Algebra & Number Theory

Volume 13
2019
No. 2

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2019 is US $385/year for the electronic version, and $590/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.251

High moments of the Estermann function

Sandro Bettin

For $a/q \in \mathbb{Q}$ the Estermann function is defined as $D(s, a/q) := \sum_{n\geq 1} d(n) n^{-s} \operatorname{e}\big(n\frac{a}{q}\big)$ if $\Re(s) > 1$ and by meromorphic continuation otherwise. For q prime, we compute the moments of $D(s, a/q)$ at the central point $s = 1/2$, when averaging over $1 \leq a < q$.

As a consequence we deduce the asymptotic for the iterated moment of Dirichlet L-functions $\sum_{\chi_1,\dots,\chi_k \pmod q} \big|L\big(\frac12, \chi_1\big)\big|^2 \cdots \big|L\big(\frac12, \chi_k\big)\big|^2 \big|L\big(\frac12, \chi_1 \cdots \chi_k\big)\big|^2$, obtaining a power saving error term.

Also, we compute the moments of certain functions defined in terms of continued fractions. For example, writing $f_\pm(a/q) := \sum_{j=0}^{r} (\pm 1)^j b_j$ where $[0; b_0, \dots, b_r]$ is the continued fraction expansion of a/q we prove that for $k \geq 2$ and q primes one has $\sum_{a=1}^{q-1} f_\pm(a/q)^k \sim 2(\zeta(k)^2/\zeta(2k))q^k$ as $q \to \infty$.

1. Introduction

Since the pioneering work of Hardy and Littlewood [1916], the study of moments of families of L-functions has gained a central role in number theory. This is mostly due their numerous applications on, e.g., nonvanishing (see [Iwaniec and Sarnak 2000; Soundararajan 2000]) and subconvexity estimates (see [Conrey and Iwaniec 2000]). Moreover, moments are also important as they highlight clearly the symmetry of each family.

In this paper we consider the moments of the Estermann function at the central point and, as a consequence, we obtain new results for moments of Dirichlet L-functions. We will describe the Estermann function in Section 1.1.2, we now focus on the family of Dirichlet L-functions. For this family only the second and fourth moments have been computed. The asymptotic for the second moment was obtained by Paley [1931], whereas Heath-Brown [1981] considered the fourth moment and showed

$$\frac{1}{\varphi^*(q)} \sideset{}{^*}\sum_{\chi \pmod q} \big|L\big(\tfrac12, \chi\big)\big|^4 \sim \frac{1}{2\pi^2} \prod_{p|q} \frac{(1-1/p)^3}{1+1/p} (\log q)^4, \tag{1-1}$$

provided that q doesn't have "too many prime divisors", a restriction that was later removed by Soundararajan [2007]. As usual, $\sum^*$ indicates that the sum is restricted to primitive characters and $\varphi^*(q)$ denotes the number of such characters. The problem of computing the full asymptotic expansion for the fourth moment was later solved by Young [2011a] in the case when q is prime. He proved

$$\frac{1}{\varphi^*(q)} \sideset{}{^*}\sum_{\chi \pmod q} \big|L\big(\tfrac12, \chi\big)\big|^4 = \sum_{i=0}^{4} c_i (\log q)^i + O(q^{-\frac{5}{512}+\varepsilon}) \tag{1-2}$$

MSC2010: primary 11M06; secondary 11A55, 11M41, 11N75.
Keywords: Estermann function, Dirichlet L-functions, divisor function, continued fractions, mean values, moments.

for some absolute constants c_i with $c_4 = (2\pi^2)^{-1}$. Recently, Blomer, Fouvry, Kowalski, Michel and Milićević [Blomer et al. 2017] introduced several improvements in Young's work improving the error term in (1-2) to $O(q^{-\frac{1}{32}+\varepsilon})$.

In this paper, we consider a variation of this problem and compute the asymptotic of

$$M_k(q) = \frac{1}{\varphi^*(q)^{k-1}} \sideset{}{^*}\sum_{\chi_1,\dots,\chi_{k-1} \ (\mathrm{mod}\ q)} \bigl|L\bigl(\tfrac{1}{2},\chi_1\bigr)\bigr|^2 \cdots \bigl|L\bigl(\tfrac{1}{2},\chi_{k-1}\bigr)\bigr|^2 \bigl|L\bigl(\tfrac{1}{2},\chi_1\cdots\chi_{k-1}\bigr)\bigr|^2, \tag{1-3}$$

where the sum has the extra restriction that $\chi_1\cdots\chi_{k-1}$ is primitive. If $k=2$, this coincides with the usual fourth moment of Dirichlet L-functions as computed by Young, whereas if $k>2$ then $M_k(q)$ should be thought of as an iterated fourth moment, since each character appears four times in the above expression. We shall prove the following theorem.

Theorem 1. *Let $k \ge 3$ and let q be prime. Then, there exists an absolute constant $A>0$ such that*

$$M_k(q) = \sum_{n=1}^{\infty} \frac{2^{\nu(n)}}{n^{k/2}} \Bigl(\bigl(\log \tfrac{q}{8n\pi}\bigr)^k + \bigl(-\tfrac{\pi}{2}\bigr)^k\Bigr) + O_\varepsilon(k^{Ak} q^{-\delta_k+\varepsilon}),$$

where $\nu(n)$ is the number of different prime factors of n, $\delta_k := (k-2-3\vartheta)/(2k+5)$ with $\vartheta = \frac{7}{64}$ being the best bound towards Selberg's eigenvalue conjecture. Also, the implicit constant depends on ε only.

Remark. Notice that δ_k is a increasing sequence such that $\delta_k \to \frac{1}{2}$ as $k \to \infty$. For $\vartheta = \frac{7}{64}$ the first few values of δ_k are $\delta_3 = \frac{43}{704}$, $\delta_4 = \frac{107}{832}$, $\delta_5 = \frac{57}{320}$.

Theorem 1 yields an asymptotic formula for $M_k(q)$ for $k < \eta(\log q)/(\log\log q)$ with $\eta > 0$ sufficiently small. Larger values of k are easier to deal with and one obtains the following corollary.

Corollary 2. *Let q be prime. Then as $q \to \infty$ we have*

$$M_k(q) \sim \frac{\zeta(k/2)^2}{\zeta(k)} (\log(q/(8\pi)) + \gamma)^k, \tag{1-4}$$

uniformly in $3 \le k = o(q^{\frac{1}{2}} \log q)$, where γ is the Euler–Mascheroni constant. Moreover this range is optimal, meaning that (1-4) is false if $k \gg q^{\frac{1}{2}} \log q$.

Remark. Notice that the main terms in (1-4) and Theorem 1 have a double pole at $k=2$. This is consistent with the fact that the main term for $M_2(q)$ has size $(\log q)^4$ rather than $(\log q)^2$. In principle one could treat the case $k=2$ together with the case $k \ge 3$. However, in order to do so one would need to include in (2-2) an extra main-term of size $q^{o(1)}$ coming from the diagonal term. For $k \ge 3$ this term is absorbed in the error term and so it is more convenient to simply exclude the case $k=2$.

A moment somewhat similar to (1-3) was previously considered by Chinta [2005] who used a multiple Dirichlet series approach to compute the asymptotic of the first moment of (roughly)

$$L\bigl(\tfrac{1}{2},\chi_{d_1}\bigr) L\bigl(\tfrac{1}{2},\chi_{d_2}\bigr) L\bigl(\tfrac{1}{2},\chi_{d_1}\chi_{d_2}\bigr), \tag{1-5}$$

where χ_d denotes the quadratic character associated to the extension $\mathbb{Q}(\sqrt{d})$ of $\mathbb{Q}$. We remark that there is a big difference between (1-3) and this case. Indeed, if χ_1, χ_2 are characters modulo q then so is $\chi_1\chi_2$, whereas if $d_1, d_2 \approx X$ then $\chi_{d_1}\chi_{d_2}$ is typically a character with conductor $\approx X^2$. This means that (1-3) roughly correspond to an iterated fourth moment, whereas the second moment of (1-5) roughly correspond to an iterated sixth moment of quadratic Dirichlet L-functions, and thus it doesn't seem to be attackable with the current technology. (As a comparison, the first moment computed by Chinta roughly correspond to an iterated third moment).

1.1. ***Twisted moments, the Estermann function, and continued fractions.*** A nice feature of Theorem 1 is that it can be essentially rephrased in terms of high moments of other functions appearing naturally in number theory. Indeed, the same computations give also the asymptotic for moments of twisted moments of Dirichlet L-functions, of the Estermann function, and of certain functions defined in terms of continued fractions. We now briefly describe each of these objects and give the corresponding version of Theorem 1.

1.1.1. *Moments of twisted moments.* Several classical methods to investigate the central values of Dirichlet L-functions pass through the study of the second moment of $L(s,\chi)$ times a Dirichlet polynomial $P_\vartheta(s,\chi) := \sum_{n\le q^\vartheta} a_n \cdot \chi(n) n^{-s}$:

$$\frac{1}{\varphi^*(q)} \sideset{}{^*}\sum_{\chi \pmod q} \left|L\left(\tfrac{1}{2},\chi\right) P_\vartheta\left(\tfrac{1}{2},\chi\right)\right|^2. \tag{1-6}$$

For example, Iwaniec and Sarnak proved that $\frac{1}{3}$ of the Dirichlet L-functions do not vanish at the central point via proving the asymptotic for such average for $\vartheta < \frac{1}{2}$ (and choosing P_ϑ to be a mollifier). Moreover, it is easy to see that if one could extend such asymptotic to all polynomials of length $\vartheta < 1$, then the Lindelöf hypothesis would follow.

Expanding the square, using the multiplicativity of Dirichlet characters, and renormalizing, one immediately sees that (1-6) can be reduced to an average of twisted moments of the form

$$M(a,q) := \frac{q^{\frac{1}{2}}}{\varphi^*(q)} \sideset{}{^*}\sum_{\chi \pmod q} \left|L\left(\tfrac{1}{2},\chi\right)^2\right| \chi(a),$$

for $(a,q)=1$. By the orthogonality of Dirichlet characters one can immediately rewrite Theorem 1 (and (1-1)) in terms of $M(a,q)$. In particular, one has

$$\sum_{a=1}^{q} M(a,q)^k = \frac{\varphi(q)}{\varphi^*(q)} q^{k/2} M_k(q) \sim \begin{cases} \dfrac{\zeta(k/2)^2}{\zeta(k)} (\log(q/(8\pi)) + \gamma)^k & \text{if } 3 \le k = o(q^{\frac{1}{2}} \log q), \\ \dfrac{1}{2\pi^2} (\log q)^4 & \text{if } k = 2, \end{cases} \tag{1-7}$$

as $q \to \infty$ with q prime, where φ is Euler's φ-function.

1.1.2. *Moments of the Estermann function.* For

$$(a,q) = 1, \quad q > 0, \quad \alpha, \beta \in \mathbb{C} \quad \text{and} \quad \Re(s) > 1 - \min(\Re(\alpha), \Re(\beta)),$$

the Estermann function is defined as

$$D_{\alpha,\beta}(s,a/q) := \sum_{n=1}^{\infty} \mathrm{e}(na/q)\frac{\tau_{\alpha,\beta}(n)}{n^s} = D_{\cos;\alpha,\beta}(s,a/q) + i\,D_{\sin;\alpha,\beta}(s,a/q), \tag{1-8}$$

where $D_{\cos}$ and $D_{\sin}$ have the same definition as D, but with $\mathrm{e}(na/q)$ replaced by $\cos(2\pi na/q)$ and $\sin(2\pi na/q)$ respectively. As usual, $\mathrm{e}(x) := e^{2\pi ix}$ and $\tau_{\alpha,\beta}(n) := \sum_{d_1 d_2 = n} d_1^{-\alpha} d_2^{-\beta}$.

$D_{\alpha,\beta}(s,a/q)$ was first introduced (with $\alpha = \beta = 0$) by Estermann who proved that it extends to a meromorphic function on $\mathbb{C}$ satisfying a functional equation

$$\text{relating } D_{\alpha,\beta}(s,a/q) \quad \text{with } D_{-\alpha,-\beta}(1-s,\pm\bar{a}/q),$$

where $\bar{a}$ denotes the multiplicative inverse of a modulo q (and similarly for $D_{\sin}$ and $D_{\cos}$ which satisfy a more symmetric functional equation given by (3-2) below).

Since the work of Estermann [1930; 1932] on the number of representations of an integer as a sum of two or more products, the Estermann function has proved itself as a valuable tool when studying additive problems of similar flavor (see, e.g., [Motohashi 1980; 1994]) and in problems related to moments of L-functions (see, e.g., [Heath-Brown 1979; Young 2011a; Conrey et al. 1986]). These applications mainly use the functional equation for D as it encodes Voronoi's summation in an analytic fashion, allowing for a simpler computation of the main terms. However, the Estermann function is an interesting object by its own right, due to its surprising symmetries (see [Bettin 2016]) and to the connections with some interesting objects in analytic number theory. For example, by the work of Ishibashi [1995] (see also [Bettin and Conrey 2013a]) one has

$$D_{\sin;1,0}(0,a/q) = \pi\,\mathrm{s}(a,q), \qquad D_{\sin;0,0}(0,a/q) = \tfrac{1}{2}\,\mathrm{c}_0(a/q),$$

where $\mathrm{s}(a,q)$ is the classical Dedekind sum and $\mathrm{c}_0(a/q)$ is a cotangent sum, related to the Nyman–Beurling criterion for the Riemann hypothesis, which has been an object of intensive studies in recent years (see, for example, [Bettin and Conrey 2013b; Maier and Rassias 2016; Bettin 2015]). Ishibashi obtained similar identities also for other values of α, β, and in particular if α is a positive odd integer one obtains that $D_{\sin;\alpha,0}(0,a/q)$ is related to certain Dedekind cotangent sums studied by Beck [2003]. All these functions satisfy certain reciprocity relations and provide examples of "quantum modular forms" (see [Zagier 2010]).

Moreover, one can also obtain formulae relating the Estermann function to twisted moments of Dirichlet L-function (see [Bettin 2016; Conrey and Ghosh 2006]) and in particular for q prime and $(a,q)=1$, one has

$$D_{\cos;0,0}\big(\tfrac{1}{2},a/q\big) + D_{\sin;0,0}\big(\tfrac{1}{2},a/q\big) = M(a,q) + \frac{2(q^{\frac{1}{2}}-1)}{\varphi(q)}\zeta\big(\tfrac{1}{2}\big)^2. \tag{1-9}$$

By this formula and (1-7), it is clear that Theorem 1 gives an asymptotic formula for the high moments of $D_{\cos;0,0}\big(\tfrac{1}{2},a/q\big) + D_{\sin;0,0}\big(\tfrac{1}{2},a/q\big)$. The method however allows one to obtain the asymptotic for the joint moments of $D_{\cos;0,0}\big(\tfrac{1}{2},a/q\big)$ and $D_{\sin;0,0}\big(\tfrac{1}{2},a/q\big)$. We shall state this in Theorem 5 below where shifts are also included (all our results will be derived from this theorem). Here we content ourselves with giving the asymptotic for the high moments of the Estermann function:

Theorem 3. *Let q be prime. Then,*

$$\frac{1}{\varphi(q)}\sum_{a=1}^{q-1} D_{0,0}\big(\tfrac12,\tfrac{a}{q}\big)^k \sim q^{k/2-1}2^{1-k/2}\frac{\zeta(k/2)^2}{\zeta(k)}\Re\Big(\big(e^{\pi i/4}\big(\log\tfrac{q}{8\pi}+\gamma\big)-e^{-\pi i/4}\tfrac{\pi}{2}\big)^k\Big)$$

as $q\to\infty$, uniformly in $3\le k=o(q^{\frac12}\log q)$. In particular, if $3\le k\ll 1$ then

$$\frac{1}{\varphi(q)}\sum_{a=1}^{q-1} D_{0,0}\big(\tfrac12,\tfrac{a}{q}\big)^k \sim q^{k/2-1}2^{1-k/2}\frac{\zeta(k/2)^2}{\zeta(k)}\Big(\cos\big(\tfrac{k\pi}{4}\big)(\log q)^k-\tfrac{\pi}{2}\sin\big(\tfrac{k\pi}{4}\big)(\log q)^{k-1}\Big)$$

as $q\to\infty$.

1.1.3. *Moments of certain functions defined in terms of continued fractions.* Finally, we discuss the relation with continued fractions. In [Bettin 2016] (see also [Young 2011b]), it was observed that $M(a,q)$, and more generally, $D_{\cos}$ and $D_{\sin}$, can be written in terms of the continued fraction expansion of a/q. Indeed, if $a,q\in\mathbb{Z}_{>0}$ and $[b_0;b_1,\dots,b_\kappa,1]$ is the continued fraction expansion of a/q, then for q prime one has

$$M(a,q)=\sum_{\substack{j=1\\ j\text{ odd}}}^{\kappa} b_j^{\frac12}\big(\log\tfrac{b_j}{8\pi}+\gamma\big)-\tfrac{\pi}{2}\sum_{\substack{j=1\\ j\text{ even}}}^{\kappa} b_j^{\frac12}+O(\log q). \tag{1-10}$$

It is therefore not surprising that Theorem 1 has an incarnation also in terms of moments for functions of the rationals defined as

$$f_{r,\pm}(a/q):=\sum_{j=1}^{\kappa}(\pm1)^j b_j^{r/2},$$

where $r\in\mathbb{Z}_{\ge1}$.

Theorem 4. *Let q be prime and let $k,r\in\mathbb{Z}_{\ge1}$ with $3\le kr=o((\log q)/(\log\log q))$. Then*

$$\sum_{a=1}^{q} f_{r,\pm}(a/q)^k\sim 2\frac{\zeta(kr/2)^2}{\zeta(kr)}q^{kr/2}$$

as $q\to\infty$.

Starting with the work of Heilbronn [1969], who considered the average value of $f_{0,+}$, there have been a very large number of papers computing the mean values of functions defined in terms of the continued fraction expansion. In particular, we cite the works [Porter 1975; Tonkov 1974] on $f_{0,+}$ and [Yao and Knuth 1975] where the asymptotic for the first moment of $f_{2,+}$ was given. However, to the knowledge of the author, Theorem 4 is the first result giving asymptotic formulae for k-th moments with $k\ge3$ without exploiting an extra average over q (as in [Hensley 1994; Baladi and Vallée 2005]). For $k=2$ the only cases previously known where obtained by Bykovskiĭ [2005] (considering the second moment of $f_{0,+}$) and by the author [Bettin 2016] (considering the second moment of a variation of $f_{2,+}$). By combining the techniques employed in [Bettin 2016] and in this paper it seems possible to extend Theorem 4 to more general functions of similar shape.

1.2. ***Brief outline of the proof of Theorem 1.*** The approximate functional equation allows one to express $M_k(q)$ roughly in the form

$$\sum_{\substack{\pm n_1 \pm n_2 \cdots \pm n_k \equiv 0 \pmod q,\\ n_1\cdots n_k \ll q^k}} \frac{d(n_1)\cdots d(n_k)}{n_1^{\frac{1}{2}}\cdots n_k^{\frac{1}{2}}}, \tag{1-11}$$

so that the problem of estimating $M_k(q)$ reduces to that of computing the asymptotic for this quadratic divisor problem. The diagonal terms (i.e., the terms with $\pm n_1 \pm n_2 \cdots \pm n_k = 0$) are a bit easier to study and give a main term; the main difficulties then lie in obtaining an asymptotic for the off-diagonal terms and in assembling the various main terms. In his proof of (1-2), which corresponds to (1-11) with $k = 2$, Young used a combination of several techniques each effective for some range of the variables n_1, n_2. In particular, when $n_1 \approx n_2$ (in the logarithmic scale) he followed an approach à la Motohashi [1997] using Kuznetsov formula, whereas when one variable is much larger than the other one, he used (new) estimates for the average value of the divisor function in arithmetic progressions.

Our approach is similar to that of Young, however there are several substantial differences which we will now discuss in some detail. First, the larger number of variables gives us the advantage of having to deal with more "flexible" sums enlarging the ranges where the various estimates are effective. For this reason, we can afford to use slightly weaker bounds employing the spectral theory only indirectly, through the bounds of Deshouilliers and Iwaniec [1982] (together with Kim and Sarnak's bound for the exceptional eigenvalues [Kim 2003]). It seems likely that one could use spectral methods in a more direct and efficient way, however the generalization of the methods in [Young 2011a] (or [Blomer et al. 2017]) to the $k \geq 3$ case is not straightforward and so we choose a simpler route as this is still sufficient for our purposes.

The larger number of variables also has a cost. Indeed, it introduces several new complications in the extraction and in the combination of the main terms, a process that requires a rather careful analysis and constitutes the central part of this paper. One of the causes of the complicated shape of the main terms (see (6-1)-(6-2)) is that with more than two variables the dichotomy "either one variable is much bigger than the other or the variables have the same size" doesn't hold for $k > 2$ and one has to (implicitly) deal also with cases such as $n_1 \approx \cdots \approx n_{k-1} \approx q^{1+1/k}$ and $n_k \approx 1$.

Another difference with Young's work arises when studying the diagonal terms. If $k = 2$, then one can handle these terms easily thanks to Ramanujan's formula $\sum_{n\geq 1} d(n)^2/n^s = \zeta(s)^4/\zeta(2s)$. If $k \geq 3$, we don't have such a nice exact formula, and we are left with the problem of showing that the series

$$\sum_{\pm n_1 \pm \cdots \pm n_k = 0} \frac{d(n_1)\cdots d(n_k)}{(n_1\cdots n_k)^s}$$

can be meromorphically continued past the line $\Re(s) = 1 - 1/k$ which is the boundary of convergence. We shall leave this problem to a different paper, [Bettin 2017], where with similar (but a bit simpler) techniques we prove that this series admits meromorphic continuation to the region $\Re(s) > 1 - 2/(k+1)$.

Last, we mention a more technical problem. One of the steps in Young's proof requires separating n_1, n_2 in expressions of the form $(n_1 \pm n_2)^{-z}$ when $\Re(z) \approx 0$. This can be easily obtained by using some classical Mellin formulae; however, whereas the Mellin integral corresponding to $(1+x)^{-z}$ converges absolutely, the Mellin integral corresponding to $(1-x)^{-z}$ converges only conditionally so that the terms containing $(n_1-n_2)^{-z}$ demand some caution. In our case this problem becomes rather more subtle as we need to apply these formulae iteratively in order to handle expressions such as $(n_1 \pm \cdots \pm n_k)^{-s}$. We overcome this difficulties by using a modification of the resulting "iterated" Mellin formula allowing us to write such expressions in terms of absolutely convergent integrals (see Section 10 for the details).

1.3. ***The structure of the paper.*** The paper is organized as follow. In Section 2 we state Theorem 5, a more general version of Theorem 3 providing the asymptotic for the mixed moments of $D_{\cos}$ and $D_{\sin}$ (as well as allowing for some small shifts). We then use this result to deduce Theorems 1, 3 and 4. In Section 3 we give some lemmas on the Estermann function which we shall need later on. It is in these lemmas that the spectral theory comes (indirectly) into play. The proof of Theorem 5 is carried out in Sections 5–9, after introducing some notation in Section 4, and constitutes the main body of the paper. Finally, in Section 10 we will prove the Mellin formula mentioned at the end of the previous section as well as some technical Lemmas needed in order to use this formula effectively.

2. Mixed moments of $D_{\cos}$ and $D_{\sin}$ and the deduction of the main theorems

Let $k \geq 1$, q be a prime and let $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k \in \mathbb{C}$. Then, for any subset $\Upsilon \subseteq \{1, \dots, k\}$ let $M_{\Upsilon,k}$ be the mixed shifted moment

$$M_{\Upsilon,k} := \frac{1}{\varphi(q)} \sum_{a=1}^{q-1} \prod_{i=1}^{k} D_{i;\alpha_i,\beta_i}\left(\tfrac{1}{2}, \tfrac{a}{q}\right),$$

where $D_{i;\alpha_i,\beta_i} := D_{\sin;\alpha_i,\beta_i}$ if $i \in \Upsilon$ and $D_i := D_{\cos;\alpha_i,\beta_i}$ otherwise. Also, let

$$\Gamma_i(s) := \begin{cases} \Gamma\left(\frac{1}{2}+s\right) & \text{if } i \in \Upsilon, \\ \Gamma(s) & \text{otherwise.} \end{cases} \tag{2-1}$$

Since $D_{\sin;\alpha_i,\beta_i}(s, -a/q) = -D_{\sin;\alpha_i,\beta_i}(s, a/q)$, then $M_{\Upsilon,k}$ is identically zero if $|\Upsilon|$ is odd. If $|\Upsilon|$ is even the asymptotic for $M_{\Upsilon,k}$ is given by the following theorem, provided that $k \geq 3$ (the corresponding theorem for $k=2$ is essentially implicit in [Young 2011a], whereas the case $k=1$ is trivial).

Theorem 5. *Let $\Upsilon \subseteq \{1, \dots, k\}$ with $|\Upsilon|$ even. Let $k \geq 3$ and let q be a prime. Let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)$, $\boldsymbol{\beta} := (\beta_1, \dots, \beta_k) \in \mathbb{C}^k$ with $|\alpha_i|, |\beta_i| \ll 1/\log q$ and $|\alpha_i|, |\beta_i| \leq \frac{1}{10}$ for all $i = 1, \dots, k$. Then, there exists an absolute constant $A > 0$ such that for any $\varepsilon > 0$ we have*

$$M_{\Upsilon,k} = \sum_{\{\alpha_i', \beta_i'\} = \{\alpha_i, \beta_i\}} \mathscr{M}_{\boldsymbol{\alpha}', \boldsymbol{\beta}'} + O_\varepsilon(k^{Ak} q^{k/2-1-\delta_k+\varepsilon}), \tag{2-2}$$

where $\delta_k := \dfrac{k-2-3\vartheta}{2k+5}$,

$$\mathscr{M}_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \frac{q^{k/2-1}}{2^{k-1}} \frac{\zeta\big(\frac{k}{2}-\sum_{i=1}^k \alpha_i\big)\zeta\big(\frac{k}{2}+\sum_{i=1}^k \beta_i\big)}{\zeta\big(k-\sum_{i=1}^k(\alpha_i-\beta_i)\big)} \prod_{i=1}^k \frac{\Gamma_i\big(\frac14-\frac{\alpha_i}{2}\big)}{\Gamma_i\big(\frac14+\frac{\alpha_i}{2}\big)}\Big(\frac{q}{\pi}\Big)^{-\alpha_i}\zeta(1-\alpha_i+\beta_i) \tag{2-3}$$

and where the implicit constant in the error term depends on ε *only.*

Remark. If $\alpha_i = \beta_i$ for some $i = 1, \dots, k$, then $\mathscr{M}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ has to be interpreted as the limit for $\alpha_i \to \beta_i$ (see (2-4) below).

As mentioned in Section 1.3, we will prove Theorem 5 in Sections 5–9. We will now deduce Theorems 1, 3, and 4 from Theorem 5.

2.1. ***Proof of Theorem 1, 3 and 4 and of Corollary 2.*** We start by observing that if $|\Upsilon|$ is even then from Theorem 5 one has

$$\sum_{a=1}^{q-1}\prod_{i=1}^k D_{i;0,0}\big(\tfrac12,\tfrac{a}{q}\big) = \frac{q^{k/2}}{2^{k-1}}\sum_{n=1}^\infty \frac{2^{\nu(n)}}{n^{k/2}}\prod_{i=1}^k\big(\log\tfrac{q}{8n\pi}+\gamma-a_i\pi\big)+O_\varepsilon(k^{Ak}q^{k/2-\delta_k+\varepsilon}), \tag{2-4}$$

where $a_i = -\frac12$ if $i \in \Upsilon$ and $a_i = \frac12$ otherwise. Indeed, if $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ satisfy the hypothesis of Theorem 5 and $\alpha_i \neq \beta_i$ for all i, then by contour integration the main term on the right hand side of (2-2) can be rewritten as

$$\sum_{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}} \mathscr{M}_{\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*} = \frac{q^{k/2-1}}{2^{k-1}}\frac{1}{(2\pi i)^k}\oint_{|s_1|=\frac14}\cdots\oint_{|s_k|=\frac14}\frac{\zeta\big(\frac{k}{2}+\sum_{i=1}^k(s_i-\alpha_i-\beta_i)\big)\zeta\big(\frac{k}{2}+\sum_{i=1}^k s_i\big)}{\zeta\big(k+\sum_{i=1}^k(2s_i-\alpha_i-\beta_i)\big)}$$
$$\times\prod_{i=1}^k\frac{\Gamma_i\big(\frac14+\frac{s-\alpha_i-\beta_i}{2}\big)}{\Gamma_i\big(\frac14-\frac{s-\alpha_i-\beta_i}{2}\big)}\Big(\frac{q}{\pi}\Big)^{s_i-\alpha_i-\beta_i}\zeta(1+s_i-\alpha_i)\zeta(1+s_i-\beta_i)\,ds_i, \tag{2-5}$$

where the circles are integrated counterclockwise. Thus, taking the limit for $\boldsymbol{\alpha},\boldsymbol{\beta}\to 0$ and expanding $\zeta(s)^2/\zeta(2s)$ as a Dirichlet series (see [Titchmarsh 1986, (1.2.8)]), we obtain

$$\sum_{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}} \mathscr{M}_{\boldsymbol{\alpha}',\boldsymbol{\beta}'} = \frac{q^{k/2-1}}{2^{k-1}}\sum_{n=1}^\infty\frac{2^{\nu(n)}}{n^{k/2}}\prod_{i=1}^k\frac{1}{2\pi i}\oint_{|s_i|=\frac14}\frac{\Gamma_i\big(\frac14+\frac{s_i}{2}\big)}{\Gamma_i\big(\frac14-\frac{s_i}{2}\big)}\Big(\frac{q}{n\pi}\Big)^{s_i}\zeta(1+s_i)^2d\,s_i$$
$$= \frac{q^{k/2-1}}{2^{k-1}}\sum_{n=1}^\infty\frac{2^{\nu(n)}}{n^{k/2}}\prod_{i=1}^k\big(\log\tfrac{q}{8n\pi}+\gamma-a_i\pi\big),$$

by the residue theorem. We remind that $\nu(n)$ is the number of distinct prime factors of n and γ is the Euler–Mascheroni constant. Equation (2-4) then follows.

To prove Theorem 1 we observe that by (2-4) we have (remember that if $|\Upsilon|$ is odd then $M_{\Upsilon,k}=0$)

$$\begin{aligned}\sum_{a=1}^{q-1}&\bigl(D_{\cos;0,0}\bigl(\tfrac12,\tfrac aq\bigr)+D_{\sin;0,0}\bigl(\tfrac12,\tfrac aq\bigr)\bigr)^k\\
&=\sum_{r=0}^{k}\binom{k}{r}\sum_{a=1}^{q-1}D_{\cos;0,0}\bigl(\tfrac12,\tfrac aq\bigr)^{k-r}D_{\sin;0,0}\bigl(\tfrac12,\tfrac aq\bigr)^{r}\\
&=\frac{q^{k/2}}{2^{k-1}}\sum_{n=1}^{\infty}\frac{2^{\nu(n)}}{n^{k/2}}\sum_{\substack{r=0\\ r\text{ even}}}\binom{k}{r}\bigl(\log\tfrac{q}{8n\pi}+\gamma-\tfrac{\pi}{2}\bigr)^{k-r}\bigl(\log\tfrac{q}{8n\pi}+\gamma+\tfrac{\pi}{2}\bigr)^{r}+\mathcal{E}_2\\
&=\frac{q^{k/2}}{2^{k}}\sum_{n=1}^{\infty}\frac{2^{\nu(n)}}{n^{k/2}}\bigl(\bigl(2\log\tfrac{q}{8n\pi}+2\gamma\bigr)^k+\bigl(-\tfrac{\pi}{2}\bigr)^k\bigr)+\mathcal{E}_2\end{aligned}$$

for some $\mathcal{E}_2\ll_\varepsilon k^{Ak}q^{k/2-\delta_k+\varepsilon}$, where in the last step we used that

$$\sum_{\substack{r=0\\ r\text{even}}}^{k}\binom{k}{r}x^{k-r}y^r=\frac{(x+y)^k+(x-y)^k}{2}\quad\text{for all } x,y\in\mathbb{R}.$$

Thus, using (1-9) one obtains Theorem 1. One easily verifies that as $q\to\infty$

$$\sum_{n=1}^{\infty}\frac{2^{\nu(n)}}{n^{k/2}}\bigl(\bigl(\log\tfrac{q}{8n\pi}+\gamma\bigr)^k+\bigl(-\tfrac{\pi}{2}\bigr)^k\bigr)\sim\frac{\zeta(k/2)^2}{\zeta(k)}\bigl(\log\tfrac{q}{8\pi}+\gamma\bigr)^k$$

uniformly in $k\geq 3$. If $k<\eta(\log q)/(\log\log q)$ with $\eta>0$ sufficiently small (but fixed), then the error term $\mathcal{E}_2$ is smaller than the above main term and so Corollary 2 follows on this range.

