sum_{P' \in C'^{(1)}} \Big(1 - \frac{1}{m_{P'}}\Big).$$

The Riemann–Hurwitz formula yields

$$2g(C') - 2 = d(2g(C) - 2) + \sum_{P' \in C'^{(1)}} (e(P'/P) - 1),$$

where $P = \theta(P')$. Hence

$$\deg K_{C',\partial'} = d(2g(C) - 2) + \sum_{P' \in C'^{(1)}} \Big(e(P'/P) - \frac{1}{m_{P'}}\Big).$$

It now follows that

$$\begin{aligned}\deg K_{C',\partial'} - d \deg K_{C,\partial} &= \sum_{P' \in C'^{(1)}} \Big(e(P'/P) - \frac{1}{m_{P'}}\Big) - d \sum_{P \in C^{(1)}} \Big(1 - \frac{1}{m_P}\Big) \\ &= \sum_{P \in C^{(1)}} \bigg[\Big(\sum_{P' \mid P} e(P'/P) - d\Big) + \Big(\frac{d}{m_P} - \sum_{P' \mid P} \frac{1}{m_{P'}}\Big)\bigg].\end{aligned}$$

Using $\sum_{P' \mid P} e(P'/P) = d$, we see that the first terms all vanish and so

$$\deg K_{C',\partial'} - d \deg K_{C,\partial} = \sum_{P \in C^{(1)}} \sum_{P' \mid P} \Big(\frac{e(P'/P)}{m_P} - \frac{1}{m_{P'}}\Big).$$

This is clearly nonnegative by (b), and we have equality if and only if condition (iii) of Definition 4.2 is satisfied at all P'. □

In the setting of this result, it follows that the points in $N_{\mathrm{loc}}(\pi, B)$ that are counted by $N_{\mathrm{loc}}(\pi_\theta, H_\theta, B)$ are expected to contribute at most to the same order of B, where H_θ is the pullback height along θ. Indeed, in Conjecture 1.1, we have

$$N_{\mathrm{loc}}(\pi_\theta, H_\theta, B) = O_\varepsilon\big((B^{1/d})^{\deg(-K_{\mathbb{P}^1,\partial_{\pi_\theta}})+\varepsilon}\big)$$

for any $\varepsilon > 0$, where we use $B^{1/d}$ since H_θ is an $\mathscr{O}(d)$-height on $\mathbb{P}^1$. Hence, in the light of Proposition 4.7 (c), we should expect no higher-order contribution from $N_{\mathrm{loc}}(\pi_\theta, H_\theta, B)$ to $N_{\mathrm{loc}}(\pi, B)$. Moreover, we should obtain the same exponent of B when θ is a finite étale orbifold morphism.

We are now ready to address the possible power of $\log B$. Let $\pi : X \to \mathbb{P}^1$ be a standard fibration, and suppose that $\theta : \mathbb{P}^1 \to \mathbb{P}^1$ is a G-torsor of orbifolds under a finite étale group scheme G of degree d, as presented in Definition 4.5. We write $\theta_v : C_v \to \mathbb{P}^1$ for the twists of θ by $v \in \mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), G)$. Finally, we shall write $\pi_v : X_v \to C_v$ for the normalisation of the pullback of π along θ_v, which is a torsor under the inner twists G_v of G [Skorobogatov 2001, p. 20]. We usually restrict to the v for which $C_v(\mathbb{Q}) \neq \varnothing$ and identify $C_v \cong \mathbb{P}^1$. For our applications, G will be abelian and we will have $G_v \cong G$.

We are now ready to compare the counting function $N^\circ_{\mathrm{loc}}(\pi, B)$ with the counting functions

$$N^\circ_{\mathrm{loc}}(\pi_v, H_v, B)$$

for various $v \in \mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), G)$, where H_v is the pullback height along θ_v. (Note that this is an $\mathscr{O}(d)$-height on the domain $C_v \cong \mathbb{P}^1$ of θ_v when $C_v(\mathbb{Q}) \neq \varnothing$.)

Proposition 4.8. *In the setting above we have the following*:

(a) *A point* $x \in \mathbb{P}^1(\mathbb{Q})$ *is counted by* $N^{\circ}_{\mathrm{loc}}(\pi, B)$ *if and only if there exists* $v \in \mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), G)$ *and* $y \in C_v(\mathbb{Q})$ *such that* $\theta_v(y) = x$ *and such that* y *is counted by* $N^{\circ}_{\mathrm{loc}}(\pi_v, H_v, B)$.

(b) *We have*

$$N^{\circ}_{\mathrm{loc}}(\pi, B) = \sum_{v \in \mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), G)} \frac{1}{\#G_v(\mathbb{Q})} N^{\circ}_{\mathrm{loc}}(\pi_v, H_v, B).$$

(c) *Let* $\theta_v^{-1}(D) = \bigcup_{1 \leqslant i \leqslant s_D} E_v^{(i)}$ *be a decomposition into irreducible components, and write* $N_v^{(i)} = \kappa(E_v^{(i)})$ *for their function fields. Then*

$$\Delta(\pi_v) = \sum_{D \in (\mathbb{P}^1)^{(1)}} \sum_{i=1}^{s_D} (1 - \delta_{D, N^{(i)}_{D,v}}(\pi)),$$

where $\delta_{D, N^{(i)}_{D,v}}$ *is given by* (1-5).

(d) *The expression* $\Delta(\pi_v)$ *only assumes finitely many values.*

Proof. (a) Let $U \subseteq \mathbb{P}^1$ be the image of the étale locus of θ. The restrictions $\theta_v : U_v \to U$ are G_v-torsors, and so we have a partition

$$U(\mathbb{Q}) = \bigsqcup_{v \in \mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), G)} \theta_v(U_v(\mathbb{Q})).$$

Furthermore, the fibre of π_v over $y \in U_v(\mathbb{Q})$ is isomorphic to the fibre of π over $x = \theta_v(y)$. Hence one of these fibres is locally soluble precisely when the other is. Finally, since $\theta : \mathbb{P}^1 \to \mathbb{P}^1$ has degree d, the pullback of the $\mathscr{O}(1)$-height pulls back to an $\mathscr{O}(d)$-height.

(b) This follows from the partition in (a) and the fact that each fibre has $\#G_v(\mathbb{Q})$ points.

(c) This directly follows from the definition of $\delta_{D,N}$ and π_v.

(d) This follows from Proposition 1.4. □

In the setting of Theorem 1.3, we consider μ_d-covers parametrised by $\mathbb{Q}^{\times}/\mathbb{Q}^{\times,d}$. The following result therefore follows from part (c) of Proposition 4.8.

Corollary 4.9. *We have* $\Delta(\pi_v) = \Theta_v(\pi)$ *in* (1-9).

In principle there might be infinitely many twists π_v for which $\Delta(\pi_v)$ differs from the expected exponent $\Delta(\pi)$ defined in (1-7). The following example illustrates an instance where the points counted by the covers for which $\Delta(\pi_v) = \Delta(\pi)$ can form a nontrivial cothin set in $\mathbb{P}^1(\mathbb{Q})$.

Example 4.10. Consider the fibration $\pi : X \to \mathbb{P}^1$ with three double fibres over 0, -1 and ∞, together with precisely one other nonsplit fibre over 1 which has multiplicity 1 and is split by a quadratic extension $K/\mathbb{Q}$. Let C_v be the conic

$$v_1 x_1^2 + v_2 x_2^2 = x_0^2$$

in $\mathbb{P}^2$ defined by $v=(v_1,v_2)\in\mathbb{Q}^\times/\mathbb{Q}^{\times,2}\times\mathbb{Q}^\times/\mathbb{Q}^{\times,2}$. We apply the partition in part (b) of Proposition 4.8 with the full family of twists

$$\theta_v : C_v \to \mathbb{P}^1, \quad [x_0 : x_1 : x_2] \mapsto [v_1 x_0^2 : v_2 x_1^2].$$

This is the finest partition in the sense of Remark 4.4 since we have $\pi_1^{\mathrm{orb}}(\mathbb{P}^1,\partial_\pi)=\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ and any θ_v is geometrically a universal orbifold cover. Consider the fibres $\theta_v^{-1}(1)$ as v varies, which on algebras are biquadratic étale $\mathbb{Q}$-algebras $\prod_i N_{1,v}^{(i)}$. Infinitely many of these contain the splitting field K of the fibre, and for such v we have

$$1-\delta_{1,\mathbb{Q}} < 1 = 1-\delta_{1,N_{1,v}^{(\alpha)}}(\pi) < \sum_i (1-\delta_{1,N_{1,v}^{(i)}}(\pi)),$$

where α is such that $K\subseteq N_{1,v}^{(\alpha)}$. However, each of these infinitely many $(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z})$-covers factors through only two $\mathbb{Z}/2\mathbb{Z}$-covers. Hence the set of points counted through the v for which

$$1-\delta_{1,\mathbb{Q}} \neq \sum_i (1-\delta_{1,N_{1,v}^{(i)}}(\pi))$$

is a thin set. In the case of a nontrivial Galois action on the components of the multiple fibres, we will need to deal with them in a similar manner to conclude that the points counted in the covers θ_v for $\Delta(\pi_v)\neq\Delta(\pi)$ form a thin set.

5. A sparsity criterion

In this section we derive a sparsity condition for the fibres of a morphism $f : X \to Y$ to have a $\mathbb{Q}_p$-point under geometric conditions on X, Y and f. After fixing a model $f : \mathscr{X} \to \mathscr{Y}$ over $\mathbb{Z}_S$ away from finitely many primes S, and after possibly enlarging S for $p\notin S$ and $\tilde{t}\in\mathscr{Y}(\mathbb{F}_p)$, we will give an exact criterion for which lifts $t\in\mathscr{Y}(\mathbb{Z}_p)$ of $\tilde{t}$ we can lift an $\mathbb{F}_p$-point on the $\mathbb{F}_p$-scheme $\mathscr{X}_{\tilde{t}}$ to a $\mathbb{Z}_p$-point on $\mathscr{X}_t$.

The exact criterion, Theorem 5.5, is based on a version of Hensel's lemma which takes into account the intersection multiplicities of a component of a fibre of f with the point in $\mathscr{X}_t(\mathbb{Z}_p)$, and the lifting is done in such a way that these required relations are preserved at each step. In order to include this information naturally, we use a logarithmic structure. For a basic introduction to log geometry with a view towards arithmetic applications, the reader is referred to [Loughran et al. 2020, §5].

5.1. ***Logarithmic lifting in families.*** Let k be a number field. Let X and Y be smooth, proper varieties over k, and let D and E be strict normal crossing divisors on X and Y, respectively, where $f^{-1}(E)\subseteq D$. Assume that the induced morphism $f:(X,D)\to(Y,E)$ is a *toroidal* morphism, i.e., a toroidal morphism between toroidal embeddings, or equivalently, a log smooth morphism of (Zariski) log regular schemes. Fix $Q\in Y(k)$. We want to understand when $f^{-1}(Q)$ is everywhere locally soluble.

Let S be a finite set of places including all *places of bad reduction for* f. This means that we have a good model $\bar{f}:(\mathscr{X},\mathscr{D})\to(\mathscr{Y},\mathscr{E})$ for f over $\mathscr{O}_{k,S}$ with the property that $\bar{f}^{-1}(\mathscr{E})\subseteq\mathscr{D}$ such that $(\mathscr{X},\mathscr{D})$ and $(\mathscr{Y},\mathscr{E})$ are still log regular, and such that $\bar{f}$ is still log smooth with respect to the divisorial log structures induced by $\mathscr{D}$ and $\mathscr{E}$.

Let $v \notin S$ be a finite place of k. Let $\mathcal{Q}_v \in \mathcal{Y}(\mathcal{O}_v)$ be the unique lift of $Q \in Y(k)$ to an $\mathcal{O}_v$-point. We will give necessary and sufficient conditions for the existence of an $\mathcal{O}_v$-point $\mathcal{P}_v$ on $\mathcal{X}$ such that $f(\mathcal{P})_v = \mathcal{Q}_v$ for $v \gg 0$.