Now assume $k\geq\eta(\log q)/(\log\log q)$. First, we observe that by (1-10) for $a\neq 1$ we have

$$|M(a,q)|\leq(q/\eta)^{\frac12}\log q \tag{2-6}$$

for any fixed $1<\eta<2$ and q sufficiently large. Indeed, this is obvious if $a=-1$, whereas if $a\neq\pm1$ then $\max_j b_j\leq(q-1)/2$ and so the above bound follows since $b_1\cdots b_\kappa\leq q$. Furthermore, from the second moment estimate $\sum_{a=1}^{q}|M(a,q)|^2\ll(\log q)^4$ it follows that for every $C>0$ there are at most $O(q(\log q)^4/C^2)$ values of a in $1<a\leq q$ such that $|M(a,q)|\geq C$. Thus, by (2-6) we have

$$\begin{aligned}\sum_{a=2}^{q}M(a,q)^k&\leq\sum_{\substack{2\leq a\leq q\\ |M(a,q)|<C}}M(a,q)^k+\sum_{\substack{2\leq a\leq q\\ |M(a,q)|\geq C}}M(a,q)^k\\
&\ll C^kq+\frac{q^{k/2+1}(\log q)^{k+4}}{\eta^{k/2}C^2}\ll\frac{q^{k/2+2/(k+2)}}{\eta^{k/2}}(\log q)^{k+2}\end{aligned}$$

for $C=\eta^{-\frac12}q^{\frac12-1/(k+2)}\log q$. Note that if $k\gg(\log q)/(\log\log q)$, then

$$\text{error term}\ll q^{k/2}\eta^{-k/4}(\log q)^k=o(q^{k/2}(\log(q/(8\pi))+\gamma)^k)\quad\text{as } q\to\infty\text{, uniformly in } k.$$

Finally, we have (see [Heath-Brown 1981])

$$M(1,q)=q^{\frac12}(\log(q/(8\pi))+\gamma)+2\zeta\left(\tfrac12\right)^2+O(q^{-\frac12})$$

so that

$$M(1,q)^k=q^{k/2}(\log(q/(8\pi))+\gamma)^k\exp\left(2\zeta\left(\tfrac12\right)^2\frac{k}{q^{\frac12}\log q}(1+O(1/\log q))\right)$$

for q large enough. Thus, if $(\log q)/(\log\log q)\ll k=o(q^{\frac12}\log q)$ we have

$$M_k(q)=q^{-k/2}\sum_{a=1}^{q}M(a,q)^k\sim(\log(q/(8\pi))+\gamma)^k\sim\frac{\zeta(k/2)^2}{\zeta(k)}(\log(q/(8\pi))+\gamma)^k$$

as $q\to\infty$, whereas this asymptotic is false if $k\gg q^{\frac12}\log q$. This concludes the proof of Corollary 2.

The proof of Theorem 3 is analogous to those of Theorem 1 and Corollary 2, with the difference that in this case we use (1-8) rather than (1-9). Indeed for some $\mathcal{E}_1\ll_\varepsilon k^{Ak}q^{k/2-\delta_k+\varepsilon}$ we have

$$\begin{aligned}\sum_{a=1}^{q-1}&D_{0,0}\left(\tfrac12,\tfrac{a}{q}\right)^k\\&=\sum_{r=0}^{k}\binom{k}{r}\sum_{a=1}^{q-1}D_{\cos;0,0}\left(\tfrac12,\tfrac{a}{q}\right)^{k-r}i^rD_{\sin;0,0}\left(\tfrac12,\tfrac{a}{q}\right)^r\\&=\frac{q^{k/2}}{2^{k-1}}\sum_{n=1}^{\infty}\frac{2^{\nu(n)}}{n^{k/2}}\sum_{\substack{r=0\\ r\text{ even}}}^{k}\binom{k}{r}\left(\log\tfrac{q}{8n\pi}+\gamma-\tfrac{\pi}{2}\right)^{k-r}i^r\left(\log\tfrac{q}{8n\pi}+\gamma+\tfrac{\pi}{2}\right)^r+\mathcal{E}_1\\&=\frac{q^{k/2}}{2^k}\sum_{n=1}^{\infty}\frac{2^{\nu(n)}}{n^{k/2}}\left(\left((1+i)\left(\log\tfrac{q}{8n\pi}+\gamma\right)-(1-i)\tfrac{\pi}{2}\right)^k+\left((1-i)\left(\log\tfrac{q}{8n\pi}+\gamma\right)-(1+i)\tfrac{\pi}{2}\right)^k\right)+\mathcal{E}_1\\&=(q/2)^{k/2}2\,\Re\left(\sum_{n=1}^{\infty}\frac{2^{\nu(n)}}{n^{k/2}}\left(e^{\pi i/4}\left(\log\tfrac{q}{8n\pi}+\gamma\right)-e^{-\pi i/4}\tfrac{\pi}{2}\right)^k\right)+\mathcal{E}_1,\\&\sim q^{k/2}2^{1-k/2}\frac{\zeta(k/2)^2}{\zeta(k)}\Re\left(\left(e^{\pi i/4}\left(\log\tfrac{q}{8\pi}+\gamma\right)-e^{-\pi i/4}\tfrac{\pi}{2}\right)^k\right)\end{aligned}$$

as $q\to\infty$ with $3\le k=o((\log q)/(\log\log q))$. One then obtains Theorem 3 on the range $3\le k=o(q^{\frac12}\log q)$ by proceeding as in the proof of Corollary 2.

2.2. ***Proof of Theorem 4.*** We compute the moments of $f_{r,+}$ only, the case of $f_{r,-}$ being analogous (using (2-8) instead of (2-7)).

We start by noticing that Corollary 11 of [Bettin 2016] gives

$$D_{\cos;0,0}\left(\tfrac12,\tfrac{a}{q}\right)=\frac12\sum_{j=1}^{\kappa}b_j^{\frac12}\left(\log\tfrac{b_j}{8\pi}+\gamma-\tfrac{\pi}{2}\right)+O(\log q),\tag{2-7}$$

$$D_{\sin;0,0}\big(\tfrac{1}{2},\tfrac{a}{q}\big)=\frac{1}{2}\sum_{j=1}^{\kappa}(-1)^j b_j^{\frac{1}{2}}\big(\log\tfrac{b_j}{8\pi}+\gamma+\tfrac{\pi}{2}\big)+O(\log q), \tag{2-8}$$

where $[0;b_1,\dots b_\kappa,1]$ is the continued fraction expansion of a/q. Moreover, since $b_1\cdots b_j\asymp q$, then if one among $b_1,\dots,b_j$, say b_{j^*}, satisfies $b_{j^*}>q/(\log q)^{100}$, and thus in particular

$$\log b_{j*}=\log q+O(\log\log q),$$

then $b_j\ll(\log q)^{100}$ for $j\neq j^*$. In particular, if $\max_j b_j>q/(\log q)^{100}$ and $1\le r=o(\log q/\log\log q)$, then

$$\begin{aligned}
f_{r,+}\big(\tfrac{a}{q}\big)&=\sum_{j=1}^{\kappa}b_j^r/2=\max_{j=1,\dots\kappa}b_j^{r/2}+O((\log q)^{50r+1})=\frac{1}{(\log q)^r}\Big(\max_{j=1,\dots\kappa}b_j^{\frac{1}{2}}\log q\Big)^r+O((\log q)^{50r+1})\\
&=\frac{1}{(\log q)^r}\bigg(\Big(\max_{j=1,\dots\kappa}b_j^{\frac{1}{2}}\big(\log\tfrac{b_j}{8\pi}+\gamma-\tfrac{\pi}{2}\big)\Big)\big(1+O(\log\log q/\log q)\big)\bigg)^r+O((\log q)^{50r+1})\\
&=\frac{2^r}{(\log q)^r}D_{\cos}\big(\tfrac{1}{2},\tfrac{a}{q}\big)^r\big(1+O(r\log\log q/\log q)\big).
\end{aligned} \tag{2-9}$$

Moreover, from (2-7) it follows easily that

$$\sum_{j=1}^{\kappa}b_j^{\frac{1}{2}}\le D_{\cos;0,0}\big(\tfrac{1}{2},\tfrac{a}{q}\big)+B\log q$$

for all a/q and some $B>0$. In particular, if $\max_j b_j\le q/(\log q)^{100}$ and q is large enough, then

$$\begin{aligned}
f_{r,+}\big(\tfrac{a}{q}\big)^k&\le\frac{q^{(k/2)(r-1)}}{(\log q)^{50k(r-1)}}\Big(\sum_{j=1}^{\kappa}b_j^{\frac{1}{2}}\Big)^k\le\frac{q^{(k/2)(r-1)}}{(\log q)^{50k(r-1)}}\big(D_{\cos;0,0}\big(\tfrac{1}{2},\tfrac{a}{q}\big)+B\log q\big)^k\\
&\ll\frac{q^{kr/2-1}}{(\log q)^{50k(r-1)+47(k-2)}}\big(D_{\cos;0,0}\big(\tfrac{1}{2},\tfrac{a}{q}\big)+B\log q\big)^2
\end{aligned} \tag{2-10}$$

for $k\ge 2$, since $\max_j b_j\le q/(\log q)^{100}$ implies $\big|D_{\cos}\big(\tfrac{1}{2},\tfrac{a}{q}\big)\big|+B\log q\le q^{\frac{1}{2}}/(\log q)^{48}$ for q large enough. Now, we have

$$\sum_{a=1}^{q}f_{r,+}(\tfrac{a}{q})^k=\sum_{\substack{1\le a<q\\ \max_j b_j>q/(\log q)^{100}}}f_{r,+}\big(\tfrac{a}{q}\big)^k+\sum_{\substack{1\le a<q\\ \max_j b_j\le q/(\log q)^{100}}}f_{r,+}\big(\tfrac{a}{q}\big)^k. \tag{2-11}$$

By (2-10) the second summand is bounded by

$$\begin{aligned}
\sum_{\substack{1\le a<q,\\ \max_j b_j\le q/(\log q)^{100}}}f_{r,+}\big(\tfrac{a}{q}\big)^k&\ll\frac{q^{kr/2-1}}{(\log q)^{50k(r-1)+48(k-2)}}\sum_{1\le a<q}\big(D_{\cos}\big(\tfrac{1}{2},\tfrac{a}{q}\big)+B\log q\big)^2\\
&\ll\frac{q^{kr/2}}{(\log q)^{50k(r-1)+48(k-2)-4}}\ll\frac{q^{kr/2}}{\log q}
\end{aligned}$$

for $kr \geq 3$ (if $k=1$ one needs to modify slightly the argument, but the final bound still holds). By (2-9) the first summand of (2-11) can be written as

$$\sum_{\substack{1\leq a<q\\ \max_j b_j>q/(\log q)^{100}}}^{q} \frac{\big(2D_{\cos;0,0}\big(\frac12,\frac aq\big)\big)^{kr}}{(\log q)^{kr}}\big(1+O(kr\log\log q/\log q)\big)$$
$$= \sum_{1\leq a<q} \frac{\big(2D_{\cos;0,0}\big(\frac12,\frac aq\big)\big)^{kr}}{(\log q)^{kr}}\big(1+O(kr\log\log q/\log q)\big)+O(q^{kr/2}/\log q)$$
$$= 2\frac{\zeta(kr)^2}{\zeta(kr/2)}q^{kr/2}\big(1+O(kr\log\log q/\log q)\big)$$

by (2-4) for $3\leq rk=o((\log q)/(\log\log q))$ and where one can complete the sum by proceeding as in the previous computation. Theorem 4 then follows.

3. The Estermann function and bounds for sums of Kloosterman sums

In this Section we give some results for the Estermann function and for the periodic zeta-function which will be needed in the proof of Theorem 5. In particular, in Section 3.1 we give the functional equation for both these functions, whereas in Section 3.2 we give a version of the approximate functional equation for the Estermann function. Finally, in Section 3.3 we give some estimates for products of the Estermann function and the periodic zeta-function, using the bounds of [Deshouillers and Iwaniec 1982] for sums of Kloosterman sums.

3.1. ***The functional equations.*** We start by giving the functional equation for the Estermann function.

Lemma 6. *For $(a,q)=1$, $q>0$ and $\alpha\in\mathbb{C}$, $D_{\alpha,\beta}(s,a/q)-q^{1-\alpha-\beta-2s}\zeta(s+\alpha)\zeta(s+\beta)$ can be extended to an entire function of s. Moreover, $D_{\alpha,\beta}(s,a/q)$ satisfies the functional equation*

$$D_{\alpha,\beta}\big(s,\tfrac aq\big)=\frac2q\Big(\frac{q}{2\pi}\Big)^{2-2s-\alpha-\beta}\Gamma(1-s-\alpha)\Gamma(1-s-\beta)$$
$$\times\Big(\cos(\pi(\alpha-\beta)/2)D_{\alpha,\beta}\big(1-s,\tfrac{\bar a}{q}\big)-\cos\big(\tfrac\pi2(2s+\alpha+\beta)\big)D_{-\alpha,-\beta}\big(1-s,-\tfrac{\bar a}{q}\big)\Big), \tag{3-1}$$

where, here and in the following, $\bar a$ denotes the multiplicative inverse of a modulo the denominator q.

Proof. This is Lemma 4 of [Conrey 1989]. □

Corollary 7. *Let*

$$\Lambda_{\cos;\alpha,\beta}\big(s,\tfrac aq\big):=\Gamma\big(\tfrac{s+\alpha}{2}\big)\Gamma\big(\tfrac{s+\beta}{2}\big)\Big(\frac q\pi\Big)^{s+(\alpha+\beta)/2}D_{\cos;\alpha,\beta}\big(s,\tfrac aq\big),$$
$$\Lambda_{\sin;\alpha,\beta}\big(s,\tfrac aq\big):=\Gamma\big(\tfrac{1+s+\alpha}{2}\big)\Gamma\big(\tfrac{1+s+\beta}{2}\big)\Big(\frac q\pi\Big)^{s+(\alpha+\beta)/2}D_{\sin;\alpha,\beta}\big(s,\tfrac aq\big).$$

Then, we have the functional equations

$$\Lambda_{\cos;\alpha,\beta}\big(s,\tfrac aq\big)=\Lambda_{\cos;-\alpha,-\beta}\big(1-s,\tfrac{\bar a}{q}\big),\qquad \Lambda_{\sin;\alpha,\beta}\big(s,\tfrac aq\big)=\Lambda_{\sin;-\alpha,-\beta}\big(1-s,\tfrac{\bar a}{q}\big). \tag{3-2}$$

Proof. These functional equations follow from (3-1), using the reflection and the duplication formulas for the Γ-function. □

We also need the basic properties of the periodic zeta-function which, for $x \in \mathbb{R}$ and $\Re(s) > 1$, is defined as

$$F(s,x) := \sum_{n=1}^{\infty} \frac{\mathrm{e}(nx)}{n^s}. \tag{3-3}$$

Notice that if $x \in \mathbb{Z}$, then $F(s,x) = \zeta(s)$.

Lemma 8. *Let $h, l \in \mathbb{Z}$ with $(h,\ell) = 1$ and $\ell > 0$, then $F(s, h/\ell)$ extends to an entire function of s with the exception of a simple pole at $s = 1$ if $\ell = 1$. Moreover, $F(s,x)$ satisfies the functional equation*

$$F(1-s, h/\ell) = \ell^{s-1} \sum_{b=1}^{\ell} \mathrm{e}(hb/\ell) \frac{\Gamma(s)}{(2\pi)^s} \bigl(e^{-\pi i s/2} F(s, b/\ell) + e^{\pi i s/2} F(s, -b/\ell)\bigr). \tag{3-4}$$

Finally, for $\ell \nmid h$ we have

$$F(0, h/\ell) = -\tfrac{1}{2} + \tfrac{i}{2} \cot(\pi h/\ell). \tag{3-5}$$

Proof. For (3-5) and the analytic continuation of F see [Apostol 1951, pp. 161, 164]. For (3-4), one divides the series for F into congruence classes modulo ℓ writing $F(s, h/\ell)$ as a sum of Hurwitz zeta-functions $\zeta(s, b/\ell)$; applying the functional equation [Apostol 1976, Theorem 12.6] for $\zeta(s,x)$ then gives (3-4). □

3.2. ***The approximate functional equation.*** Next, we give an approximate functional equation allowing us to express a product of k Estermann functions as a sum of total length about $q^{k/2}$.

Lemma 9. *Let $k \geq 1$ and $\Upsilon \subseteq \{1, \dots, k\}$. Let $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)$ be an entire function satisfying $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(-s) = G_{-\boldsymbol{\alpha},-\boldsymbol{\beta}}(s)$, $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(0) = 1$ and $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}\bigl(\frac{1}{2} - \alpha_i\bigr) = G_{\boldsymbol{\alpha},\boldsymbol{\beta}}\bigl(\frac{1}{2} - \beta_i\bigr) = 0$ for $i = 1, \dots, k$ and decaying faster than any power of s on vertical strips. Let*

$$g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) := \pi^{-ks} \prod_{j=1}^{k} \frac{\Gamma_i\bigl(\bigl(\frac{1}{2}+s+\alpha_i\bigr)/2\bigr)\Gamma_i\bigl(\bigl(\frac{1}{2}+s+\beta_i\bigr)/2\bigr)}{\Gamma_i\bigl(\bigl(\frac{1}{2}+\alpha_i\bigr)/2\bigr)\Gamma_i\bigl(\bigl(\frac{1}{2}+\beta_i\bigr)/2\bigr)}, \tag{3-6}$$

$$X_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \prod_{j=1}^{k} \frac{\Gamma_i\bigl(\bigl(\frac{1}{2}-\alpha_i\bigr)/2\bigr)\Gamma_i\bigl(\bigl(\frac{1}{2}-\beta_i\bigr)/2\bigr)}{\Gamma_i\bigl(\bigl(\frac{1}{2}+\alpha_i\bigr)/2\bigr)\Gamma_i\bigl(\bigl(\frac{1}{2}+\beta_i\bigr)/2\bigr)} \Bigl(\frac{q}{\pi}\Bigr)^{-\alpha_i-\beta_i}$$

and for any $c_s > 0$ let

$$V_{\boldsymbol{\alpha},\boldsymbol{\beta}}(x) := \frac{1}{2\pi i} \int_{(c_s)} G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) x^{-s} \frac{ds}{s},$$

where, as usual, $\int_{(c)} \cdot \, ds$ indicates that the integral is taken along the vertical line from $c - i\infty$ to $c + i\infty$. Then for $a, q \in \mathbb{Z}$, with $q > 1$ and $(a,q) = 1$ we have

$$\prod_{i=1}^{k} D_{i;\alpha_i,\beta_i}\bigl(\tfrac{1}{2}, \tfrac{a}{q}\bigr) = S_{\boldsymbol{\alpha},\boldsymbol{\beta}}(a,q) + X_{\boldsymbol{\alpha},\boldsymbol{\beta}} S_{-\boldsymbol{\alpha},-\boldsymbol{\beta}}(\bar{a}, q), \tag{3-7}$$

where $\bar{a}$ is the inverse of a modulo q and

$$S_{\boldsymbol{\alpha},\boldsymbol{\beta}}(a,q) := \frac{i^{-|\Upsilon|}}{2^k} \sum_{\varepsilon=(\pm_1 1,\dots,\pm_k 1)\in\{\pm 1\}^k} \sum_{n_1,\dots,n_k\geq 1} \rho_\Upsilon(\varepsilon) \frac{\tau_{\alpha_1,\beta_1}(n_1)\cdots\tau_{\alpha_k,\beta_k}(n_k)}{(n_1\cdots n_k)^{\frac12}} \times \mathrm{e}\bigg(\frac{a(\pm_1 n_1 \pm_2 \cdots \pm_k n_k)}{q}\bigg) V_{\boldsymbol{\alpha},\boldsymbol{\beta}}\bigg(\frac{n_1\cdots n_k}{q^k}\bigg),$$

with $\rho_\Upsilon(\varepsilon) := \prod_{i\in\Upsilon}(\pm_i 1)$.

Proof. By contour integration and the functional equation, we have

$$\begin{aligned}
\prod_{i=1}^k \Lambda_{i;\alpha_i,\beta_i}\big(\tfrac12,\tfrac{a}{q}\big)
&= \frac{1}{2\pi i}\bigg(\int_{(2)} - \int_{(-2)}\bigg) \prod_{i=1}^k \Lambda_{i;\alpha_i,\beta_i}\big(\tfrac12+s,\tfrac{a}{q}\big)\cdot G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)\,\frac{ds}{s}\\
&= \frac{1}{2\pi i}\int_{(2)} \prod_{i=1}^k \Lambda_{i;\alpha_i,\beta_i}\big(\tfrac12+s,\tfrac{a}{q}\big)\cdot G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)\,\frac{ds}{s} + \frac{1}{2\pi i}\int_{(2)} \prod_{i=1}^k \Lambda_{i;-\alpha_i,-\beta_i}\big(\tfrac12+s,\tfrac{\bar{a}}{q}\big)\cdot G_{-\boldsymbol{\alpha},-\boldsymbol{\beta}}(s)\,\frac{ds}{s}.
\end{aligned}$$

Now, expanding the Estermann functions into their Dirichlet series, we see that

$$\begin{aligned}
&\frac{1}{2\pi i}\int_{(2)} \prod_{i=1}^k \frac{\Lambda_{i;\alpha_i,\beta_i}\big(\tfrac12+s,\tfrac{a}{q}\big)}{\Gamma_i\big(\big(\tfrac12+\alpha_i\big)/2\big)\Gamma_i\big(\big(\tfrac12+\beta_i\big)/2\big)\big(\tfrac{q}{\pi}\big)^{\frac12+(\alpha_i+\beta_i)/2}}\cdot G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)\,\frac{ds}{s}\\
&\quad= \frac{i^{-|\Upsilon|}}{2^k}\sum_{n_1,\dots,n_k\in\mathbb{Z}\setminus\{0\}} \operatorname{sgn}\bigg(\prod_{i\in\Upsilon} n_i\bigg) \frac{\tau_{\alpha_1,\beta_1}(|n_1|)\cdots\tau_{\alpha_k,\beta_k}(|n_k|)}{|n_1\cdots n_k|^{\frac12}}\,\mathrm{e}\bigg(\frac{a(n_1+\cdots+n_k)}{q}\bigg)\\
&\qquad\times \frac{1}{2\pi i}\int_{(2)} G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)\bigg(\frac{|n_1\cdots n_k|}{q^k}\bigg)^{-s}\frac{ds}{s}
\end{aligned}$$

and the lemma follows. □

3.3. ***Estimates for the Estermann function.*** In this section we give two bounds for certain averages of products the Estermann function and the periodic zeta-function. Both bounds depend on estimates for Kloosterman sums, more specifically on Weil's bound and on (a minor modification of) a bound by Deshouilliers and Iwaniec [1982]. We recall that the classical Kloosterman sum is defined as

$$S(m,n;\ell) := \sideset{}{^*}\sum_{c \pmod{\ell}} \mathrm{e}\bigg(\frac{mc+n\bar{c}}{\ell}\bigg)$$

for any $c,m,n\in\mathbb{Z}$, $c\geq 1$, where $\sum^*$ indicates that the sum is over $c \pmod{\ell}$ such that $(c,\ell)=1$. Also, we recall that Weil's bound gives $S(m,n;\ell)\ll d(\ell)(m,n,\ell)^{\frac12}\ell^{\frac12}$. Using this bound we obtain the following lemma.

Lemma 10. *Let* $r > 0$, $0 < \delta < 1$, $C \geq 2$, $\eta_0 \neq 0$ *and* $(\eta_1, \ldots, \eta_r) \in \{\pm 1\}^r$. *Let* $|a| \leq 2C\delta$, $|b| \leq C\delta$ *and* $|a_j|, |b_j| < \delta$ *for* $j = 1, \ldots, r$. *Then, for some* $A > 0$ *we have*

$$\sum_{\ell \geq 1} \frac{1}{\ell^{C+a}} \sideset{}{^*}\sum_{h \pmod{\ell}} F\big(1 + C\big(s - \tfrac{1}{2}\big) + b, \tfrac{\eta_0 h}{\ell}\big) \prod_{j=1}^{r} D_{a_j, b_j}\big(\tfrac{1}{2}, \tfrac{\eta_j h}{\ell}\big) \ll_\delta (AC/\delta)^{A(r+C)} (1 + |s|)^{A(r+C)} \tag{3-8}$$

in the strip

$$-\frac{1}{2} + \frac{r + \frac{3}{2}}{C + r - \frac{1}{2}} + 8\delta < \Re(s) < \frac{1}{2} - 2\delta,$$

where F *is the periodic zeta-function defined in* (3-3). *Moreover, the left hand side of* (3-8) *is meromorphic in the half plane* $\Re(s) > -\frac{1}{2} + (r + \frac{3}{2})/(C + r - \frac{1}{2}) + 8\delta$ *with poles at* $s = \frac{1}{2} - a_j$ *and* $s = \frac{1}{2} - b_j$ *for* $j = 1, \ldots, r$ *and* $s = \frac{1}{2} - b/C$ *and these poles are simple if* $a_1, \ldots, a_r, b_1, \ldots, b_r$ *and* b/C *are all distinct.*

Proof. For $L \geq 1$, let

$$H_L(s) := \sum_{L < \ell \leq 2L} \frac{1}{\ell^{C+a}} \sideset{}{^*}\sum_{h \pmod{\ell}} F\big(1 + C\big(s - \tfrac{1}{2}\big) + b, \tfrac{\eta_0 h}{\ell}\big) \prod_{j=1}^{r} D_{a_j, b_j}\big(\tfrac{1}{2} + s, \tfrac{\eta_j h}{\ell}\big),$$

$$K(s) := \prod_{j=1}^{r} \big(s - \tfrac{1}{2} + a_j\big)\big(s - \tfrac{1}{2} + b_j\big).$$

Notice that if $\ell \neq 1$ and $(h, \ell) = 1$ then $F(x, h/\ell)$ is entire and thus so is $H_L(s)K(s)$ for all $L \geq 1$. Now, if $\Re(s) = \frac{1}{2} + 2\delta$, then a trivial bound gives

$$H_L(s)K(s) \ll (1 + |s|^{2r})(A/\delta)^{2r+1} L^{-C+2+2\delta C}, \tag{3-9}$$

where, here and in the following, A denotes a sufficiently large positive constant, which might change from line to line.

Next, take $\Re(s) = -\frac{1}{2} - 2\delta$. Then, applying the functional equations (3-1) and (3-4) to D and F, expanding D and F into their Dirichlet series, and using Stirling's formula in the crude form

$$\Gamma(\sigma + it) \ll c^{-1}(1 + A|\sigma|)^{|\sigma|}(1 + |t|)^{\sigma - \frac{1}{2}} e^{-(\pi/2)|t|}, \quad \sigma \geq c > 0, \tag{3-10}$$

we see that

$$\begin{aligned} H_L(s)K(s) \ll A^r C^{AC} (1 + |s|)^{A(r+C)} L^{r-1+(5C+6r)\delta} \sum_{L < \ell \leq 2L} \sum_{u=1}^{\ell} \big| F\big(C\big(\tfrac{1}{2} - s\big) - b, u/\ell\big)\big| \\ \times \sum_{n_1, \ldots, n_r \in \mathbb{Z}_{\neq 0}} \frac{|\tau_{-a_1, -b_1}(|n_1|) \cdots \tau_{-a_r, -b_r}(|n_r|)|}{|n_1^{1+2\delta} \cdots n_r^{1+2\delta}|} \big| S(\eta_0 u, n_1 + \cdots + n_k; \ell)\big|, \end{aligned}$$

and thus by Weil's bound we obtain

$$H_L(s)K(s) \ll (A/\delta)^{2r+5} C^{AC} (1 + |s|)^{A(r+C)} L^{r + \frac{3}{2} + 6(r+C)\delta}, \tag{3-11}$$

when $\Re(s)=-\frac{1}{2}-2\delta$. Thus, by (3-9), (3-11) and the Phragmén–Lindelöf principle, if $-\frac{1}{2}-2\delta\le\Re(s)\le\frac{1}{2}+2\delta$ we have

$$H_L(s)K(s)\ll(A/\delta)^{2r+5}C^{AC}(1+|s|)^{A(r+C)}L^{r+\frac{3}{2}-(C+r-\frac{1}{2})(\Re(s)+\frac{1}{2})+5\delta(r+C)}.$$

Moreover, if $\left|s-\frac{1}{2}\right|>2\delta$ then $K(s)\gg\delta^{2r}$ and thus, if $-\frac{1}{2}-2\delta\le\Re(s)\le\frac{1}{2}-2\delta$, we have

$$H_L(s)\ll(A/\delta)^{4r+5}C^{AC}(1+|s|)^{A(r+C)}L^{r+\frac{3}{2}-(C+r-\frac{1}{2})(\Re(s)+\frac{1}{2})+5\delta(r+C)}.$$

It follows that if

$$-\frac{1}{2}+\frac{r+\frac{3}{2}+6\delta(r+C)}{C+r-\frac{1}{2}}\le\Re(s)\le\frac{1}{2}-2\delta \tag{3-12}$$

then

$$\sum_{\ell>1}\frac{1}{\ell^{C+a}}\sideset{}{^*}\sum_{h\ (\mathrm{mod}\ \ell)}F\big(1+C\big(s-\tfrac{1}{2}\big)+b,\tfrac{\eta_0 h}{\ell}\big)\prod_{j=1}^{r}D_{a_j,b_j}\big(\tfrac{1}{2}+s,\tfrac{\eta_j h}{\ell}\big)\ll_\delta(A/\delta)^{4r+6}C^{AC}(1+|s|)^{A(r+C)}.$$

Finally, the contribution of the $\ell=1$ term to the left hand side of (3-8) is

$$\zeta\big(1+C\big(s-\tfrac{1}{2}\big)+b\big)\prod_{j=1}^{r}\zeta\big(\tfrac{1}{2}+s+a_j\big)\zeta\big(\tfrac{1}{2}+s+b_j\big)\ll(A/\delta)^{2r+1}C^{AC}(1+|s|)^{A(r+C)}$$

when s satisfies (3-12) and thus (3-8) follows. We conclude by remarking that the above computations also give the meromorphicity of the left hand side of (3-8) on $\Re(s)\ge-\frac{1}{2}+\big(r+\frac{3}{2}+5\delta(r+C)\big)/\big(C+r-\frac{1}{2}\big)$. $\square$

We now states a variation of a bound by Deshouilliers–Iwaniec for sums of Kloosterman sums (see Theorem 9 and (1.52) of [Deshouillers and Iwaniec 1982]), which is also essentially implicit in the more general bounds given in [Blomer et al. 2007; Harman et al. 2004] (see [Watt 2005, Theorem 1.4]).

Lemma 11. *Let W be a smooth function supported in $[1,2]$ and satisfying $W^{(i)}(x)\ll C^i$ for $i=0,1,2$ and some $C>1$. Let a_m, $b_n\ll 1$ be sequences of complex numbers supported in $[M,2M]$ and $[N,2N]$ respectively. Then, for $q\ge 1$ and $\eta\in\{\pm1\}$ we have*

$$\sum_{m,n,\ell\ge1}W(\ell/L)a_m b_n S(qm,\eta n;\ell)\ll_\varepsilon q^{\vartheta+\varepsilon}C^{\frac{9}{2}+\varepsilon}(L^{1+\varepsilon}+q^{\frac{1}{2}})MN, \tag{3-13}$$

where $\vartheta=\frac{7}{64}$.

Proof. First we observe that we can assume that a_m is supported on integers which are coprime with q. Indeed, if (3-13) holds in the coprime case, then since $\vartheta<\frac{1}{2}$ we have

$$\begin{aligned}\sum_{m,n,\ell\ge1}W(\ell/L)a_m b_n S(qm,\eta n;\ell)&=\sum_{d\,|\,q^\infty}\sum_{\substack{m,n,\ell\ge1\\(m,q)=1}}W(\ell/L)a_{dm}b_n S(qdm,\eta n;\ell)\\&\ll\sum_{d\,|\,q^\infty}\frac{q^{\vartheta+\varepsilon}}{d^{1-\vartheta-\varepsilon}}C^{\frac{9}{2}+\varepsilon}(L^{1+\varepsilon}+(dq)^{\frac{1}{2}})MN\\&\ll q^{\vartheta+\varepsilon}C^{\frac{9}{2}+\varepsilon}(L^{1+\varepsilon}+q^{\frac{1}{2}})MN,\end{aligned}$$

as claimed. To prove (3-13) in the coprime case, one proceeds as in the proof of Theorem 9 of [Deshouillers and Iwaniec 1982] applying Kuznetsov's formula. Then one uses the multiplicativity of Hecke-eigenvalues to separate q and m and applies the Kim–Sarnak bound [Kim 2003] for Hecke eigenvalues to deal with the contribution of the q-coefficient. The rest of the proof carries on as in [Deshouillers and Iwaniec 1982] essentially unchanged other than for the parameter X which is now multiplied by $q^{\frac{1}{2}}$. We remark that the multiplicativity of Hecke eigenvalues holds since we are in the case of level 1 for which there are only new-forms.

The above argument was carried out in detail in [Blomer et al. 2007, Theorem 4], where the authors deal with the more general case of arbitrary level which introduces several difficulties especially when dealing with the contribution of the Eisenstein spectrum. In some ranges [Blomer et al. 2007, Theorem 4] gives a weaker bound than (3-13), but one can easily modify their proof to obtain (3-13). Indeed, for $D = 1$ the bound on the last display of [Blomer et al. 2007, p. 75] can be modified to give (in the same notation as in [Blomer et al. 2007])

$$\ll_{\varepsilon,p_1,p_2} ((1+X)Zq)^{4\varepsilon}\left(\frac{Z}{|\xi_1|M}\right)^{p_1}\left(\frac{Z}{|\xi_2|N}\right)^{p_2} MNq^{2\vartheta}\,\frac{Z^{\frac{3}{2}}+ZX+X^2+M/q}{1+X/Z}\|a_2\|_2^2. \tag{3-14}$$

If $Z^{1+\varepsilon} \geq X$ then this is obvious since this bound is weaker than the bound in [Blomer et al. 2007], aside from the fact that we removed the factors $(1+C/\sqrt{MN})^{2\vartheta}$ and $Z^{2\vartheta}$ since Selberg's eigenvalue conjecture holds when the level is $D = 1$. If $X > Z^{1+\varepsilon}$, then $Z^{1+\varepsilon} < T_0 = 16X$ and so only the summands with $|t_j| \leq 1$ and $1 \leq |t_j| \leq T_0 = 16X$ give a nonnegligible contribution. The terms with $|t_j| \leq 1$ then are bounded as in the first display of [Blomer et al. 2007, p. 75] without ignoring the extra saving $(1+X/Z)^{-1}$ as done there, whereas for the terms with $1 \leq |t_j| \leq T_0$ we use the bound in the first line of the second display of [Blomer et al. 2007, p. 75] using $T_0 = 16X$.

In the case $D = 1$ the contribution of both the holomorphic and the continuous spectrum can be treated in the same way without extra difficulties, obtaining that also their contribution is bounded by (3-14). Using these bounds we then obtain that the left hand side of (3-13) is

$$\begin{aligned}&\ll_\varepsilon q^{\vartheta+\varepsilon}C^{2+\varepsilon}\frac{(C^{\frac{3}{2}}+C\sqrt{qMN}/L+qMN/L^2+M)^{\frac{1}{2}}(C^{\frac{3}{2}}+C\sqrt{qMN}/L+qMN/L^2+N)^{\frac{1}{2}}}{1+\sqrt{qMN}/(LC)}\\ &\qquad\qquad\times L^{1+\varepsilon}\sqrt{MN}\\ &\ll q^{\vartheta+\varepsilon}C^{\frac{9}{2}+\varepsilon}(L^{1+\varepsilon}+q^{\frac{1}{2}})MN.\end{aligned}$$

□

Remark. Using the variation of the spectral large sieve given by Blomer and Milićević [2015, Theorem 8], one obtains a bound which improves upon (3-13) when the parameters are in certain ranges. It is likely that the use of such a bound in combination with (3-13) would lead to a better bound for the error term in Theorem 5. However for simplicity we choose to use (3-13) in all ranges, since this is sufficient for our purposes.

Using Lemma 11 we obtain the following result.

Lemma 12. *For $r \geq 1$, let $t_0, \dots, t_r \in \mathbb{R}$, $(\eta_1, \dots, \eta_r) \in \{\pm 1\}^r$, and let $\eta_0 \neq 0$. Furthermore, let $|a_j|, |b_j| < \delta$ for $j = 1, \dots, r$ and some $0 < \delta < 1$. Finally, let $L > 0$ and let $W(x)$ be a smooth function supported on $[1, 2]$ with $W^{(i)}(x) \ll 1$ for $i = 0, 1, 2$. Then, if $w \in \mathbb{C}$ and $\sigma \geq 2\delta$, we have that*

$$\mathfrak{S} := \sum_{\ell \geq 1} \frac{W(\ell/L)}{\ell^{1+w}} \sideset{}{^*}\sum_{h \pmod{\ell}} F\big(1+\sigma+it_0, \tfrac{\eta_0 h}{\ell}\big) \prod_{j=1}^{r} D_{a_j,b_j}\big(-\sigma+it_j, \tfrac{\eta_j h}{\ell}\big)$$

is bounded by

$$\mathfrak{S} \ll_\delta L^{r(3\sigma+1)-\Re(w)} \frac{A^{r(\sigma+1)}}{\delta^{2r}} K_r(\sigma, w, t_1, \dots, t_j) \times \begin{cases} |\eta_0|^{\vartheta+\delta} & \text{if } L \geq |\eta_0|^{\frac{1}{2}}, \\ |\eta_0|^{\frac{1}{6}+\frac{\vartheta}{3}+\delta} & \text{always}, \end{cases} \tag{3-15}$$

for some absolute $A > 0$ and where

$$K_r(s, w, t_1, \dots, t_j) := (1+\sigma)^{2r(2\sigma+1)}(1+|w|)^4 \prod_{j=1}^{r}(1+|s|+|t_j|)^{1+4\sigma}$$

Proof. Applying the functional equation (3-1), expanding D and F into their Dirichlet series, and using (3-10) we obtain

$$\mathfrak{S} \ll_\delta A^r(1+A\sigma)^{2r(\sigma+\delta+1)} \bigg(\prod_{j=1}^{r}(1+|t_j|)^{2(\sigma+\delta)+1}\bigg) \bigg|\sum_{\ell \geq 1}\sum_{m \geq 1}\sum_{n \in \mathbb{Z}} \frac{W_0(\ell) f_n}{m^{1+\sigma+it_0}} S(\eta_0 m, n; \ell)\bigg|,$$

where

$$W_0(x) := W(x/L) x^{r(2\sigma+1)-1-w-\sum_{j=1}^{r}(t_j+a_j+b_j)} \qquad \ll (2L)^{r(3\sigma+1)-1-\Re(w)},$$

$$f_n := \sum_{\substack{n_1,\dots,n_r \in \mathbb{Z}_{\neq 0} \\ n_1+\dots+n_r = n}} \frac{\tau_{-a_1,-b_1}(|n_1|)}{n_1^{1+s-it_1}} \cdots \frac{\tau_{-a_r,-b_r}(|n_r|)}{n_r^{1+s-it_r}} \ll_\delta (A/\delta)^{2r} \frac{1}{|n|^{1+\delta/2}}.$$

Splitting the sums over n and m into dyadic blocks and applying (3-13) one easily gets the bound

$$\mathfrak{S} \ll_\delta L^{r(3\sigma+1)-\Re(w)} \frac{A^{r(\sigma+1)}}{\delta^{2r}} K_r(\sigma, w, t_1, \dots, t_j)|\eta_0|^{\vartheta+\delta}(1+|\eta_0|^{\frac{1}{2}}/L), \tag{3-16}$$

which gives (3-15) in the case $L \geq |\eta_0|^{\frac{1}{2}}$. Applying Weil's bound rather than (3-13), one obtains

$$\mathfrak{S} \ll_\delta L^{r(3\sigma+1)-\Re(w)} \frac{A^{r(\sigma+1)}}{\delta^{2r}} K_r(\sigma, w, t_1, \dots, t_j) L^{\frac{1}{2}}, \tag{3-17}$$

and taking the minimum between (3-16) and (3-17) one gets (3-15) also in the case $L < |\eta_0|^{\frac{1}{2}}$. □

4. Some assumptions

In this section we set up some notation and make some simplifying assumptions, which we will use throughout the rest of the paper.

First, q will always denote a prime, k an integer greater than 2, and Υ a subset of $\{1, \dots, k\}$ of even cardinality. Moreover we shall use the convention that A and ε denote respectively a sufficiently large and an arbitrarily small positive constant on which the implicit bounds are allowed to depend and whose value might change from line to line.

Also, we assume $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k) \in \mathbb{A}^k_{2C}$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k) \in \mathbb{A}^k_{C/2}$ for some constant $C > 0$ (with $4C/\log q \le \frac{1}{10}$), where $\mathbb{A}_r$ denotes the annulus $\{s \in \mathbb{C} \mid r/\log q \le |s| \le 2r/\log q\}$. This assumption can then be removed in the proof of Theorem 5 by analytic continuation and the maximum modulus principle, since both the left hand side and, by (2-5), the main term on the right hand side of (2-2) are analytic functions of the shifts in $|\alpha_i|, |\beta_i| \le 4C/\log q$. We remark in particular, that with the above assumption, we have $|\alpha_i|, |\beta_i|, |\alpha_i - \beta_i| \asymp 1/\log q$.

Moreover, for the rest of the paper we fix an entire function $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)$ as follow:

$$G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) := \frac{Q_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)}{Q_{\boldsymbol{\alpha},\boldsymbol{\beta}}(0)} \frac{\xi\left(\frac{1}{2}+s\right)}{\xi\left(\frac{1}{2}\right)}, \tag{4-1}$$

where $\xi(s) := \frac{1}{2}s(s-1)\pi^{-\pi/2}\Gamma\left(\frac{1}{2}s\right)\zeta(s)$ is the Riemann ξ-function and

$$Q_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) := \prod_{i=1}^{k}\left(\left(s^2 - (\alpha_i - \beta_i)^2\right)\left(\tfrac{1}{4} - (s+\alpha_i)^2\right)\left(\tfrac{1}{4} - (s+\beta_i)^2\right)\right).$$

By the functional equation for the Riemann zeta-function we have $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(-s) = G_{-\boldsymbol{\alpha},-\boldsymbol{\beta}}(s)$ and so $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)$ satisfies the hypotheses of Lemma 9. Moreover, using Stirling's formula (3-10) we also obtain

$$G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) \ll (A\log q)^{2k} e^{-C_1|t|}(1+|\sigma|)^{A(|\sigma|+k)}, \tag{4-2}$$

for all $s = \sigma + it \in \mathbb{C}$ and some $C_1 > 0$.

Finally, we notice that from the functional equations (3-2), for $i = 1, \dots, k$ we have the convexity bound

$$D_{i,\alpha_i,\beta_i}\left(\tfrac{1}{2}, \tfrac{a}{q}\right) \ll q^{\frac{1}{2}}(\log q)^2$$

and so trivially $M_{\Upsilon,k} \ll (Aq^{\frac{1}{2}}(\log q)^2)^k$. Also, from (2-5) it is easy to see that one also has

$$\sum_{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}} \mathscr{M}_{\boldsymbol{\alpha}',\boldsymbol{\beta}'} \ll q^{k/2-1}(A\log q)^k.$$

It follows that Theorem 5 is trivial if $k \gg \log q/\log\log q$ since in this case $(Ak)^{Ak} \gg q^{A/2}$. Thus, we will assume $k = o(\log q/\log\log q)$. In particular, for q large enough we have $|\alpha_i|, |\beta_i| \le 4C/\log q < 1/(k\log\log q) < \frac{\varepsilon}{2k}$ and a fortiori

$$|\alpha_1| + \cdots + |\alpha_k| + |\beta_1| + \cdots + |\beta_k| < \varepsilon.$$

Moreover, notice that under these assumptions we also have the inequality $(k/\varepsilon)^{Ak} \ll (\log q)^{Ak} \ll q^{\varepsilon}$, which we shall often use.

5. Dividing into diagonal and off-diagonal terms and structure of the proof

By the approximate functional Equation (3-7) and the orthogonality of additive characters, we can decompose $M_{\Upsilon,k}$ into diagonal and off-diagonal terms:

$$M_{\Upsilon,k} := \frac{1}{\varphi(q)} \sum_{a=1}^{q-1} \prod_{i=1}^{k} D_{i;\alpha_i,\beta_i}\big(\tfrac{1}{2}, \tfrac{a}{q}\big) = \mathcal{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} + X_{\boldsymbol{\alpha},\boldsymbol{\beta}} \mathcal{D}_{-\boldsymbol{\alpha},-\boldsymbol{\beta}} + \mathcal{O}_{\boldsymbol{\alpha},\boldsymbol{\beta}} + X_{\boldsymbol{\alpha},\boldsymbol{\beta}} \mathcal{O}_{-\boldsymbol{\alpha},-\boldsymbol{\beta}},$$

where

$$\mathcal{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \frac{i^{|\Upsilon|}}{2^k} \sum_{\varepsilon \in \{\pm 1\}^k} \rho_\Upsilon(\varepsilon) \sum_{\pm_1 n_1 \pm_2 \cdots \pm_k n_k = 0} \frac{\tau_{\alpha_1,\beta_1}(n_1) \cdots \tau_{\alpha_k,\beta_k}(n_k)}{(n_1 \cdots n_k)^{\frac{1}{2}}} V_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left(\frac{n_1 \cdots n_k}{q^k}\right),$$

$$\mathcal{O}_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \frac{i^{|\Upsilon|}}{2^k} \sum_{\varepsilon \in \{\pm 1\}^k} \rho_\Upsilon(\varepsilon) \mathcal{O}'_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}},$$

$$\mathcal{O}'_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{d \mid q} d \frac{\mu(q/d)}{\varphi(q)} \sum_{\substack{d \mid (\pm_1 n_1 \pm_2 \cdots \pm_k n_k) \\ \pm_1 n_1 \pm_2 \cdots \pm_k n_k \neq 0}} \frac{\tau_{\alpha_1,\beta_1}(n_1) \cdots \tau_{\alpha_k,\beta_k}(n_k)}{(n_1 \cdots n_k)^{\frac{1}{2}}} V_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left(\frac{n_1 \cdots n_k}{q^k}\right),$$

and the sum over ε is a sum over $\varepsilon = \{\pm_1 1, \ldots, \pm_k 1\} \in \{1, -1\}^k$.

The diagonal term $\mathcal{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ will be treated in Section 6, using the results of [Bettin 2017]. The terms with $d = 1$ in $\mathcal{O}'_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}$ could be easily dealt with in a simple way, however it is more convenient to keep them together with the other off-diagonal terms.

Lemma 13. *We have*

$$\mathcal{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \mathscr{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} + O(q^{k/2-2k/(k+1)+\varepsilon}), \tag{5-1}$$

where $\mathscr{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is as defined in (6-1).

For the off-diagonal terms we introduce partitions of unity. We need a function $P : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, satisfying

$$\sum_N{}^{\dagger} P(x/N) = 1, \quad \forall x > 0,$$

where by $\sum^{\dagger}$ we mean that the index runs through the elements of a certain (fixed) set of positive real numbers such that $\sum^{\dagger}_{X^{-1} \leq N \leq X} 1 \ll \log X$. Also, we require that $P(x)$ is supported on $1 \leq x \leq 2$ and $P^{(j)}(x) \ll j^{Aj}$ for some $A > 0$. It is not difficult to construct such a partition.[1] Notice that under these conditions, the Mellin transform of $P(x)$,

$$\widetilde{P}(s) := \int_0^\infty P(x) x^{s-1}\, dx,$$

is entire and satisfies

$$\widetilde{P}(\sigma + it) \ll (1 + j + |\sigma|)^{Aj} A^{|\sigma|} (1 + |t|)^{-j}, \quad \forall j \geq 0. \tag{5-2}$$

[1] For example take the set of indexes in $\sum^{\dagger}$ to be $\left\{\left(\frac{3}{2}\right)^n \mid n \in \mathbb{Z}\right\}$ and $P(x) = \int_1^{\frac{3}{2}} \eta(xy)\, \frac{dy}{y}$, where $\eta(x) = Ce^{-1/(1-(4x-7)^2)}$ for $\left|x - \frac{7}{4}\right| < \frac{1}{4}$ and $\eta(x) = 0$ otherwise, and where C is such that $\int_{\mathbb{R}} \eta(y)\, \frac{dy}{y} = 1$.

Using partitions of unity we can decompose $\mathcal{O}'_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ into

$$\mathcal{O}'_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{N_1,\dots,N_k}^{\dagger} \mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}(N_1,\dots,N_k),$$

where $\mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}(N_1,\dots,N_k)$ is defined as $\mathcal{O}'_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}$, with the only difference that the summands are multiplied by $P(n_1/N_1)\cdots P(n_k/N_k)$. In the following we will often omit to indicate the dependencies from $N_1,\dots N_k$ for ease of notation.

The following two Lemmas summarize our results on the off-diagonal terms. The first Lemma, which is effective when $N_1,\dots,N_k$ are close together, uses the spectral theory of automorphic forms (via the bounds proven in Section 3.3) and is proven in Section 7. The second lemma, which is effective when one of the N_i is considerably larger than the others, uses the bounds for sums of Kloosterman sums proven by Young [2011a] and is proven in Section 9.

Lemma 14. *Let $N_{\max}$ be the maximum among $N_1,\dots,N_k$. Then*

$$\sum_{\varepsilon\in\{\pm1\}^k} \rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}(N_1,\dots,N_k) = \mathscr{M}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(N_1,\dots,N_k) + \mathcal{E}_{1;\boldsymbol{\alpha},\boldsymbol{\beta}}(N_1,\dots,N_k), \tag{5-3}$$

where

$$\mathcal{E}_{1;\boldsymbol{\alpha},\boldsymbol{\beta}} \ll \frac{N_{\max}^{\varepsilon}}{q^{1-\varepsilon}}\left(\frac{q^{\vartheta}N_{\max}^{k/2+\frac12}}{(N_1\cdots N_k)^{\frac12}} + \frac{q^{k/2-\frac13+\vartheta/3}N_{\max}^{\frac12}}{(N_1\cdots N_k)^{\frac12}} + \frac{q^{\frac16+\vartheta/3}(N_1\cdots N_k)^{\frac12}}{N_{\max}} + \frac{(N_1\cdots N_k)^{\frac12}}{N_{\max}^{\frac12}}\right) \tag{5-4}$$

and $\mathscr{M}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(N_1,\dots,N_k)$ is defined in (7-38). Moreover,

$$\mathscr{M}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(N_1,\dots,N_k) \ll q^{\varepsilon}(N_1\cdots N_k)^{\frac12}N_{\max}^{-1+\varepsilon}. \tag{5-5}$$

Lemma 15. *Let $N_{\max}$ be the maximum among $N_1,\dots,N_k$. Then*

$$\mathcal{O}''_{\varepsilon,\alpha,\beta}(N_1,\dots,N_k) \ll q^{\varepsilon}\left(\left(\frac{N_1\cdots N_k}{N_{\max}}\right)^{\frac12-1/(2(k-1))}\left(\frac{q^{\frac34}}{N_{\max}^{\frac12}}+1\right) + \frac{(N_1\cdots N_k)^{\frac12}}{N_{\max}^{3/4}}\right).$$

Notice that in the crucial case $N_1\cdots N_k \asymp q^k$ Lemma 14 is nontrivial for $N_{\max} \ll q^{2-2(\vartheta+1)/(k+1)-\delta}$ for any fixed $\delta>0$, whereas Lemma 15 is nontrivial as long as $N_{\max} \gg_k q^{\frac43+\delta}$. In particular, in order to have a nontrivial bound for all ranges we need $\vartheta < \frac{k-2}{3}$ and so for $k=3$ we need $\vartheta < \frac13$.

The following lemma, which we shall prove in Section 8, allows us to combine the various main terms.

Lemma 16. *We have*

$$\frac{i^{|\Upsilon|}}{2^k}\sum_{\substack{N_1,\dots,N_k\\ N_1\cdots N_k\ll q^{k+\varepsilon}}}^{\dagger} (\mathscr{M}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(N_1,\dots,N_k) + X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathscr{M}_{-\boldsymbol{\beta},-\boldsymbol{\alpha}}(N_1,\dots,N_k))$$
$$= -(\mathscr{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} + X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathscr{D}_{-\boldsymbol{\beta},-\boldsymbol{\alpha}}) + \sum_{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}} \mathscr{M}_{\boldsymbol{\alpha}',\boldsymbol{\beta}'} + O(q^{k/2-\frac32+\iota_k+\varepsilon}),$$

where $\mathscr{M}_{\alpha,\beta}$ is as defined in (2-3) and $\iota_k = \frac{3}{14}$ if $k=4$ and $\iota_k=0$ otherwise.

We conclude the section with the deduction of Theorem 5 from the above lemmas.

Proof of Theorem 5. Lemma 14 gives us the asymptotic for $\sum_\varepsilon \rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}$ in the range where the variables are close together. If one variable is much larger than the others then (5-5) and Lemma 15 give us that both $\mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}$ and the main term are small and so we obtain a second formula for $\sum_\varepsilon \rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}$,

$$\sum_{\varepsilon\in\{\pm1\}^k} \rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\alpha,\beta}(N_1,\dots,N_k) = \mathscr{M}_{\alpha,\beta}(N_1,\dots,N_k) + \mathcal{E}_{2;\alpha,\beta}(N_1,\dots,N_k), \tag{5-6}$$

where

$$\mathcal{E}_{2;\alpha,\beta} \ll q^\varepsilon\biggl(\biggl(\frac{N_1\cdots N_k}{N_{\max}}\biggr)^{\frac{1}{2}-1/(2(k-1))}\biggl(\frac{q^{\frac{3}{4}}}{N_{\max}^{\frac{1}{2}}}+1\biggr) + \frac{(N_1\cdots N_k)^{\frac{1}{2}}}{N_{\max}^{\frac{3}{4}}}\biggr).$$

Finally, in the range where $N_1\cdots N_k$ is much smaller than q^k one can improve upon (5-3) and (5-6) by simply bounding trivially $O''_{\varepsilon,\alpha,\beta}$ and $\mathscr{M}_{\alpha,\beta}$ by $q^{-1+\varepsilon}(N_1\cdots N_k)^{\frac{1}{2}}$. We then record here the following third formula for $\sum_\varepsilon \rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}}$:

$$\sum_{\varepsilon\in\{\pm1\}^k} \rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\alpha,\beta}(N_1,\dots,N_k) = \mathscr{M}_{\alpha,\beta}(N_1,\dots,N_k) + \mathcal{E}_{3;\alpha,\beta}(N_1,\dots,N_k), \tag{5-7}$$

with $\mathcal{E}_{3;\alpha,\beta}(N_1,\dots,N_k) \ll q^{-1+\varepsilon}(N_1\cdots N_k)^{\frac{1}{2}}$.

Combining (5-3), (5-6) and (5-7), and adding the condition $N_1\cdots N_k \ll q^{k+\varepsilon}$ at a negligible cost, Lemma 16 gives

$$\begin{aligned}
&\mathcal{O}_{\boldsymbol{\alpha},\boldsymbol{\beta}} + X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathcal{O}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \\
&\qquad= \frac{i^{|\Upsilon|}}{2^k}\sum_{\varepsilon\in\{\pm1\}^k}\rho_\Upsilon(\varepsilon)\sideset{}{^\dagger}\sum_{\substack{N_1,\dots,N_k\\ N_1\cdots N_k\ll q^{k+\varepsilon}}}\bigl(\mathcal{O}'_{\alpha,\beta}(N_1,\dots,N_k) + X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathcal{O}'_{-\boldsymbol{\beta},-\boldsymbol{\alpha}}(N_1,\dots,N_k)\bigr) + O(1) \\
&\qquad= -(\mathscr{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} + X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathscr{D}_{-\boldsymbol{\beta},-\boldsymbol{\alpha}}) + \sum_{\{\alpha'_i,\beta'_i\}=\{\alpha_i,\beta_i\}}\mathscr{M}_{\boldsymbol{\alpha}',\boldsymbol{\beta}'} + \mathscr{E}_{\boldsymbol{\alpha},\boldsymbol{\beta}} + O(q^{k/2-\frac{3}{2}+\iota_k+\varepsilon}),
\end{aligned}$$

where

$$\mathscr{E}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \ll \max_{\substack{N_1,\dots,N_k\\ N_1\cdots N_k\ll q^{k+\varepsilon}}}\bigl(\min(\mathscr{E}_1,\mathscr{E}_2,\mathscr{E}_3)\bigr).$$

Thus, since the term $-(\mathscr{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} + X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathscr{D}_{-\boldsymbol{\beta},-\boldsymbol{\alpha}})$ cancels out with the main term of the diagonal term given by (5-1), to conclude the proof of Theorem 5 we just need to show that $\mathscr{E}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \ll q^{k/2-1-\delta_k+\varepsilon}$. Writing $N_{\max} = q^a$ and $N_1\cdots N_k = q^b$ (and considering only the contribution from the first summand in (5-4), since it easy to see the other terms produce a contribution which is $O(q^{k/2-\frac{3}{2}+\varepsilon})$), we have that it is sufficient to show that

$$\max_{i=1,2,3}\max_{\substack{a\le b\le k\\ ka\ge b}}\min\Bigl(\frac{k+1}{2}a-\frac{b}{2}-1+\vartheta,\, L_i(a,b),\, \frac{b}{2}-1\Bigr) = \frac{k}{2}-\frac{3}{2}+\frac{3(3+2\vartheta)}{2(2k+5)} = \frac{k}{2}-1-\delta_k,$$

for $k\ge 3$, where

$$L_1(a,b) := \frac{3}{4}-\frac{a}{2}+(b-a)\Bigl(\frac{1}{2}-\frac{1}{2(k-1)}\Bigr),\quad L_2(a,b) := (b-a)\Bigl(\frac{1}{2}-\frac{1}{2(k-1)}\Bigr),\quad L_3(a,b) := \frac{b}{2}-\frac{3}{4}a.$$

If the maximum is attained at the interior of $\{a \le b \le k, ka \ge b\}$, then it must occur when $\frac{k+1}{2}a - \frac{b}{2} - 1 = \frac{b}{2} - 1 = L_i(a,b)$ for $i = 1, 2$, or 3 and so it would be $\frac{7k}{20} - \frac{13}{20}$, $\frac{k}{3} - \frac{2}{3}$ and $\frac{k}{3} - \frac{2}{3}$ respectively. Along the lines $a = b$, $ka = b$ and $b = k$ we have

$$\begin{aligned}
&\max_{i=1,2,3}\max_{0\le a\le k}\min\Big(\frac{k}{2}a - 1 + \vartheta,\, L_i(a,a),\, \frac{a}{2} - 1\Big) = \max_{0\le a\le k}\min\Big(L_i(a,a),\, \frac{a}{2} - 1\Big) = 0,\\
&\max_{i=1,2,3}\max_{0\le a\le 1}\min\Big(\frac{a}{2} - 1 + \vartheta,\, L_i(a,ka),\, \frac{ka}{2} - 1\Big) \le -\frac{1}{2} + \vartheta \le 0,\\
&\max_{i=1,2,3}\max_{1\le a\le k}\min\Big(\frac{k+1}{2}a - \frac{k}{2} - 1 + \vartheta,\, L_i(a,k),\, \frac{k}{2} - 1\Big)\\
&\qquad = \max\Big(\frac{k}{2} - \frac{7}{4} + \frac{8k(2+\vartheta) - 19 - 12\vartheta}{4(k^2+2k-4)},\, \frac{k}{2} - \frac{3}{2} + \frac{2(k+\vartheta) - 5 - 4\vartheta}{2(k^2+k-3)},\, \frac{k}{2} - \frac{3}{2} + \frac{3(3+2\vartheta)}{2(2k+5)}\Big)\\
&\qquad = \frac{k}{2} - \frac{3}{2} + \frac{3(3+2\vartheta)}{2(2k+5)},
\end{aligned}$$

for $k \ge 3$ and $\vartheta \le \frac{1}{3}$. Theorem 5 then follows. □

6. The diagonal terms

In this section we prove Lemma 13 deducing it from the following Lemma in [Bettin 2017]. We recall that in Section 4 we assumed $|\alpha_i|, |\beta_i| < \frac{\varepsilon}{2k}$ for all $i = 1, \dots, k$.

Lemma 17. *For $\Re(s) > 1 - \frac{1}{k} - \frac{1}{k}\sum_{i=1}^{k}\min(\Re(\alpha_i), \Re(\beta_i))$, let*

$$\mathcal{W}_{\alpha,\beta}(s) := \frac{i^{|\Upsilon|}}{2^k}\sum_{\varepsilon\in\{\pm1\}^k}\rho_\Upsilon(\varepsilon)\sum_{\pm_1 n_1 \pm_2 \cdots \pm_k n_k = 0}\frac{\tau_{\alpha_1,\beta_1}(n_1)\cdots\tau_{\alpha_k,\beta_k}(n_k)}{(n_1\cdots n_k)^s}.$$

Also, let

$$\begin{aligned}
\mathcal{W}^{\dagger}_{\alpha,\beta}(s) := \sum_{(\mathcal{I},\boldsymbol{\alpha}',\boldsymbol{\beta}')\in S_{\boldsymbol{\alpha},\boldsymbol{\beta}}} \frac{2^{|\mathcal{J}|+1}\pi^{|\mathcal{I}|/2-1}}{|\mathcal{I}|(s-1) + s_{\mathcal{I};\boldsymbol{\alpha}'} + 1}\Bigg(\prod_{i\in\mathcal{I}}\zeta(1-\alpha_i'+\beta_i')\frac{\Gamma_i\big(-\alpha_i'/2 + (1+s_{\mathcal{I};\boldsymbol{\alpha}'})/(2|\mathcal{I}|)\big)}{\Gamma_i\big(\frac{1}{2} + \alpha_i'/2 - (1+s_{\mathcal{I};\boldsymbol{\alpha}'})/(2|\mathcal{I}|)\big)}\Bigg)\\
\times\sum_{\ell\ge1}\sum_{h \pmod{\ell}}^{*}\frac{1}{\ell^{|\mathcal{I}'|-\sum_{i\in\mathcal{I}}(\alpha_i'-\beta_i')}}\prod_{i\notin\mathcal{I}}D_i\big(1+\alpha_i' - (1+s_{\mathcal{I};\boldsymbol{\alpha}'})/|\mathcal{I}|,\, \alpha_i'-\beta_i',\, h/\ell\big).
\end{aligned}$$

where $s_{\mathcal{I};\boldsymbol{\alpha}'} := \sum_{i\in\mathcal{I}}\alpha_i'$ and

$$S_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \left\{(\mathcal{I},\boldsymbol{\alpha}',\boldsymbol{\beta}')\;\middle|\; \begin{array}{l}\mathcal{I}\subseteq\{1,\dots k\},\ |\mathcal{I}| > |\mathcal{J}|+1,\ |\mathcal{I}\cap\Upsilon|\ \text{even},\\ \{\alpha_i',\beta_i'\} = \{\alpha_i,\beta_i\}\ \forall i\in\mathcal{I},\ (\alpha_i',\beta_i') = (\alpha_i,\beta_i)\ \forall i\notin\mathcal{I}\end{array}\right\}.$$

Then for any $\varepsilon > 0$, $\mathcal{W}_{\alpha,\beta}(s) - \mathcal{W}^{\dagger}_{\alpha,\beta}(s)$ extends to a holomorphic function on $\Re(s) \ge 1 - \frac{2-4\varepsilon}{k+1}$ and in such a half plane it satisfies $\mathcal{W}_{\alpha,\beta}(s) - \mathcal{W}^{\dagger}_{\alpha,\beta}(s) \ll \big(\frac{k}{\varepsilon}(1+|s|)\big)^{Ak}$.