If $\mathcal{Q}_v \nsubseteq \mathcal{E}$, then the $\mathcal{O}_v$-point $\mathcal{Q}_v$ can be seen as a morphism

$$\mathcal{Q}_v : (\operatorname{Spec} \mathcal{O}_v)^\dagger \to (\mathcal{Y}, \mathcal{E})$$

of log schemes, where $(\operatorname{Spec} \mathcal{O}_v)^\dagger$ is the scheme $\operatorname{Spec} \mathcal{O}_v$ equipped with the divisorial log structure induced by the closed point. This morphism induces a morphism of associated *Kato fans*

$$F(\mathcal{Q}_v) : \operatorname{Spec} \mathbb{N} \cong F((\operatorname{Spec} \mathcal{O}_v)^\dagger) \to F(\mathcal{Y}, \mathcal{E}).$$

In other words, we get an $\mathbb{N}$-valued point $F(\mathcal{Q}_v) \in F(\mathcal{Y}, \mathcal{E})(\mathbb{N})$.

If $\mathcal{Q}_v$ is the image of $\mathcal{P}_v \in \mathcal{X}(\mathcal{O}_v)$, then clearly $F(\mathcal{Q}_v)$ cannot lie anywhere in $F(\mathcal{Y}, \mathcal{E})(\mathbb{N})$; it needs to be an element of the potentially smaller set

$$\operatorname{image}(F(\mathcal{X}, \mathcal{D})(\mathbb{N}) \to F(\mathcal{Y}, \mathcal{E})(\mathbb{N})).$$

This means that if $F(\mathcal{Q}_v)$ does *not* lie in the image of $F(\mathcal{X}, \mathcal{D})(\mathbb{N})$, then surely $\mathcal{Q}_v$ cannot lift to an $\mathcal{O}_v$-point on $\mathcal{X}$. This is a *sparsity criterion* in the sense of [Loughran and Smeets 2016, §2] but still a rather naive one, since it does not take important arithmetic information into account.

Definition 5.1. Let $\overline{\mathcal{P}}_v$ be an $\mathbb{F}_p$-point on $\mathcal{X}$. With the notation above, we define $F(\mathcal{X}, \mathcal{D})(\mathbb{N})_{\overline{\mathcal{P}}_v}$ as the subset of $F(\mathcal{X}, \mathcal{D})(\mathbb{N})$ with the property that $\overline{\mathcal{P}}_v$ lies in the logarithmic stratum associated to the image of the closed point $\mathbb{N}_{>0}$ of $\operatorname{Spec} \mathbb{N}$.

Proposition 5.2. *With notation as above, let $\overline{\mathcal{P}}_v$ be an $\mathbb{F}_p$-point on $\mathcal{X}_{\mathcal{Q}_v}$, and assume that $F(\mathcal{Q}_v)$ does **not** lie in*

$$\operatorname{image}(F(\mathcal{X}, \mathcal{D})(\mathbb{N})_{\overline{\mathcal{P}}_v} \to F(\mathcal{Y}, \mathcal{E})(\mathbb{N})).$$

Then $\overline{\mathcal{P}}_v \in \mathcal{X}_{\mathcal{Q}_v}(\mathbb{F}_p)$ does not lift to $\mathcal{P}_v \in \mathcal{X}_{\mathcal{Q}_v}(\mathcal{O}_v)$.

Proof. Assume that $\overline{\mathcal{P}}_v$ lifts; i.e., $\mathcal{Q}_v = \bar{f}(\mathcal{P}_v)$ for some $\mathcal{P}_v \in \mathcal{X}(\mathcal{O}_v)$ with $\overline{\mathcal{P}}_v = \mathcal{P}_v \bmod v$ (which is the image of $\operatorname{Spec} \mathbb{F}_v$ under $\mathcal{P}_v$). Therefore the image of $F(\mathcal{P}_v) \in F(\mathcal{X}, \mathcal{D})(\mathbb{N})$ under the map $F(\mathcal{X}, \mathcal{D})(\mathbb{N}) \to F(\mathcal{Y}, \mathcal{E})(\mathbb{N})$ comes from $F(\mathcal{X}, \mathcal{D})(\mathbb{N})_{\overline{\mathcal{P}}_v}$, as desired. □

Remark 5.3. In fact, the above sparsity condition can often be phrased in a more classical way. Let $\pi : X \to \mathbb{P}^1$ be a standard fibration, and suppose we have a projective model $\pi : \mathcal{X} \to \mathbb{P}^1_{\mathcal{O}_S}$ with $\mathcal{X}$ smooth over $\mathcal{O}_S$. Let $h \in \mathcal{O}_S[x, y]$ be an irreducible binary form; in practice we will only need to consider the finitely many h for which the fibre of π over $V(h) \subseteq \mathbb{P}^1$ is nonsplit. After possibly enlarging S, we may argue as follows. Suppose that we have a point $\mathcal{Q}_v$ in the fibre $\mathcal{X}_{\mathcal{P}_v}$ over a point $\mathcal{P}_v \in \mathbb{P}^1(\mathcal{O}_v)$ for $v \notin S$. Let $\mathcal{Z}_i$ be the geometrically irreducible components of $\mathcal{X}_{V(h)}$ which contain $\overline{\mathcal{Q}}_v \in \mathcal{X}_{\overline{\mathcal{P}}_v}$. Then we may conclude that $v(h(\mathcal{P}_v))$ lies in the positive linear span of the multiplicities m_i of $\mathcal{Z}_i$. Indeed, in the local ring of $\overline{\mathscr{Q}}_v$, we have the factorisation $\pi^* h = \prod_i h_i^{m_i}$. Hence

$$v(h(\mathscr{P}_v)) = v(\pi^* h(\mathscr{Q}_v)) = \sum m_i v(h_i(\mathscr{Q}_v)).$$

Moreover, since we are only considering the components containing $\overline{\mathscr{Q}}_v$, we clearly have $v(h_i(\mathscr{Q}_v)) > 0$ for each i. In this way one can use Proposition 5.2 to give a sparsity criterion for general fibrations that generalises [Loughran and Smeets 2016, Theorem 2.8] after excluding a subscheme of codimension at least 2 on the base.

We are considering the log smooth setting since then we can provide a converse result to Proposition 5.2 using the logarithmic Hensel lemma [Loughran et al. 2020, Proposition 5.13].

Proposition 5.4. *If $\overline{\mathscr{P}}_v$ is an $\mathbb{F}_v$-point on $\mathscr{X}_{\mathscr{Q}_v}$, the following are equivalent*:

(a) *$\overline{\mathscr{P}}_v$ lifts to an $\mathscr{O}_v$-point on $\mathscr{X}_{\mathscr{Q}_v}$*;

(b) $F(\mathscr{Q}_v) \in \operatorname{image}(F(\mathscr{X}, \mathscr{D})(\mathbb{N})_{\overline{\mathscr{P}}_v} \to F(\mathscr{Y}, \mathscr{E})(\mathbb{N}))$.

Proof. Since we have already shown that (a) implies (b), it remains to prove the reverse implication. This is an application of [Loughran et al. 2020, Proposition 5.13]. Indeed, let $s^\dagger = \operatorname{Spec} \mathbb{F}_v$, with the standard log structure of rank 1, and $S^\dagger = \operatorname{Spec} \mathscr{O}_v$. Let $j : s^\dagger \to S^\dagger$ be the canonical closed immersion.

By assumption there is an element $p_v \in F(\mathscr{X}, \mathscr{D})(\mathbb{N})_{\overline{\mathscr{P}}_v}$ which maps to $F(\mathscr{Q}_v) \in F(\mathscr{Y}, \mathscr{E})(\mathbb{N})$, and there is an $\mathbb{F}_v$-point $u : \operatorname{Spec} \ \mathbb{F}_v \to X$ on the associated stratum of $(\mathscr{X}, \mathscr{D})$. We can uniquely make u into a morphism of log schemes $s^\dagger \to (\mathscr{X}, \mathscr{D})$ such that $F(u) = p_v$ under the identification $F(\mathbb{N}) = F(s^\dagger)$, similar to the proof of Proposition 6.1 in [Loughran et al. 2020].

Since $F(f)$ maps $F(u)$ to $F(\mathscr{Q}_v)$, we have a commutative diagram

$$\begin{array}{ccc} s^\dagger & \xrightarrow{u} & (\mathscr{X}, \mathscr{D}) \\ {\scriptstyle j}\downarrow & & \downarrow{\scriptstyle \bar{f}} \\ S^\dagger & \xrightarrow[\mathscr{Q}_v]{} & (\mathscr{Y}, \mathscr{E}) \end{array}$$

Now [Loughran et al. 2020, Proposition 5.13] provides a lift $S^\dagger \to (\mathscr{X}, \mathscr{D})$ of $\mathscr{Q}_v$. The morphism of schemes which underlies this lift is the $\mathscr{O}_v$-point $\mathscr{P}_v$ we are looking for. □

5.2. ***Sparsity criteria.*** Using Proposition 5.4, we can give precise conditions for locally solubility. We allow ourselves to work over a general number field $k/\mathbb{Q}$ and so define a *standard fibration* to be a dominant morphism $\pi : X \to \mathbb{P}^1$ with geometrically integral generic fibre such that X is a smooth, proper, geometrically irreducible k-variety.

Let E be the reduced divisor of $\mathbb{P}^1$ of the nonsplit fibres of π. Let D be the reduced divisor underlying $\pi^{-1}(E)$. By embedded resolutions of singularities, there exists a birational morphism $X' \to X$ such that the pullback D' of D has strict normal crossings. Since $X \setminus D \cong X' \setminus D'$ over $\mathbb{P}^1$, we see that $N_{\mathrm{loc}}(\pi', B)$ differs by a constant from $N_{\mathrm{loc}}(\pi, B)$, where $\pi' : X' \to X \to Y$ is the composition. Thus, for the purposes of upper and lower bounds, we can assume without loss of generality that the reduced subschemes of the nonsplit fibres of π have strict normal crossings.

Theorem 5.5. *Let $X \to \mathbb{P}^1$ be a standard fibration whose nonsplit fibres in their reduced subscheme structure are sncd. There exists a finite set of primes S and a model $\mathscr{X} \to \mathbb{P}^1_{\mathscr{O}_S}$ such that the following holds for $v \notin S$. Fix a point $\mathscr{Q} \in \mathbb{P}^1(\mathscr{O}_S)$ for which the fibre X_Q is split. Then any $\mathbb{F}_v$-point $\overline{\mathscr{P}}_v \in \mathscr{X}_{\mathscr{Q}}(\mathbb{F}_v)$ lifts to a point $\mathscr{P}_v \in \mathscr{X}_{\mathscr{Q}}(\mathscr{O}_v)$ precisely if, for every closed point $V(h) \in (\mathbb{P}^1)^{(1)}$, we have that $v(h(\mathscr{Q}))$ lies in the positive linear span of the multiplicities m_i of the components of $\mathscr{X}_{V(h)}$ that contain $\overline{\mathscr{P}}_v$.*

From now on, when we have fixed a place $v \notin S$, we will assume that a closed point $V(h) \in (\mathbb{P}^1)^{(1)}$ is given by an $\mathscr{O}_v$-primitive irreducible form $h \in \mathscr{O}_v[x, y]$.

Note that the last condition in Theorem 5.5 is trivially satisfied for all closed points $V(h)$ for which $v(h(\mathscr{Q})) = 0$ and also for those for which $X_{V(h)}$ is split. By restricting S further, we can assume that there is at most one nonsplit fibre $X_{V(h)}$ for which we have to check this condition.

Proof of Theorem 5.5. By the definition of D on X and E on $\mathbb{P}^1$, we see that $(X, D) \to (\mathbb{P}^1, E)$ is log smooth. For a suitable finite set of primes S, this extends to $\mathscr{O}_S$-schemes and divisors such that $\mathscr{D} \subseteq \mathscr{X}$ and $\mathscr{E} \subseteq \mathbb{P}^1_{\mathscr{O}_S}$ still have strict normal crossings and $(\mathscr{X}, \mathscr{D}) \to (\mathbb{P}^1_{\mathscr{O}_S}, \mathscr{E})$ is also log smooth. We will check that this model satisfies the condition.