Proof. Theorem 3 of [Bettin 2017] gives the meromorphic continuation and the bound for each ε. Thus, one obtains the lemma by summing over ε (for the simplification of the polar term one proceeds as in Lemma 23; see also Remark 2 of [Bettin 2017]). □

Proof of Lemma 13. Writing $V_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ in terms of it's Mellin transform we have

$$\mathcal{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \frac{1}{2\pi i}\int_{(2)} G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)\mathcal{W}_{\alpha,\beta}\big(\tfrac{1}{2}+s\big)q^{ks}\,\frac{ds}{s}.$$

We write $\mathcal{W}_{\alpha,\beta}\big(\frac{1}{2}+s\big)$ as $\mathcal{W}^{\dagger}_{\alpha,\beta}\big(\frac{1}{2}+s\big)+\big(\mathcal{W}_{\alpha,\beta}\big(\frac{1}{2}+s\big)-\mathcal{W}^{\dagger}_{\alpha,\beta}\big(\frac{1}{2}+s\big)\big)$. For the second term we move the line of integration to $\Re(s)=\frac{1}{2}-(2-4\varepsilon)/(k+1)$ and bound trivially using (4-2) obtaining an error of size $O(k^{Ak}q^{k/2-2k/(k+1)+\varepsilon}) = O(q^{k/2-2k/(k+1)+\varepsilon})$. For the first term we move the line of integration to $\Re(s)=-\frac{1}{2}$ picking up the residues from the poles. We obtain (5-1) with

$$\mathcal{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{\substack{\mathcal{I}\cup\mathcal{J}=\{1,\dots,k\},\ \mathcal{I}\cap\mathcal{J}=\varnothing\\ |\mathcal{I}|>|\mathcal{J}|+1,\ |\mathcal{I}\cap\Upsilon|\text{even}}}\ \sum_{\substack{\{\alpha'_i,\beta'_i\}=\{\alpha_i,\beta_i\}\ \forall i\in\mathcal{I}\\ (\alpha'_j,\beta'_j)=(\alpha_j,\beta_j)\ \forall j\in\mathcal{J}}} \mathscr{D}_{\mathcal{I};\boldsymbol{\alpha}',\boldsymbol{\beta}'} \tag{6-1}$$

where

$$\begin{aligned}\mathscr{D}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} := 2^k\frac{G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s_{\mathcal{I};\boldsymbol{\alpha}})}{s_{\mathcal{I};\boldsymbol{\alpha}}\pi^{|\mathcal{I}|}} g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s_{\mathcal{I};\boldsymbol{\alpha}})q^{ks_{\mathcal{I};\boldsymbol{\alpha}}}\Bigg(\prod_{i\in\mathcal{I}}\frac{\pi^{\frac{1}{2}}\zeta(1-\alpha_i+\beta_i)}{2^{\frac{1}{2}+\alpha_i+s_{\mathcal{I},\boldsymbol{\alpha}}}}\frac{\Gamma_i\big(\frac{1}{4}-(\alpha_i+s_{\mathcal{I},\boldsymbol{\alpha}})/2\big)}{\Gamma_i\big(\frac{1}{4}+(\alpha_i+s_{\mathcal{I},\boldsymbol{\alpha}})/2\big)}\Bigg)\\ \times\sum_{\ell}\ \sideset{}{^*}\sum_{h\ (\mathrm{mod}\ \ell)}\frac{1}{\ell^{|\mathcal{I}|-\sum_{i\in\mathcal{I}}(\alpha_i-\beta_i)}}\Bigg(\prod_{j\in\mathcal{J}}D_j\big(\tfrac{1}{2}+\alpha_j+s_{\mathcal{I},\boldsymbol{\alpha}},\ \alpha_j-\beta_j,\ \pm_j\tfrac{h}{\ell}\big)\Bigg)\end{aligned} \tag{6-2}$$

and $s_{\mathcal{I};\boldsymbol{\alpha}} := \sum_{i\in\mathcal{I}}\alpha_i$. □

7. The terms close to the diagonal

In this section we prove Lemma 14. First, we assume that N_1 is the maximum of $N_1,\dots,N_k$, as we can do since both the main term and the error terms in Lemma 14 are symmetric in the indexes. Moreover, since we assumed that $|\Upsilon|$ is even, then we have

$$\sum_{\varepsilon\in\{\pm1\}^k}\rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\alpha,\beta} = 2\sum_{\substack{\varepsilon\in\{\pm1\}^k\\ \pm_1 1=-1}}\rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}},$$

where here and in the following $\varepsilon=(\pm_1 1,\pm_2 1,\dots,\pm_k 1)$. We split $\mathcal{O}''_{\varepsilon,\alpha,\beta}$ further, depending on the sign and the size of $\pm_* f := -n_1\pm_2 n_2\pm_3\cdots\pm_k n_k$ (with $f>0$), introducing another partition of unity controlling the size of f:

$$\sum_{\varepsilon\in\{\pm1\}^k}\rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\boldsymbol{\alpha},\boldsymbol{\beta}} = 2\sideset{}{^\dagger}\sum_{N_*\ll kN_1q^{\varepsilon/k}}\ \sum_{\substack{\varepsilon\in\{\pm1\}^k\\ \pm_1=-1}}\rho_\Upsilon(\varepsilon)\sum_{\pm_*1\in\{\pm1\}}\mathcal{K}_{\varepsilon,\pm_*,\boldsymbol{\alpha},\boldsymbol{\beta}}, \tag{7-1}$$

where

$$\begin{aligned}\mathcal{K}_{\varepsilon,\pm_*;\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{d\mid q} d\frac{\mu(q/d)}{\varphi(q)}\sum_{\substack{f\geq1,\\ f\equiv0\ (\mathrm{mod}\ d)}}\ \sum_{\substack{n_1,\dots,n_k\geq1\\ n_1=\pm_2n_2\pm_3\cdots\pm_kn_k\pm_*f}}\frac{\tau_{\alpha_1,\beta_1}(n_1)\cdots\tau_{\alpha_k,\beta_k}(n_k)}{(n_1\cdots n_k)^{\frac{1}{2}}}\\ \times V_{\boldsymbol{\alpha},\boldsymbol{\beta}}\bigg(\frac{n_1\cdots n_k}{q^k}\bigg)P\bigg(\frac{n_1}{N_1}\bigg)\cdots P\bigg(\frac{n_k}{N_k}\bigg)P\bigg(\frac{f}{N_*}\bigg).\end{aligned}$$

Notice that in (7-1) we truncated the sum over N_* at $N_* \ll kN_1 q^{\varepsilon/k}$, as we clearly could.

7.1. ***Separating the variables arithmetically.*** We wish to separate the variables in

$$\tau_{\alpha_1,\beta_1}(n_1) = \tau_{\alpha_1,\beta_1}(\pm_2 n_2 \pm_3 \cdots \pm_k n_k \pm_* f).$$

One can achieve this goal by using Ramanujan's identity

$$\tau_{a,b}(n) = n^{-a}\tau_{0,b-a}(n) = n^{-a}\zeta(1-a+b)\sum_{\ell=1}^{\infty}\frac{c_\ell(n)}{\ell^{1-a+b}}, \tag{7-2}$$

which holds for $n \neq 0$ and $\Re(a-b) < 0$. The coefficient $c_\ell(n)$ denotes the Ramanujan sum

$$c_\ell(n) := \sideset{}{^*}\sum_{h \pmod{\ell}} \mathrm{e}\Big(\frac{nh}{\ell}\Big).$$

However, since (7-2) doesn't hold in a neighborhood of $a = b = 0$, it is more convenient to follow Young's approach and use the following lemma, which rephrases (7-2) as an approximate functional equation for $\tau_{a,b}(n)$.

Lemma 18. *Let $n \in \mathbb{Z}_{>0}$ and let $a, b \in \mathbb{C}$. Then,*

$$\tau_{a,b}(n) = n^{-a}\sum_{\ell}\frac{c_\ell(n)}{\ell^{1-a+b}}\upsilon_{a-b}\Big(\frac{\ell^2}{n}\Big) + n^{-b}\sum_{\ell}\frac{c_\ell(n)}{\ell^{1+a-b}}\upsilon_{b-a}\Big(\frac{\ell^2}{n}\Big) \tag{7-3}$$

where

$$\upsilon_a(x) = \int_{(c_w)} x^{-w/2}\zeta(1-a+w)\frac{G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w)}{w}\,dw,$$

where $c_w > |\Re(a-b)|$ and $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w)$ is as defined in (4-1).

Proof. See Lemma 5.4 of Young [2011a]. □

Applying (7-3) and splitting the resulting sum over ℓ using another partition of unity (and adding the restriction $L \geq \frac{1}{2}$ as we can do since P is supported on $[1, 2]$), we rewrite $\mathcal{K}_{\varepsilon,\pm_*;\boldsymbol{\alpha},\boldsymbol{\beta}}$ as

$$\mathcal{K}_{\varepsilon,\pm_*;\boldsymbol{\alpha},\boldsymbol{\beta}} = \sum_{\substack{\{\alpha_1',\beta_1'\}=\{\alpha_1,\beta_1\}\\ (\alpha_j',\beta_j')=(\alpha_j,\beta_j)\ \forall j\neq 1}}\ \sideset{}{^\dagger}\sum_{L\geq\frac{1}{2}} \mathcal{L}_{\boldsymbol{\alpha}',\boldsymbol{\beta}'}, \tag{7-4}$$

where $\boldsymbol{\alpha}' := (\alpha_1', \ldots, \alpha_k')$, $\boldsymbol{\beta}' := (\beta_1', \ldots, \beta_k')$ and

$$\mathcal{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{\substack{n_1,\ldots,n_k,\ f\geq 1\\ n_1=\pm_2 n_2\pm_3\cdots\pm_k n_k\pm_* f}}\ \sum_{\substack{d\mid q\\ d\mid f}}\frac{\mu(q/d)d}{\varphi(q)}\sum_{\ell}\ \sideset{}{^*}\sum_{h \pmod{\ell}}\frac{c_\ell(\pm_2 n_2\pm_3\cdots\pm_k n_k\pm_* f)}{\ell^{1-\alpha+\beta}}\upsilon_{\alpha_1-\beta_1}\Big(\frac{\ell^2}{n_1}\Big)$$
$$\times\frac{\tau_{\alpha_2,\beta_2}(n_2)\cdots\tau_{\alpha_k,\beta_k}(n_k)}{n_1^{\frac{1}{2}}\cdots n_k^{\frac{1}{2}}}V_{\boldsymbol{\alpha},\boldsymbol{\beta}}\Big(\frac{n_1\cdots n_k}{q^k}\Big)P\Big(\frac{n_1}{N_1}\Big)\cdots P\Big(\frac{n_k}{N_k}\Big)P\Big(\frac{f}{N_*}\Big)P\Big(\frac{\ell}{L}\Big).$$

Notice that we have omitted to indicate the dependency of $\mathcal{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ from ε and $\pm_*$ in order to save notation.

Expressing P, $\upsilon_{\alpha_1-\beta_1}$ and V in terms of their Mellin transform and making the change of variables $u_i \to u_i - s$, for $i = 1, \dots, k$, we see that $\mathcal{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ can be written as

$$\mathcal{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \sum_{d\,|\,q} \frac{\mu(q/d)d}{\varphi(q)} \sum_{\substack{n_2,\dots,n_k,\, f\geq 1,\, d\,|\,f\\ \pm_2 n_2 \pm_3 \cdots \pm_k n_k \pm_* f>0}} \sum_{\ell} \sum_{h \pmod{\ell}}^{*} \frac{1}{(2\pi i)^{k+3}}$$
$$\times \int_{(c_s, c_w, \boldsymbol{c_u}, c_{u_*})} \frac{N_*^{u_*}}{f^{u_*}} P\Big(\frac{\ell}{L}\Big) \frac{N_1^{u_1-s}\cdots N_k^{u_k-s}}{\ell^{1-\alpha_1+\beta_1+w}} \frac{\tau_{\alpha_2,\beta_2}(n_2)\cdots\tau_{\alpha_k,\beta_k}(n_k) c_\ell(\pm_2 n_2 \pm_3 \cdots \pm_k n_k \pm_* f)}{(\pm_2 n_2 \pm_3 \cdots \pm_k n_k \pm_* f)^{\frac12+\alpha_1+u_1-w/2} n_2^{\frac12+u_2}\cdots n_k^{\frac12+u_k}}$$
$$\times \widetilde{P}(u_*)\widetilde{P}(u_1-s)\cdots\widetilde{P}(u_k-s) q^{ks} \frac{H_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)}{ws}\, dw\, ds\, \boldsymbol{du}\, du_*, \tag{7-5}$$

where $\boldsymbol{du} := du_1 \cdots du_k$, $\boldsymbol{c_u}$ denotes the lines of integration $c_{u_1}, \dots, c_{u_k}$ and

$$H_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s) := \zeta(1+w-\alpha_1+\beta_1) G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w) g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s).$$

Notice that, by the definitions (4-1) and (3-6) of $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)$ and $g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)$, $H_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)$ is entire and decays rapidly in both variables w and s:

$$H_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s) \ll e^{-C_2(|\Im(s)|+|\Im(w)|)}(1+|\Re(s)|+|\Re(w)|)^{A(|\Re(s)|+|\Re(w)|+k)}, \tag{7-6}$$

for some $C_2 > 0$. As lines of integration, we take

$$c_s := \varepsilon/k, \quad c_{u_1} = -3k - \tfrac12 - \alpha_1 + 7\varepsilon, \quad c_{u_*} = c_{u_2} = \cdots = c_{u_k} = 4k, \quad c_w = 10\varepsilon.$$

The real parts of the lines are chosen to be large enough so that the various sums are absolutely convergent.

7.2. ***Separating the variables analytically.*** To complete the separation of the variables, we need also to deal with the factor $(\pm_2 n_2 \pm_3 \cdots \pm_k n_k \pm_* f)^{\frac12+\alpha_1+u_1-w/2}$ in (7-5). In order to do so, we use Lemma 27, in Section 10. We apply the lemma with $\kappa := k+1$, $B := 3k$ and $\upsilon_1 = \frac12 - \alpha_1 - u_1 + \frac{w}{2}$, so that $\Re(\upsilon_1) = B + 1 - 2\varepsilon$. We get

$$\mathcal{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \sum_{\nu} \frac{B!}{\nu_2! \cdots \nu_k! \nu_*!} (\mathcal{N}_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} + \mathcal{N}'_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}), \tag{7-7}$$

where the sum is over $\nu = (\nu_2, \dots, \nu_k, \nu_*) \in \mathbb{Z}^k_{\geq 0}$ satisfying

$$\nu_2 + \cdots + \nu_k + \nu_* = B, \qquad \nu_i = 0 \quad \text{if } \pm_i 1 = -1, \qquad \nu_* = 0 \quad \text{if } \pm_* 1 = -1,$$

and $\mathcal{N}_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}$ is defined by

$$\mathcal{N}_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{d|q} \frac{\mu(q/d)d}{\varphi(q)} \sum_{\substack{n_1,\dots,n_k,\, \ell\geq 1\\ f\geq 1,\, d\,|\,f}} \frac{P(\ell/L)}{(2\pi i)^{2k+3}} \int_{(c_s, c_w, \boldsymbol{c_u}, c_{u_*}, c_{\upsilon_*})} \frac{c_\ell(\pm_2 n_2 \pm_3 \cdots \pm_k n_k \pm_* f)}{f^{\upsilon_*+u_*-\nu_*}} N_*^{u_*}$$
$$\times \frac{q^{ks} N_1^{u_1-s}}{\ell^{1-\alpha_1+\beta_1+w}} \widetilde{P}(u_*)\widetilde{P}(u_1-s) \bigg(\prod_{i=2}^{k} \int_{(c_{\upsilon_i})} \frac{\tau_{\alpha_i,\beta_i}(n_i) N_i^{u_i-s}}{n_i^{\frac12+u_i+\upsilon_i-\nu_i}} \widetilde{P}(u_i-s)\bigg)$$
$$\times \Psi_{\varepsilon^*,B}\big(\tfrac12 - \alpha_1 - u_1 + w/2, \boldsymbol{\upsilon}, \upsilon_*\big) \frac{H_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)}{ws}\, ds\, dw\, \boldsymbol{du}\, \boldsymbol{d\upsilon}\, du_*\, d\upsilon_*, \tag{7-8}$$

with $c_{v_2} = \cdots = c_{v_k} = c_{v_*} = \varepsilon/k$, and $\mathcal{N}'_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}$ is defined in the same way with lines of integrations $c'_{v_2} = \cdots = c'_{v_k} = c'_{v_*} = \frac{1}{2}$ in place of $c_{v_2}, \ldots, c_{v_k}, c_{v_*}$. Also, in (7-8) we used the notation $\boldsymbol{v} := (v_2, \ldots, v_k)$, $\boldsymbol{dv} := dv_2 \cdots dv_k$ and $\varepsilon^* := (\pm_1 1, \ldots, \pm_k 1, \pm_* 1)$ and $\Psi_{\varepsilon^*,B}$ is as in (10-5).

The contribution of $\mathcal{N}'_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}$ can be bounded by moving the lines of integration c_{u_i} to $c_{u_i} = 2\varepsilon + \nu_i$ for $i = 2, \ldots, k$ and c_{u_*} to $c_{u_*} = \frac{1}{2} + \nu_* + \varepsilon$ and bounding trivially. We obtain

$$\mathcal{N}'_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} \ll q^{-1+\varepsilon} N_1^{-B-\frac{1}{2}+A\varepsilon} N_*^{\frac{1}{2}+\nu_*+2\varepsilon} N_2^{\nu_2+2\varepsilon} \cdots N_k^{\nu_k+2\varepsilon} L^{-\varepsilon}$$

and thus

$$\sum_{\nu} \frac{B!}{\nu_2! \cdots \nu_k! \nu_*!} \mathcal{N}'_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} \ll q^{-1+A\varepsilon} N_1^{A\varepsilon} L^{-\varepsilon},$$

since N_1 is the maximum among $N_1, \ldots, N_k$ and $N_* \ll k N_1 q^{\varepsilon/k}$.

Next, we open the Ramanujan sum in (7-8) and we execute the sums over $n_2, \ldots, n_k, f$ as we can do since the integrals and sums are absolutely convergent since $\nu_i, \nu_* \le B = 3k$ and $c_{u_i} = c_{u_*} = 4k$ for all i. We obtain

$$\mathcal{N}_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} = \sum_{d|q} \frac{\mu(q/d)}{\varphi(q)} \sum_{\ell} \frac{P(\ell/L)}{(2\pi i)^{2k+3}} \int_{(c_s, c_w, \boldsymbol{c_u}, c_{u_*}, c_{v_*})} \sum_{h \ (\mathrm{mod}\ \ell)}^{*} \frac{d^{1-\nu_*-u_*+v_*}}{\ell^{1-\alpha_1+\beta_1+w}} F\big(v_* + u_* - \nu_*, \pm_* \tfrac{dh}{\ell}\big) q^{ks}$$
$$\times N_*^{u_*} \widetilde{P}(u_*) \Bigg(\prod_{i=2}^{k} \int_{(c_{v_i})} D_{\alpha_i,\beta_i}\big(\tfrac{1}{2} + u_i + v_i - \nu_i, \tfrac{\pm_i h}{\ell}\big) \widetilde{P}(u_i - s) N_i^{u_i - s} \Bigg)$$
$$\times N_1^{u_1-s} \widetilde{P}(u_1 - s) \Psi_{\varepsilon^*,B}\big(\tfrac{1}{2} - \alpha_1 - u_1 + \tfrac{w}{2}, \boldsymbol{v}, v_*\big) \frac{H_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)}{ws} \, ds \, dw \, \boldsymbol{du} \, \boldsymbol{dv} \, du_* \, dv_*,$$

where, after moving the lines of integration $c_{u_2}, \ldots, c_{u_k}, c_{u_*}$, we have

$$\begin{gathered} c_s := \varepsilon/k, \quad c_{u_1} = -B - \tfrac{1}{2} - \Re(\alpha_1) + 7\varepsilon, \quad c_w = 10\varepsilon, \\ c_{v_2} = \cdots = c_{v_k} = c_{v_*} = \varepsilon/k, \quad c_{u_*} = 1 + \nu_* + 2\varepsilon - \varepsilon/k, \end{gathered} \tag{7-9}$$

and $c_{u_i} = \frac{1}{2} + \nu_i + \varepsilon/k$, for $i = 2, \ldots, k$.

Remarks. (1), Thanks to (7-6) and to Lemma 28 in Section 10, the integrals in $\mathcal{N}_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}$ are all absolutely convergent when the line of integration are chosen so that $\Re(v_2) = \cdots = \Re(v_k) = \Re(v_*) = \varepsilon/k$ and $\Re(v_1) := \Re\big(\frac{1}{2} - \alpha_1 - u_1 + \frac{w}{2}\big) = B + 1 - 2\varepsilon$ (and even if an extra factor of $\prod_{i=2}^{k}(1 + |u_i| + |v_i|)^{1+4\varepsilon}$ is introduced inside the integrals, as will be relevant later on in the argument). In the following computations, until Lemma 20, we will (almost) always arrange the lines of integration in a way such that $\Re(v_1)$ is kept equal to $B + 1 - 2\varepsilon$.[2] This ensures the absolute convergence of the integrals in all the bounds we give.

(2) We also observe that, by the definition (10-5), the poles of $\Psi_{\varepsilon^*,B}(v_1, \boldsymbol{v}, v_*)$ are contained in the set

$$\big\{(v_1, \boldsymbol{v}, v_*) \in \mathbb{C}^{k+1} \mid v_i \in \mathbb{Z}_{\le 0} \text{ for some } i \in \{1, \ldots, k\} \text{ or } v_1 + \cdots + v_k + v_* = B + 1\big\}.$$

[2] The only exception is in the proof of (7-19), where we need to take $\Re(v_1) = B - 2\varepsilon$. One can easily verify however that the integrals are all absolutely convergent also in that case.

(3) One should morally think of having $B = \nu_* = \nu_2 = \cdots = \nu_k = 0$, as their presence is just an artificial effect of forcing the various integrals over v_i to be absolutely convergent. Also, we chose and shall keep c_{u_*} in a way so that we stay just to the right of the pole of F. Aside from this, in the following computations our goal will typically be that of moving $c_{u_2}, \dots, c_{u_k}$ to the left thus obtaining savings in $N_2, \dots, N_k$. Since $D\big(s, \frac{h}{\ell}\big)$ grows roughly like $\ell^{1-\Re(s)}$ when $0 < \Re(s) < 1$, we then need to move w to the right to insure the convergence of the sum over ℓ. This in turn forces us to move c_{u_1} to the right since we need $\Re\big(\frac12 - \alpha_1 - u_1 + \frac{w}{2} + v_2 + \cdots + v_k + v_*\big) < B + 1$ to avoid a pole of $\Psi_{\varepsilon^*,B}$. Doing so we lose a power of N_1; however, since in the first argument of $\Psi_{\varepsilon^*,B}$ u_1 appears with a coefficient which is (negative the) double of that of w, we have that the gain in the exponents of $N_2, \dots, N_k$ is superior to the loss in the exponent of N_1. This will then produce a saving when the variables are close to the diagonal, that is when N_1 is not much larger than $(N_2 \cdots N_k)^{1/(k-1)}$.

7.3. ***Picking up the residues of the Estermann function.*** For each $i = 2, \dots, k$ we move the line of integration c_{u_i} to $c_{u_i} = -\frac12 + \nu_i - 2\varepsilon$, passing through the poles of the Estermann function at

$$u_i = \tfrac12 - \alpha_i - v_i + \nu_i \quad \text{and} \quad u_i = \tfrac12 - \beta_i - v_i + \nu_i .$$

By Lemma 6 and the residue theorem, we obtain

$$\mathcal{N}_{\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} = \sum_{\substack{I \cup J = \{2,\dots,k\} \\ I \cap J = \varnothing}} \sum_{\substack{\{\alpha'_i,\beta'_i\} = \{\alpha_i,\beta_i\}\ \forall i \in I \\ (\alpha'_j,\beta'_j) = (\alpha_j,\beta_j)\ \forall j \in J \cup \{1\}}} \mathcal{P}_{I;\nu;\boldsymbol{\alpha}',\boldsymbol{\beta}'}, \tag{7-10}$$

where, for $I \cup J = \{2, \dots, k\},\ I \cap J = \varnothing$,

$$\begin{aligned}\mathcal{P}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{d \mid q} \frac{\mu(q/d)}{\varphi(q)} \sum_{\ell} \frac{P(\ell/L)}{(2\pi i)^5} \int_{(c_s,c_w,c_{u_1},c_{u_*},c_{v_*})} \frac{q^{ks} d^{1-v_*-u_*+\nu_*}}{\ell^{\sum_{i \in I \cup \{1\}}(1-\alpha_i+\beta_i)+w}} \sideset{}{^*}\sum_{h \ (\mathrm{mod}\ \ell)} F\big(v_* + u_* - \nu_*, \pm_* \tfrac{dh}{\ell}\big) \\ \times \bigg(\prod_{j \in J} \frac{1}{(2\pi i)^2} \int_{(c_{u_j},c_{v_j})} D_{\alpha_j,\beta_j}\big(\tfrac12 + u_j + v_j - \nu_j, \tfrac{\pm_j h}{\ell}\big) \widetilde{P}(u_j - s) N_j^{u_j - s}\, du_j\bigg) \\ \times \bigg(\prod_{i \in I} \frac{1}{2\pi i} \int_{(c_{v_i})} \widetilde{P}\big(\tfrac12 - \alpha_i - v_i + \nu_i - s\big) N_i^{\frac12 - \alpha_i - v_i + \nu_i - s}\bigg) N_*^{u_*} \widetilde{P}(u_*) N_1^{u_1 - s} \widetilde{P}(u_1 - s) \\ \times \Psi_{\varepsilon^*,B}\big(\tfrac12 - \alpha_1 - u_1 + \tfrac{w}{2}, \boldsymbol{v}, v_*\big) \frac{H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)}{ws}\, ds\, dw\, du_1\, \boldsymbol{dv}\, du_*\, dv_*\end{aligned}$$

and

$$H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s) := G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(w) g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) \zeta(1 + w - \alpha_1 + \beta_1) \prod_{i \in I} \zeta(1 - \alpha_i + \beta_i), \tag{7-11}$$

so that

$$H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s) \ll (A \log q)^{|I|} e^{-C_2(|\Im(s)| + |\Im(w)|)} (1 + |\Re(s)| + |\Re(w)|)^{A(|\Re(s)| + |\Re(w)| + k)}. \tag{7-12}$$

We remind also that the lines of integrations are given by (7-9) and

$$c_{u_j} = -\tfrac{1}{2} + \nu_j - 2\varepsilon, \quad \text{for all } j \in J. \tag{7-13}$$

7.4. ***Applying the bounds on sums of Kloosterman sums.*** In this section, we apply Lemma 12 to give a bound for $\mathcal{P}_{I,\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}$ under certain conditions.

Lemma 19. *Let $I \subseteq \{2, \dots, k\}$ and let $J := \{2, \dots, k\} \setminus I$. Then, if $|I| \le |J|$ we have*

$$\mathcal{P}_{I,\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} \ll q^{-1+A\varepsilon} N_1^{A\varepsilon} \big(q^{\vartheta} N_1^{(k+1)/2} + q^{k/2-\frac{1}{3}+\vartheta/3} N_1^{\frac{1}{2}}\big)(N_1 \cdots N_k)^{-\frac{1}{2}} L^{-\varepsilon}, \tag{7-14}$$

whereas if $|I| > |J|$ and $\nu_j > 0$ for some $j \in J$, then

$$\mathcal{P}_{I,\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} \ll q^{-\frac{5}{6}+\vartheta/3+A\varepsilon} N_1^{-1+A\varepsilon} (N_1 \cdots N_k)^{\frac{1}{2}} L^{-\varepsilon}. \tag{7-15}$$

Proof. First, we bound the sums over h and ℓ by Lemma 12 and we bound trivially the integrals which are all convergent by (5-2), (7-12) and (10-6) when the lines of integrations are given by (7-9) and (7-13). Doing so, we obtain

$$\begin{aligned}
\mathcal{P}_{I,\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} &\ll q^{-\frac{5}{6}+\vartheta/3+A\varepsilon} N_1^{-B-\frac{1}{2}+A\varepsilon} N_*^{1+\nu_*+\varepsilon} \Big(\prod_{i\in I} N_i^{\frac{1}{2}+\nu_i}\Big)\Big(\prod_{j\in J} N_j^{-\frac{1}{2}+\nu_j}\Big) L^{|J|-|I|-\varepsilon} \\
&\quad \times \int_{(c_s,c_w,c_{u_1},c_{u_*},c_{v_*})} \Big(\prod_{j\in J} \int_{(c_{u_j})} (1+|v_j|+|u_j|)^{1+4\varepsilon} |\widetilde{P}(u_j - s)||du_j|\Big) \\
&\quad \times \Big(\prod_{i\in I} \int_{(c_{v_i})} \big|\widetilde{P}\big(\tfrac{1}{2} - \alpha_i - v_i + \nu_i - s\big)\big|\Big) |\widetilde{P}(u_*)||\widetilde{P}(u_1 - s)| \\
&\quad \times \big|\Psi_{\varepsilon^*,B}\big(\tfrac{1}{2} - \alpha_1 - u_1 + \tfrac{w}{2}, \boldsymbol{v}, v_*\big)\big| \frac{|H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)|}{|ws|} |\, ds\, dw\, du_1\, \boldsymbol{dv}\, du_*\, dv_*| \\
&\ll q^{-\frac{5}{6}+\frac{\vartheta}{3}+A\varepsilon} N_1^{-B-\frac{1}{2}+A\varepsilon} N_*^{1+\nu_*} \Big(\prod_{i\in I} N_i^{\frac{1}{2}+\nu_i}\Big)\Big(\prod_{j\in J} N_j^{-\frac{1}{2}+\nu_j}\Big) L^{|J|-|I|-\varepsilon}.
\end{aligned} \tag{7-16}$$

If $|I| - |J| > 0$ and at least one of the ν_j is greater than zero, then this is bounded by

$$\begin{aligned}
\mathcal{P}_{I,\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} &\ll q^{-\frac{5}{6}+\vartheta/3+A\varepsilon} N_1^{A\varepsilon} \frac{N_*^{\nu_*+1}}{N_1^{\nu_*+1}} (N_1 \cdots N_k)^{\frac{1}{2}} \Big(\prod_{i\in I} \frac{N_i^{\nu_i}}{N_1^{\nu_i}}\Big)\Big(\prod_{j\in J} \frac{N_j^{\nu_j-1}}{N_1^{\nu_j}}\Big) L^{-\varepsilon} \\
&\ll q^{-\frac{5}{6}+\vartheta/3+A\varepsilon} N_1^{-1+A\varepsilon} (N_1 \cdots N_k)^{\frac{1}{2}} L^{-\varepsilon},
\end{aligned}$$

since $B = \nu_1 + \cdots + \nu_k$ and $N_2, \dots, N_k \le N_1$, $N_* \ll k N_1 q^{\varepsilon/k}$.

Now assume $|I|-|J|\le 0$ and let $L\le q^{\frac12}$. In this case (7-16) gives

$$\begin{aligned}\mathcal{P}_{I,\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} &\ll q^{-\frac56+\vartheta/3+A\varepsilon}N_1^{\frac12+|I|+A\varepsilon}\frac{N_*^{\nu_*+1}}{N_1^{\nu_*+1}}(N_1\cdots N_k)^{-\frac12}\biggl(\prod_{i\in I}\frac{N_i^{1+\nu_i}}{N_1^{1+\nu_i}}\biggr)\biggl(\prod_{j\in J}\frac{N_j^{\nu_j}}{N_1^{\nu_j}}\biggr)q^{\frac12(|J|-|I|)-\varepsilon}\\ &\ll q^{-\frac56+\vartheta/3+\frac12(|J|-|I|)+A\varepsilon}N_1^{\frac12+|I|+A\varepsilon}(N_1\cdots N_k)^{-\frac12}\\ &\ll q^{-\frac13+\vartheta/3+A\varepsilon}(N_1^{k/2+A\varepsilon}q^{-\frac12}+q^{k/2-1}N_1^{\frac12})(N_1\cdots N_k)^{-\frac12},\end{aligned}\tag{7-17}$$

since $|I|=k-1-|J|$ and $\frac{k-1}{2}\le|J|\le k-1$.

Finally, if $|I|-|J|\le 0$ and $L>q^{\frac12}$, then we move the lines of integration c_w and c_{u_1} to

$$c_w=|J|-|I|+10\varepsilon=k-1-2|I|+10\varepsilon,\quad c_{u_1}=-1-B+\tfrac{k}{2}-\Re(\alpha_1)-|I|+7\varepsilon.$$

Then, we use Lemma 12 and bound trivially the integrals (using (5-2), (7-12) and (10-6)) and we obtain

$$\begin{aligned}\mathcal{P}_{I,\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} &\ll q^{-1+\vartheta+A\varepsilon}N_1^{-1-B+k/2-|I|+A\varepsilon}N_*^{1+\nu_*}\biggl(\prod_{i\in I}N_i^{\frac12+\nu_i}\biggr)\biggl(\prod_{j\in J}N_j^{-\frac12+\nu_j}\biggr)L^{-\varepsilon}\\ &\ll\frac{q^{-1+\vartheta+A\varepsilon}}{L^{\varepsilon}}N_1^{\frac{k}{2}+A\varepsilon}\frac{N_*^{\nu_*+1}}{N_1^{\nu_*+1}}\biggl(\prod_{i\in I}\frac{N_i}{N_1}\biggr)\biggl(\prod_{i=2}^{k}\frac{N_i^{-\frac12+\nu_i}}{N_1^{\nu_i}}\biggr)\ll\frac{q^{-1+\vartheta+A\varepsilon}N_1^{k/2+\frac12+A\varepsilon}}{(N_1\cdots N_k)^{\frac12}L^{\varepsilon}}.\end{aligned}\tag{7-18}$$

Thus, since $N_1^{k/2}q^{-\frac12}\ll q^{k/2-1}N_1^{\frac12}+q^{-1}N_1^{k/2+\frac12}$, we have that (7-17) and (7-18) imply (7-14). □

7.5. *Reassembling the sum over ν and further manipulations.* By the previous section, we only need to consider the $\mathcal{P}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}$ with $|I|>|J|$ and $\nu_j=0$ for all $j\in J$ (and lines of integration given in (7-9) and (7-13)). For each $j\in J$, we move c_{u_j} to $\frac12+\nu_j-2\varepsilon$ and simultaneously c_{v_j} to $c_{v_j}=-1+\varepsilon/k$, passing through the pole of $\Psi_{\varepsilon^*,B}$ at $v_j=0$. The contribution of the integral on the new line of integration can be bounded by

$$\ll q^{-\frac56+\vartheta/3+A\varepsilon}N_1^{-1+A\varepsilon}(N_1\cdots N_k)^{\frac12}L^{-\varepsilon},\tag{7-19}$$

as can be see by moving c_{u_1} to $c_{u_1}=-B-\frac32-\alpha_1+7\varepsilon$ and bounding the sums and integrals as in the proof of (7-15). Thus we only need to consider the residue at $v_j=0$ for all $j\in J$.