Consider $\overline{\mathscr{P}}_v \in \mathscr{X}_{\mathscr{Q}}(\mathbb{F}_v)$, and let $V(h) \subseteq \mathbb{P}^1_{\mathscr{O}_S}$ be the unique nonsplit fibre containing $\overline{\mathscr{Q}}_v = \pi(\overline{\mathscr{P}}_v)$. Suppose that we can write $v(h(\mathscr{Q})) = \sum_i a_i m_i$, with $a_i > 0$ integers and m_i the multiplicities of the r components of $X_{v(h)}$ which contain $\overline{\mathscr{P}}_v$. Around $\overline{\mathscr{P}}_v$ and $\overline{\mathscr{Q}}_v$ the Kato fans have affine charts $\mathbb{N}^r$ and $\mathbb{N}$. Under this identification, we have $F(\mathscr{Q}_v) = v(h(\mathscr{Q})) \in \mathbb{N}$, and $F(\mathscr{X}, \mathscr{D})(\mathbb{N}) \to F(\mathbb{P}^1_{\mathscr{O}_S}, \mathscr{E})(\mathbb{N})$ is given by $(u_i) \mapsto \sum m_i u_i$. Hence the result follows from Proposition 5.4. □

Remark 5.6. In [Loughran and Smeets 2016, §2] the following was proven: if $v(h(\mathscr{Q})) = 1$, then $\mathscr{X}_{\mathscr{Q}}$ is a regular scheme. This implies that any $\mathbb{F}_v$-point on $\mathscr{X}_{\mathscr{Q}}$ which lies on the intersection of at least two components of the reduction $\mathscr{X}_{\mathscr{Q},v}$ does not lift to a $\mathbb{Q}_v$-point on $\mathscr{X}_{\mathscr{Q}}$. This last statement directly follows from our criterion above since then the valuation $v(h(\mathscr{Q})) = 1$ cannot possibly lie in the positive linear span of two positive integers.

The above conditions make it easy to check if an $\mathbb{F}_v$-point lifts. However, one cannot deduce the existence of $\mathbb{F}_v$-points purely from valuations and multiplicities, as explained by Loughran and Matthiesen [2024, Lemma 6.2]. In general, this only allows us to give necessary conditions for local solubility.

Corollary 5.7. *Let $X \to \mathbb{P}^1$ be a standard fibration, and let $Q \in \mathbb{P}^1(k)$. There exists a finite set of places S such that, for each $v \notin S$ with $X_Q(k_v) \neq \varnothing$, the following condition is satisfied: for every closed point $D = V(h) \in (\mathbb{P}^1)^{(1)}$, we have either $v(h(Q)) > m_D$, or $v(h(Q)) = m_D$ and v belongs to*

$$T_D = \{v \notin S : \mathrm{Frob}_v \text{ fixes an element of } S_D\}.$$

(*Recall that S_D is the set of geometric components of X_D of minimum multiplicity m_D.*)

In the special case that the nonsplit fibres all lie above k-rational points in $\mathbb{P}^1$, we can (after possibly extending the set S again) make this even more precise, as follows.

Corollary 5.8. *Let $X \to \mathbb{P}^1$ be a standard fibration, and let $Q \in \mathbb{P}^1(k)$. Assume that the nonsplit fibres of $X \to \mathbb{P}^1$ all lie above k-rational points. Then $X_Q(k_v) \neq \varnothing$ precisely if, for every $V(h) \in (\mathbb{P}^1)^{(1)}$, the fibre $X_{V(h)}$ has a stratum fixed by* Frob_v *which lies on the intersection of components of multiplicity m_i such that $v(h(Q))$ lies in the positive linear span of the m_i.*

Proof. We will start with S and $\mathscr{X} \to \mathbb{P}^1_{\mathscr{O}_S}$ as above. By the results above we have that $P_v \in X_Q(k_v)$ reduces to an $\mathbb{F}_v$-point on $\mathscr{X}$. Since this $\mathbb{F}_v$-points lifts we get the result.

For the inverse implication we will need to enlarge S, as follows. Firstly we do so to assume that all fibres of $\mathscr{X} \setminus \mathscr{D} \to \mathbb{P}^1_{\mathscr{O}_S} \setminus \mathscr{E}$ are geometrically integral. Consider an $\mathbb{F}_v$-point on $\mathbb{P}^1_{\mathscr{O}_S} \setminus \mathscr{E}$ for $v \notin S$. Its fibre in $\mathscr{X}$ lies in the open stratum $\mathscr{X} \setminus \mathscr{D}$ and contains a smooth $\mathbb{F}_v$-point by Lang–Weil. Now let W be a geometric component of a nonopen stratum of (X, D), which is defined over k'/k. The closure $\mathscr{W}$ of W will have geometrically irreducible fibres over all but finitely many places of k'. Hence after enlarging S we see that $\mathscr{W}$ has an $\mathbb{F}_{v'}$-point for all $v' \mid v$ for $v \notin S$. Since there are only finitely many strata and each has again finitely many geometric components, we can enlarge S to make this true for all possible W.

Suppose now that Frob_v fixes a geometric component W of a stratum which has k'/k as its field of definition. Consider the multiplicities m_i of the components that contain W. Since Frob_v fixes W, we conclude that there is a place $v' \mid v$ of k' of residue degree 1. For this v', we see that W contains an $\mathbb{F}_{v'} = \mathbb{F}_v$-point. We can lift this point under the conditions in Theorem 5.5. □

6. Multiple fibres via the large sieve

We place ourselves in the setting of Theorems 1.2 and 1.3. Let $\pi : X \to \mathbb{P}^1$ be a standard fibration with orbifold divisor

$$\partial_\pi = \Big(1 - \frac{1}{m_0}\Big)[0] + \Big(1 - \frac{1}{m_\infty}\Big)[\infty],$$

in the notation of (1-4) for $m_0, m_\infty \in \mathbb{N}$. Note that $2 - \deg \partial_\pi = 1/m_0 + 1/m_\infty$. We define $d = \gcd(m_0, m_\infty)$. We shall apply the theory from Section 4 to the family of μ_d-torsors

$$\theta_v : \mathbb{P}^1 \to \mathbb{P}^1, \quad [x_0 : x_1] \to [v_0 x_0^d : v_1 x_1^d],$$

which are parametrised by $v = v_1/v_0 \in \mathbb{Q}^\times/\mathbb{Q}^{\times,d} = \mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \mu_d)$. Let $\pi_v : X_v \to \mathbb{P}^1$ be the normalisation of the pullback of π along θ_v.

The main result of this section is the following, which pertains to the density of locally soluble fibres on the standard fibration $\pi_v : X_v \to \mathbb{P}^1$ relative to the pullback height H_v along θ_v. We denote by $\mathrm{rad}(n) = \prod_{p \mid n} p$ the square-free radical of any $n \in \mathbb{N}$.

Proposition 6.1. *Let $\varepsilon > 0$, and let $v = v_1/v_0 \in \mathbb{Q}^\times/\mathbb{Q}^{\times,d}$. Then*

$$N_{\mathrm{loc}}(\pi_v, H_v, B) \ll_\varepsilon c_{v,\varepsilon} B^{1/m_0 + 1/m_\infty},$$

where

$$c_{v,\varepsilon} = \frac{|v_0 v_1|^\varepsilon}{\mathrm{rad}(v_0)|v_0|^{1/m_0}\, \mathrm{rad}(v_1)|v_1|^{1/m_\infty}}. \tag{6-1}$$

Furthermore, if $|v_0v_1| \leqslant B^\varepsilon$, *then*

$$N_{\mathrm{loc}}(\pi_v, H_v, B) \ll_\varepsilon c_{v,\varepsilon} \frac{B^{1/m_0+1/m_\infty}}{(\log B)^{\Theta_v(\pi)}},$$

where $\Theta_v(\pi)$ *is given by* (1-9).

We shall begin the proof of this result in Section 6.2. Our argument is based on the large sieve, which is recalled in Section 6.1. Taking the result on faith for the moment, we proceed to show how it can be used to establish Theorems 1.2 and 1.3.

Remark 6.2. Proposition 6.1 is consistent with Conjecture 1.5 for a fixed choice of $v \in \mathbb{Q}^\times/\mathbb{Q}^{\times,d}$. Indeed, we have $H_v(x) = H(x)^d$, where $H(x)$ is an $\mathscr{O}(1)$-height on $\mathbb{P}^1$. It follows that

$$N_{\mathrm{loc}}(\pi_v, B) = N_{\mathrm{loc}}(\pi_v, H, B) = N_{\mathrm{loc}}(\pi_v, H_v, B^d) \ll_v \frac{B^{d/m_0+d/m_\infty}}{(\log B)^{\Theta_v(\pi)}}.$$

The orbifold base of π_v is (X_v, ∂_{π_v}), with

$$\partial_{\pi_v} = \Big(1 - \frac{d}{m_0}\Big)[0] + \Big(1 - \frac{d}{m_\infty}\Big)[\infty]$$

by part (b) of Proposition 4.7. It follows from (4-1) and part (c) of Proposition 4.7 that

$$\frac{d}{m_0} + \frac{d}{m_\infty} = -d \deg K_{\pi,\partial_\pi} = 2 - \deg \partial_{\pi_v}.$$

Moreover, $\Theta_v(\pi) = \Delta(\pi_v)$ by part (c) of Proposition 4.8.

Proof of Theorem 1.2. In this case there is only one multiple fibre above 0, and so $m_\infty = 1$ and $d = 1$. Thus $\mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \mu_d)$ is the trivial group, and it follows directly from Proposition 6.1 that

$$N_{\mathrm{loc}}(\pi, B) \ll \frac{B^{1/m_0+1}}{(\log B)^{\Theta_1(\pi)}}.$$

We have already seen that $1/m_0 + 1 = 2 - \deg \partial_\pi$. Moreover, we saw that $\Theta_1(\pi) = \Delta(\pi)$ in (1-11). □

Proof of Theorem 1.3. We appeal to the decomposition in part (b) of Proposition 4.8. This gives

$$N_{\mathrm{loc}}(\pi, B) \ll \sum_{v=v_1/v_0 \in \mathbb{Q}^\times/\mathbb{Q}^{\times,d}} N_{\mathrm{loc}}(\pi_v, H_v, B).$$

For any $\delta > 0$, we clearly have

$$\sum_{n>x} \frac{1}{\mathrm{rad}(n)n^\delta} < \sum_{n=1}^\infty \frac{(n/x)^{\delta/2}}{\mathrm{rad}(n)n^\delta} = x^{-\delta/2} \prod_p \Big(1 + \sum_{k=1}^\infty \frac{1}{p^{1+k\delta/2}}\Big)$$
$$\ll_\delta x^{-\delta/2}.$$

Let $\varepsilon > 0$. In the light of the latter bound, it follows from the first part of Proposition 6.1 that there exists $\delta(\varepsilon) > 0$ such that the terms with $|v_0v_1| > B^\varepsilon$ make an overall contribution $O_\varepsilon(B^{1/m_0+1/m_\infty-\delta(\varepsilon)})$ to $N_{\mathrm{loc}}(\pi, B)$. For the terms with $|v_0v_1| \leqslant B^\varepsilon$, we apply the second part of Proposition 6.1.

This easily leads to the conclusion that

$$N_{\mathrm{loc}}(\pi, B) \ll_{\varepsilon} B^{1/m_0+1/m_\infty-\delta(\varepsilon)} + \sum_{|v_0v_1|\leqslant B^{\varepsilon}} c_{v,\varepsilon} \frac{B^{1/m_0+1/m_\infty}}{(\log B)^{\Theta_v(\pi)}} \ll_{\varepsilon} \frac{B^{1/m_0+1/m_\infty}}{(\log B)^{\Theta(\pi)}},$$

where $\Theta(\pi)$ is given by (1-10). The statement of the theorem follows, since we have already remarked that $1/m_0 + 1/m_\infty = 2 - \deg \partial_\pi$. □

6.1. ***The large sieve.*** We begin by stating the version of the large sieve that we shall use in this paper.

Lemma 6.3. *Let $m \in \mathbb{N}$, let $B_0, B_1 \geqslant 1$ and let $\Omega \subseteq \mathbb{Z}^2$. For each prime p, assume that there exists $\bar{\omega}(p) \in [0, 1)$ such that the reduction modulo p^m of Ω has cardinality at most $(1-\bar{\omega}(p))p^{2m}$. Then*

$$\#\{\boldsymbol{x} \in \Omega : |x_i| \leqslant B_i \text{ for } i = 0, 1\} \ll \frac{(B_0 + Q^{2m})(B_1 + Q^{2m})}{L(Q)}$$

for any $Q \geqslant 1$, where

$$L(Q) = \sum_{q\leqslant Q} \mu^2(q) \prod_{p\mid q} \frac{\bar{\omega}(p)}{1-\bar{\omega}(p)}.$$

Proof. When $m = 1$, this is a straightforward rephrasing of the multidimensional large sieve worked out in [Kowalski 2008, Theorem 4.1]. The extension to $m > 1$ is routine and will not be explained here. □

6.2. ***Preliminary steps.*** Recall that $d = \gcd(m_0, m_\infty)$. Henceforth, we usually write $\boldsymbol{v} = (v_0, v_1) \in \mathbb{Z}^2_{\mathrm{prim}}$ for the point $v = v_1/v_0 \in \mathbb{Q}^\times/\mathbb{Q}^{\times,d}$. We may clearly proceed under the assumption that v_0 and v_1 are both free of d-th powers.

Let S be a large enough finite set of primes, as required for the arguments in Section 5 to go through. Suppose that $E_1, \dots, E_r \in (\mathbb{P}^1)^{(1)}$ are the closed points distinct from 0 and ∞, where π_v is not smooth. For each $1 \leqslant j \leqslant r$, assume that $E_j = V(h_j)$ for a square-free binary form $h_j \in \mathbb{Z}_S[x_0, x_1]$. We may further assume that h_j is irreducible over $\mathbb{Q}$ and coprime to the monomial x_0x_1, and that the coefficients of h_j are relatively coprime.

We proceed by defining the sets

$$T_0 = \{p \notin S : \mathrm{Frob}_p \text{ fixes an element of } S_0\},$$
$$T_\infty = \{p \notin S : \mathrm{Frob}_p \text{ fixes an element of } S_\infty\},$$
$$U_j = \{p \notin S : \mathrm{Frob}_p \text{ fixes an element of } S_{E_j}\}$$

for $1 \leqslant j \leqslant r$. The fibre $X_{v,y}$ of the fibration $\pi_v : X_v \to \mathbb{P}^1$ has a $\mathbb{Q}_p$-point precisely if $X_{\pi_v(y)}$ does, and thus we can apply the sparsity conditions in Corollary 5.7. This yields the upper bound $N_{\mathrm{loc}}(\pi_v, B) \leqslant M_{\boldsymbol{v}}(B)$, where $M_{\boldsymbol{v}}(B)$ is defined to be the number of $\boldsymbol{y} = (y_0, y_1) \in \mathbb{Z}^2$ such that $\gcd(v_0y_0, v_1y_1) = 1$ and $\max\{|v_0y_0^d|, |v_1y_1^d|\} \leqslant B$, with

$$p \notin S \Rightarrow \begin{cases} [v_p(x_0) = m_0 \text{ and } p \in T_0] \text{ or } v_p(x_0) > m_0, \\ [v_p(x_1) = m_\infty \text{ and } p \in T_\infty] \text{ or } v_p(x_1) > m_\infty, \\ p \| h_j(\boldsymbol{x}) \Rightarrow p \in U_j, \end{cases}$$

where $(x_0, x_1) = (v_0 y_0^d, v_1 y_1^d)$. We write $v_0 = a_0 w_0'$ and $y_0 = b_0 z_0$, where $w_0' z_0$ is coprime to all the primes in S and $p \mid a_0 b_0 \Rightarrow p \in S$. Let $p \notin S$. Then $v_p(w_0' z_0^d) = m_0$ if and only if $v_p(w_0') = 0$ and $v_p(z_0) = m_0/d$, since w_0' is free of d-th powers. Similarly, if $v_p(w_0' z_0^d) > m_0$, then either $v_p(z_0) > m_0/d$, or $v_p(z_0) = m_0/d$ and $p \mid w_0'$. This suggests that we may write

$$v_0 = a_0 w_0, \quad y_0 = b_0 s_0^{m_0/d} t_0^{m_0/d} u_0,$$

where

- $p \mid a_0 b_0 \Rightarrow p \in S$;
- $p \mid s_0 w_0 u_0 \Rightarrow p \notin S$;
- s_0 and t_0 are square-free;
- $p \mid w_0 \Rightarrow p \mid s_0$;
- $p \mid t_0 \Rightarrow p \in T_0$; and
- u_0 is $(m_0/d + 1)$-full.

Similarly, we have a factorisation

$$v_1 = a_1 w_1, \quad y_1 = b_1 s_1^{m_\infty/d} t_1^{m_\infty/d} u_1,$$

where

- $p \mid a_1 b_1 \Rightarrow p \in S$;
- $p \mid s_1 w_1 u_1 \Rightarrow p \notin S$;
- s_1 and t_1 are square-free;
- $p \mid w_1 \Rightarrow p \mid s_1$;
- $p \mid t_1 \Rightarrow p \in T_\infty$; and
- u_1 is $(m_\infty/d + 1)$-full.

There are $O_\varepsilon(|v_0 v_1|^\varepsilon)$ choices for $a_i, s_i, w_i \in \mathbb{Z}$ for $i = 0, 1$ by the standard estimate for the divisor function. We fix a choice of $b_0,\ b_1,\ u_0,\ u_1$ and write

$$A_0 = a_0 b_0^d s_0^{m_0} u_0^d w_0 \quad \text{and} \quad A_1 = a_1 b_1^d s_1^{m_\infty} u_1^d w_1. \tag{6-2}$$

Note that we have $\gcd(A_0, A_1) = 1$. Moreover, let

$$R_0 = \left(\frac{B}{|A_0|}\right)^{1/m_0}, \quad R_1 = \left(\frac{B}{|A_1|}\right)^{1/m_\infty},$$

and

$$g_j(\boldsymbol{t}) = h_j(A_0 t_0^{m_0}, A_1 t_1^{m_\infty}) \quad \text{for } 1 \leqslant j \leqslant r. \tag{6-3}$$

The binary form $g_j(t_0, t_1)$ is square-free and coprime to the monomial t_0t_1 since $h_j(t_0, t_1)$ satisfies these properties. For a (possibly infinite) set T of primes, let

$$\mathbf{1}_T(n) = \begin{cases} 1 & \text{if } p \mid n \Rightarrow p \in T, \\ 0 & \text{otherwise.} \end{cases}$$

In what follows, we will also write T^c to denote the complement of T in the full set of primes.

Then, with all this notation in mind, we have

$$M_{\boldsymbol{v}}(B) \ll \sum_{v_0=a_0w_0} \sum_{v_1=a_1w_1} \sum_{\substack{b_0,b_1 \\ p \mid b_0b_1 \Rightarrow p \in S}} \sum_{u_0,u_1 \in \mathbb{Z}} L(R_0, R_1),$$

where

$$L(R_0, R_1) = \sum_{\substack{(t_0,t_1)\in\mathbb{Z}^2 \\ |t_0|\leqslant R_0, |t_1|\leqslant R_1}} \mu^2(t_0t_1)\mathbf{1}_{T_0}(t_0)\mathbf{1}_{T_\infty}(t_1) \prod_{j=1}^{r} \mathbf{1}^{\sharp}_{U_j}(t_0, t_1) \tag{6-4}$$

and where

$$\mathbf{1}^{\sharp}_{U_j}(t_0, t_1) = \begin{cases} 1 & \text{if } p \| g_j(\boldsymbol{t}) \Rightarrow p \in U_j, \\ 0 & \text{otherwise.} \end{cases}$$

The trivial bound for $L(R_0, R_1)$ is

$$L(R_0, R_1) \ll \frac{B^{1/m_0+1/m_\infty}}{|A_0|^{1/m_0}|A_1|^{1/m_\infty}} \ll \frac{B^{1/m_0+1/m_\infty}}{|s_0|\,|v_0|^{1/m_0}|s_1|\,|v_1|^{1/m_\infty}|b_0u_0|^{d/m_0}|b_1u_1|^{d/m_\infty}}$$

by (6-2). Clearly

$$|s_i| \gg \operatorname{rad}(v_i) \quad \text{for } i = 0, 1 \tag{6-5}$$

for a suitable implied constant depending only on S. Note that

$$\sum_{\substack{|b_0|>J \\ p \mid b_0 \Rightarrow p \in S}} |b_0|^{-d/m_0} \ll \frac{1}{J^{d/m_0}}$$

for any $J \geqslant 1$. Similarly,

$$\sum_{\substack{|u_0|>J \\ u_0 \text{ is } (m_0/d+1)\text{-full}}} |u_0|^{-d/m_0} \ll \frac{1}{J^{d^2/(m_0(m_0+d))}}.$$

Let $\varepsilon > 0$. In what follows it will be convenient to recall the notation (6-1) for $c_{v,\varepsilon}$ in the statement of Proposition 6.1. It now follows that the overall contribution to $M_{\boldsymbol{v}}(B)$ from parameters $b_0,\ u_0$ in the range $\min(|b_0|, |u_0|) > B^{\varepsilon}$ or parameters $b_1,\ u_1$ in the range $\min(|b_1|, |u_1|) > B^{\varepsilon}$ is clearly

$$\ll_{\varepsilon} c_{v,\varepsilon} B^{1/m_0+1/m_\infty-\varepsilon/(m_0^2m_\infty^2)}$$

since we have seen that there are $O_{\varepsilon}(|v_0v_1|^{\varepsilon})$ choices for $a_i, s_i, w_i \in \mathbb{Z}$ associated to a particular choice of $\boldsymbol{v}$. Thus we deduce that

$$M_{\boldsymbol{v}}(B) \ll_{\varepsilon} \sum_{v_0=a_0w_0} \sum_{v_1=a_0w_1} \sum_{\substack{|b_0|,|b_1|\leqslant B^{\varepsilon} \\ p \mid b_0b_1 \Rightarrow p \in S}} \sum_{|u_0|,|u_1|\leqslant B^{\varepsilon}} L(R_0, R_1) + c_{v,\varepsilon} B^{1/m_0+1/m_\infty-\varepsilon/(m_0^2m_\infty^2)}. \tag{6-6}$$

6.3. ***Application of the large sieve.*** We shall now apply Lemma 6.3 to estimate (6-4), which we shall apply with $m=2$. Let $\Omega \subseteq \mathbb{Z}^2$ be the set of vectors $\boldsymbol{t} \in \mathbb{N}^2$ such that $\mathbf{1}_{T_0}(t_0)\mathbf{1}_{T_\infty}(t_1)=1$ and for which $p \in U_j$ whenever there exists an index j such that $p \| g_j(\boldsymbol{t})$. For any prime $p \notin S$, let

$$A_0(p)=\{\boldsymbol{t}\in(\mathbb{Z}/p^2\mathbb{Z})^2 : p \mid t_0 \text{ and } p \notin T_0\}$$

and

$$A_\infty(p)=\{\boldsymbol{t}\in(\mathbb{Z}/p^2\mathbb{Z})^2 : p \mid t_1 \text{ and } p \notin T_\infty\}.$$

Similarly, let

$$B_j(p)=\{\boldsymbol{t}\in(\mathbb{Z}/p^2\mathbb{Z})^2 : p \| g_j(\boldsymbol{t}) \text{ and } p \notin U_j\}$$

for $1\leqslant j\leqslant r$. Then $\#\Omega \bmod p^2 \leqslant (1-\bar{\omega}(p))p^4$, where

$$\bar{\omega}(p)=\frac{\#(A_0(p)\cup A_\infty(p)\cup B_1(p)\cup\cdots\cup B_r(p))}{p^4}.$$

In particular, we have $\bar{\omega}(p)\in[0,1)$. The following result is concerned with estimating this quantity.

Lemma 6.4. *Let $p\notin S$, and let $d=\gcd(m_0,m_\infty)$. Then*

$$\bar{\omega}(p)=\frac{\mathbf{1}_{T_0^c}(p)}{p}+\frac{\mathbf{1}_{T_\infty^c}(p)}{p}+\sum_{j=1}^{r}\frac{\mathbf{1}_{U_j^c}(p)\nu_j(p;\boldsymbol{v})}{p^2}+O\left(\frac{\gcd(p,A_0A_1)}{p^2}\right),$$

where

$$\nu_j(p;\boldsymbol{v})=\#\{\boldsymbol{t}\in\mathbb{F}_p^2 : h_j(v_0t_0^d, v_1t_1^d)=0\}.$$

Proof. Recall that $\gcd(A_0,A_1)=1$, that $g_j(t_0,t_1)$ is defined in (6-3), and that $g_j(t_0,t_1)$ is square-free and coprime to the monomial t_0t_1. If $p\mid A_0A_1$, we take the trivial upper bound

$$\#(A_0(p)\cup A_\infty(p)\cup B_1(p)\cup\cdots\cup B_r(p))=O(p^3),$$

whence $\bar{\omega}(p)=O(1/p)$, which is satisfactory.