In the same way, we move the line of integration c_{v_*} to $c_{v_*}=1+\varepsilon/k$ and c_{u_*} to $c_{u_*}=\nu_*+2\varepsilon-\varepsilon/k$, passing through the pole of $\Psi_{\varepsilon^*,B}$ at $v_*=B+1-\bigl(\frac12-\alpha_1-u_1+\frac{w}{2}\bigr)-\sum_{i\in I}v_i$. The contribution of the new line of integration can be bounded by (7-19) in a similar way, so again we only need to consider the contribution of the residue. Thus, summarizing (and recalling (7-7) and (7-10)), we arrive at

$$\begin{aligned}\mathcal{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}=&\sum_{\substack{I\cup J=\{2,\dots,k\}\\ I\cap J=\varnothing,\,|I|>|J|}}\ \sum_{\substack{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}\ \forall i\in I\\ (\alpha_j',\beta_j')=(\alpha_j,\beta_j)\ \forall j\in J\cup\{1\}}}\ \sum_{\nu}\frac{B!}{\nu_2!\cdots\nu_k!\nu_*!}\mathcal{Q}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}\\ &+O\biggl(\frac{q^{-1+A\varepsilon}N_1^{A\varepsilon}}{(N_1\cdots N_k)^{\frac12}}\bigl(q^{\vartheta}N_1^{(k+1)/2}+q^{k/2-\frac13+\vartheta/3}N_1^{\frac12}+q^{\frac16+\vartheta/3}N_2\cdots N_k\bigr)L^{-\varepsilon}\biggr),\end{aligned}\tag{7-20}$$

where the sum over ν is now over $\nu = (\nu_2, \dots, \nu_k, \nu_*) \in \mathbb{Z}^k_{\geq 0}$ satisfying

$$\nu_2 + \dots + \nu_k + \nu_* = B, \qquad \nu_i = 0 \quad \text{if } \pm_i 1 = -1 \text{ or } i \in J, \qquad \nu_* = 0 \quad \text{if } \pm_* 1 = -1,$$

and where

$$\begin{aligned}\mathcal{Q}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{d \mid q} \frac{\mu(q/d)}{\varphi(q)} \sum_{\ell} \sideset{}{^*}\sum_{h \pmod{\ell}} \frac{P(\ell/L)}{(2\pi i)^{4+|I|}} \int_{(c_s, c_w, c_{u_1}, c_{u_*}, c_{v_i} \forall i \in I)} \frac{d^{\frac{1}{2}-B-\alpha_1-u_1+w/2+\sum_{i\in I} v_i - u_* + \nu_*}}{\ell^{\sum_{i \in I \cup \{1\}}(1-\alpha_i+\beta_i)+w}} \\ \times q^{ks} F\big(B + \tfrac{1}{2} + \alpha_1 + u_1 - \tfrac{w}{2} - \textstyle\sum_{i \in I} v_i + u_* - \nu_*, \pm_* \tfrac{dh}{\ell}\big) \frac{H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)}{ws} \\ \times \bigg(\prod_{j \in J} \frac{1}{2\pi i} \int_{(c_{u_j})} D_{\alpha_j,\beta_j}\big(\tfrac{1}{2} + u_j, \tfrac{\pm_j h}{\ell}\big) \widetilde{P}(u_j - s) N_j^{u_j - s}\, du_j \bigg) \\ \times N_*^{u_*} \widetilde{P}(u_*) N_1^{u_1 - s} \widetilde{P}(u_1 - s) \Psi'_{I;\varepsilon_I^*,B}\big(\tfrac{1}{2} - \alpha_1 - u_1 + \tfrac{w}{2}, \boldsymbol{v}_I\big) \\ \times \bigg(\prod_{i \in I} \widetilde{P}(\tfrac{1}{2} - \alpha_i - v_i + \nu_i - s) N_i^{\frac{1}{2} - \alpha_i - v_i + \nu_i - s} d^{\nu_i} dv_i \bigg) dw\, du_1\, ds\, du_*, \end{aligned} \tag{7-21}$$

for $\boldsymbol{v}_I := (v_i)_{i \in I}$, $\varepsilon_I^* = (\pm_i 1)_{i \in I \cup \{*\}}$ and

$$\begin{aligned} \Psi'_{I;\varepsilon_I^*,B}(v_1, \boldsymbol{v}_I) &:= \frac{\Gamma\big(B + 1 - v_1 - \sum_{i \in I} v_i\big) \prod_{i \in I \cup \{1\}} \Gamma(v_i)}{\Gamma(V_{\mp_*;\varepsilon_I}(v_1, \boldsymbol{v}_I))\Gamma(B + 1 - V_{\mp_*;\varepsilon_I}(v_1, \boldsymbol{v}_I))}, \\ V_{\pm,\varepsilon_I}(v_1, \boldsymbol{v}_I) &:= \sum_{\substack{i \in I \cup \{1\} \\ \pm_i 1 = \pm 1}} v_i. \end{aligned} \tag{7-22}$$

We also remind that the line of integrations are

$$\begin{aligned} c_s &:= \varepsilon/k, \quad c_w = 10\varepsilon, \quad c_{u_*} = 1 + \nu_* + 2\varepsilon - \varepsilon/k, \\ c_{u_j} &= -\tfrac{1}{2} + \nu_j - 2\varepsilon \quad \forall j \in J, \qquad c_{v_i} = \frac{\varepsilon}{k} \quad \forall i \in I \end{aligned} \tag{7-23}$$

and $c_{u_1} = -B - \frac{1}{2} - \Re(\alpha_1) + 7\varepsilon$.

Remark. Notice that the integrand in (7-21) decays rapidly along a vertical strip in each of the variables of integration. In particular, in the following computations we will always be able to bound the integrals trivially.

At this point, we wish to execute the sum over the partitions of unity N_*. However, first we need to remove the truncation $N_* \ll kN_1 q^{\varepsilon/k}$. This can be done at a negligible cost, as shown in the following lemma.

Lemma 20. *We have*

$$\sideset{}{^\dagger}\sum_{N_* \ll kN_1 q^{\varepsilon/k}} \mathcal{Q}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} = \sideset{}{^\dagger}\sum_{N_*} \mathcal{Q}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} + O\big(q^{-2+A\varepsilon} N_1^{A\varepsilon} (N_1 \cdots N_k)^{\frac{1}{2}} L^{-\varepsilon}\big). \tag{7-24}$$

Proof. Assume $N_* \gg kN_1 q^{\varepsilon/k}$. Given a large positive integer Δ, we move c_{v_i} to $c_{v_i} = -\Delta k + \frac{1}{2}$ for all $i \in I$ in (7-21). Doing so, we pass through the poles of

$$\Psi'_{I;\varepsilon_I^*,B}\big(\tfrac{1}{2} - \alpha_1 - u_1 + \tfrac{w}{2}, \boldsymbol{v}_I\big) \quad \text{at } v_i \in S_\Delta := \{-r \mid r \in \mathbb{Z}, 0 \le r < \Delta k\},$$

so that we have to deal with a sum of $(\Delta k + 1)^{|I|}$ terms coming from the contribution of the residues and of the integrals on the new lines of integration.[3] Then, for each of these terms, we move the line of integration c_{u_1} to $c_{u_1} = \Delta k$. This can be done without crossing any pole of $\Psi'_{I;\varepsilon_I^*,B}$ if the term was coming from picking up a residue in each of the variables v_i for all $i \in I$, since in this case the Γ factor in the denominator of $\Psi'_{I;\varepsilon_I^*,B}$ cancel the poles of $\Gamma(v_1) = \Gamma\big(\frac{1}{2} - \alpha_1 - u_1 + \frac{w}{2}\big)$. Otherwise, we also have to consider the residues of $\Psi'_{I;\varepsilon_I^*,B}$ at $\frac{1}{2} - \alpha_1 - u_1 + \frac{w}{2} \in S_\Delta$. In all cases, however, all the terms will still have at least one integral left (besides the w and u_* integrals) with line of integration $c_{v_i} = -\Delta k + \frac{1}{2}$ or $c_{u_1} = \Delta k$. Finally, for each of these terms we place the line of integration c_{u_*} so that the real part of the argument of the function F in (7-21) is still $1 + (4 - |I|/k)\varepsilon$ (we can do this without crossing poles).

Thus, bounding these terms by using Lemma 12 and the estimates (5-2), (7-12) and (10-6), we obtain

$$\begin{aligned}\mathcal{Q}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} &\ll q^{-\frac{5}{6}+\vartheta/3+A\varepsilon} \sum_{\substack{r_1,\dots r_k \in (\{0,1,\dots,\Delta-1\}\cup\{\Delta k - \frac{1}{2}\}) \\ r_i = \Delta k - \frac{1}{2} \text{ for some } i \in (\{1\}\cup I) \\ r_i = 0 \text{ if } i \in J}} \frac{N_1^{\frac{1}{2}+r_1+A\varepsilon} N_2^{r_2+\frac{1}{2}+\nu_2} \cdots N_k^{r_k+\frac{1}{2}+\nu_k}}{N_*^{B+\sum_{i=1}^k r_i - \nu_* + (4-|I|/k)\varepsilon} L^\varepsilon} \prod_{j\in J} N_j^{-1} \\ &\ll q^{-1+A\varepsilon} \frac{N_1^{A\varepsilon}(N_1\cdots N_k)^{\frac{1}{2}}}{q^{\Delta\varepsilon} N_*^\varepsilon} \bigg(\prod_{j\in J} N_j^{-1}\bigg) L^{-\varepsilon} \ll q^{-2+A\varepsilon} N_1^{A\varepsilon} (N_1\cdots N_k)^{\frac{1}{2}} (LN_*)^{-\varepsilon},\end{aligned}$$

if Δ is large enough with respect to ε. Equation (7-24) then follows. □

We now move the line of integration c_{u_1} to $c_{u_1} = \frac{1}{2} + 3\varepsilon$ and then execute the sum over N_*, which we do by using the following lemma.

Lemma 21. *Let $K(s)$ be a function which is analytic and grows at most polynomially on a strip $|\Re(s)| < c$ for some $c > 0$. Then, for any $-c < c_u < c$ we have*

$$\sum_N^{\dagger} \frac{1}{2\pi i} \int_{(c_u)} K(u)\widetilde{P}(u) N^u \, du = K(0).$$

Proof. For $\varepsilon > 0$, let $\hat{g}_\varepsilon(x)$ be the Mellin transform of $g_\varepsilon(u) := K(u)e^{\varepsilon u^2}$. We have

$$\sum_N^{\dagger} \frac{1}{2\pi i} \int_{(c_u)} K(u)\widetilde{P}(u) N^u \, du = \lim_{\varepsilon\to 0} \sum_N^{\dagger} \frac{1}{2\pi i} \int_{(0)} K(u) e^{\varepsilon u^2} \widetilde{P}(u) N^u \, du$$

as can be seen by splitting the sum according to whether $N \ge 1$ or $N < 1$ and moving the line of integration accordingly to $-c/2$ or $c/2$. We then write $g_\varepsilon(u)$ in terms of its Mellin transform. Exchanging the order

[3] In total there are $(\Delta k + 1)^{|I|}$ terms because for each v_i we have the possibility of taking the residue at $v_i = -r_i \in S_\Delta$ or to take the integral at $c_{v_i} = -\Delta k + \frac{1}{2}$.

of the integrals, as allowed by the bound $\hat{g}_\varepsilon(x) \ll \min(x^{c/2}, x^{-c/2})$, and executing the integral over u we obtain

$$\lim_{\varepsilon\to 0} \sum_N^{\dagger} \int_0^\infty g_\varepsilon(x) P(1/Nx)\,\frac{dx}{x} = \lim_{\varepsilon\to 0} \int_0^\infty g_\varepsilon(x)\,\frac{dx}{x} = F(0). \qquad \square$$

We move the line of integration c_{u_1} to $c_{u_1} = \frac{1}{2} + 3\varepsilon$ without crossing poles and apply the above lemma obtaining

$$\begin{aligned}\sum_{N_*}^{\dagger} \mathcal{Q}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}} = \sum_{d\,|\,q} \frac{\mu(q/d)}{\varphi(q)} \sum_{\ell} \sum_{h \pmod{\ell}}^{*} \frac{P(\ell/L)}{(2\pi i)^{3+|I|}} \int_{(c_s,c_w,c_{u_1},\, c_{v_i}\forall i\in I)} \frac{d^{\frac{1}{2}-B-\alpha_1-u_1+w/2+\sum_{i\in I} v_i+\nu_*}}{\ell^{\sum_{i\in I\cup\{1\}}(1-\alpha_i+\beta_i)+w}}&\\
\times q^{ks} F\big(B+\tfrac{1}{2}+\alpha_1+u_1-\tfrac{w}{2}-\textstyle\sum_{i\in I} v_i-\nu_*,\ \pm_*\tfrac{h}{\ell}\big)&\\
\times \Bigg(\prod_{j\in J} \frac{1}{2\pi i}\int_{(c_{u_j})} D_{\alpha_j,\beta_j}\big(\tfrac{1}{2}+u_j, \tfrac{\pm_j h}{\ell}\big)\widetilde{P}(u_j-s)N_j^{u_j-s}\,du_j\Bigg)&\\
\times N_1^{u_1-s}\widetilde{P}(u_1-s)\Psi'_{I;\varepsilon_I^*,B}\big(\tfrac{1}{2}-\alpha_1-u_1+\tfrac{w}{2}, \boldsymbol{v}_I\big)\frac{H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)}{ws}&\\
\times \prod_{i\in I}\Big(\widetilde{P}(\tfrac{1}{2}-\alpha_i-v_i+\nu_i-s)N_i^{\frac{1}{2}-\alpha_i-v_i+\nu_i-s}d^{v_i}\,dv_i\Big)\,dw\,du_1\,ds,&\qquad (7\text{-}25)\end{aligned}$$

with lines of integrations that we can take to be given by (7-23) and $c_{u_1} = \frac{1}{2} + 3\varepsilon$.

We are finally ready to execute the sum over ν. We do this in the following lemma, which also summarizes the previous computations.

Lemma 22. *We have*

$$\begin{aligned}\sum_{\varepsilon\in\{\pm1\}^k} \rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\alpha,\beta} = 2\sum_L^{\dagger} \sum_{\substack{I\cup J=\{2,\dots,k\}\\ I\cap J=\varnothing,\,|I|>|J|}} \sum_{\substack{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}\,\forall i\in I_1\\ (\alpha_j',\beta_j')=(\alpha_j,\beta_j)\,\forall j\in J}} \sum_{\substack{\varepsilon\in\{\pm1\}^k\\ \pm_1 1=-1}} \rho_\Upsilon(\varepsilon) \sum_{\pm_* 1\in\{\pm1\}} \mathcal{R}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}}&\\
+O\bigg(\frac{q^{-1+A\varepsilon}N_1^{A\varepsilon}}{(N_1\cdots N_k)^{\frac{1}{2}}L^\varepsilon}\Big(q^{\vartheta}N_1^{(k+1)/2}+q^{k/2-\frac{1}{3}+\vartheta/3}N_1^{\frac{1}{2}}+q^{\frac{1}{6}+\vartheta/3}N_2\cdots N_k\Big)\bigg),&\qquad (7\text{-}26)\end{aligned}$$

where $I_1 := I\cup\{1\}$ *and*

$$\begin{aligned}\mathcal{R}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{d\,|\,q} \frac{\mu(q/d)}{\varphi(q)} \sum_{\ell} \sum_{h \pmod{\ell}}^{*} \frac{P(\ell/L)}{(2\pi i)^{3+|I|}} \int_{(c_s,c_w,c_{u_1},\, c_{v_i}\forall i\in I)} \frac{q^{ks} d^{\frac{1}{2}-\alpha_1-u_1+w/2+\sum_{i\in I} v_i}N_1^{u_1-s}}{\ell^{\sum_{i\in I_1}(1-\alpha_i+\beta_i)+w}}&\\
\times \Bigg(\prod_{j\in J} \frac{1}{2\pi i}\int_{(c_{u_j})} D_{\alpha_j,\beta_j}\big(\tfrac{1}{2}+u_j, \tfrac{\pm_j h}{\ell}\big)\widetilde{P}(u_j-s)N_j^{u_j-s}du_j\Bigg)&\\
\times \widetilde{P}(u_1-s)\times F\big(\tfrac{1}{2}+\alpha_1+u_1-\tfrac{w}{2}-\textstyle\sum_{i\in I} v_i, \pm_*\tfrac{dh}{\ell}\big)\Psi'_{I;\varepsilon_I^*,0}\big(\tfrac{1}{2}-\alpha_1-u_1+\tfrac{w}{2}, \boldsymbol{v}_I\big)&\\
\times \Bigg(\prod_{i\in I}\widetilde{P}\big(\tfrac{1}{2}-\alpha_i-v_i-s\big)N_i^{\frac{1}{2}-\alpha_i-v_i-s}d^{v_i}\,dv_i\Bigg)\frac{H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(w,s)}{ws}\,dw\,du_1\,ds&\end{aligned}$$

and lines of integrations given by (7-23) *and* $c_{u_1} = \frac{1}{2} + 3\varepsilon$.

Proof. Using (7-1), (7-4), (7-20) and (7-24) we obtain (7-26), with $\mathcal{R}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}}$ replaced by

$$\mathcal{R}'_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{\nu} \frac{B!}{\nu_2! \cdots \nu_k! \nu_*!} \sum_{N_*}^{\dagger} \mathcal{Q}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}$$

and $\sum^{\dagger} \mathcal{Q}_{I;\nu;\boldsymbol{\alpha},\boldsymbol{\beta}}$ as in (7-25). Thus, the lemma reduces to showing that $\mathcal{R}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}} = \mathcal{R}'_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}}$. This is an immediate consequence of Lemma 29 below, which is applicable since the pole of F is canceled by the sum over d. □

7.6. ***Reassembling the sum over ε.*** Now, we can also get rid of the integral over w. To do this, first we move the lines of integration c_{u_1} and c_{u_j} for $j \in J$ (without passing through poles), so that we have

$$c_{u_1} = -\Re(\alpha_1) + 7\varepsilon, \qquad c_{u_j} = \tfrac{1}{2} - \Re(\alpha_j) - 2\varepsilon \quad \forall j \in J \qquad c_s := \varepsilon/k, \qquad c_{v_i} = \frac{\varepsilon}{k} \quad \forall i \in I \tag{7-27}$$

and then we move c_w to $c_w = -1 + 10\varepsilon$ passing through a pole at $w = 0$. The contribution of the new line of integration is trivially bounded by

$$\ll q^{-1+A\varepsilon} N_1^{-\frac{1}{2}+A\varepsilon} (N_1 \cdots N_k)^{\frac{1}{2}} L^{-\varepsilon}. \tag{7-28}$$

since we have the convexity bound $D\big(1-2\varepsilon+it, \alpha_j - \beta_j, \frac{h}{\ell}\big) \ll \ell^{3\varepsilon}(1+|t|)^{3\varepsilon}$ and $|I| \geq 2$ (since $|I| > |J|$ and $k \geq 3$). Thus, we only need to consider the contribution from the residue at $w = 0$.

By the convexity bound

$$F\big(\tfrac{1}{2} + 7\varepsilon - |I|\varepsilon/k + it, \tfrac{h}{\ell}\big) \ll \ell^{\frac{1}{2}}(1+|t|)^{\frac{1}{2}},$$

the contribution of the $d = 1$ term is also bounded by (7-28). Thus, using also that $\varphi(q)^{-1} = q^{-1} + O(q^{-2})$ for q prime, we have

$$\mathcal{R}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}} = \mathcal{S}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}} + O\big(q^{-1+A\varepsilon} N_1^{-\frac{1}{2}+A\varepsilon} (N_1 \cdots N_k)^{\frac{1}{2}} L^{-\varepsilon}\big) \tag{7-29}$$

with

$$\begin{aligned}\mathcal{S}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}} = \sum_{\ell} \sum_{h \pmod{\ell}}^{*} \frac{P(\ell/L)}{(2\pi i)^{2+|I|}} \int_{(c_s, c_{u_1}, c_{v_i} \forall i \in I)} & \frac{q^{ks-\frac{1}{2}-\alpha_1-u_1+\sum_{i\in I} v_i}}{\ell^{\sum_{i\in I\cup\{1\}}(1-\alpha_i+\beta_i)}} N_1^{u_1-s} \widetilde{P}(u_1-s) \\ \times \Big(\prod_{j\in J} \frac{1}{2\pi i} \int_{(c_{u_j})} & D_{\alpha_j,\beta_j}\big(\tfrac{1}{2}+u_j, \pm_* \tfrac{\pm_j h}{\ell}\big) \widetilde{P}(u_j - s) N_j^{u_j - s}\, du_j\Big) \\ \times F\big(\tfrac{1}{2} + \alpha_1 + u_1 - {\textstyle\sum_{i\in I}} v_i, \tfrac{qh}{\ell}\big) & \Psi'_{I;\varepsilon_I^*,0}\big(\tfrac{1}{2} - \alpha_1 - u_1, \boldsymbol{v}_I\big) \\ \times \Big(\prod_{i\in I} \widetilde{P}\big(\tfrac{1}{2} - \alpha_i - v_i - s\big) & N_i^{\frac{1}{2}-\alpha_i-v_i-s} q^{v_i}\, dv_i\Big) \frac{H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(0,s)}{s}\, du_1\, ds\end{aligned}$$

and lines of integration given by (7-27). Notice that we made the change of variable $h \to \pm_* h$.

We are ready to reassemble the sum over ε. To do this, first we split ε into ε_{I_1} and ε_J, where $\varepsilon_S := (\pm_i 1)_{i\in S}$; in particular, $\rho_\Upsilon(\varepsilon) = \rho_\Upsilon(\varepsilon_{I_1})\rho_\Upsilon(\varepsilon_J)$ where, with a slight abuse of notation, we write

$\rho_\Upsilon(\varepsilon_S) := \prod_{i\in\Upsilon\cap S}(\pm_i 1)$. Then, we observe that

$$\sum_{\varepsilon_J\in\{\pm1\}^{|J|}} \rho_\Upsilon(\varepsilon_J)\prod_{j\in J} D_{\alpha_j,\beta_j}\big(s_j, \pm_* \tfrac{\pm_j h}{\ell}\big) = 2^{|J|}(\pm_* i)^{|\Upsilon\cap J|}\prod_{j\in J} D_{j;\alpha_j,\beta_j}\big(s_j, \tfrac{h}{\ell}\big).$$

Thus,

$$\begin{aligned}\sum_{\substack{\varepsilon\in\{\pm1\}^k\\ \pm_1 1=-1}} \rho_\Upsilon(\varepsilon) &\sum_{\pm_* 1\in\{\pm1\}} \mathcal{S}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}}\\ &= 2^{|J|} i^{|\Upsilon\cap J|}\sum_\ell \sideset{}{^*}\sum_{h \pmod{\ell}} \frac{P(\ell/L)}{(2\pi i)^{2+|I|}} \frac{q^{ks-\frac12-\alpha_1-u_1+\sum_{i\in I} v_i}}{\ell^{\sum_{i\in I\cup\{1\}}(1-\alpha_i+\beta_i)}}\\ &\quad\times \int_{(c_s, c_{u_1},\, c_{v_i}\forall i\in I)} \bigg(\prod_{j\in J}\frac{1}{2\pi i}\int_{(c_{u_j})} D_{j;\alpha_j,\beta_j}\big(\tfrac12+u_j, \tfrac{h}{\ell}\big)\widetilde{P}(u_j-s)N_j^{u_j-s}\,du_j\bigg)\\ &\qquad\times N_1^{u_1-s}\widetilde{P}(u_1-s)F\big(\tfrac12+\alpha_1+u_1-\textstyle\sum_{i\in I} v_i, \tfrac{qh}{\ell}\big)\mathcal{X}_I\big(\tfrac12-u_1-\alpha_1, \boldsymbol{v}_I\big)\\ &\qquad\qquad\times\bigg(\prod_{i\in I}\widetilde{P}(\tfrac12-\alpha_i-v_i-s)N_i^{\frac12-\alpha_i-v_i-s} q^{v_i}\,dv_i\bigg)\frac{H'_{I;\boldsymbol{\alpha},\boldsymbol{\beta}}(0,s)}{s}\,du_1\,ds, \end{aligned}\tag{7-30}$$

with

$$\begin{aligned}\mathcal{X}_I(v_1,\boldsymbol{v}_I) &:= \sum_{\pm_* 1\in\{\pm1\}}(\pm_* 1)^{|\Upsilon\cap J|}\sum_{\substack{\varepsilon_{I_1}\in\{\pm1\}^{|I_1|}\\ \pm_1 1=-1}}\rho_\Upsilon(\varepsilon_{I_1})\Psi'_{I;\varepsilon_I^*,0}(v_1,\boldsymbol{v}_I),\\ &= \Gamma\bigg(1-\sum_{i\in I_1} v_i\bigg)\prod_{i\in I_1}\Gamma(v_i)\sum_{\pm_* 1\in\{\pm1\}}(\pm_* 1)^{|\Upsilon\cap I_1|}\\ &\qquad\qquad\times\sum_{\substack{\varepsilon_{I_1}\in\{\pm1\}^{|I_1|}\\ \pm_1 1=-1}}\rho_\Upsilon(\varepsilon_{I_1})\bigg(\Gamma\bigg(\sum_{\substack{i\in I_1\\ \pm_i 1=\mp_* 1}} v_i\bigg)\Gamma\bigg(1-\sum_{\substack{i\in I_1\\ \pm_i 1=\mp_* 1}} v_i\bigg)\bigg)^{-1},\end{aligned}$$

by the definition (7-22) of $\Psi'_{I;\varepsilon_I^*,0}$ and since $(\pm_* 1)^{|\Upsilon\cap J|} = (\pm_* 1)^{|\Upsilon\cap(I\cup\{1\})|}$ for $|\Upsilon|$ even (as we have assumed). Also, we remind that we defined $\varepsilon_I^* := (\pm_i 1)_{i\in I\cup\{*\}}$ and $I_1 := I\cup\{1\}$.

We will now give a Γ-function identity, which we will use to give a symmetric expression for $\mathcal{X}_I(v_1,\boldsymbol{v}_I)$.

Lemma 23. *Let $r\geq 1$, $\Theta\subseteq\{1,\dots,r\}$ and $(s_1,\dots,s_r)\in\mathbb{C}^r$. For $\varepsilon_r = (\pm_1,\dots,\pm_r 1)\in\{\pm1\}^r$ let $\rho_\Theta(\varepsilon) := \prod_{i\in\Theta}(\pm_i 1)$. Then,*[4]

$$\begin{aligned}\prod_{i=1}^r\Gamma(s_i)\sum_{\pm_* 1\in\{\pm1\}}(\pm_* 1)^{|\Theta|}\sum_{\substack{\varepsilon\in\{\pm1\}^r\\ \pm_1 1=-1}}\rho_\Theta(\varepsilon)\bigg(\Gamma\bigg(\sum_{\pm_i 1=\mp_* 1} s_i\bigg)\Gamma\bigg(1-\sum_{\pm_i 1=\mp_* 1} s_i\bigg)\bigg)^{-1}&\\ = \frac{2^{s_1+\dots+s_r}}{\pi^{1-r/2}}\bigg(\prod_{i\in\Theta}\frac{\Gamma\big(\frac12+\frac{s_i}{2}\big)}{\Gamma\big(1-\frac{s_i}{2}\big)}\bigg)\bigg(\prod_{i\notin\Theta}\frac{\Gamma\big(\frac{s_i}{2}\big)}{\Gamma\big(\frac12-\frac{s_i}{2}\big)}\bigg)\sin\Big(\frac{\pi}{2}(s_1+\dots+s_r)-\frac{\pi}{2}|\Theta|\Big).&\end{aligned}$$

[4]The identity has to be interpreted as an identity between meromorphic functions.

Proof. First, we observe that, by analytic continuation, we can assume that $s_1,\dots,s_r\in\mathbb{R}\setminus\mathbb{Z}$. Thus, using the reflection formula for the Gamma function to have

$$\Bigl(\Gamma\Bigl(\sum_{\pm_i 1=\mp_* 1} s_i\Bigr)\Gamma\Bigl(1-\sum_{\pm_i 1=\mp_* 1} s_i\Bigr)\Bigr)^{-1}=\pi^{-1}\sin\Bigl(\pi\sum_{\pm_i 1=\mp_* 1} s_i\Bigr)=\pi^{-1}\Im\Bigl(\exp\Bigl(\pi i\sum_{\pm_i 1=\mp_* 1} s_i\Bigr)\Bigr).$$

It follows that

$$\begin{aligned}S&:=\sum_{\pm_* 1\in\{\pm1\}}(\pm_* 1)^{|\Theta|}\sum_{\substack{\varepsilon\in\{\pm1\}^r\\ \pm_1 1=-1}}\rho_\Theta(\varepsilon)\Bigl(\Gamma\Bigl(\sum_{\pm_i 1=\mp_* 1} s_i\Bigr)\Gamma\Bigl(1-\sum_{\pm_i 1=\mp_* 1} s_i\Bigr)\Bigr)^{-1}\\ &=\pi^{-1}\Im\bigl(e^{\pi i s_1}A_+ +(-1)^{|\Theta|}A_-\bigr),\end{aligned}$$

where

$$A_\pm:=\sum_{\substack{\varepsilon\in\{\pm1\}^r\\ \pm_1 1=-1}}\rho_\Theta(\varepsilon)\exp\Bigl(\pi i\sum_{\substack{i=2\\ \pm_i 1=\mp 1}}^r s_i\Bigr).$$

Now, since $\rho_\Theta(\varepsilon)=\prod_{i\in\Theta}(\pm_i 1)=(-1)^{|\Theta\cap\{1\}|}\prod_{i\in\Theta\setminus\{1\}}(\pm_i 1)$, we have

$$\begin{aligned}A_\pm&=(-1)^{|\Theta\cap\{1\}|}\Bigl(\prod_{\substack{i=2\\ i\in\Theta}}^r(\pm1\mp e^{\pi i s_i})\Bigr)\Bigl(\prod_{\substack{i=2\\ i\notin\Theta}}^r(1+e^{\pi i s_i})\Bigr)\\ &=(\pm1)^{|\Theta\cap\{1\}|}(\mp1)^{|\Theta|}i^{|\Theta\setminus\{1\}|}2^{r-1}\exp\Bigl(\frac{\pi i}{2}\sum_{i=2}^r s_i\Bigr)\Bigl(\prod_{\substack{i=2\\ i\in\Theta}}^r\sin\Bigl(\frac{\pi s_i}{2}\Bigr)\Bigr)\Bigl(\prod_{\substack{i=2\\ i\notin\Theta}}^r\cos\Bigl(\frac{\pi s_i}{2}\Bigr)\Bigr)\end{aligned}$$

and thus

$$\begin{aligned}S&=\frac{2^{r-1}}{\pi}\Bigl(\prod_{\substack{i\in\Theta\\ i\neq1}}\sin\Bigl(\frac{\pi s_i}{2}\Bigr)\Bigr)\Bigl(\prod_{\substack{i\notin\Theta\\ i\neq1}}\cos\Bigl(\frac{\pi s_i}{2}\Bigr)\Bigr)(-1)^{|\Theta|}\Im\bigl((e^{\pi i s_1}+(-1)^{|\Theta\cap\{1\}|})e^{(\pi i/2)|\Theta\setminus\{1\}|+(\pi i/2)\sum_{i=2}^r s_i}\bigr)\\ &=\frac{2^r}{\pi}\Bigl(\prod_{i\in\Theta}\sin\Bigl(\frac{\pi s_i}{2}\Bigr)\Bigr)\Bigl(\prod_{i\notin\Theta}\cos\Bigl(\frac{\pi s_i}{2}\Bigr)\Bigr)(-1)^{|\Theta|}\Im\bigl(i^{|\Theta\cap\{1\}|}e^{(\pi i/2)|\Theta\setminus\{1\}|+(\pi i/2)\sum_{i=1}^r s_i}\bigr)\\ &=\frac{2^r}{\pi}\Bigl(\prod_{i\in\Theta}\sin\Bigl(\frac{\pi s_i}{2}\Bigr)\Bigr)\Bigl(\prod_{i\notin\Theta}\cos\Bigl(\frac{\pi s_i}{2}\Bigr)\Bigr)\sin\Bigl(-\frac{\pi}{2}|\Theta|+\frac{\pi}{2}\sum_{i=1}^r s_i\Bigr).\end{aligned}$$

By the duplication and the reflection formula for the Γ-function we have

$$\sin\Bigl(\frac{\pi s}{2}\Bigr)\Gamma(s)=\pi^{\frac12}2^{s-1}\frac{\Gamma\bigl(\frac12+\frac s2\bigr)}{\Gamma\bigl(1-\frac s2\bigr)},\qquad \cos\Bigl(\frac{\pi s}{2}\Bigr)\Gamma(s)=\pi^{\frac12}2^{s-1}\frac{\Gamma\bigl(\frac s2\bigr)}{\Gamma\bigl(\frac12-\frac s2\bigr)}$$

and thus the lemma follows. □

Applying Lemma 23 with $r=|I_1|$ and using the definition (2-1) of Γ_i, we obtain

$$\mathcal{X}_I(v_1,\boldsymbol{v}_I)=\Gamma\Bigl(1-\sum_{i\in I_1}v_i\Bigr)\frac{2^{\sum_{i\in I_1}v_i}}{\pi^{1-|I_1|/2}}\Bigl(\prod_{i\in I_1}\frac{\Gamma_i\bigl(\frac{v_i}{2}\bigr)}{\Gamma_i\bigl(\frac12-\frac{v_i}{2}\bigr)}\Bigr)\sin\Bigl(-\frac{\pi}{2}|I_1\cap\Upsilon|+\frac{\pi}{2}\sum_{i\in I_1}v_i\Bigr).$$