Suppose henceforth that $p\nmid A_0A_1$. We proceed by noting that the intersection of any two sets in the union $A_0(p)\cup A_\infty(p)\cup B_1(p)\cup\cdots\cup B_r(p)$ contains $O(p^2)$ elements of $(\mathbb{Z}/p^2\mathbb{Z})^2$. Thus

$$\bar{\omega}(p)=\frac{\mathbf{1}_{T_0^c}(p)}{p}+\frac{\mathbf{1}_{T_\infty^c}(p)}{p}+\sum_{j=1}^{r}\frac{\#B_j(p)}{p^4}+O\left(\frac{1}{p^2}\right).$$

Turning to $\#B_j(p)$ for $j\in\{1,\dots,r\}$, we write $\boldsymbol{u}=\boldsymbol{x}+p\boldsymbol{y}$ for $\boldsymbol{x},\boldsymbol{y}\in\mathbb{F}_p^2$. Thus

$$\#\{\boldsymbol{t}\in(\mathbb{Z}/p^2\mathbb{Z})^2 : p^2\mid g_j(\boldsymbol{t})\}=\sum_{\substack{\boldsymbol{x}\in\mathbb{F}_p^2\\ g_j(\boldsymbol{x})=0}}\#\{\boldsymbol{y}\in\mathbb{F}_p^2 : \boldsymbol{y}.\nabla g_j(\boldsymbol{x})=-g_j(\boldsymbol{x})/p\}.$$

On enlarging S, we can assume that $\nabla g_j(\boldsymbol{x})\neq\mathbf{0}$ for any $\boldsymbol{x}$ in the sum. Thus each of the $O(p)$ values of $\boldsymbol{x}$ produces $O(p)$ choices of $\boldsymbol{y}$, giving

$$\#\{\boldsymbol{t}\in(\mathbb{Z}/p^2\mathbb{Z})^2 : p^2\mid g_j(\boldsymbol{t})\}=O(p^2).$$

Hence

$$\#B_j(p) = \mathbf{1}_{U_j^c}(p)p^2\#\{\boldsymbol{t} \in \mathbb{F}_p^2 : g_j(\boldsymbol{t}) = 0\} + O(p^2).$$

Putting this together, we have shown that

$$\bar{\omega}(p) = \frac{\mathbf{1}_{T_0^c}(p)}{p} + \frac{\mathbf{1}_{T_\infty^c}(p)}{p} + \sum_{j=1}^{r} \frac{\mathbf{1}_{U_j^c}(p)\lambda_j(p; A_0, A_1)}{p^2} + O\left(\frac{1}{p^2}\right),$$

where

$$\lambda_j(p; A_0, A_1) = \#\{\boldsymbol{t} \in \mathbb{F}_p^2 : h_j(A_0 t_0^{m_0}, A_1 t_1^{m_\infty}) = 0\}$$

for $1 \leqslant j \leqslant r$. In order to complete the proof of the lemma, it will suffice to prove that

$$\lambda_j(p; A_0, A_1) = \nu_j(p; \boldsymbol{v}) + O(1) \tag{6-7}$$

for $1 \leqslant j \leqslant r$, in the notation of the lemma.

To see this, let e be the least common multiple of m_0 and m_∞, so that $e = m_0 m_\infty / d$. We pick a generator $\alpha \in \mathbb{F}_p^*$ of $\mathbb{F}_p^*/(\mathbb{F}_p^*)^e$. Then it is easily confirmed that

$$\langle \alpha^{de/m_0} \rangle = (\mathbb{F}_p^*)^d/(\mathbb{F}_p^*)^{m_0} \quad \text{and} \quad \langle \alpha^{de/m_\infty} \rangle = (\mathbb{F}_p^*)^d/(\mathbb{F}_p^*)^{m_\infty}$$

on noting that $(\mathbb{F}_p^*)^{m_0}$ and $(\mathbb{F}_p^*)^{m_\infty}$ are subgroups of $(\mathbb{F}_p^*)^d$. (Indeed, to check the first equality, for example, it suffices to confirm that α^{de/m_0} has order m_0/d in $\mathbb{F}_p^*$.) The group $(\mathbb{F}_p^*)^d/(\mathbb{F}_p^*)^{m_0}$ has order $N_0 = \gcd(m_0, p-1)$ and, likewise, $(\mathbb{F}_p^*)^d/(\mathbb{F}_p^*)^{m_\infty}$ has order $N_\infty = \gcd(m_\infty, p-1)$. It follows from this that any nonzero d-th power in $\mathbb{F}_p$ can be represented as $u^{m_0}\alpha^{edk/m_0}$ for some $k \in \mathbb{Z}/N_0\mathbb{Z}$, and such a representation is unique up to multiplication of u by one of the m_0-th roots of unity in $\mathbb{F}_p$, of which there are N_0.

Similarly, we can represent any nonzero d-th power in $\mathbb{F}_p$ as $u^{m_\infty}\alpha^{ed\ell/m_\infty}$ for some $\ell \in \mathbb{Z}/N_\infty\mathbb{Z}$ in exactly N_∞ ways.

We will use this to partition the counting function ν_j and remember to divide by $N_0 N_\infty$ when collecting the parts. Define

$$\lambda_j(p; A_0, A_1; k, \ell) = \#\{\boldsymbol{t} \in \mathbb{F}_p^2 : h_j(A_0 t_0^{m_0}\alpha^{edk/m_0}, A_1 t_1^{m_\infty}\alpha^{ed\ell/m_\infty}) = 0\}$$

for any $k \in \mathbb{Z}/N_0\mathbb{Z}$ and $\ell \in \mathbb{Z}/N_\infty\mathbb{Z}$. Let $\beta = \alpha^{-edk/m_0 - ed\ell/m_\infty}$. On multiplying through by $\beta^{\deg(h_j)}$ and recalling that h_j is homogeneous, we obtain

$$\begin{aligned}\lambda_j(p; A_0, A_1; k, \ell) &= \#\{\boldsymbol{t} \in \mathbb{F}_p^2 : h_j(A_0 t_0^{m_0}\alpha^{edk/m_0}\beta, A_1 t_1^{m_\infty}\alpha^{ed\ell/m_\infty}\beta) = 0\} \\ &= \#\{\boldsymbol{t} \in \mathbb{F}_p^2 : h_j(A_0 t_0^{m_0}\alpha^{-ed\ell/m_\infty}, A_1 t_1^{m_\infty}\alpha^{-edk/m_0}) = 0\}.\end{aligned}$$

But $ed/m_\infty = m_0$ and $ed/m_0 = m_\infty$. Hence a simple change of variables yields

$$\lambda_j(p; A_0, A_1; k, \ell) = \lambda_j(p; A_0, A_1; 0, 0). \tag{6-8}$$

Let $\nu_j^*(p; A_0, A_1)$ denote the contribution to $\nu_j(p; A_0, A_1)$ from $t_0t_1 \neq 0$, and also similarly for $\lambda_j^*(p; A_0, A_1; k, \ell)$. Then we may write

$$\begin{aligned}\nu_j(p; A_0, A_1) &= \nu_j^*(p; A_0, A_1) + O(1)\\ &= \frac{1}{N_0N_\infty}\sum_{k\in\mathbb{Z}/N_0\mathbb{Z}}\sum_{\ell\in\mathbb{Z}/N_\infty\mathbb{Z}} \lambda_j^*(p; A_0, A_1; k, \ell) + O(1)\\ &= \lambda_j^*(p; A_0, A_1; 0, 0) + O(1)\end{aligned}$$

by (6-8). Noting that $\lambda_j^*(p; A_0, A_1; 0, 0) = \lambda_j(p; A_0, A_1) + O(1)$, we have therefore shown that

$$\lambda_j(p; A_0, A_1) = \nu_j(p; A_0, A_1) + O(1).$$

At this point we recall the factorisation (6-2) together with the fact that $v_i = a_i s_i w_i$ for $i = 0, 1$. Hence, since $p \nmid A_0A_1$, a simple change of variables shows that

$$\nu_j(p; A_0, A_1) = \#\{\boldsymbol{t} \in \mathbb{F}_p^2 : h_j(v_0(b_0 s_0^{m_0/d} t_0)^d, v_1(b_1 s_1^{m_\infty/d} t_1)^d) = 0\} = \nu_j(p; \boldsymbol{v}),$$

from which the claim (6-7) follows. □

We will need to study the average size of $\bar{\omega}(p)$ as p varies. We break this into the following results.

Lemma 6.5. *We have*

$$\sum_{\substack{p\leqslant x\\ p\notin T_0}} \frac{1}{p} = (1 - \delta_{0,\mathbb{Q}}(\pi)) \log\log x + O(1)$$

and

$$\sum_{\substack{p\leqslant x\\ p\notin T_\infty}} \frac{1}{p} = (1 - \delta_{\infty,\mathbb{Q}}(\pi)) \log\log x + O(1)$$

in the notation of (1-5).

Proof. This is a straightforward consequence of the Chebotarev density theorem in the form presented in [Serre 2012, Theorem 3.4], for example. □

Our next result concerns the average behaviour of the function $\nu_j(p; \boldsymbol{v})$ in Lemma 6.4, as we average over primes $p \notin U_j$. This is more difficult and requires the use of notation introduced at the start of Section 2.2, which we recall here. For a number field $F/\mathbb{Q}$, let $\mathscr{P}_F$ denote the set of primes $p \in \mathbb{Z}$ that are unramified in F and for which there exists a prime ideal $\mathfrak{p} \mid p\mathfrak{o}_F$ of residue degree 1. For any positive integer $m \leqslant [F : \mathbb{Q}]$, we write $\mathscr{P}_{F,m}$ for the subset of $p \in \mathscr{P}_F$ for which there are precisely m prime ideals above p of residue degree 1.

For each $j \in \{1, \dots, r\}$, define the étale algebra

$$N_{E_j,d,v_1/v_0} = \mathbb{Q}[x]/(r_j(x)),$$

where $r_j(x) = h_j(x^d, v_1/v_0)$. As in (1-8), this has a factorisation into number fields

$$N_{E_j,d,v_1/v_0} = N^{(1)} \times \cdots \times N^{(s)},$$

where $N^{(k)} = N^{(k)}_{E_j,d,v}$ for $1 \leqslant k \leqslant s$, where the dependency of s on j is suppressed for legibility.

Lemma 6.6. *For each $j \in \{1, \dots, r\}$, we have*

$$\sum_{\substack{p \leqslant x \\ p \notin U_j}} \frac{\nu_j(p; \boldsymbol{v})}{p^2} = \sum_{k=1}^{s} (1 - \delta_{D, N^{(k)}}(\pi)) \log\log x + O(1 + \omega(v_0 v_1))$$

in the notation of (1-5), *where $\omega(n)$ denotes the number of distinct prime factors of $n \in \mathbb{Z}$.*

Proof. We have

$$\sum_{\substack{p \leqslant x \\ p \notin U_j}} \frac{\nu_j(p; \boldsymbol{v})}{p^2} = \sum_{\substack{p \leqslant x \\ p \notin U_j \\ p \nmid v_0 v_1}} \frac{\nu_j(p; \boldsymbol{v})}{p^2} + \sum_{\substack{p \leqslant x \\ p \notin U_j \\ p \mid v_0 v_1}} \frac{\nu_j(p; \boldsymbol{v})}{p^2}.$$

Since $\gcd(v_0, v_1) = 1$, the second term is seen to be

$$\ll \sum_{\substack{p \leqslant x \\ p \mid v_0 v_1}} \frac{1}{p} \ll \omega(v_0 v_1).$$

Next, we see that

$$\sum_{\substack{p \leqslant x \\ p \notin U_j \\ p \nmid v_0 v_1}} \frac{\nu_j(p; \boldsymbol{v})}{p^2} = \sum_{\substack{p \leqslant x \\ p \notin U_j \\ p \nmid v_0 v_1}} \frac{\#\{t \in \mathbb{F}_p : h_j(t^d, v_1/v_0) = 0\}}{p} + O(1).$$

Write $r_j(t) = h_j(t^d, v_1/v_0)$, and let $r_j(t) = r_j^{(1)}(t) \cdots r_j^{(s)}(t)$ be its factorisation into irreducible factors over $\mathbb{Q}$. Then $N^{(k)}$ is the number field $\mathbb{Q}[t]/(r_j^{(k)})$ for $1 \leqslant k \leqslant s$. We have

$$\sum_{\substack{p \leqslant x \\ p \notin U_j \\ p \nmid v_0 v_1}} \frac{\nu_j(p; \boldsymbol{v})}{p^2} = \sum_{k=1}^{s} \sum_{\substack{p \leqslant x \\ p \notin U_j \\ p \nmid v_0 v_1}} \frac{\#\{t \in \mathbb{F}_p : r_j^{(k)}(t) = 0\}}{p} + O(1).$$

To begin with, it follows from the prime ideal theorem that

$$\sum_{p \leqslant x} \frac{\#\{t \in \mathbb{F}_p : r_j^{(k)}(t) = 0\}}{p} = \log\log x + O(1 + \omega(v_0 v_1)).$$

Next, we note that $p \in U_j$ if and only if Frob_p fixes a component of S_{E_j}. Let $\mathscr{F}_j$ denote the set of fields of definition of the elements of S_{E_j}. Then, for any $p \notin S$, the condition $p \in U_j$ is equivalent to the condition $p \in \mathscr{P}_{\mathscr{F}_j} := \bigcup_{F \in \mathscr{F}_j} \mathscr{P}_F$. Likewise, for any positive integer $m \leqslant [N^{(k)} : \mathbb{Q}]$, we will have $\#\{t \in \mathbb{F}_p : r_j^{(k)}(t) = 0\} = m$ if and only if $p \in \mathscr{P}_{N^{(k)}, m}$. Hence

$$\sum_{\substack{p \leqslant x \\ p \notin U_j}} \frac{\nu_j(p; \boldsymbol{v})}{p^2} = \sum_{k=1}^{s} \left(\log\log x - \sum_{m=1}^{[N^{(k)}:\mathbb{Q}]} m \sum_{\substack{p \leqslant x \\ p \in \mathscr{P}_{N^{(k)}, m} \cap \mathscr{P}_{\mathscr{F}_j}}} \frac{1}{p} \right) + O(1 + \omega(v_0 v_1)).$$

The remaining sum over primes is susceptible to a further application of the Chebotarev density theorem. Once coupled with Theorem 2.3 and (2-1), this leads to the statement of the lemma. ☐

We may combine the previous two results to produce a lower bound for the quantity $L(Q)$ in Lemma 6.3, with the choice of $\bar{\omega}(p)$ from Lemma 6.4.

Lemma 6.7. *For any $\varepsilon > 0$, we have the lower bound*

$$L(Q) \gg_{\varepsilon} \frac{(\log Q)^{\Theta_v(\pi)}}{|A_0A_1|^{\varepsilon}},$$

where $\Theta_v(\pi)$ is given by (1-9).

Proof. Since $1-\bar{\omega}(p) \leqslant 1$, we have

$$L(Q) \geqslant \sum_{q\leqslant Q} \mu^2(q) \prod_{p\mid q} \bar{\omega}(p).$$

There are many results in the literature concerning mean values of nonnegative arithmetic functions. However, we can get by with the relatively crude lower bound found in [Friedlander and Iwaniec 2010, Theorem A.3], which is based on an application of Rankin's trick. Let $\gamma : \mathbb{N} \to \mathbb{R}_{\geqslant 0}$ be a multiplicative arithmetic function that is supported on square-free integers and which satisfies

$$\sum_{y<p\leqslant x} \gamma(p)\log p \leqslant a\log(x/y)+b \tag{6-9}$$

for any $x > y > 2$ for appropriate constants $a, b > 0$. Then it follows from [Friedlander and Iwaniec 2010, Theorem A.3] that

$$\sum_{n\leqslant x} \gamma(n) \gg \prod_{p\leqslant x}(1+\gamma(p)), \tag{6-10}$$

where the implied constant is allowed to depend on a and b. We seek to apply this with

$$\gamma(n) = \mu^2(n)\prod_{p\mid n}\bar{\omega}(p).$$

It is clear from Lemma 6.4 that $\bar{\omega}(p) = O(1/p)$. Hence

$$\sum_{y<p\leqslant x} \gamma(p)\log p \ll 1+\sum_{y<p\leqslant x}\frac{\log p}{p} \ll 1+\log(x/y)$$

uniformly in v_0 and v_1. Hence (6-9) holds for $a, b = O(1)$, and it follows from (6-10) that

$$L(Q) \gg \prod_{p\leqslant Q}(1+\bar{\omega}(p))$$

for an absolute implied constant. On appealing once more to Lemma 6.4, we find that

$$\log\biggl(\prod_{p\leqslant Q}(1+\bar{\omega}(p))\biggr) = \sum_{\substack{p\leqslant Q\\ p\notin T_0}}\frac{1}{p}+\sum_{\substack{p\leqslant Q\\ p\notin T_\infty}}\frac{1}{p}+\sum_{j=1}^{r}\sum_{\substack{p\leqslant Q\\ p\notin U_j}}\frac{\nu_j(p;\boldsymbol{v})}{p^2}+O(1+\omega(A_0A_1)).$$

These sums are estimated using Lemmas 6.5 and 6.6, leading to the conclusion that

$$\log\biggl(\prod_{p\leqslant Q}(1+\bar{\omega}(p))\biggr) = \widetilde{\Theta}(\pi, v_1/v_0)\log\log Q+O(1+\omega(A_0A_1)),$$

where

$$\widetilde{\Theta}(\pi, v_1/v_0) = 2 - \delta_{0,\mathbb{Q}}(\pi) - \delta_{\infty,\mathbb{Q}}(\pi) + \sum_{j=1}^{r}\sum_{k=1}^{s}(1 - \delta_{E_j, N^{(k)}_{E_j,d,v_1/v_0}}(\pi))$$

in the notation of (1-5). Clearly $\widetilde{\Theta}(\pi, v_1/v_0) = \Theta_v(\pi)$, the latter being defined in (1-9). Hence the statement of the lemma follows on exponentiating and using the fact that $\omega(n) \ll \log|n|/\log\log|n|$ for any nonzero $n \in \mathbb{Z}$. □

6.4. ***Completion of the proof of Proposition 6.1.*** We begin by focussing on the estimation of the quantity $L(R_0, R_1)$ that was defined in (6-4). In view of (6-2), we see that

$$A_0 = v_0(b_0 s_0^{m_0/d} u_0)^d \quad \text{and} \quad A_1 = v_1(b_1 s_1^{m_\infty/d} u_1)^d.$$

Recall that $s_i \mid v_i$ for $i = 0, 1$. Taking $Q = B^\varepsilon$, we note that

$$R_0^{m_0} = \frac{B}{|A_0|} \geqslant \frac{B}{|v_0(s_0 b_0 u_0)^{m_0}|} \geqslant B^{1-(1+3m_0)\varepsilon} \geqslant Q^{4m_0},$$

provided that $\varepsilon \leqslant 1/(1+7m_0)$. Similarly, we can assume that $R_1 \geqslant Q^4$ if $\varepsilon > 0$ is chosen to be sufficiently small. Hence, with these choices, we have

$$(R_0 + Q^4)(R_1 + Q^4) \ll R_0 R_1 \ll \frac{B^{1/m_1+1/m_\infty}}{|A_0|^{1/m_0}|A_1|^{1/m_\infty}}.$$

We may now apply Lemma 6.7 in Lemma 6.3 to deduce that

$$L(R_0, R_1) \ll_\varepsilon \frac{B^{1/m_0+1/m_\infty}}{|A_0|^{1/m_0}|A_1|^{1/m_\infty}} \cdot \frac{|A_0 A_1|^\varepsilon}{(\log B)^{\Theta_v(\pi)}}.$$

Substituting into (6-6), recalling (6-5) and summing over b_0, b_1, u_0, u_1, the statement of Proposition 6.1 easily follows.

7. Examples: lower bounds and asymptotics

Let $\pi : X \to \mathbb{P}^1$ be a standard fibration. It is clear from the constructions in Section 5 that we are only able to interpret local solubility conditions outside a finite set of places of S which depends on π. This set S should contain a set of places for which Corollary 5.8 holds, and such a set can be determined explicitly. With more work one might be able to incorporate local solubility at places in S, but this should not change the order of growth, which is the main interest in this paper. Accordingly, for any finite set S of primes, we introduce the counting function

$$N_{\mathrm{loc},S}(\pi, B) = \#\{x \in \mathbb{P}^1(\mathbb{Q}) \cap \pi(X(\mathbf{A}_\mathbb{Q}^S)) : H(x) \leqslant B\},$$

where H is the usual height function on $\mathbb{P}^1(\mathbb{Q})$ and $\mathbf{A}_\mathbb{Q}^S$ is the set of adèles away from S. We clearly have $N_{\mathrm{loc},S}(\pi, B) \geqslant N_{\mathrm{loc}}(\pi, B)$, and we expect these two counting functions to have the same order of magnitude.

We shall prove several results about Halphen surfaces. Let $m > 1$ be an integer. A *Halphen pencil* is a geometrically irreducible pencil of plane curves of degree $3m$ with multiplicity m at nine base points $P_1, \dots, P_9$. We let X be the *Halphen surface of order m* obtained by blowing up $\mathbb{P}^2$ at these nine points, as introduced in [Halphen 1882]. We shall assume that $P_1, \dots, P_9$ are globally defined over $\mathbb{Q}$, so that X is a smooth, proper, geometrically integral surface defined over $\mathbb{Q}$. In fact, X is a rational elliptic surface and we obtain a standard morphism $\pi : X \to \mathbb{P}^1$ such that there exists a unique fibre of multiplicity m. In particular, π does not admit a section.

7.1. ***Lower bounds.*** In this section we establish an array of lower bounds for $N_{\mathrm{loc},S}(\pi, B)$. The following result demonstrates that Conjecture 1.5 would be false with the exponent $\Delta(\pi)$ and that it is indeed sometimes necessary to take a smaller exponent.

Theorem 7.1. *Let $\pi : X \to \mathbb{P}^1$ be a standard fibration. Assume it only has nonsplit fibres above* 0, 1 *and* ∞, *comprising geometrically irreducible double fibres over* 0 *and* ∞, *and a nonsplit fibre of multiplicity* 1 *above* 1 *that is split by a quadratic extension. Then there is a finite set of places S such that*

$$B \ll N_{\mathrm{loc},S}(\pi, B) \ll B.$$

Proof. Suppose that $F = \mathbb{Q}(\sqrt{d})$ is the quadratic extension that splits the fibre above 1 for square-free $d \in \mathbb{Z}$. Then it is clear that

$$0 \leqslant \Theta(\pi) = \min_{K/\mathbb{Q}\ \text{quadratic}} (1 - \delta_{1,K}(\pi)) \leqslant 1 - \delta_{1,F}(\pi) = 0.$$

Hence the upper bound is a direct consequence of Theorem 1.3.

For the lower bound, we compose the exact counting problem using Corollary 5.8. Thus there exists a finite set of places S, containing the prime divisors of $2d$, such that

$$N_{\mathrm{loc},S}(\pi, B) = \frac{1}{2}\#\left\{(a, b) \in \mathbb{Z}^2_{\mathrm{prim}} : \begin{array}{r} |a|, |b| \leqslant B,\ p \notin S \Rightarrow [2 \mid v_p(a) \text{ and } 2 \mid v_p(b)], \\ [p \notin S \text{ and } p \mid a - b] \Rightarrow p \in \mathscr{P}_F \end{array}\right\}.$$

The lower bound is provided by taking pairs (a, b) of the form (u^2, dv^2). □

In this result we have $2 - \deg \partial_\pi = 1$, so that the exponent of B matches the predicted exponent of B in Conjectures 1.1 and 1.5. We also have

$$\delta_{0,\mathbb{Q}}(\pi) = \delta_{\infty,\mathbb{Q}}(\pi) = 1 \quad \text{and} \quad \delta_{1,\mathbb{Q}}(\pi) = \tfrac{1}{2},$$

so that $\Delta(\pi) = \frac{1}{2}$. However, we saw in the proof that $\Theta(\pi) = 0$. Thus Theorem 7.1 is in agreement with Conjecture 1.5.

Let us describe what is going on geometrically. Consider the finite étale orbifold μ_2-cover $\theta_v : \mathbb{P}^1 \to \mathbb{P}^1$ given by $(x : y) \mapsto (x^2 : vy^2)$ and the pullback fibrations $\pi_v : X_v \to \mathbb{P}^1$ obtained from normalisation of the pullback of π along θ_v. By Proposition 4.7, we see that the two double fibres of π pull back to components of multiplicity 1 on π_v. Also, all fibres which do not lie over 1 in the composition $X_v \xrightarrow{\pi_v} \mathbb{P}^1 \xrightarrow{\theta_v} \mathbb{P}^1$ are split. We proceed by studying the fibres over 1.