Thus, plugging this expression into (7-30), making the change of variables

$$u_i = \tfrac{1}{2} - \alpha_i - v_i - s \quad \forall i \in I, \qquad u_j \to u_j + s \quad \forall j \in (J \cup \{1\}),$$

and moving slightly the lines of integration, we obtain

$$\sum_{\substack{\varepsilon \in \{\pm 1\}^k \\ \pm_1 1 = -1}} \rho_\Upsilon(\varepsilon) \sum_{\pm_* 1 \in \{\pm 1\}} \mathcal{S}_{I;\varepsilon^*;\boldsymbol{\alpha},\boldsymbol{\beta}} = \mathcal{U}_{I_1;\boldsymbol{\alpha},\boldsymbol{\beta}}, \tag{7-31}$$

where for $\mathcal{I} \subseteq \{1, \dots, k\}$ (with $|\mathcal{I}| \geq 2$), $\mathcal{J} := \{1, \dots, k\} \setminus \mathcal{I}$, we define

$$\begin{aligned}\mathcal{U}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} := -2^{|\mathcal{J}|} i^{|\Upsilon \cap \mathcal{J}|} \sum_{\ell} \sideset{}{^*}\sum_{h \pmod{\ell}} \frac{P(\ell/L)}{(2\pi i)^{1+k}} \int_{(c_s, \boldsymbol{c_u})} \frac{q^{ks-1}}{\pi \ell^{\sum_{i\in\mathcal{I}}(1-\alpha_i+\beta_i)}} \frac{H''_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}(s)}{s} \\ \times \Bigg(\prod_{j \in \mathcal{J}} D_{j;\alpha_j,\beta_j}\big(\tfrac{1}{2} + u_j + s, \tfrac{h}{\ell}\big) \widetilde{P}(u_j) N_j^{u_j} \Bigg) \Gamma\big(1 + |\mathcal{I}|\big(s - \tfrac{1}{2}\big) + \textstyle\sum_{i\in\mathcal{I}}(u_i + \alpha_i)\big) \\ \times \sin\Bigg(\frac{\pi}{2}|\Upsilon \cap \mathcal{I}| + \frac{\pi}{2}|\mathcal{I}|\big(s - \tfrac{1}{2}\big) + \frac{\pi}{2} \sum_{i \in \mathcal{I}} (u_i + \alpha_i) \Bigg) \\ \times F\big(1 + |\mathcal{I}|(s - \tfrac{1}{2}) + \textstyle\sum_{i\in\mathcal{I}}(u_i + \alpha_i), \tfrac{qh}{\ell}\big) \Bigg(\displaystyle\prod_{i \in \mathcal{I}} \frac{\pi^{\frac{1}{2}} \widetilde{P}(u_i) N_i^{u_i}}{(2q)^{u_i+\alpha_i+s-\frac{1}{2}}} \frac{\Gamma_i\big(\frac{1}{4} - \frac{u_i+\alpha_i+s}{2}\big)}{\Gamma_i\big(\frac{1}{4} + \frac{u_i+\alpha_i+s}{2}\big)} \Bigg) \, ds \, d\boldsymbol{u},\end{aligned} \tag{7-32}$$

with lines of integration

$$c_{u_i} = \tfrac{1}{2} - \Re(\alpha_i) - 3\frac{\varepsilon}{k} - \varepsilon \quad \forall i \in \{1, \dots, k\}, \qquad c_s := \varepsilon/k, \tag{7-33}$$

and, recalling (7-11),

$$H''_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}(s) := G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s) \prod_{i \in \mathcal{I}} \zeta(1 - \alpha_i + \beta_i). \tag{7-34}$$

For future use we remark that if we move $c_{u'_i}$ to $c_{u'_i} = -\frac{1}{2} - \Re(\alpha_1) - 5\varepsilon$ for some $i' \in \mathcal{I}$ we get

$$\mathcal{U}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} \ll N_{i'}^{-1+A\varepsilon} q^{A\varepsilon} (N_1 \cdots N_k)^{\frac{1}{2}} L^{-\varepsilon}. \tag{7-35}$$

Also, if $|\mathcal{I}| > |\mathcal{J}|$, then moving the line of integration to c_{u_j} to $c_{u_j} = -\frac{1}{2} + 5\frac{\varepsilon}{k} - \varepsilon$ for all $j \in \mathcal{J}$ (leaving the other lines of integration as in (7-33)), we obtain

$$\mathcal{U}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} \ll q^{-1+A\varepsilon} (N_1 \cdots N_k)^{\frac{1}{2}} L^{-\varepsilon} \prod_{j \in \mathcal{J}} N_j^{-1+A\varepsilon}. \tag{7-36}$$

Finally, moving c_s to $c_s = \frac{1}{2} + B - 3\frac{\varepsilon}{k}$ and c_{u_i} to $c_{u_i} = -B$ for all $i = 1, \dots, k$ we obtain

$$\mathcal{U}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} \ll_B (N_1 \cdots N_k / q^k)^{-B} q^{k/2-1+A\varepsilon}, \tag{7-37}$$

if $|\mathcal{I}| \geq 2$.

We are now ready to complete the proof of Lemma 14. Using (7-26), (7-29) and (7-31) we obtain

$$\sum_{\varepsilon\in\{\pm 1\}^k} \rho_\Upsilon(\varepsilon)\mathcal{O}''_{\varepsilon,\alpha,\beta} = \sum_L^\dagger \sum_{\substack{I\cup J=\{2,\dots,k\}\\ I\cap J=\varnothing,\ |I|>|J|}} \sum_{\substack{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}\ \forall i\in I_1\\ (\alpha_j',\beta_j')=(\alpha_j,\beta_j)\ \forall j\in J}} 2\mathcal{U}_{I_1;\boldsymbol{\alpha},\boldsymbol{\beta}} + O\biggl(\frac{N_1^{A\varepsilon}}{q^{1-A\varepsilon}}\biggl(\frac{q^\vartheta N_1^{k/2}+q^{k/2-\frac13+\vartheta/3}}{(N_2\cdots N_k)^{\frac12}} + (q^{\frac16+\vartheta/3}N_1^{-\frac12}+1)(N_2\cdots N_k)^{\frac12}\biggr)\biggr)$$
$$= \mathcal{M}_{\alpha,\beta} + O\biggl(\frac{N_1^{A\varepsilon}}{q^{1-A\varepsilon}}\biggl(\frac{q^\vartheta N_1^{k/2}+q^{k/2-\frac13+\vartheta/3}}{(N_2\cdots N_k)^{\frac12}} + (q^{\frac16+\vartheta/3}N_1^{-\frac12}+1)(N_2\cdots N_k)^{\frac12}\biggr)\biggr)$$

where

$$\mathcal{M}_{\alpha,\beta}(N_1,\dots,N_k) := \sum_L^\dagger \sum_{\substack{\mathcal{I}\cup\mathcal{J}=\{1,\dots,k\}\\ \mathcal{I}\cap\mathcal{J}=\varnothing,\ |\mathcal{I}|>|\mathcal{J}|+1}} \sum_{\substack{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}\ \forall i\in\mathcal{I}\\ (\alpha_j',\beta_j')=(\alpha_j,\beta_j)\ \forall j\in\mathcal{J}}} 2\mathcal{U}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}. \tag{7-38}$$

and $\mathcal{U}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}$ as defined in (7-32). Noticed that in the last step we used (7-36) to extend the sum over the subsets of $\{1\ \dots, k\}$ to include also the sets $\mathcal{I}$ that do not contain 1. Moreover, by (7-35) and (7-36) we also have

$$\mathcal{M}_{\alpha,\beta}(N_1,\dots,N_k) \ll q^{A\varepsilon}(N_1\cdots N_k)^{\frac12}N_1^{-1+A\varepsilon}.$$

and thus the proof of Lemma 14 is complete. Also, by (7-37) for any $B>0$ we have

$$\mathcal{M}_{\alpha,\beta}(N_1,\dots,N_k) \ll_B q^{k/2-1+A\varepsilon}(N_1\cdots N_k/q^k)^{-B} \tag{7-39}$$

We remark that we reached a formula for $\mathcal{M}_{\alpha,\beta}$ which is completely symmetric in the $N_1,\dots,N_k$. This is important in order to remove the assumption that N_1 is the largest of $N_1,\dots,N_k$, so that we can sum over the partitions of unity.

8. Assembling the main terms

In this section we prove Lemma 16.

We start by moving c_{u_i} to $c_{u_i}=0$ for all $i\in\mathcal{I}$ and c_s to $\frac12-3\frac{\varepsilon}{k}$ (we can do this without passing through any pole nor having a problem of convergence). Then, after extending the sum over the partitions of unity $L, N_1,\dots,N_k$ using (7-39) and summing over them using Lemma 21 we obtain

$$\sum^\dagger_{\substack{N_1,\dots,N_k\\ N_1\cdots N_k\ll q^{k+\varepsilon}}} \mathcal{M}_{\alpha,\beta}(N_1,\dots,N_k) = \sum_{\substack{\mathcal{I}\cup\mathcal{J}=\{1,\dots,k\}\\ \mathcal{I}\cap\mathcal{J}=\varnothing,\,|\mathcal{I}|>|\mathcal{J}|+1}} \sum_{\substack{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}\ \forall i\in\mathcal{I}\\ (\alpha_j',\beta_j')=(\alpha_j,\beta_j)\ \forall j\in\mathcal{J}}} 2\,\mathcal{V}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} + O(1)$$

with

$$\mathscr{V}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} := -2^{|\mathcal{J}|} i^{|\Upsilon\cap\mathcal{J}|} \frac{1}{2\pi i}\int_{(c_s)} \sum_{\ell} \sideset{}{^*}\sum_{h \pmod{\ell}} \frac{q^{\frac12|\mathcal{I}|+|\mathcal{J}|s-1}}{\ell^{\sum_{i\in\mathcal{I}}(1-\alpha_i+\beta_i)}} \Gamma\big(1+|\mathcal{I}|\big(s-\tfrac12\big)+\textstyle\sum_{i\in\mathcal{I}}\alpha_i\big)$$
$$\times\bigg(\prod_{j\in\mathcal{J}} D_{j;\alpha_j,\beta_j}\big(\tfrac12+s,\tfrac{h}{\ell}\big)\bigg)\bigg(\prod_{i\in\mathcal{I}} \frac{(2\pi)^{\frac12}}{2^{\alpha_i+s}q^{\alpha_i}} \frac{\Gamma_i\big(\frac14-\frac{\alpha_i+s}{2}\big)}{\Gamma_i\big(\frac14+\frac{\alpha_i+s}{2}\big)}\bigg)\frac{H''_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}(s)}{\pi s}$$
$$\times F\big(1+|\mathcal{I}|\big(s-\tfrac12\big)+\textstyle\sum_{i\in\mathcal{I}}\alpha_i,\tfrac{h}{\ell}\big)\sin\bigg(\dfrac{\pi}{2}|\Upsilon\cap\mathcal{I}|+\dfrac{\pi}{2}|\mathcal{I}|\big(s-\tfrac12\big)+\dfrac{\pi}{2}\displaystyle\sum_{i\in\mathcal{I}}\alpha_i\bigg)\,ds, \tag{8-1}$$

and line of integration $c_s=\frac12-3\frac{\varepsilon}{k}$.

Now, for each integral we move the line of integration c_s to

$$c_s=\max\bigg(0,-\frac12+\frac{|\mathcal{J}|+\frac32}{|\mathcal{I}|+|\mathcal{J}|-\frac12}\bigg)+9\frac{\varepsilon}{k}=\begin{cases}\frac{3}{4k-2}+9\frac{\varepsilon}{k} & \text{if } |\mathcal{I}|=|\mathcal{J}|+2=\frac{k}{2}+1 \text{ with } k \text{ even},\\ 9\frac{\varepsilon}{k} & \text{if } |\mathcal{I}|>|\mathcal{J}|+2.\end{cases}$$

picking up the residue of the pole of the Γ-function at

$$s'=s'(\boldsymbol{\alpha})=\frac12-\frac{1+\sum_{i\in\mathcal{I}}\alpha_i}{|\mathcal{I}|}$$

(unless $k=4$, $|\mathcal{I}|=3$ in which case we stay on the right of such pole). Notice that Lemma 10 guarantees the convergence of the sum over ℓ on the new line of integration. Also, a quick computation shows that if $\mathcal{I}\neq I_k:=\{1,\dots,k\}$ (and $|\mathcal{I}|>|\mathcal{J}|+1$) then

$$\frac12|\mathcal{I}|+|\mathcal{J}|c_s-1\le\frac{k}{2}-\frac32+\iota_k+9\frac{\varepsilon}{k}$$

where $\iota_k=\frac{3}{14}$ if $k=4$ and $\iota_k=0$ otherwise. In particular, if $\mathcal{I}\neq I_k$, then by (3-8) the contribution of the integral on the new line of integration is $O(q^{k/2-\frac32+\iota_k+A\varepsilon})$ and we obtain

$$\mathscr{V}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}=\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}+O(q^{k/2-\frac32+\iota_k+A\varepsilon}), \tag{8-2}$$

where

$$\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} := -\frac{2^{|\mathcal{J}|}i^{|\Upsilon\cap\mathcal{J}|}}{|\mathcal{I}|}\sum_{\ell}\sideset{}{^*}\sum_{h \pmod{\ell}}\frac{q^{\frac12|\mathcal{I}|+|\mathcal{J}|s'-1}}{\ell^{\sum_{i\in\mathcal{I}}(1-\alpha_i+\beta_i)}}F\big(0,\tfrac{h}{\ell}\big)\sin\Big(\frac{\pi}{2}(|\Upsilon\cap\mathcal{I}|-1)\Big)$$
$$\times\bigg(\prod_{j\in\mathcal{J}} D_{j;\alpha_j,\beta_j}\big(\tfrac12+s',\tfrac{h}{\ell}\big)\bigg)\bigg(\prod_{i\in\mathcal{I}}\frac{(2\pi)^{\frac12}}{2^{s'+\alpha_i}q^{\alpha_i}}\frac{\Gamma_i\big(\frac14-\frac{\alpha_i+s'}{2}\big)}{\Gamma_i\big(\frac14+\frac{\alpha_i+s'}{2}\big)}\bigg)\frac{H''_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}(s')}{\pi s'}.$$

(Notice that (8-2) holds also in the case $k=4$, $|\mathcal{I}|=3$, since from (3-5) and a trivial bound it follows that $\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}$ is convergent and $O(q^{\frac23+A\varepsilon})=O(q^{k/2-\frac32+\iota_k+A\varepsilon})$).

If $\mathcal{I}=I_k$, then

$$\mathscr{V}_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}}=\mathscr{X}_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}}+\mathscr{V}'_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}},$$

where $\mathscr{V}'_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}}$ is as in (8-1), but with the line of integration $c_s=9\varepsilon/k$.

Now, notice that if $|\Upsilon\cap\mathcal{J}|$ is odd (and thus so is $|\Upsilon\cap\mathcal{I}|$ since $|\Upsilon|$ is even), then the sine in the expression defining $\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}$ is equal to 0 and thus so is $\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}$. If $|\Upsilon\cap\mathcal{J}|$ is even, then the product of the Estermann functions in the definition of $\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}$ is invariant under the change $h\mapsto -h$; in particular, using the identity $F(0,h/\ell)+F(0,-h/\ell)=-1$ (which follows immediately from (3-5)), we obtain $\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}=-\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}+\mathscr{K}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}$ and so $\mathscr{X}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}=\frac{1}{2}\mathscr{K}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}$, where

$$\begin{aligned}\mathscr{K}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}} &:= -\frac{2^{|\mathcal{J}|}}{|\mathcal{I}|}\sum_{\ell}\sideset{}{^*}\sum_{h \pmod{\ell}}\frac{q^{\frac{1}{2}|\mathcal{I}|+|\mathcal{J}|s'-1}}{\ell^{\sum_{i\in\mathcal{I}}(1-\alpha_i+\beta_i)}}\biggl(\prod_{j\in\mathcal{J}}D_{j;\alpha_j,\beta_j}\bigl(\tfrac{1}{2}+s',\tfrac{h}{\ell}\bigr)\biggr)\\ &\qquad\qquad\times\biggl(\prod_{i\in\mathcal{I}}\frac{(2\pi)^{\frac{1}{2}}}{2^{s+\alpha_i}q^{\alpha_i}}\frac{\Gamma_i\bigl(\frac{1}{4}-\frac{\alpha_i+s'}{2}\bigr)}{\Gamma_i\bigl(\frac{1}{4}+\frac{\alpha_i+s'}{2}\bigr)}\biggr)\frac{H''_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}(s')}{\pi s'}\\ &= -\mathscr{D}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}},\end{aligned}$$

where $\mathscr{D}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}$ is as in (6-2), since $\frac{1}{2}|\mathcal{I}|+|\mathcal{J}|s'-1=ks'+\bigl(\frac{1}{2}-s'\bigr)|\mathcal{I}|-1=ks'+\sum_{i\in\mathcal{I}}\alpha_i$ and by the definition (7-34) of $H''_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}(s')$. It follows that

$$\sum_{\substack{\mathcal{I}\cup\mathcal{J}=\{1,\dots,k\}\\ \mathcal{I}\cap\mathcal{J}=\varnothing\\ \frac{k}{2}+\frac{3}{4}<|\mathcal{I}|}}\ \sum_{\substack{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}\ \forall i\in\mathcal{I}\\ (\alpha_j',\beta_j')=(\alpha_j,\beta_j)\ \forall j\in\mathcal{J}}}2\mathscr{V}_{\mathcal{I};\boldsymbol{\alpha},\boldsymbol{\beta}}=\sum_{\{\alpha_i',\beta_i'\}=\{\alpha_i,\beta_i\}\ \forall i\in I_k}2\mathscr{V}'_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}}-\mathscr{D}_{\boldsymbol{\alpha},\boldsymbol{\beta}}+O(q^{k/2-\frac{1}{2}+\iota_k+A\varepsilon})$$

and thus to conclude the proof of Lemma 16, we just need to show $\mathscr{V}'_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}}+X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathscr{V}'_{I_k;-\boldsymbol{\beta},-\boldsymbol{\alpha}}=\mathscr{M}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ with $\mathscr{M}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ as in (2-3). First, we need the following lemma.

Lemma 24. *For $\Re(s_1+s_1)>2$ and $\Re(s_2)>1$ we have*

$$\sum_{\ell}\frac{1}{\ell^{s_2}}\sideset{}{^*}\sum_{h \pmod{\ell}}F\bigl(s_1,\tfrac{h}{\ell}\bigr)=\frac{\zeta(s_1)\zeta(s_1+s_2-1)}{\zeta(s_2)}. \tag{8-3}$$

Proof. From the functional equation for $F(s,x)$ and the Phragmén–Lindelöf principle one sees that if $|s_1-1|>\varepsilon'>0$, then if

$$\sideset{}{^*}\sum_{h \pmod{\ell}}\bigl|F\bigl(s_1,\tfrac{h}{\ell}\bigr)\bigr|\ll_{s,\varepsilon,\varepsilon'}1+\ell^{1-\Re(s_1)+\varepsilon}$$

for all $\varepsilon>0$. It follows that the left hand side of (8-3) defines a meromorphic function in s_1,s_2 on the region $\Re(s_2)>1$, $\Re(s_1+s_2)>2$. Now, assume $\Re(s_1),\Re(s_2)>1$. Expanding F into its Dirichlet expansion (3-3), executing the sum over h, and using (7-2), we obtain

$$\sum_{\ell}\frac{1}{\ell^{s_2}}\sideset{}{^*}\sum_{h \pmod{\ell}}F\bigl(s_1,\tfrac{h}{\ell}\bigr)=\sum_{\ell}\frac{1}{\ell^{s_2}}\sum_{n}\frac{c_\ell(n)}{n^{s_1}}=\frac{1}{\zeta(s_2)}\sum_{n}\frac{\sigma_{1-s_2}(n)}{n^{s_1}}=\frac{\zeta(s_1)\zeta(s_1+s_2-1)}{\zeta(s_2)}.$$

The lemma then follows by analytic continuation. $\square$

Applying this Lemma, we see that

$$\mathscr{V}'_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}} = -\frac{q^{k/2-1}}{2\pi i}\int_{(c_s)} \Gamma\Big(1+ks-\frac{k}{2}+\sum_{i=1}^{k}\alpha_i\Big)\frac{\zeta\big(1+ks-\frac{k}{2}+\sum_{i=1}^{k}\alpha_i\big)\zeta\big(\frac{k}{2}+ks+\sum_{i=1}^{k}\beta_i\big)}{\zeta\big(k-\sum_{i=1}^{k}(\alpha_i-\beta_i)\big)} \times \sin\Big(\frac{\pi}{2}\Big(|\Upsilon|-\frac{k}{2}+ks+\sum_{i=1}^{k}\alpha_i\Big)\Big)\Big(\prod_{i=1}^{k}\frac{\Gamma_i\big(\frac14-\frac{\alpha_i+s}{2}\big)}{\Gamma_i\big(\frac14+\frac{\alpha_i+s}{2}\big)}\frac{\pi^{\frac12}q^{-\alpha_i}}{2^{s-\frac12+\alpha_i}}\Big)\frac{H''_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}}(s)}{\pi s}\,ds,$$

so that by the functional equation (using that $|\Upsilon|$ is even) and the definition (7-34) of H'' we obtain

$$\mathscr{V}'_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}} = (-1)^{|\Upsilon|/2}\frac{q^{\frac{k}{2}-1}}{2\pi i}\int_{(c_s)} \frac{\zeta\big(\frac{k}{2}-ks-\sum_{i=1}^{k}\alpha_i\big)\zeta\big(\frac{k}{2}+ks+\sum_{i=1}^{k}\beta_i\big)}{\zeta\big(k-\sum_{i=1}^{k}(\alpha_i-\beta_i)\big)} \times G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)\Big(\prod_{i=1}^{k}\zeta(1-\alpha_i+\beta_i)\frac{\Gamma_i\big(\frac14-\frac{\alpha_i+s}{2}\big)\Gamma_i\big(\frac14+\frac{\beta_i+s}{2}\big)}{\Gamma_i\big(\big(\frac12+\alpha_i\big)/2\big)\Gamma_i\big(\big(\frac12+\beta_i\big)/2\big)}\Big(\frac{q}{\pi}\Big)^{-\alpha_i}\Big)\frac{ds}{s}.$$

Notice that changing s into $-s$ we obtain exactly $-X_{\boldsymbol{\beta},\boldsymbol{\alpha}}$ times the analogous term coming from $\mathscr{V}'_{I_k;-\boldsymbol{\beta},-\boldsymbol{\alpha}}$, but with line of integration $c_s = -9\frac{\varepsilon}{k}$. Thus, $\mathscr{V}'_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}} + X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathscr{V}'_{I_k;-\boldsymbol{\beta},-\boldsymbol{\alpha}}$ coincides with the residue of the above integral at $s=0$, that is

$$\mathscr{V}'_{I_k;\boldsymbol{\alpha},\boldsymbol{\beta}} + X_{\boldsymbol{\beta},\boldsymbol{\alpha}}\mathscr{V}'_{I_k;-\boldsymbol{\beta},-\boldsymbol{\alpha}} = (-1)^{|\Upsilon|/2}q^{k/2-1}\frac{\zeta\big(\frac{k}{2}-\sum_{i=1}^{k}\alpha_i\big)\zeta\big(\frac{k}{2}+\sum_{i=1}^{k}\beta_i\big)}{\zeta\big(k-\sum_{i=1}^{k}(\alpha_i-\beta_i)\big)}\prod_{i=1}^{k}\zeta(1-\alpha_i+\beta_i)\frac{\Gamma_i\big(\frac14-\frac{\alpha_i}{2}\big)}{\Gamma_i\big(\frac14+\frac{\alpha_i}{2}\big)}\Big(\frac{q}{\pi}\Big)^{-\alpha_i}.$$

Thus, Lemma 16 follows.

9. The terms far from the diagonal

We will use the following result of Young to prove Lemma 15.

Lemma 25. *Let q be prime and let L, $K \ll q^{1+\varepsilon}$ and let W be a smooth function with compact support on $\mathbb{R}_{>0}$. Then,*

$$\sum_{0<\ell<L}\Big|\sum_{(k,q)=1} \mathrm{e}\Big(\frac{\ell\bar{k}}{q}\Big)W\Big(\frac{k}{K}\Big)\Big| \ll L^{\frac12}q^{\frac34+\varepsilon} + q^{\varepsilon}K^{\frac12}L.$$

Proof. This is Proposition 4.3 of [Young 2011a] with the extra condition requiring q to be prime which easily allows us to remove the condition $(q,\ell)=1$ from the first sum. □

Proof of Lemma 15. For simplicity we shall take $\boldsymbol{\alpha}=\boldsymbol{\beta}=\mathbf{0} := (0,\dots,0)$, as the shifts don't play any role in this lemma and the same argument with obvious modifications works also when $\alpha_i, \beta_i \ll 1/\log q$.

By symmetry, we can assume that N_1 is the maximum of the N_i and that N_2 is the second largest. Also, we assume $N_1\cdots N_k \ll q^{k+\varepsilon}$ and $N_1 \gg q^{1+3\varepsilon}$, since otherwise the result is trivial.

Now, we start by observing that we can remove the condition $\pm_1 n_1 \pm_2 \cdots \pm_k n_k \neq 0$ in $\mathcal{O}''_\varepsilon := \mathcal{O}''_{\varepsilon,\mathbf{0},\mathbf{0}}$ at a cost of an admissible error:

$$\mathcal{O}''_\varepsilon = \sum_{d\,|\,q} d\frac{\mu(q/d)}{\varphi(q)} \sum_{d|(\pm_1 n_1\pm_2\cdots\pm_k n_k)} \frac{d(n_1)\cdots d(n_k)}{(n_1\cdots n_k)^{\frac12}} V_{\boldsymbol{\alpha},\boldsymbol{\beta}}\Big(\frac{n_1\cdots n_k}{q^k}\Big) P\Big(\frac{n_1}{N_1}\Big)\cdots P\Big(\frac{n_k}{N_k}\Big) + O\big(q^{A\varepsilon}(N_1\cdots N_k)^{\frac12}/N_1\big).$$

Next, we decompose n_1 into $n_1 = f_1 g_1$ and attach to the new variables two partitions of unity so that $f_1 \asymp F_1$, $g_1 \asymp G_1$ with $F_1 \geq G_1$ and $F_1 G_1 \asymp N_1$. We shall also add the condition $(g_1, q) = 1$ at a cost of an error which is easily seen to be $O(q^{k/2-2+A\varepsilon})$. Writing V and $P(n_1/N_1)$ in terms of their Mellin transform, we obtain

$$\mathcal{O}''_\varepsilon = \sideset{}{^\dagger}\sum_{\substack{F_1G_1\asymp N_1\\ G_1\le F_1}} \int'_{(c_s,c_{u_1})} q^{ks} N_1^{u_1} \sum_{0\le|m|<M} w_m(s)\mathscr{A}_m(s+u_1)G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)g_{\boldsymbol{\alpha},\boldsymbol{\beta}}\widetilde{P}(u_1)\,\frac{ds}{s}\,du_1 + O\big(q^{A\varepsilon}(N_1\cdots N_k)^{\frac12}/N_1 + q^{k/2-2+A\varepsilon}\big), \tag{9-1}$$

where $M := \min(2kN_2, q)$, the $\int'$ indicates that the integrals are truncated at $|u_1|, |s| \le q^\varepsilon$, the lines of integrations are $c_{u_1} = 0$, $c_s = \varepsilon/k$, and

$$\mathscr{A}_m(s) := \sum_{d\,|\,q} d\frac{\mu(q/d)}{\varphi(q)} \sum_{(g_1,q)=1}\ \sum_{f_1g_1\equiv m \pmod d} \frac{1}{(f_1g_1)^{\frac12+s}} P\Big(\frac{f_1}{F_1}\Big)P\Big(\frac{g_1}{G_1}\Big),$$
$$w_m(s) := \sum_{\pm_2 n_2\pm_3\cdots\pm_k n_k\equiv -m \pmod q} \frac{d(n_2)\cdots d(n_k)}{(n_2\cdots n_k)^{\frac12+s}} P\Big(\frac{n_2}{N_2}\Big)\cdots P\Big(\frac{n_k}{N_k}\Big).$$

Now, we apply Poisson's summation formula to the sum over f_1 and we see that for $\Re(s) = \varepsilon/k$

$$\begin{aligned}\mathscr{A}_m(s) &= \sum_{d\,|\,q}\frac{\mu(q/d)}{\varphi(q)} \sum_{(g_1,q)=1} \frac{P(g_1/G_1)}{g_1^{\frac12+s}} \sum_{0\le|\ell|\le\frac{dq^{A\varepsilon}}{F_1}} \mathrm{e}\Big(\frac{\ell m\overline{g_1}}{d}\Big)\int_0^\infty \frac{P(x/F_1)}{x^{\frac12+s}}\,\mathrm{e}\Big(-\frac{\ell x}{d}\Big)\,dx + O(q^{-1})\\ &= \mathscr{A}^*_m(s) + O(q^{-1}),\end{aligned}$$

where

$$\mathscr{A}^*_m(s) = \frac{F_1^{\frac12-s}}{\varphi(q)} \sum_{(g_1,q)=1} \frac{P(g_1/G_1)}{g_1^{\frac12+s}} \sum_{0<|\ell|\le L} \mathrm{e}\Big(\frac{\ell m\overline{g_1}}{q}\Big)\int_0^\infty \frac{P(x)}{x^{\frac12+s}}\,\mathrm{e}\Big(-\frac{\ell F_1 x}{q}\Big)\,dx$$

and $L = q^{1+A\varepsilon}/F_1$. Indeed, the sum over ℓ in the $d = 1$ summands contains only the term $\ell = 0$ which cancel with the $\ell = 0$ term from $d = q$. Thus, (9-1) becomes

$$\mathcal{O}''_\varepsilon = \sideset{}{^\dagger}\sum_{\substack{F_1G_1\asymp N_1,\\ G_1\le F_1}} \int'_{(c_s,c_{u_1})} q^{ks} N_1^{u_1} \sum_{0\neq|m|<M} w_m(s)\mathscr{A}^*_m(s+u_1)G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)g_{\boldsymbol{\alpha},\boldsymbol{\beta}}\widetilde{P}(u_1)\,\frac{ds}{s}\,du_1 + O(q^{k/2-\frac32+A\varepsilon} + q^{A\varepsilon}(N_1\cdots N_k)^{\frac12}/N_1),$$

since the contribution of the terms with $m=0$ can be bounded trivially by

$$\begin{aligned}&\ll \frac{L(F_1G_1)^{\frac12}}{q}(1+N_2/q)N_2^{-\frac12}(N_3\cdots N_k)^{\frac12}q^{A\varepsilon}\\ &\ll q^{A\varepsilon}N_2^{-\frac12}(N_3\cdots N_k)^{\frac12}+q^{-1+A\varepsilon}(N_2\cdots N_k)^{\frac12}\ll q^{k/2-\frac32+A\varepsilon}.\end{aligned}$$

since $N_2^{-1}N_3\cdots N_k\ll q^{k-3+A\varepsilon}$ and $N_2\cdots N_k\ll q^{k-2+A\varepsilon}$. Also, we assume $N_1\le F_1^2\le q^{2+2A\varepsilon}$, since otherwise $\mathscr{A}_m^*(s)$ is identically zero.