First we study the fibre of 1 in θ_v. For $v \in \mathbb{Q}^\times/\mathbb{Q}^{\times,2}$, we have $\theta^{-1}(1) = \operatorname{Spec} A$, where A is the degree-2 étale algebra $\mathbb{Q}(\sqrt{v})$ if $v \notin \mathbb{Q}^{\times,2}$ and $\mathbb{Q} \times \mathbb{Q}$ for $v \in \mathbb{Q}^{\times,2}$. This gives

$$\Delta(\pi_v) = \sum_{D' \mid D} (1 - \delta_{D'}(\pi_v)) = \begin{cases} 0 & \text{if } v \equiv d, \\ \frac{1}{2} + \frac{1}{2} = 1 & \text{if } v \equiv 1, \\ \frac{1}{2} & \text{otherwise,} \end{cases}$$

where the sum ranges over all points D' lying above $D = 1 \in (\mathbb{P}^1)^{(1)}$. In the first case, the fibre over 1 (which is split by F) pulls back to an F-point and becomes split. In the second case, the fibre pulls back to two $\mathbb{Q}$-points. In the last case, the fibre is irreducible and its residue field is linearly disjoint from the splitting field, and we obtain $\Delta(\pi_v) = \Delta(\pi)$, in general.

Theorem 7.1 indicates that the main contribution to the point count comes from the single cover π_d. If we were to exclude the thin set of points coming from this cover, we are left with infinitely many covers π_v, with $\Delta(\pi_v) = \Delta(\pi)$ for $v \neq 1$. Proposition 4.7 (b) implies that the covers have no multiple fibres, since it gives

$$m_{P'} = \frac{m_P}{\gcd(m_P, e(P'/P))} = \frac{2}{\gcd(2, 2)} = 1$$

for each $P' \mid 0, \infty$. Hence, in the light of the original Loughran–Smeets conjecture [2016, Conjecture 1.6], we expect the remaining covers to contribute order $B/\sqrt{\log B}$ to the counting function, apart from the cover corresponding to 1, which should contribute order $B/\log B$.

Our second lower bound deals with the case of precisely two nonsplit fibres and is consistent with Conjecture 1.5 since $\deg \partial_\pi = 2 - 1/m_0 - 1/m_\infty$.

Theorem 7.2. *Let $\pi : X \to \mathbb{P}^1$ be a standard fibration for which the only nonsplit fibres lie over* 0 *and* ∞. *Then there is a finite set of places S such that*

$$N_{\mathrm{loc},S}(\pi, B) \gg \frac{B^{1/m_0 + 1/m_\infty}}{(\log B)^{\Delta(\pi)}}.$$

Proof. We begin by using Corollary 5.8 to give explicit conditions for local solubility away from S after passing to an sncd model $X' \to \mathbb{P}^1$. This leads to the conclusion that $N_{\mathrm{loc},S}(\pi, B)$ is equal to the number of $x = (x_0 : x_1) \in \mathbb{P}^1(\mathbb{Q})$ with $H(x) \leqslant B$ such that, for each $i \in \{0, 1\}$ and every $p \notin S$, Frob_p fixes a collection of intersecting components Z_j of X'_{D_i} such that $v_p(x_i) \in \langle m(Z_j) \rangle_{\mathbb{N}}$, where $D_i = V(x_i)$. The following is clearly a sufficient condition for the fibre over x to have a $\mathbb{Q}_p$-point: for all i, the Frobenius Frob_p fixes a component of Z of minimal multiplicity in X'_{D_i} and $m(Z) \mid v_p(x_i)$. The density ∂_i of rational primes p for which Frob_p fixes an element of S_{D_i} is equal to $\delta_{D_i}(\pi) = \delta_{D_i,\kappa(D_i)}(\pi)$ in the notation of (1-5). Hence the statement of the theorem now follows from Proposition 3.1 and (1-7). □

7.2. ***Halphen surfaces with one nonsplit fibre.*** Generically, a Halphen surface has no other nonsplit fibres apart from the multiple one. Even in these cases the counting problem still depends on the Galois action on the components of the multiple fibres and how these components intersect. We record some results which illustrate this phenomenon; it will be convenient to keep in mind the notation (3-1).

We begin with the following result, which agrees with Conjecture 1.5 since $\deg \partial_\pi = 1 - 1/m$ and $\Delta(\pi) = 0$.

Theorem 7.3. *Let $X \to \mathbb{P}^1$ be a Halphen surface with a single nonsplit fibre over* 0, *that is the fibre of multiplicity m. Suppose that this fibre has a geometric component fixed by* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. *Then there exists a finite set S such that*

$$N_{\mathrm{loc},S}(\pi, B) \sim c_{\pi,S} B^{1+1/m},$$

where

$$c_{\pi,S} = \frac{2c_S\left(1+\frac{1}{m}\right)}{\zeta(2)c_S\left(\frac{1}{m}\right)} \prod_{p\in S}\left(1+\frac{1}{p}\right)^{-1}.$$

Proof. By Corollary 5.8, we see that there is a finite set of places S such that

$$N_{\mathrm{loc},S}(\pi, B) = \tfrac{1}{2}\#\{(a, b) \in \mathbb{Z}^2_{\mathrm{prim}} : |a|, |b| \leqslant B,\ p \notin S \Rightarrow m \mid v_p(a)\}.$$

We may apply Proposition 3.1 with $m_0 = m$ and $m_1 = 1$, and with $\mathscr{P}_0 = \mathscr{P}_1$ equal to the full set of rational primes. In particular $\partial_0 = \partial_1 = 1$, and it follows that $N_{\mathrm{loc},S}(\pi, B) \sim c_{\pi,S} B^{1+1/m}$ as $B \to \infty$, where

$$c_{\pi,S} = \frac{2c_S\left(1+\frac{1}{m}\right)}{c_S(1)c_S\left(\frac{1}{m}\right)} \prod_{p\notin S}\left(1-\frac{1}{p^2}\right) \prod_{p\in S}\left(1-\frac{1}{p}\right)^2$$

in the notation of (3-1). The statement easily follows on simplifying the expression for the constant. □

The following two results agree with Conjecture 1.5 since in both cases we have $\deg \partial_\pi = 1 - 1/m$ and $\Delta(\pi) = \frac{2}{3}$. Moreover, in these two examples, we have multiple fibres which do not have a geometrically integral component. This demonstrates the need to define (1-5) in terms of S_D for each divisor D, which allows us to work with the Galois action on the components of a fibre of minimum multiplicity.

Theorem 7.4. *Let $X \to \mathbb{P}^1$ be a Halphen surface with a single nonsplit fibre over* 0, *that is the fibre of multiplicity m. Suppose that this fibre consists of three conjugate lines split by a cubic Galois extension $K/\mathbb{Q}$ that do* ***not*** *all meet in a point. Then there exists a finite set S such that*

$$N_{\mathrm{loc},S}(\pi, B) \sim c_{\pi,S} \frac{B^{1+1/m}}{(\log B)^{2/3}},$$

where

$$c_{\pi,S} = \frac{2m^{2/3}c_S(1)^{1/3}c_S\left(1+\frac{1}{m}\right)}{\Gamma\left(\frac{1}{3}\right)c_S\left(\frac{1}{m}\right)} \prod_{\substack{p\in\mathscr{P}_K\\ p\notin S}}\left(1+\frac{1}{p}\right)\left(1-\frac{1}{p}\right)^{1/3} \prod_{\substack{p\notin\mathscr{P}_K\\ p\notin S}}\left(1-\frac{1}{p}\right)^{1/3}.$$

Proof. Suppose that the three conjugate lines are split by the cubic Galois extension $K/\mathbb{Q}$. By Corollary 5.8, we see that there is a finite set of places S such that $N_{\mathrm{loc},S}(\pi, B)$ is equal to

$$\tfrac{1}{2}\#\{(a, b) \in \mathbb{Z}^2_{\mathrm{prim}} : |a|, |b| \leqslant B,\ [p \notin S \text{ and } p \mid a] \Rightarrow [m \mid v_p(a) \text{ and } p \in \mathscr{P}_K]\},$$

where $\mathscr{P}_K$ is the set of rational primes p that are unramified in K and split completely. We may apply Proposition 3.1 with $m_0 = m$ and $m_1 = 1$, and with $\mathscr{P}_0 = \mathscr{P}_K$ and $\mathscr{P}_1$ equal to the full set of rational primes. In particular, $\partial_0 = \frac{1}{3}$ and $\partial_1 = 1$. It follows that

$$N_{\mathrm{loc},S}(\pi, B) \sim c_{\pi,S} \frac{B^{1+1/m}}{(\log B)^{2/3}},$$

where

$$c_{\pi,S} = \frac{2m^{2/3}}{\Gamma\left(\frac{1}{3}\right)} \cdot \frac{c_S\left(1+\frac{1}{m}\right)}{c_S\left(\frac{1}{m}\right)} \prod_{\substack{p\in\mathscr{P}_K \\ p\notin S}} \left(1-\frac{1}{p^2}\right) \prod_{p\in\mathscr{P}_K\cap S} \left(1-\frac{1}{p}\right) \prod_{p\in\mathscr{P}_K} \left(1-\frac{1}{p}\right)^{-2/3} \prod_{p\notin\mathscr{P}_K} \left(1-\frac{1}{p}\right)^{1/3}.$$

The statement of the proposition follows on simplifying this expression. □

The next result agrees with Conjecture 1.5 since

$$\deg \partial_\pi = 1 - \frac{1}{m} \quad \text{and} \quad \Delta(\pi) = \frac{2}{3}.$$

Theorem 7.5. *Let $X \to \mathbb{P}^1$ be a Halphen surface with a single nonsplit fibre over* 0, *that is the fibre of multiplicity m. Suppose that this fibre consists of three conjugate lines split by a cubic Galois extension $K/\mathbb{Q}$ that* ***do*** *meet in a point. Then there exists a finite set S such that*

$$\frac{B^{1+1/m}}{(\log B)^{2/3}} \ll N_{\mathrm{loc},S}(\pi, B) \ll \frac{B^{1+1/m}}{(\log B)^{2/3}}.$$

Proof. The upper bound follows from Theorem 1.2. The lower bound was proven in Theorem 7.2. □

Theorem 7.5 illustrates the need for the nonsplit fibres to be sncd; the counting problem for this setting will be

$$[p \notin S \text{ and } p \mid a] \Rightarrow [(3m \mid v_p(a)) \text{ or } (m \mid v_p(a) \text{ and } p \in \mathscr{P}_K)].$$

The condition $3m \mid v_p(a)$ comes from a Galois fixed component of multiplicity $3m$ on the multiple fibre of the sncd-model of X. However, no such component exists on the multiple fibre of X itself.

7.3. ***Halphen surfaces with two nonsplit fibres.*** In practice, it can be difficult to construct Halphen surfaces with more than one nonsplit fibre. We present two such examples, both of which verify Conjecture 1.5.

Theorem 7.6. *There exists a Halphen surface $X \to \mathbb{P}^1$ of degree* 2 *with two nonsplit fibres: the multiple fibre is geometrically irreducible and has multiplicity* 2, *and the other is an sncd divisor of Kodaira classification I_6 split by a cubic Galois extension $K/\mathbb{Q}$. Moreover, there exists a finite set of places S, and an explicit constant $c_{\pi,S} > 0$ such that*

$$N_{\mathrm{loc},S}(\pi, B) \sim c_{\pi,S} \frac{B^{1+1/2}}{(\log B)^{2/3}}.$$

Proof. Let us first fix the cyclic cubic number field $K/\mathbb{Q}$. Now choose two sets of three conjugate points $P_i, Q_i \in \mathbb{P}^2(K)$, indexed by $i \in \mathbb{Z}/3\mathbb{Z}$. We let R_i be the intersecting point of the lines $P_{i+1}P_{i+2}$ and $Q_{i+1}Q_{i+2}$. For generic choices of P_i and Q_i, the R_i are well-defined and there is a unique smooth cubic through the nine points P_i, Q_i and R_i.

We will consider $X = \mathrm{Bl}_{P_i,Q_i,R_i}\mathbb{P}^2$. The two nonsplit fibres of X come from the double cubic passing through these nine points, and the sextic curve which is geometrically the union of the six lines $P_{i+1}P_{i+2}$ and $Q_{i+1}Q_{i+2}$. Under blowup the first curve turns into a geometrically integral fibre of multiplicity 2, and the other into six lines meeting in a cycle. The three lines P_1P_2, P_2P_3 and P_3P_1 are permuted by $\mathrm{Gal}(K/\mathbb{Q})$ and no longer meet on X. For a generic choice of P_i and Q_i, there will be no other nonsplit fibres.