Changing the order of summation and integration and bounding trivially $G_{\boldsymbol{\alpha},\boldsymbol{\beta}}(s)g_{\boldsymbol{\alpha},\boldsymbol{\beta}}\widetilde{P}(u_1)$, we see that

$$\mathcal{O}_\varepsilon''\ll \sideset{}{^\dagger}\sum_{\substack{F_1G_1\asymp N_1\\ G_1\le F_1}}\int'_{(c_s,c_{u_1})}q^\varepsilon|E(s,u_1)|\,|ds\,du_1|+q^{k/2-\frac32+A\varepsilon}+q^{A\varepsilon}\frac{(N_1\cdots N_k)^{\frac12}}{N_1},\tag{9-2}$$

where

$$\begin{aligned}E(s,u_1)&:=\frac{F_1^{\frac12}}{\varphi(q)}\sum_{0<|\ell|\le L}\sum_{0<|m|<M}|w_m(s)|\Bigg|\sum_{(g_1,q)=1}\frac{P(g_1/G_1)}{g_1^{\frac12+s+u_1}}\,\mathrm{e}\,\frac{\ell m\overline{g_1}}{q}\Bigg|\\ &\ll \frac{F_1^{\frac12}}{qG_1^{\frac12}}\max_{0<|r|\le R}c_r\sum_{0\ne|r|\le R}\Bigg|\sum_{(g_1,q)=1}P\Big(\frac{g_1}{G_1}\Big)\Big(\frac{G_1}{g_1}\Big)^{\frac12+s+u_1}\mathrm{e}\Big(\frac{r\overline{g_1}}{q}\Big)\Bigg|,\end{aligned}$$

with $R:=\min(2kLN_2,q)\le 2kq^{1+A\varepsilon}\min(N_2/F_1,1)$ and

$$\begin{aligned}c_r&:=\sum_{\substack{\ell m\equiv r \pmod q\\ 0<|m|\le M,\ 0<|\ell|\le L}}|w_m(s)|\ll\sum_{\substack{0<|\ell|\le L,\ n_2\asymp N_2,\dots,n_k\asymp N_k\\ (\pm_2 n_2\pm_3\cdots\pm_k n_k)\ell\equiv -r \pmod q}}q^{A\varepsilon}d(n_2)\cdots d(n_2)(N_2\cdots N_k)^{-\frac12}\\ &\ll\sum_{\substack{0<|\ell|\le L,\ |n|\ll kN_2\\ n\ell\equiv -r \pmod q}}q^{A\varepsilon}N_2^{-\frac12}(N_3\cdots N_k)^{\frac12}\ll\sum_{\substack{|a|\ll kLN_2\\ a\equiv -r \pmod q}}d(a)q^{A\varepsilon}N_2^{-\frac12}(N_3\cdots N_k)^{\frac12}\\ &\ll q^{A\varepsilon}N_2^{-\frac12}(N_3\cdots N_k)^{\frac12}(1+LN_2/q).\end{aligned}$$

Thus, by Lemma 25, for $|s|,|u_1|\ll q^{A\varepsilon}$, $\Re(s)=\varepsilon/k$, $\Re(u_1)=0$ we have

$$\begin{aligned}E(s,u_1)&\ll q^{-1+A\varepsilon}F_1^{\frac12}(G_1N_2)^{-\frac12}(N_3\cdots N_k)^{\frac12}(1+LN_2/q)(R^{\frac12}q^{\frac34}+G_1^{\frac12}R)\\ &\ll q^{A\varepsilon}(F_1/G_1N_2)^{\frac12}(N_3\cdots N_k)^{\frac12}(1+N_2/F_1)\min\big(F_1^{-\frac12}N_2^{\frac12}q^{\frac14}+G_1^{\frac12}N_2/F_1,\ q^{\frac14}+G_1^{\frac12}\big)\\ &=q^{A\varepsilon}(N_3\cdots N_k)^{\frac12}(1+N_2/F_1)\min\Bigg(\frac{q^{\frac14}}{G_1^{\frac12}}+\frac{N_2^{\frac12}}{F_1^{\frac12}},\ \frac{F_1^{\frac12}}{N_2^{\frac12}}\frac{q^{\frac14}}{G_1^{\frac12}}+\frac{F_1^{\frac12}}{N_2^{\frac12}}\Bigg).\end{aligned}$$

For $x, y > 0$ we have $(1+x^2)\min(y+x, y/x+1/x) \le (x+1)(y+1)$, whence

$$
\begin{aligned}
E(s,u_1) &\ll q^{A\varepsilon}(N_3\cdots N_k)^{\frac12}\big(N_2^{\frac12}q^{\frac14}N_1^{-\frac12}+N_2^{\frac12}F_1^{-\frac12}+q^{\frac14+\varepsilon}G_1^{-\frac12}+1\big)\\
&\ll q^{A\varepsilon}(N_3\cdots N_k)^{\frac12}\big(N_2^{\frac12}N_1^{-\frac14}+q^{\frac34}N_1^{-\frac12}+1\big)\\
&\ll q^{A\varepsilon}\big((N_2\cdots N_k)^{\frac12}N_1^{-\frac14}+(N_2\cdots N_k)^{\frac12-1/(2(k-1))}(q^{\frac34}N_1^{-\frac12}+1)\big),
\end{aligned}
\tag{9-3}
$$

where in the second inequality we used that $N_1 \gg q$, $F_1^{-\frac12} \le N_1^{-\frac14}$ and $G_1^{-\frac12} \asymp (F_1/N_1)^{\frac12} \le N_1^{-\frac12}q^{\frac12+A\varepsilon}$, and in the third one that $N_3\cdots N_k \le N_2^{k-2}$ (so that $N_3\cdots N_k \le (N_2\cdots N_k)^{(k-2)/(k-1)}$). The lemma then follows by inserting (9-3) in (9-2). □

10. A Mellin formula

In this section we prove a formula to separate the variables in expressions of the form $(\pm_1 x_1 \pm_2 \cdots \pm_\kappa x_\kappa)^{-s}$ which generalizes the Mellin transforms given in the following lemma.

Lemma 26. *Let $x, y > 0$. Then*

$$
(x+y)^{-b} = \frac{1}{2\pi i}\int_{(c_v)} \frac{\Gamma(v)\Gamma(b-v)}{\Gamma(b)} x^{v-b}y^{-v}\,dv, \tag{10-1}
$$

for $0 < c_v < \Re(b)$. Moreover, for $\Re(b) < 0 < c_w$, we have

$$
(x-y)^{-b}\chi_{\mathbb{R}_{>0}}(x-y) = \frac{1}{2\pi i}\int_{(c_w)} \frac{\Gamma(w)\Gamma(1-b)}{\Gamma(1-b+w)} x^{w-b}y^{-w}\,dw, \tag{10-2}
$$

where $\chi_X(x)$ is the indicator function of the set X.

Equation (10-1) can be used repeatedly to give a formula for $(x_1+\cdots+x_\kappa)^{-s}$ valid for $\Re(s) > 0$. However, it is not straightforward to obtain a satisfactory formula valid in the case when there are some minus signs, as the integrals obtained by repeatedly applying (10-1) and (10-2) are not absolutely convergent. The following Lemma overcomes this problem by introducing an extra integration.

Lemma 27. *Let $\kappa \ge 2$ and $x_1, \dots x_\kappa > 0$. Let $\varepsilon = (\pm_1, \dots, \pm_\kappa 1) \in \{\pm 1\}^\kappa$, with $\pm_1 1 = -1$. Let $B \in \mathbb{Z}_{\ge 0}$ be such that $\frac{\kappa}{2}+\frac12 < \Re(v_1) < B+1$. Moreover, let $c_{v_2}, \dots, c_{v_\kappa}, c'_{v_2}, \dots, c'_{v_\kappa} > 0$ be such that*

$$
\Re(v_1) + c_{v_2} + \cdots + c_{v_\kappa} < B+1 < \Re(v_1) + c'_{v_2} + \cdots + c'_{v_\kappa}. \tag{10-3}
$$

Then

$$
\begin{aligned}
&(\pm_2 x_2 \pm_3 \cdots \pm_\kappa x_\kappa)^{v_1-1}\chi_{\mathbb{R}_{>0}}(\pm_2 x_2 \pm_3 \cdots \pm_\kappa x_\kappa)\\
&= \sum_{\substack{\nu=(\nu_2,\dots,\nu_\kappa)\in\mathbb{Z}_{\ge 0}^{\kappa-1}\\ \nu_2+\cdots+\nu_\kappa=B\\ \nu_i=0 \text{ if } \pm_i=-1}} \frac{B!}{\nu_2!\cdots\nu_\kappa!}\frac{1}{(2\pi i)^{\kappa-1}}\bigg(\int_{(c_{v_2},\dots,c_{v_\kappa})} - \int_{(c'_{v_2},\dots,c'_{v_\kappa})}\bigg)\frac{\Psi_{\varepsilon,B}(v_1,\dots,v_\kappa)}{x_2^{v_2-\nu_2}\cdots x_\kappa^{v_\kappa-\nu_\kappa}}\,dv_2\cdots dv_\kappa,
\end{aligned}
\tag{10-4}
$$

where

$$\Psi_{\varepsilon,B}(s_1,\dots,s_\kappa) := \frac{\Gamma(s_1)\cdots\Gamma(s_\kappa)}{\Gamma(V_{+;\varepsilon}(s_1,\dots,s_\kappa))\Gamma(V_{-;\varepsilon}(s_1,\dots,s_\kappa))}\,\frac{G(B+1-s_1-\cdots-s_\kappa)}{B+1-s_1-\cdots-s_\kappa},$$
$$V_{\pm;\varepsilon}(v_1,\dots,v_\kappa) := \sum_{\substack{1\le i\le\kappa\\ \pm_i 1=\pm 1}} v_i \tag{10-5}$$

and $G(s)$ is any entire function such that $G(0)=1$ and $G(\sigma+it)\ll e^{-C_1|t|}(1+|\sigma|)^{C_2|\sigma|}$ for some fixed $C_1, C_2>0$.

Remarks. (1) If $\varepsilon=(-1,\dots,-1)$, then Ψ_ε has to be interpreted as being identically zero.

(2) If $\xi(s)$ is the Riemann ξ-function, then $G(s)=\xi(s)/\xi(0)$ satisfies the hypothesis of the lemma.

Before giving a proof for Lemma 27, we give the following lemma which implies that the integrals in (10-4) are absolutely convergent.

Lemma 28. *Let $s_i=\sigma_i+it_i$ for $i=1,\dots,\kappa$. Then, for some $A>0$ we have*

$$\Psi_{\varepsilon,B}(s_1,\dots,s_\kappa) \ll \frac{1}{\delta^\kappa}\,\frac{(1+B+|\sigma_1|+\cdots+|\sigma_\kappa|)^{A(1+B+|\sigma_1|+\cdots+|\sigma_\kappa|)}}{(1+|t_1|)^{\frac12-\sigma_1}\cdots(1+|t_\kappa|)^{\frac12-\sigma_\kappa}(1+|t_1|+\cdots+|t_\kappa|)^{\sigma_1+\cdots+\sigma_\kappa-1}}, \tag{10-6}$$

provided that the s_i are located at a distance greater than $\delta>0$ from the poles of Ψ_ε.

Proof. By Stirling's formula (and the reflection's formula for the Gamma function), if the distance of $s=\sigma+it$ from the poles of $\Gamma(s)$ is greater than δ, then we have

$$\Gamma(s)\ll\frac{1}{\delta}(1+A_1|\sigma|)^{|\sigma|}(1+|t|)^{\sigma-\frac12}e^{-(\pi/2)|t|},\qquad \Gamma(s)^{-1}\ll(1+A_1|\sigma|)^{|\sigma|}(1+|t|)^{-\sigma+\frac12}e^{(\pi/2)|t|},$$

for some $A_1>0$. It follows that

$$\begin{aligned}\Psi_{\varepsilon,B}&(s_1,\dots,s_\kappa)\\ &\ll\frac{1}{\delta^\kappa}\,\frac{(1+B+|\sigma_1|+\cdots+|\sigma_\kappa|)^{A_2(1+B+|\sigma_1|+\cdots+|\sigma_\kappa|)}}{(1+|t_1|)^{\frac12-\sigma_1}\cdots(1+|t_\kappa|)^{\frac12-\sigma_\kappa}}\\ &\qquad\times\frac{e^{-(\pi/2)(|t_1|+\cdots+|t_\kappa|-|V_{+;\varepsilon}(t_1,\dots,t_\kappa)|-|V_{-;\varepsilon}(t_1,\dots,t_\kappa)|)-C_1|t_1+\cdots+t_\kappa|}}{(1+|V_{+;\varepsilon}(t_1,\dots,t_\kappa)|)^{V_{+;\varepsilon}(\sigma_1,\dots,\sigma_\kappa)-\frac12}(1+|V_{-;\varepsilon}(t_1,\dots,t_\kappa)|)^{V_{-;\varepsilon}(\sigma_1,\dots,\sigma_\kappa)-\frac12}},\end{aligned} \tag{10-7}$$

for some $A_2>0$. Now, we have

$$\frac{e^{-C_1|x+y|}}{(1+|x|)^{\eta_1}(1+|y|)^{\eta_2}}\ll\frac{(1+|\eta_1|+|\eta_2|)^{A_3(|\eta_1|+|\eta_2|)}}{(1+|x|+|y|)^{\eta_1+\eta_2}},$$

for some $A_3>0$ (depending on C_1). Thus, the factor on the second line of (10-7) is

$$\begin{aligned}&\ll(1+|\sigma_1|+\cdots+|\sigma_\kappa|)^{A_4(|\sigma_1|+\cdots+|\sigma_\kappa|)}\frac{e^{-(\pi/2)(|t_1|+\cdots+|t_\kappa|-|V_{+;\varepsilon}(t_1,\dots,t_\kappa)|-|V_{-;\varepsilon}(t_1,\dots,t_\kappa)|)}}{\big(1+|V_{+;\varepsilon}(t_1,\dots,t_\kappa)|+|V_{-;\varepsilon}(t_1,\dots,t_\kappa)|\big)^{\sigma_1+\cdots+\sigma_\kappa-1}}\\ &\ll\frac{(1+|\sigma_1|+\cdots+|\sigma_\kappa|)^{A_5(|\sigma_1|+\cdots+|\sigma_\kappa|)}}{(1+|t_1|+\cdots+|t_\kappa|)^{\sigma_1+\cdots+\sigma_\kappa-1}},\end{aligned}$$

and (10-6) follows. $\square$

Proof of Lemma 27. First, we remark that the estimate (10-6) implies the absolute convergence of the integrals on the right hand side of (10-4) and justifies the following computations.

we prove the lemma by induction. First we consider the case $\kappa = 2$. From (10-1) we have

$$\begin{aligned}(x_2+x_3)^{v_1-1} &= (x_2+x_3)^B(x_2+x_3)^{v_1-1-B}\\ &= \sum_{\substack{\nu_2,\nu_3\in\mathbb{Z}_{\geq 0}\\ \nu_2+\nu_3=B}} \frac{B!}{\nu_2!\nu_3!} x^{\nu_2} x_3^{\nu_3} \frac{1}{2\pi i}\int_{(c_{v_3})} \frac{\Gamma(v_3)\Gamma(1+B-v_1-v_3)}{\Gamma(1+B-v_1)x_2^{B+1-v_1-v_3}x_3^{v_3}}\,dv_3,\end{aligned} \tag{10-8}$$

for $0 < c_{v_3} < 1 + B - \Re(v_1)$. Now, by contour integration,

$$\frac{\Gamma(1+B-v_1-v_3)}{\Gamma(1+B-v_1)} x_2^{v_1+v_3-B-1} = \frac{1}{2\pi i}\left(\int_{(c_{v_2})} - \int_{(c'_{v_2})}\right) \frac{\Gamma(v_2)x_2^{-v_2}}{\Gamma(v_2+v_3)} \frac{G(B+1-v_1-v_2-v_3)}{B+1-v_1-v_2-v_3}\,dv_2,$$

where $c_{v_2}, c'_{v_2} > 0$ and $c_{v_2} < -\Re(v_1+v_3)+B+1 < c'_{v_2}$. Inserting this into (10-8) we obtain (10-4) in the case $\varepsilon = (-1, 1, 1)$.

The case $\varepsilon = (-1, 1, -1)$ (and thus its permutation $\varepsilon = (-1, -1, 1)$) follows in the same way from (10-2).

Now, let $\varepsilon = (-1, \pm_2, \dots, \pm_{\kappa+1}) \in \{\pm 1\}^{\kappa+1}$ with $\kappa \geq 2$ and suppose (10-4) holds for all $\varepsilon' \in \{\pm 1\}^{\kappa}$ with $\pm'_1 1 = -1$. Since $\kappa + 1 \geq 3$ there are two indexes $2 \leq i < j \leq \kappa + 1$ such that $\pm_i 1 = \pm_j 1$ and without loss of generality we can assume $i = \kappa,\ j = \kappa+1$. Then, letting $\varepsilon' = (-1, \pm_2, \dots, \pm_\kappa)$, we have

$$\begin{aligned}&(\pm_2 x_2 \pm_3 \cdots \pm_{\kappa+1} x_{\kappa+1})^{v_1-1}\chi_{\mathbb{R}_{>0}}(\pm_2 x_2 \pm_3 \cdots \pm_{\kappa+1} x_{\kappa+1})\\ &\quad= \sum_{\substack{\nu=(\nu_2,\dots,\nu_\kappa)\in\mathbb{Z}^{\kappa}_{\geq 0}\\ \nu_2+\cdots+\nu_{\kappa+1}=B\\ \nu_i=0 \text{ if } \pm_i=-1}} \frac{B!}{\nu_2!\cdots\nu_\kappa!}\frac{1}{(2\pi i)^{\kappa-1}}\\ &\qquad\qquad\times\left(\int_{(c_{v_2},\dots,c_{v_\kappa})} - \int_{(c'_{v_2},\dots,c'_{v_\kappa})}\right)\frac{\Psi_{\varepsilon',B}(v_1,\dots,v_\kappa)}{x_2^{v_2-\nu_2}\cdots x_{\kappa-1}^{v_{\kappa-1}-\nu_{\kappa-1}}}(x_\kappa+x_{\kappa+1})^{-v_\kappa+\nu_\kappa}\,dv_2\cdots dv_\kappa\end{aligned}$$

where $c_{v_2}, \dots, c_{v_\kappa}, c'_{v_2}, \dots, c'_{v_\kappa} > 0$ satisfy (10-3). Then, we expand the binomial $(x_\kappa + x_{\kappa+1})^{\nu_\kappa}$ and apply (10-1) to $(x_\kappa + x_{\kappa+1})^{-v_\kappa}$. We obtain

$$\begin{aligned}&(\pm_2 x_2 \pm_3 \cdots \pm_{\kappa+1} x_{\kappa+1})^{v_1-1}\chi_{\mathbb{R}_{>0}}(\pm_2 x_2 \pm_3 \cdots \pm_{\kappa+1} x_{\kappa+1})\\ &\quad= \sum_{\substack{\nu=(\nu_2,\dots,\nu_{\kappa+1})\in\mathbb{Z}^{\kappa}_{\geq 0}\\ \nu_2+\cdots+\nu_{\kappa+1}=B\\ \nu_i=0 \text{ if } \pm_i=-1}} \frac{B!}{\nu_2!\cdots\nu_{\kappa+1}!}\frac{1}{(2\pi i)^{\kappa}}\left(\int_{(c_{v_2},\dots,c_{v_{\kappa+1}})} - \int_{(c'_{v_2},\dots,c'_{v_{\kappa+1}})}\right)\\ &\qquad\qquad\times\frac{\Psi_{\varepsilon',B}(v_1,\dots,v_\kappa)}{x_2^{v_2-\nu_2}\cdots x_{\kappa-1}^{v_{\kappa-1}-\nu_{\kappa-1}}x_\kappa^{v_\kappa-v_{\kappa+1}-\nu_\kappa}x_{\kappa+1}^{v_{\kappa+1}-\nu_{\kappa+1}}}\frac{\Gamma(v_{\kappa+1})\Gamma(v_\kappa-v_{\kappa+1})}{\Gamma(v_\kappa)}\,dv_2\cdots dv_\kappa,\end{aligned}$$

where $c_{v_{\kappa+1}}, c'_{v_{\kappa+1}} > 0$ are such that $0 < c_{v_{\kappa+1}} < c_{v_\kappa}$. We make the change of variables $v_\kappa \to v_\kappa + v_{\kappa+1}$ and the lemma follows. $\square$

Lemma 29. *Let $\kappa \geq 2$ and let $\varepsilon^* = (\pm_1 1, \dots, \pm_\kappa 1, \pm_* 1) \in \{\pm 1\}^{\kappa+1}$, with $\pm_1 1 = -1$. For $B \geq 0$, let*

$$\Psi'_{\varepsilon^*,B}(v_1, \dots, v_\kappa) := \frac{\Gamma(B+1-v_1-\dots-v_\kappa)\Gamma(v_1)\cdots\Gamma(v_\kappa)}{\Gamma(V_{\mp_*;\varepsilon}(v_1,\dots,v_\kappa))\Gamma(B+1-V_{\mp_*;\varepsilon}(v_1,\dots,v_\kappa))},$$

where $V_{\pm;\varepsilon}$ is defined in (10-5).

Let $\mathcal{F}(v_0, \dots, v_\kappa)$ be analytic on $\{(v_0,\dots,v_\kappa) \in \mathbb{C}^{\kappa+1} \mid 0 < \Re(v_0) < B+1\}$ and assume that for $0 < \Re(v_0) < B+1$ and any $A > 0$ one has that $\mathcal{F}$ satisfies

$$\mathcal{F}(v_0,\dots,v_\kappa) \ll \prod_{i=2}^{\kappa} (1+|v_i|)^{-A}$$

where the implicit constant may depend on A, v_1 and $\Re(v_0)$. Then for any $v_1 \in \mathbb{C}$ and $c_{v_2}, \dots, c_{v_\kappa} > 0$ satisfying $0 < \Re(v_1) + c_{v_2} + \dots + c_{v_\kappa} < 1$ we have

$$\begin{aligned}\sum_{\nu} \frac{B!}{\nu_*!\nu_2!\cdots\nu_\kappa!} \int_{(c_{v_2},\dots,c_{v_\kappa})} &\Psi'_{\varepsilon,B}(v_1,\dots,v_\kappa) \\ &\times \mathcal{F}\big(B+1-\nu_*-v_1-\dots-v_\kappa, v_1, v_2-\nu_2, \dots, v_\kappa-\nu_\kappa\big)\, dv_2\cdots dv_\kappa \\ &= \int_{(c_{v_2},\dots,c_{v_\kappa})} \Psi'_{\varepsilon,0}(v_1,\dots,v_\kappa)\mathcal{F}\big(1-v_1-\dots-v_\kappa, v_1, \dots, v_\kappa\big)\, dv_2\cdots dv_\kappa, \qquad (10\text{-}9)\end{aligned}$$

where the sum on the left is taken over $\nu = (\nu_2, \dots, \nu_\kappa, \nu_*) \in \mathbb{Z}^\kappa_{\geq 0}$ satisfying

$$\nu_2 + \dots + \nu_\kappa + \nu_* = B, \qquad \nu_i = 0 \quad \text{if } \pm_i = -1 \text{ or } i \in J, \qquad \nu_* = 0 \quad \text{if } \pm_* = -1.$$

Proof. Making the change of variables $v_i \to v_i + \nu_i$, for $i = 2, \dots, \kappa$, moving back the lines of integration to c_{v_i} (as we can do without crossing any pole), and switching the order of summation and integration, we see that the left hand side of (10-9) is equal to

$$\int_{(c_{v_2},\dots,c_{v_\kappa})} \sum_{\nu} \frac{B!}{\nu_*!\nu_2!\cdots\nu_\kappa!} \Psi'_\varepsilon(v_1, v_2+\nu_2, \dots, v_\kappa+\nu_\kappa)\mathcal{F}(1-v_1-\dots-v_\kappa, v_1, \dots, v_\kappa)\, dv_2\cdots dv_\kappa.$$

Now, the identity $\mathrm{B}(s_1+1, s_2) + \mathrm{B}(s_1, s_2+1) = \mathrm{B}(s_1, s_2)$, satisfied by the Beta function $\mathrm{B}(s_1, s_2) := \Gamma(s_1)\Gamma(s_2)\Gamma(s_1+s_2)^{-1}$, can be generalized to

$$\sum_{\substack{(r_1,\dots,r_m)\in\mathbb{Z}^m_{\geq 0} \\ r_1+\dots+r_m = r}} \frac{r!}{r_1!\cdots r_m!} \frac{\Gamma(s_1+r_1)\cdots\Gamma(s_m+r_m)}{\Gamma(r+s_1+\dots+s_m)} = \frac{\Gamma(s_1)\cdots\Gamma(s_m)}{\Gamma(s_1+\dots+s_m)},$$

for $m, r \geq 1$, $s_1, \dots, s_m \in \mathbb{C}$. Thus, we have

$$\sum_{\substack{\nu=(\nu_2,\dots,\nu_\kappa,\nu_*)\in\mathbb{Z}^\kappa_{\geq 0} \\ \nu_2+\dots+\nu_\kappa+\nu_*=B \\ \nu_i=0 \text{ if } \pm_i=-1 \\ \nu_*=0 \text{ if } \pm_*=-1}} \frac{B!}{\nu_*!\nu_2!\cdots\nu_\kappa!} \Psi_{\varepsilon,B}(v_1, v_2+\nu_2 \dots, v_\kappa+\nu_\kappa) = \Psi'_{\varepsilon,0}(v_1,\dots,v_\kappa)$$

and the lemma follows. □

Acknowledgments

The author would like to thank Sary Drappeau, Dimitris Koukoulopoulos and Maksym Radziwiłł for helpful discussions and the referee for several useful suggestions.

This paper was started while the author was a postdoc at the Centre de Recherche Mathématiques (CRM) in Montréal, and was completed during another visit at the same institution. The author is grateful to the CRM for the hospitality and for providing a stimulating working environment.

The author is member of the INdAM group GNAMPA and is partially supported by PRIN "Number Theory and Arithmetic Geometry".

References

[Apostol 1951] T. M. Apostol, "On the Lerch zeta function", *Pacific J. Math.* **1** (1951), 161–167. MR Zbl

[Apostol 1976] T. M. Apostol, *Introduction to analytic number theory*, Springer, 1976. MR Zbl

[Baladi and Vallée 2005] V. Baladi and B. Vallée, "Euclidean algorithms are Gaussian", *J. Number Theory* **110**:2 (2005), 331–386. MR Zbl

[Beck 2003] M. Beck, "Dedekind cotangent sums", *Acta Arith.* **109**:2 (2003), 109–130. MR Zbl

[Bettin 2015] S. Bettin, "On the distribution of a cotangent sum", *Int. Math. Res. Not.* **2015**:21 (2015), 11419–11432. MR Zbl

[Bettin 2016] S. Bettin, "On the reciprocity law for the twisted second moment of Dirichlet L-functions", *Trans. Amer. Math. Soc.* **368**:10 (2016), 6887–6914. MR Zbl

[Bettin 2017] S. Bettin, "Linear correlations of the divisor function", preprint, 2017. To appear in *Acta Arith.* arXiv

[Bettin and Conrey 2013a] S. Bettin and B. Conrey, "Period functions and cotangent sums", *Algebra Number Theory* **7**:1 (2013), 215–242. MR Zbl

[Bettin and Conrey 2013b] S. Bettin and J. B. Conrey, "A reciprocity formula for a cotangent sum", *Int. Math. Res. Not.* **2013**:24 (2013), 5709–5726. MR Zbl

[Blomer and Milićević 2015] V. Blomer and D. Milićević, "The second moment of twisted modular L-functions", *Geom. Funct. Anal.* **25**:2 (2015), 453–516. MR Zbl

[Blomer et al. 2007] V. Blomer, G. Harcos, and P. Michel, "A Burgess-like subconvex bound for twisted L-functions", *Forum Math.* **19**:1 (2007), 61–105. MR Zbl

[Blomer et al. 2017] V. Blomer, É. Fouvry, E. Kowalski, P. Michel, and D. Milićević, "On moments of twisted L-functions", *Amer. J. Math.* **139**:3 (2017), 707–768. MR Zbl

[Bykovskiĭ 2005] V. A. Bykovskiĭ, "An estimate for the dispersion of lengths of finite continued fractions", *Fundam. Prikl. Mat.* **11**:6 (2005), 15–26. In Russian; translation in *J. Math. Sci* **146**:2 (2007), 5634–5643. MR

[Chinta 2005] G. Chinta, "Mean values of biquadratic zeta functions", *Invent. Math.* **160**:1 (2005), 145–163. MR Zbl

[Conrey 1989] J. B. Conrey, "More than two fifths of the zeros of the Riemann zeta function are on the critical line", *J. Reine Angew. Math.* **399** (1989), 1–26. MR Zbl

[Conrey and Ghosh 2006] J. B. Conrey and A. Ghosh, "Remarks on the generalized Lindelöf hypothesis", *Funct. Approx. Comment. Math.* **36** (2006), 71–78. MR Zbl

[Conrey and Iwaniec 2000] J. B. Conrey and H. Iwaniec, "The cubic moment of central values of automorphic L-functions", *Ann. of Math.* (2) **151**:3 (2000), 1175–1216. MR Zbl

[Conrey et al. 1986] J. B. Conrey, A. Ghosh, and S. M. Gonek, "Large gaps between zeros of the zeta-function", *Mathematika* **33**:2 (1986), 212–238. MR Zbl

[Deshouillers and Iwaniec 1982] J.-M. Deshouillers and H. Iwaniec, "Kloosterman sums and Fourier coefficients of cusp forms", *Invent. Math.* **70**:2 (1982), 219–288. MR Zbl

[Estermann 1930] T. Estermann, "On the representations of a number as the sum of two products", *Proc. London Math. Soc.* (2) **31**:2 (1930), 123–133. MR Zbl

[Estermann 1932] T. Estermann, "On the representations of a number as the sum of three or more products", *Proc. London Math. Soc.* (2) **34**:3 (1932), 190–195. MR Zbl

[Hardy and Littlewood 1916] G. H. Hardy and J. E. Littlewood, "Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes", *Acta Math.* **41**:1 (1916), 119–196. MR JFM

[Harman et al. 2004] G. Harman, N. Watt, and K. Wong, "A new mean-value result for Dirichlet L-functions and polynomials", *Q. J. Math.* **55**:3 (2004), 307–324. MR Zbl

[Heath-Brown 1979] D. R. Heath-Brown, "The fourth power moment of the Riemann zeta function", *Proc. London Math. Soc.* (3) **38**:3 (1979), 385–422. MR Zbl

[Heath-Brown 1981] D. R. Heath-Brown, "The fourth power mean of Dirichlet's L-functions", *Analysis* **1**:1 (1981), 25–32. MR Zbl

[Heilbronn 1969] H. Heilbronn, "On the average length of a class of finite continued fractions", pp. 87–96 in *Number theory and analysis*, edited by P. Turán, Plenum, New York, 1969. MR Zbl

[Hensley 1994] D. Hensley, "The number of steps in the Euclidean algorithm", *J. Number Theory* **49**:2 (1994), 142–182. MR Zbl

[Ishibashi 1995] M. Ishibashi, "The value of the Estermann zeta functions at $s = 0$", *Acta Arith.* **73**:4 (1995), 357–361. MR Zbl

[Iwaniec and Sarnak 2000] H. Iwaniec and P. Sarnak, "The non-vanishing of central values of automorphic L-functions and Landau–Siegel zeros", *Israel J. Math.* **120**:A (2000), 155–177. MR Zbl

[Kim 2003] H. H. Kim, "Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2", *J. Amer. Math. Soc.* **16**:1 (2003), 139–183. MR Zbl

[Maier and Rassias 2016] H. Maier and M. T. Rassias, "Generalizations of a cotangent sum associated to the Estermann zeta function", *Commun. Contemp. Math.* **18**:1 (2016), art. id. 1550078. MR Zbl

[Motohashi 1980] Y. Motohashi, "An asymptotic series for an additive divisor problem", *Math. Z.* **170**:1 (1980), 43–63. MR Zbl

[Motohashi 1994] Y. Motohashi, "The binary additive divisor problem", *Ann. Sci. École Norm. Sup.* (4) **27**:5 (1994), 529–572. MR Zbl

[Motohashi 1997] Y. Motohashi, *Spectral theory of the Riemann zeta-function*, Cambridge Tracts in Math. **127**, Cambridge Univ. Press, 1997. MR Zbl

[Paley 1931] R. E. A. C. Paley, "On the k-analogues of some theorems in the theory of the Riemann sigma-function", *Proc. London Math. Soc.* (2) **32**:1 (1931), 273–311. MR Zbl

[Porter 1975] J. W. Porter, "On a theorem of Heilbronn", *Mathematika* **22**:1 (1975), 20–28. MR Zbl

[Soundararajan 2000] K. Soundararajan, "Nonvanishing of quadratic Dirichlet L-functions at $s = \frac{1}{2}$", *Ann. of Math.* (2) **152**:2 (2000), 447–488. MR Zbl

[Soundararajan 2007] K. Soundararajan, "The fourth moment of Dirichlet L-functions", pp. 239–246 in *Analytic number theory* (Göttingen, 2005), edited by W. Duke and Y. Tschinkel, Clay Math. Proc. **7**, Amer. Math. Soc., Providence, RI, 2007. MR Zbl

[Titchmarsh 1986] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., Oxford Univ. Press, 1986. MR Zbl

[Tonkov 1974] T. Tonkov, "The mean length of finite continued fractions", *Math. Balkanica* **4** (1974), 617–629. In Russian. MR Zbl

[Watt 2005] N. Watt, "Fourier coefficients of modular forms and eigenvalues of a Hecke operator", *Funct. Approx. Comment. Math.* **34** (2005), 27–146. MR Zbl

[Yao and Knuth 1975] A. C. Yao and D. E. Knuth, "Analysis of the subtractive algorithm for greatest common divisors", *Proc. Nat. Acad. Sci. U.S.A.* **72**:12 (1975), 4720–4722. MR Zbl

[Young 2011a] M. P. Young, "The fourth moment of Dirichlet L-functions", *Ann. of Math.* (2) **173**:1 (2011), 1–50. MR Zbl

[Young 2011b] M. P. Young, "The reciprocity law for the twisted second moment of Dirichlet L-functions", *Forum Math.* **23**:6 (2011), 1323–1337. MR Zbl

[Zagier 2010] D. Zagier, "Quantum modular forms", pp. 659–675 in *Quanta of maths* (Paris, 2007), edited by E. Blanchard et al., Clay Math. Proc. **11**, Amer. Math. Soc., Providence, RI, 2010. MR Zbl

Communicated by Philippe Michel
Received 2017-02-14 Revised 2018-05-08 Accepted 2018-08-10

sandro.bettin@gmail.com *Dipartimento di Matematica, Università di Genova, 16146 Genova, Italy*

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.301

Le théorème de Fermat sur certains corps de nombres totalement réels

Alain Kraus

Soit K un corps de nombres totalement réel. Pour tout nombre premier $p \geq 5$, notons F_p la courbe de Fermat d'équation $x^p + y^p + z^p = 0$. Sous l'hypothèse que 2 est totalement ramifié dans K, on établit quelques résultats sur l'ensemble $F_p(K)$ des points de F_p rationnels sur K. On obtient un critère pour que le théorème de Fermat asymptotique soit vrai sur K, critère relatif à l'ensemble des newforms modulaires paraboliques de Hilbert sur K, de poids parallèle 2 et de niveau l'idéal premier au-dessus de 2. Il peut souvent se tester simplement numériquement, notamment quand le nombre de classes restreint de K vaut 1. Par ailleurs, en utilisant la méthode modulaire, on démontre le théorème de Fermat de façon effective, sur certains corps de nombres dont les degrés sur $\mathbb{Q}$ sont 3, 4, 5, 6 et 8.