Let us assume the multiple fibre lies above 0 and the other nonsplit fibre over ∞. The fibres of $X \to \mathbb{P}^1$ are all sncd, so we can directly compose the counting problem to find that

$$N_{\mathrm{loc},S}(\pi, B) = \tfrac{1}{2}\#\{(a,b) \in \mathbb{Z}^2_{\mathrm{prim}} : |a|, |b| \leqslant B,\ p \notin S \Rightarrow 2 \mid v_p(a),\ [p \notin S \text{ and } p \mid b] \Rightarrow p \in \mathscr{P}_K\}.$$

Such a counting problem is dealt with by Proposition 3.1. □

Theorem 7.7. *There exists a Halphen surface $X \to \mathbb{P}^1$ of degree 3 with two nonsplit fibres: the multiple fibre is geometrically irreducible and has multiplicity 3, and the other is a non-sncd divisor of Kodaira classification I_3 split by a cubic Galois extension $K/\mathbb{Q}$. Moreover, there exists a finite set of places S such that*

$$\frac{B^{1+1/3}}{(\log B)^{2/3}} \ll N_{\mathrm{loc},S}(\pi, B) \ll \frac{B^{1+1/3}}{(\log B)^{2/3}}.$$

We will return to this surface in Section 7.4 to create another interesting example. There we will assume that the multiple fibre lies over 0 and the remaining nonsplit fibre lies over ∞.

Proof of Theorem 7.7. Let $E/\mathbb{Q}$ be an elliptic curve with $E(\mathbb{Q})_{\mathrm{tors}} = \mathbb{Z}/9\mathbb{Z}$. Let $K/\mathbb{Q}$ be a cyclic cubic number field $K/\mathbb{Q}$ such that $\operatorname{rank} E(\mathbb{Q}) < \operatorname{rank} E(K)$. We will fix

(i) a generator $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$,

(ii) a generator $A \in E(\mathbb{Q})_{\mathrm{tors}}$,

(iii) $B \in E(K) \setminus E(\mathbb{Q})$ such that $B + \sigma(B) + \sigma^2(B) = O \in E(\mathbb{Q})$, and any

(iv) $C \in E(K) \setminus E(\mathbb{Q})$.

With this notation in mind, consider the nine points

$$\begin{aligned} P_i &= \sigma^i(C), \\ Q_i &= \sigma^i(-2C + B + A), \\ R_i &= \sigma^i(C - 3A). \end{aligned}$$

For general choices, we find that $\mathrm{Bl}_{P_i,Q_i,R_i}\mathbb{P}^2$ is a Halphen surface of degree 3. In particular, there is a smooth cubic through the nine points, which becomes the geometrically irreducible triple fibre on X.

Moreover, we have

$$\sum_i (P_i + Q_i + R_i) - P_j + R_j = O,$$

so that there is a cubic curve which passes through all nine points except P_j and has a singularity at R_j. The union of these three curves becomes the nonsplit I_3-fibre split by K.

For the lower bound we may apply Theorem 7.2, and the upper bound follows from Theorem 1.2. □

7.4. ***A nonsplit fibre over a point of higher degree.*** Our final result concerns a surface of Halphen type with a fibration over $\mathbb{P}^1$ that has one multiple fibre and a nonsplit fibre over a degree-2 point. Our local solubility criteria do not apply to this case in general, but we are nonetheless able to deduce explicit criteria.

Consider the Halphen surface $\pi : X \to \mathbb{P}^1$ from Theorem 7.7 with $m = 3$, with a multiple fibre over 0 and a nonsplit fibre over ∞ split by a Galois cubic extension $K/\mathbb{Q}$. Let $\pi' : X' \to \mathbb{P}^1$ be the normalisation of the pullback of π along the morphism $\theta : \mathbb{P}^1 \to \mathbb{P}^1$ given by $[u : v] \mapsto [u^2 : u^2 + v^2]$. We claim that the surface X' has a unique multiple fibre over $u = 0$, whose multiplicity is 3, and that the only other nonsplit fibre lies over the degree-2 point $u^2 + v^2 = 0$ and is split by K. To see this we note that the fibres of the pullback of X are precisely the fibres of X', and normalisation only changes the fibres over 0 and ∞. The multiplicities of the new fibres can then be computed using Proposition 4.3. Note that $\partial_{\pi'} = \frac{2}{3}[0]$ and $\Delta(\pi') = 1 - \delta_{u^2+v^2}(\pi') = \frac{2}{3}$. We shall now prove the following result, which is easily seen to agree with the prediction in Conjecture 1.5.

Theorem 7.8. *For the surface $\pi' : X' \to \mathbb{P}^1$ as above, there exists a finite set S such that*

$$\frac{B^{4/3}}{(\log B)^{2/3}} \ll N_{\mathrm{loc},S}(\pi', B) \ll \frac{B^{4/3}}{(\log B)^{2/3}}.$$

Proof. The upper bound follows directly from Theorem 1.2. To prove the lower bound, we note that, for all but finitely many points $x \in \mathbb{P}^1(\mathbb{Q})$, the fibre of $X' \to \mathbb{P}^1$ is isomorphic to the fibre of $X \to \mathbb{P}^1$ over $\theta(x) \in \mathbb{P}^1(\mathbb{Q})$. Hence we can apply the criterion in Corollary 5.8 to determine local solubility for X. Noting that $v_p(u^2)$ is divisible by 3 precisely if this is true for $v_p(u)$, we find that $N_{\mathrm{loc},S}(\pi', B)$ is

$$\frac{1}{2}\#\left\{(u, v) \in \mathbb{Z}^2_{\mathrm{prim}} : \begin{array}{r} |u|, |v| \leqslant B,\ p \notin S \Rightarrow 3 \mid v_p(u), \\ [p \notin S \text{ and } p \mid u^2 + v^2] \Rightarrow p \in \mathscr{P}_K \end{array}\right\} + O(1).$$

On restricting to positive coprime u and v and demanding that u is a cube, we arrive at the lower bound

$$N_{\mathrm{loc},S}(\pi', B) \geqslant \tfrac{1}{2} M(B) + O(1),$$

where

$$M(B) = \#\{(u, v) \in \mathbb{Z}^2_{\mathrm{prim}} : 0 \leqslant u^3, v \leqslant B,\ p \mid u^6 + v^2 \Rightarrow p \in \mathscr{P}_K\}.$$

Note that $u^3, v \leqslant B$ whenever $u^6 + v^2 \leqslant B^2$. Hence

$$M(B) \geqslant \#\{(u, v) \in \mathbb{Z}^2_{\geqslant 0} : \gcd(u, v) = 1,\ u^6 + v^2 \leqslant B^2,\ p \mid u^6 + v^2 \Rightarrow p \in \mathscr{P}_K\}.$$

The right-hand side is exactly the quantity estimated via the β-sieve in [Friedlander and Iwaniec 2010, Theorem 11.31], with the outcome that

$$M(B) \gg \left(\frac{B^2}{\log(B^2)}\right)^{2/3}.$$

The statement of the theorem now follows. □

Acknowledgements

We are very grateful to Tim Santens for useful conversations and to the anonymous referees for numerous pertinent remarks. While working on this paper, Browning was supported by a FWF grant (DOI 10.55776/P32428), Lyczak was supported by UKRI MR/V021362/1, and Smeets was supported by grant G0B1721N of the Fund for Scientific Research – Flanders.

References

[Arango-Piñeros et al. 2022] S. Arango-Piñeros, D. Keliher, and C. Keyes, "Mertens' theorem for Chebotarev sets", *Int. J. Number Theory* **18**:8 (2022), 1823–1842. MR

[Browning and Van Valckenborgh 2012] T. D. Browning and K. Van Valckenborgh, "Sums of three squareful numbers", *Exp. Math.* **21**:2 (2012), 204–211. MR

[Browning et al. 2023] T. Browning, J. Lyczak, and R. Sarapin, "Local solubility for a family of quadrics over a split quadric surface", *Involve* **16**:2 (2023), 331–342. MR

[Campana 2004] F. Campana, "Orbifolds, special varieties and classification theory", *Ann. Inst. Fourier* (*Grenoble*) **54**:3 (2004), 499–630. MR

[Campana 2005] F. Campana, "Fibres multiples sur les surfaces: aspects geométriques, hyperboliques et arithmétiques", *Manuscripta Math.* **117**:4 (2005), 429–461. MR

[Campana 2011] F. Campana, "Orbifoldes géométriques spéciales et classification biméromorphe des variétés kählériennes compactes", *J. Inst. Math. Jussieu* **10**:4 (2011), 809–934. MR

[Colliot-Thélène et al. 1997] J.-L. Colliot-Thélène, A. N. Skorobogatov, and P. Swinnerton-Dyer, "Double fibres and double covers: paucity of rational points", *Acta Arith.* **79**:2 (1997), 113–135. MR

[Friedlander and Iwaniec 2010] J. Friedlander and H. Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications **57**, Amer. Math. Soc., Providence, RI, 2010. MR

[Graber et al. 2003] T. Graber, J. Harris, and J. Starr, "Families of rationally connected varieties", *J. Amer. Math. Soc.* **16**:1 (2003), 57–67. MR

[Halphen 1882] G. Halphen, "Sur les courbes planes du sixième degré à neuf points doubles", *Bull. Soc. Math. France* **10** (1882), 162–172. MR

[Kowalski 2008] E. Kowalski, *The large sieve and its applications*, Cambridge Tracts in Mathematics **175**, Cambridge University Press, 2008. MR

[Loughran 2018] D. Loughran, "The number of varieties in a family which contain a rational point", *J. Eur. Math. Soc.* (*JEMS*) **20**:10 (2018), 2539–2588. MR

[Loughran and Matthiesen 2024] D. Loughran and L. Matthiesen, "Frobenian multiplicative functions and rational points in fibrations", *J. Eur. Math. Soc.* (*JEMS*) **26**:12 (2024), 4779–4830. MR

[Loughran and Smeets 2016] D. Loughran and A. Smeets, "Fibrations with few rational points", *Geom. Funct. Anal.* **26**:5 (2016), 1449–1482. MR

[Loughran et al. 2020] D. Loughran, A. N. Skorobogatov, and A. Smeets, "Pseudo-split fibers and arithmetic surjectivity", *Ann. Sci. Éc. Norm. Supér.* (4) **53**:4 (2020), 1037–1070. MR

[Pieropan and Schindler 2024] M. Pieropan and D. Schindler, "Hyperbola method on toric varieties", *J. Éc. polytech. Math.* **11** (2024), 107–157. MR

[Pieropan et al. 2021] M. Pieropan, A. Smeets, S. Tanimoto, and A. Várilly-Alvarado, "Campana points of bounded height on vector group compactifications", *Proc. Lond. Math. Soc.* (3) **123**:1 (2021), 57–101. MR

[Serre 1990] J.-P. Serre, "Spécialisation des éléments de $\mathrm{Br}_2(\mathbb{Q}(T_1, \dots, T_n))$", *C. R. Acad. Sci. Paris Sér. I Math.* **311**:7 (1990), 397–402. MR

[Serre 2012] J.-P. Serre, *Lectures on $N_X(p)$*, Research Notes in Mathematics **11**, CRC Press, Boca Raton, FL, 2012. MR

[Skorobogatov 2001] A. Skorobogatov, *Torsors and rational points*, Cambridge Tracts in Mathematics **144**, Cambridge University Press, 2001. MR

[Smeets 2017] A. Smeets, "Insufficiency of the étale Brauer–Manin obstruction: towards a simply connected example", *Amer. J. Math.* **139**:2 (2017), 417–431. MR

[Stoppino 2011] L. Stoppino, "Fibrations of Campana general type on surfaces", *Geom. Dedicata* **155** (2011), 69–80. MR

[Wirsing 1967] E. Wirsing, "Das asymptotische Verhalten von Summen über multiplikative Funktionen, II", *Acta Math. Acad. Sci. Hungar.* **18** (1967), 411–467. MR

Communicated by Jean-Louis Colliot-Thélène
Received 2024-01-19 Revised 2024-10-14 Accepted 2024-11-12

tdb@ist.ac.at *Institute of Science and Technology Austria, Klosterneuburg, Austria*

julian.lyczak@uantwerpen.be *Department of Mathematics, University of Antwerp, Antwerpen, Belgium*

arne.smeets@kuleuven.be *Department of Mathematics, KU Leuven, Heverlee, Belgium*

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 19 No. 10 2025