Let K be a totally real number field. For all prime number $p \geq 5$, let us denote by F_p the Fermat curve of equation $x^p + y^p + z^p = 0$. Under the assumption that 2 is totally ramified in K, we establish some results about the set $F_p(K)$ of points of F_p rational over K. We obtain a criterion so that the asymptotic Fermat's last theorem is true over K, criterion related to the set of Hilbert modular cuspidal newforms over K, of parallel weight 2 and of level the prime ideal above 2. It is often simply testable numerically, particularly if the narrow class number of K is 1. Furthermore, using the modular method, we prove Fermat's last theorem effectively, over some number fields whose degrees over $\mathbb{Q}$ are 3, 4, 5, 6 and 8.

1. Introduction

Soient K un corps de nombres totalement réel et $p \geq 5$ un nombre premier. Notons

$$F_p : x^p + y^p + z^p = 0 \tag{1-1}$$

la courbe de Fermat d'exposant p. Dans le cas où 2 est totalement ramifié dans K, on se propose de faire quelques remarques sur la description de l'ensemble $F_p(K)$ des points de F_p rationnels sur K. On se préoccupera en particulier de cette description quand de plus le nombre de classes restreint de K vaut 1.

Adoptons la terminologie selon laquelle $F_p(K)$ est trivial, si pour tout point $(x, y, z) \in F_p(K)$ on a $xyz = 0$. Tel est le cas si $K = \mathbb{Q}$ [Wiles 1995]. On dira que le théorème de Fermat asymptotique est vrai sur K, si $F_p(K)$ est trivial dès que p est plus grand qu'une constante qui ne dépend que de K. Parce que K ne contient pas les racines cubiques de l'unité, la conjecture *abc* sur K implique le théorème de Fermat asymptotique sur K (cf. [Browkin 2006]).

MSC2010 : primary 11D41; secondary 11G05, 11R37.
Mots-clefs : Fermat equation, number fields, elliptic curves, modular method.

Au cours de ces dernières années, les principaux résultats qui ont été établis concernant l'équation de Fermat sur les corps totalement réels sont dus à Freitas et Siksek. Ils ont notamment obtenu un critère permettant parfois de démontrer le théorème de Fermat asymptotique sur un corps totalement réel [Freitas et Siksek 2015a, Theorem 3]. En particulier, ils en ont déduit le théorème de Fermat asymptotique pour une proportion de $5/6$ de corps quadratiques réels. Ils ont par ailleurs démontré que pour $K=\mathbb{Q}(\sqrt{m})$, où $m\le 23$ est un entier sans facteurs carrés, autre que 5 et 17, l'ensemble $F_p(K)$ est trivial pour tout $p\ge 5$ [Freitas et Siksek 2015b]. Le cas où $m=2$ avait déjà été établi dans [Jarvis et Meekin 2004].

1.1. ***Le critère de Freitas et Siksek.*** Énonçons leur résultat dans le cas où 2 est totalement ramifié dans K.

On notera dans toute la suite, d le degré de K sur $\mathbb{Q}$, O_K l'anneau d'entiers de K et $\mathscr{L}$ l'idéal premier de O_K au-dessus de 2. On a $2O_K=\mathscr{L}^d$.

Soient $v_{\mathscr{L}}$ la valuation sur K associée à $\mathscr{L}$ et $U_{\mathscr{L}}$ le groupe des $\{\mathscr{L}\}$-unités de K. Posons

$$S=\{a\in U_{\mathscr{L}}\mid 1-a\in U_{\mathscr{L}}\}.$$

Désignons par (**FS**) la condition suivante :

(**FS**) pour tout $a\in S$, on a

$$|v_{\mathscr{L}}(a)|\le 4d. \tag{1-2}$$

Leur critère est le suivant :

Théorème 1. *Supposons que la condition* (**FS**) *soit satisfaite par K. Alors, le théorème de Fermat asymptotique est vrai sur K.*

L'ensemble S est fini [Siegel 1929]. Avec les travaux de Smart, on dispose d'algorithmes permettant d'expliciter S, sous réserve que le groupe $U_{\mathscr{L}}$ soit connu [Smart 1998]. Dans ce cas, la condition (**FS**) est donc en principe testable sur K. Cela étant, la détermination de S n'est pas pour l'instant implémentée dans des logiciels de calcul et expliciter S reste un travail généralement important. Par exemple, pour le sous-corps totalement réel maximal $\mathbb{Q}(\mu_{16})^+$ du corps cyclotomique des racines 16-ièmes de l'unité, S est de cardinal 585 [Freitas et Siksek 2015a, 1.3]. On peut vérifier que la condition (**FS**) est satisfaite, en particulier le théorème de Fermat asymptotique est vrai sur le corps $\mathbb{Q}(\mu_{16})^+$ [loc. cit.].

Dans l'objectif de démontrer le théorème de Fermat asymptotique sur certains corps de nombres, dans lesquels 2 est totalement ramifié, on va introduire ci-dessous une nouvelle condition, qui d'un point de vue numérique a l'avantage, à ce jour, de pouvoir se tester souvent simplement sur machine. On établit dans le théorème 2 qu'elle est équivalente à (**FS**), moyennant une hypothèse de modularité pour certaines courbes elliptiques sur K.

1.2. ***La condition* (C).** Le nombre premier 2 étant supposé totalement ramifié dans K, désignons par $\mathcal{H}$ l'ensemble des newforms modulaires paraboliques de Hilbert sur K, de poids parallèle 2 et de niveau $\mathscr{L}$. C'est un système fini libre sur $\mathbb{C}$. Pour tout $\mathfrak{f}\in\mathcal{H}$ et tout idéal premier non nul $\mathfrak{q}$ de O_K, notons $a_{\mathfrak{q}}(\mathfrak{f})$ le coefficient de Fourier de $\mathfrak{f}$ en $\mathfrak{q}$. C'est un entier algébrique. Le sous-corps $\mathbb{Q}_{\mathfrak{f}}$ de $\mathbb{C}$ engendré par les coefficients $a_{\mathfrak{q}}(\mathfrak{f})$ est une extension finie de $\mathbb{Q}$. C'est un corps totalement réel ou un corps CM. (Voir par exemple [Cremona et Dembélé 2014; Dembélé et Voight 2013].)

Pour tout idéal premier $\mathfrak{q}$ de O_K, notons $\mathrm{Norm}(\mathfrak{q})$ sa norme sur $\mathbb{Q}$.

La condition est la suivante :

(**C**) pour tout $\mathfrak{f} \in \mathcal{H}$ tel que $\mathbb{Q}_{\mathfrak{f}} = \mathbb{Q}$, il existe un idéal premier $\mathfrak{q}$ de O_K, distinct de $\mathscr{L}$, tel que l'on ait

$$a_{\mathfrak{q}}(\mathfrak{f}) \not\equiv \mathrm{Norm}(\mathfrak{q}) + 1 \pmod 4. \tag{1-3}$$

Elle est en apparence moins simple que la condition (**FS**), mais on peut généralement la tester en utilisant le logiciel de calcul `Magma` [Bosma et al. 1997], disons si $d \leq 8$ et si le discriminant de K n'est pas trop grand. Dans le cas où $d \leq 6$, on dispose également de tables de newforms décrivant $\mathcal{H}$, qui sont directement implémentées dans [LMFDB 2013].

Signalons par ailleurs que pour établir le théorème de Fermat asymptotique sur K de façon effective, la détermination de $\mathcal{H}$ est, comme on le verra, a priori indispensable dans la mise en œuvre de la méthode modulaire.

À titre indicatif, le corps quadratique réel de plus petit discriminant pour lequel la condition (**C**) n'est pas satisfaite est $K = \mathbb{Q}(\sqrt{114})$. Il existe une courbe elliptique sur K, de conducteur $\mathscr{L}$, ayant tous ses points d'ordre 2 rationnels sur K [LMFDB 2013], ce qui explique pourquoi (**C**) n'est pas réalisée sur ce corps (théorème 2, lemme 12 et [Freitas et al. 2015]) ; il en est ainsi pour une infinité de corps quadratiques. Par exemple, en utilisant la table 1 de [Freitas et Siksek 2015a], on peut démontrer qu'il existe une infinité de corps quadratiques réels $\mathbb{Q}(\sqrt{m})$, avec m sans facteurs carrés congru à 7 modulo 8, pour lesquels la condition (**C**) n'est pas satisfaite.

Voyons un autre exemple illustrant la condition (**C**). Soit K le sous-corps totalement réel maximal du corps cyclotomique $\mathbb{Q}(\mu_{48})$. C'est le corps totalement réel de degré 8 sur $\mathbb{Q}$ de plus petit discriminant, dans lequel 2 soit totalement ramifié (on peut le vérifier avec les tables de [Voight]). On constate avec `Magma` que l'on a $|\mathcal{H}| = 16$ et que la condition (**C**) est satisfaite par K, car il n'existe pas de newforms $\mathfrak{f} \in \mathcal{H}$ telles que $\mathbb{Q}_{\mathfrak{f}} = \mathbb{Q}$; pour tout $\mathfrak{f} \in \mathcal{H}$, on a $[\mathbb{Q}_{\mathfrak{f}} : \mathbb{Q}] = 4$.

1.3. *Hypothèse sur le nombre de classes restreint de K.* Notons h_K^+ le nombre de classes restreint de K. Rappelons que h_K^+ est le degré sur K de l'extension abélienne de K non ramifiée aux places finies maximale.

Malgré de nombreux essais expérimentaux, je ne suis pas parvenu à trouver un exemple de corps de nombres totalement réel K tel que 2 soit totalement ramifié dans K et que $h_K^+ = 1$, sans que la condition (**C**) soit satisfaite. En particulier, a-t-on toujours l'implication

$$2O_K = \mathscr{L}^d \quad \text{et} \quad h_K^+ = 1 \quad \Rightarrow \quad (\mathbf{C}) \quad ?$$

En fait, la réponse est positive si on a $d \in \{1, 2, 4, 8\}$. La raison étant que pour tout entier n, il existe au plus un corps totalement réel K pour lequel on a $d = 2^n$, $2O_K = \mathscr{L}^d$ et $h_K^+ = 1$, à savoir le sous-corps totalement réel maximal du corps cyclotomique des racines 2^{n+2}-ièmes de l'unité (théorème 6). On indiquera par ailleurs quelques constatations numériques en faveur de cette implication.

Signalons que, comme conséquence du théorème 4, si cette implication était toujours vraie, cela impliquerait le théorème de Fermat asymptotique sur tout corps de nombres totalement réels pour lesquels $2\mathcal{O}_K = \mathscr{L}^d$ et $h_K^+ = 1$.

Les hypothèses selon lesquelles $2\mathcal{O}_K = \mathscr{L}^d$ et $h_K^+ = 1$ sont très favorables dans l'application de la méthode modulaire pour obtenir des versions effectives du théorème de Fermat sur K. Cela est notamment dû au fait qu'avec ces hypothèses, on peut normaliser toute solution de l'équation de Fermat (1-1) de façon simple (proposition 16). On illustrera cette méthode pour certains corps de nombres de degré $d \in \{3, 4, 5, 6, 8\}$. On établira par exemple que pour le corps $K = \mathbb{Q}(\mu_{16})^+$, l'ensemble $F_p(K)$ est trivial pour tout $p \geq 5$.

Tous les calculs numériques que cet article a nécessités ont été effectués avec les logiciels de calcul `Pari` [PARI 2015] et `Magma`.

Remarque. Pendant la période de l'examen de cet article par le referee, Freitas et Siksek ont démontré une version généralisée de l'implication suggérée ci-dessus ; voir [Freitas et Siksek 2018]. En particulier, compte tenu du théorème 4, si 2 est totalement ramifié dans K et si $h_K^+ = 1$, le théorème de Fermat asymptotique est vrai sur K.

Partie I. Énoncé des résultats

Soit K un corps de nombres totalement réel.

Rappelons qu'une courbe elliptique E/K est dite modulaire s'il existe une newform modulaire parabolique de Hilbert sur K, de poids parallèle 2 et de niveau le conducteur de E, ayant la même fonction L que celle de E.

Conjecturalement, toute courbe elliptique définie sur K est modulaire. Cela est démontré si $K = \mathbb{Q}$ [Wiles 1995; Taylor et Wiles 1995; Breuil et al. 2001] et si K est un corps quadratique [Freitas et al. 2015]. Par ailleurs, à $\overline{K}$-isomorphisme près, l'ensemble des courbes elliptiques sur K qui ne sont pas modulaires est fini [loc. cit.].

2. Les conditions (C) et (FS)

Théorème 2. *Supposons que les deux conditions suivantes soient satisfaites :*

(1) 2 *est totalement ramifié dans* K.

(2) *Toute courbe elliptique définie sur* K, *de conducteur* $\mathscr{L}$, *ayant tous ses points d'ordre* 2 *rationnels sur* K, *est modulaire.*

Alors, les conditions **(C)** *et* **(FS)** *sont équivalentes.*

Si K est un corps quadratique dans lequel 2 est totalement ramifié, les conditions **(C)** et **(FS)** sont donc équivalentes.

3. Théorème de Fermat asymptotique

Comme conséquence directe des théorèmes 1 et 2, on obtient l'énoncé suivant :

Théorème 3. *Supposons que les trois conditions suivantes soient satisfaites :*

(1) 2 *est totalement ramifié dans* K.

(2) *Toute courbe elliptique définie sur* K, *de conducteur* $\mathscr{L}$, *ayant tous ses points d'ordre* 2 *rationnels sur* K, *est modulaire.*

(3) *La condition* **(C)** *est satisfaite.*

Alors, le théorème de Fermat asymptotique est vrai sur K.

Remarque. Dans les limites des tables de newforms modulaires de Hilbert sur un corps totalement réel K figurant dans [LMFDB 2013], pour lesquelles 2 est totalement ramifié dans K, on constate les données numériques suivantes. Les corps intervenant dans ces tables sont de degré $d \leq 6$.

(1) Pour $d = 3$, il y a treize tels corps de nombres (à isomorphisme près). Leurs discriminants sont

$$148, 404, 564, 756, 788, 1076, 1300, 1396, 1492, 1524, 1556, 1620, 1940.$$

Pour chacun d'eux, la condition **(C)** est satisfaite.

Cela étant, la condition **(C)** n'est pas toujours satisfaite si $d = 3$. Par exemple, elle ne l'est pas pour le corps $K = \mathbb{Q}(\alpha)$ où $\alpha^3 - 32\alpha + 2 = 0$. En effet, posons $a = 16\alpha$; cet entier est dans S et on a $v_{\mathscr{L}}(a) = 13$. La courbe elliptique E/K d'équation

$$y^2 = x(x-a)(x+1-a)$$

est de conducteur $\mathscr{L}$ et a tous ses points d'ordre 2 rationnels sur K (cf. la démonstration du lemme 12, équation (6-1)). En utilisant le théorème 18, on peut démontrer qu'elle est modulaire. Il existe donc $\mathfrak{f} \in \mathcal{H}$ ayant la même fonction L que celle de E. On a $\mathbb{Q}_{\mathfrak{f}} = \mathbb{Q}$ et $\mathfrak{f}$ ne vérifie pas la condition (1-3), d'où notre assertion.

(2) Pour $d = 4$, il y a quinze corps de nombres, dont les discriminants sont

$$2048, 2304, 4352, 6224, 7168, 7488, 11344, 12544, 13824, 14336, 14656, 15952, 16448, 18432, 18688.$$

Pour chacun de ces corps, la condition **(C)** est satisfaite. Excepté pour le corps de discriminant 16448, on a $|\mathcal{H}| = 0$.

Signalons à titre indicatif que pour le corps $K = \mathbb{Q}(\alpha)$ où $\alpha^4 - 12\alpha^2 - 18\alpha - 5 = 0$, la condition **(C)** n'est pas satisfaite. On peut le vérifier comme dans l'alinéa précédent, en considérant l'entier $a = 16(1+\alpha)^2(4+4\alpha+\alpha^2)$. Il appartient à S et on a $v_{\mathscr{L}}(a) = 18$. On constate alors que la courbe elliptique sur K, d'équation $y^2 = x(x-a)(x+1-a)$, est modulaire et que son conducteur est $\mathscr{L}$.

(3) Pour $d = 5$, il y a trois corps de nombres, dont les discriminants sont

$$126032, 153424, 179024,$$

et ils vérifient la condition (C). Pour $d = 6$, il n'y a pas de tels corps dans les tables de [LMFDB 2013].

Dans le cas où l'on a $h_K^+ = 1$, on peut s'affranchir de l'hypothèse de modularité :

Théorème 4. *Supposons que les trois conditions suivantes soient satisfaites :*

(1) 2 *est totalement ramifié dans* K.

(2) *On a* $h_K^+ = 1$.

(3) *La condition* (C) *est satisfaite.*

Alors, le théorème de Fermat asymptotique est vrai sur K.

4. Question

Certaines constatations numériques concernant les hypothèses faites dans le théorème 4 suggèrent la question suivante :

Question 4.1. Supposons 2 totalement ramifié dans K et $h_K^+ = 1$. La condition (C) est-elle toujours satisfaite ? [1]

Proposition 5. *La réponse est positive si on a* $d \in \{1, 2, 4, 8\}$.

C'est une conséquence du résultat qui suit. Sa démonstration repose sur la première assertion du théorème 39 de l'Appendice. Pour tout $n \geq 1$, soit $\mu_{2^{n+2}}$ le groupe des racines 2^{n+2}-ièmes de l'unité. Notons $\mathbb{Q}(\mu_{2^{n+2}})^+$ le sous-corps totalement réel maximal de $\mathbb{Q}(\mu_{2^{n+2}})$.

Théorème 6. *Soient n un entier et K un corps de nombres totalement réel, de degré 2^n sur $\mathbb{Q}$, satisfaisant les conditions suivantes :*

(1) 2 *est totalement ramifié dans* K.

(2) *On a* $h_K^+ = 1$.

Alors, on a $K = \mathbb{Q}(\mu_{2^{n+2}})^+$.

Le corps $\mathbb{Q}(\mu_{2^{n+2}})^+$ satisfait la première condition. Pour $n \leq 5$, son nombre de classes restreint vaut 1. On conjecture que pour tout n, son nombre de classes vaut 1. Certains résultats récents ont été démontrés dans cette direction [Fukuda et Komatsu 2011].

On peut vérifier directement avec `Magma` que pour $n \leq 3$, la condition (C) est satisfaite pour $\mathbb{Q}(\mu_{2^{n+2}})^+$, ce qui implique la proposition 5.

1. Comme je l'ai signalé dans la remarque à la fin de l'introduction (page 304), il est maintenant démontré que la réponse à cette question est positive [Freitas et Siksek 2018].

Faits expérimentaux. Indiquons quelques constatations numériques en faveur d'une réponse positive à la question 4.1. En utilisant les tables de Voight, j'ai dressé une liste de corps totalement réels pour lesquels :

(1) $d \in \{3, 5, 6, 7\}$,

(2) 2 est totalement ramifié dans K,

(3) $h_K^+ = 1$,

(4) le discriminant D_K de K est pair plus petit qu'une borne fixée.

Dans le tableau ci-dessous, l'entier N est le nombre de corps totalement réels de degré d et de discriminant D_K pair plus petit que la borne que l'on s'est fixée (à isomorphisme près). Dans la dernière colonne se trouve le nombre de corps pour lesquels 2 est totalement ramifié et $h_K^+ = 1$.

d	Borne sur D_K	N	$2O_K = \mathscr{L}^d$ et $h_K^+ = 1$
3	$21 \cdot 10^3$	378	80
5	$17 \cdot 10^5$	315	23
6	$21 \cdot 10^6$	361	7
7	$207 \cdot 10^6$	32	2

Il y a cent-douze corps K intervenant dans ce tableau pour lesquels $2O_K = \mathscr{L}^d$ et $h_K^+ = 1$. Pour chacun d'eux, on constate avec `Magma` que la condition (C) est satisfaite.

Les discriminants de ces cent-douze corps sont explicités ci-dessous. Des éléments primitifs de chacun de ces corps sont déterminés dans les tables de Voight.

$d = 3$										$d = 5$			$d = 6$	$d = 7$
148	2708	4628	7668	9076	10324	12852	15252	16532	19252	126032	629584	1197392	2803712	46643776
404	2804	4692	7700	9204	10580	13172	15284	17556	19348	153424	708944	1280592	4507648	196058176
564	3124	4852	7796	9300	10868	13684	15380	17684	19572	179024	747344	1284944	5163008	
756	3252	5172	8308	9460	11060	13748	15444	17716	20276	207184	970448	1395536	6637568	
1300	3508	5204	8372	9812	11092	13972	15700	17780	20436	223824	981328	1550288	7718912	
1524	3540	5940	8628	10164	11476	14420	16084	18292	20724	394064	1034192	1664592	10766336	
1620	3604	6420	8692	10260	12660	14516	16116	18644	20788	453712	1104464	1665360	20891648	
2228	3892	7028	9044	10292	12788	14964	16180	18740	20948	535120	1172304			

On en déduit avec le théorème 4 l'énoncé suivant :

Proposition 7. *Pour chacun des corps de nombres K indiqués ci-dessus, le théorème de Fermat asymptotique est vrai sur K.*

Signalons que, pour $d \neq 6$, le groupe de Galois sur $\mathbb{Q}$ de la clôture galoisienne de K est isomorphe à $\mathbb{S}_d$. Pour $d = 6$, il est non abélien d'ordre 12 ou 72.

5. Théorème de Fermat effectif – exemples

On établit le théorème de Fermat asymptotique de façon effective pour quelques corps de nombres figurant dans ces tables, ainsi que pour les corps $\mathbb{Q}(\mu_{16})^+$ et $\mathbb{Q}(\mu_{32})^+$.

Théorème 8. *Soit K un corps cubique réel de discriminant $D_K \in \{148, 404, 564\}$. Pour tout $p \geq 5$, l'ensemble $F_p(K)$ est trivial.*

Théorème 9. *Posons $K = \mathbb{Q}(\alpha)$ avec*

$$\alpha^5 - 6\alpha^3 + 6\alpha - 2 = 0. \tag{5-1}$$

(C'est le corps de plus petit discriminant intervenant dans les colonnes "$d = 5$" du tableau ci-dessus.) *Pour tout p distinct de 5, 13, 17, 19, l'ensemble $F_p(K)$ est trivial.*

Théorème 10. *Posons $K = \mathbb{Q}(\alpha)$ avec*

$$\alpha^6 + 2\alpha^5 - 11\alpha^4 - 16\alpha^3 + 15\alpha^2 + 14\alpha - 1 = 0. \tag{5-2}$$

(C'est le corps de plus petit discriminant intervenant dans la colonne "$d = 6$" du tableau ci-dessus.) *Pour tout $p \geq 29$, distinct de 37, l'ensemble $F_p(K)$ est trivial.*

Théorème 11. (1) *Pour tout $p \geq 5$, l'ensemble $F_p(\mathbb{Q}(\mu_{16})^+)$ est trivial.*

(2) *Pour tout $p > 6724$, l'ensemble $F_p(\mathbb{Q}(\mu_{32})^+)$ est trivial.*

On a $6724 = (1+3^4)^2$, qui est, pour $d = 8$, la borne obtenue dans [Oesterlé 1996] concernant les points de p-torsion des courbes elliptiques sur les corps de nombres de degré d (voir aussi [Derickx 2016]).

Partie II. Les théorèmes 2 et 6

Pour tout idéal premier $\mathfrak{q}$ de O_K, notons $v_\mathfrak{q}$ la valuation sur K qui lui est associée, et pour toute courbe elliptique E/K, notons j_E son invariant modulaire.

6. Démonstration du théorème 2

Lemme 12. *Les trois assertions suivantes sont équivalentes :*

(1) *La condition* **(FS)** *est satisfaite.*

(2) *Il n'existe pas de courbe elliptique E/K telle que l'on ait :*

  (i) $v_{\mathscr{L}}(j_E) < 0$,

  (ii) *pour tout idéal premier $\mathfrak{q}$ de O_K, distinct de $\mathscr{L}$, on a $v_\mathfrak{q}(j_E) \geq 0$,*

  (iii) *E a tous ses points d'ordre 2 rationnels sur K.*

(3) *Il n'existe pas de courbe elliptique sur K, de conducteur $\mathscr{L}$, ayant tous ses points d'ordre 2 rationnels sur K.*

Démonstration. Vérifions l'implication (1) $\Rightarrow$ (2). Supposons pour cela qu'il existe une courbe elliptique E/K satisfaisant les trois conditions de la deuxième assertion. D'après la condition (iii), à torsion quadratique près, il existe $\lambda \in K$ tel que E/K possède une équation de la forme de Legendre (voir [Silverman 2009, p. 49, Proposition 1.7(a)] et sa démonstration)

$$y^2 = x(x-1)(x-\lambda).$$

Posons

$$\mu = 1 - \lambda.$$

On a les égalités

$$j_E = 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(1-\lambda)^2} = 2^8 \frac{(1-\lambda\mu)^3}{(\lambda\mu)^2}.$$

Soit $O_{\mathscr{L}}$ l'anneau des $\{\mathscr{L}\}$-entiers de K. D'après la condition (ii), on a

$$j_E \in O_{\mathscr{L}}.$$

De plus, λ, $\frac{1}{\lambda}$, μ et $\frac{1}{\mu}$ sont racines d'un polynôme unitaire (de degré 6) à coefficients dans $O_{\mathscr{L}}$. Par suite, λ et μ appartiennent à S. Posons

$$t = \max\{|v_{\mathscr{L}}(\lambda)|, |v_{\mathscr{L}}(\mu)|\}.$$

Il résulte de la condition (i) que l'on a $t > 0$. L'égalité $\lambda + \mu = 1$ implique alors que l'on a

$$v_{\mathscr{L}}(\lambda) = v_{\mathscr{L}}(\mu) = -t \quad \text{ou} \quad v_{\mathscr{L}}(\lambda) = 0,\ v_{\mathscr{L}}(\mu) = t \quad \text{ou} \quad v_{\mathscr{L}}(\lambda) = t,\ v_{\mathscr{L}}(\mu) = 0.$$

On obtient dans tous les cas

$$v_{\mathscr{L}}(j_E) = 8v_{\mathscr{L}}(2) - 2t = 8d - 2t.$$

La condition (i) implique $t > 4d$ et donc la condition **(FS)** n'est pas satisfaite (inégalité (1-2)). Cela prouve la première implication.

L'implication (2) $\Rightarrow$ (3) est immédiate.

Démontrons l'implication (3) $\Rightarrow$ (1). Supposons la condition **(FS)** non satisfaite, autrement dit qu'il existe $a \in S$ tel que l'on ait

$$|v_{\mathscr{L}}(a)| > 4d.$$

Posons $b = 1 - a$.

Supposons $v_{\mathscr{L}}(a) > 4d$. Vérifions que la courbe elliptique E/K d'équation

$$y^2 = x(x-a)(x+b) \tag{6-1}$$

est de conducteur $\mathscr{L}$, ce qui établira l'implication dans ce cas. Notons $c_4(E)$ et $\Delta(E)$ les invariants standard associés à cette équation. On a

$$c_4(E) = 16(a^2 + ab + b^2) \quad \text{et} \quad \Delta(E) = 16(ab)^2.$$

Pour tout idéal premier $\mathfrak{q}$ de O_K distinct de $\mathscr{L}$, on a $v_{\mathfrak{q}}(\Delta(E)) = 0$, donc E/K a bonne réduction en $\mathfrak{q}$. Posons

$$x = 4X \quad \text{et} \quad y = 8Y + 4X.$$

On obtient comme nouveau modèle de E/K

$$(W) : \qquad Y^2 + XY = X^3 - \frac{a}{2}X^2 - \frac{ab}{16}X.$$

On a $v_{\mathscr{L}}(a) > 4d = 4v_{\mathscr{L}}(2)$, donc ce modèle est entier. Par ailleurs, on a

$$v_{\mathscr{L}}(c_4(E)) = 4d \quad \text{et} \quad v_{\mathscr{L}}(\Delta(E)) = 4d + 2v_{\mathscr{L}}(a) > 12d,$$
$$c_4(E) = 2^4 c_4(W) \quad \text{et} \quad \Delta(E) = 2^{12}\Delta(W),$$

d'où

$$v_{\mathscr{L}}(c_4(W)) = 0 \quad \text{et} \quad v_{\mathscr{L}}(\Delta(W)) > 0.$$

Ainsi E a réduction de type multiplicatif en $\mathscr{L}$, d'où notre assertion.

Supposons $v_{\mathscr{L}}(a) < -4d$. Posons

$$a' = \frac{1}{a} \quad \text{et} \quad b' = 1 - a'.$$

On a $b' = -b/a$ donc a' est dans S et on a $v_{\mathscr{L}}(a') > 4d$. Comme ci-dessus, on vérifie que la courbe elliptique d'équation $y^2 = x(x - a')(x + b')$ est de conducteur $\mathscr{L}$. Cela établit l'implication. □

Lemme 13. *Supposons que toute courbe elliptique sur K, de conducteur $\mathscr{L}$, ayant tous ses points d'ordre 2 sur K, soit modulaire. Les deux assertions suivantes sont équivalentes :*

(1) *Il n'existe pas de courbe elliptique sur K, de conducteur $\mathscr{L}$, ayant tous ses points d'ordre 2 rationnels sur K.*

(2) *La condition (C) est satisfaite.*

Démonstration. Pour tout idéal premier $\mathfrak{q} \neq \mathscr{L}$ de O_K, posons $\mathbb{F}_{\mathfrak{q}} = O_K/\mathfrak{q}$.

Supposons que la première condition soit réalisée. Vérifions que la seconde l'est aussi. Soit $\mathfrak{f}$ une newform de $\mathcal{H}$ telle que $\mathbb{Q}_{\mathfrak{f}} = \mathbb{Q}$. Procédons par l'absurde en supposant que pour tout idéal premier $\mathfrak{q} \neq \mathscr{L}$ de O_K la condition (1-3) ne soit pas satisfaite. Parce que le niveau de $\mathfrak{f}$ est $\mathscr{L}$, il existe une courbe elliptique E/K, de conducteur $\mathscr{L}$, ayant la même fonction L que celle de $\mathfrak{f}$ [Freitas et Siksek 2015a, Theorem 8]. Pour tout idéal premier $\mathfrak{q} \neq \mathscr{L}$ de O_K, l'ordre de $E(\mathbb{F}_{\mathfrak{q}})$ est donc multiple de 4. Il en résulte que E/K est liée par une isogénie de degré ≤ 2 à une courbe elliptique F/K ayant tous ses points d'ordre 2 sur K [Şengün et Siksek 2018, Lemma 7.5]. Le conducteur de F/K est $\mathscr{L}$, d'où une contradiction. (On n'a pas utilisé ici l'hypothèse de modularité.)

Inversement, supposons qu'il existe une courbe elliptique E/K, de conducteur $\mathscr{L}$, ayant tous ses points d'ordre 2 sur K. Par hypothèse, E étant modulaire, il existe une newform $\mathfrak{f} \in \mathcal{H}$ ayant la même fonction L que celle de E. On a $\mathbb{Q}_{\mathfrak{f}} = \mathbb{Q}$. Soit $\mathfrak{q}$ un idéal premier de O_K distinct de $\mathscr{L}$. La courbe elliptique E a bonne réduction en $\mathfrak{q}$ et l'application $E(K)[2] \to E(\mathbb{F}_{\mathfrak{q}})$ est injective (cf. [Silverman 2009, Propopsition 3.1,

p. 192]). Par suite, 4 divise l'ordre de $E(\mathbb{F}_{\mathfrak{q}})$, autrement dit, on a $a_{\mathfrak{q}}(\mathfrak{f}) \equiv \mathrm{Norm}(\mathfrak{q})+1 \pmod 4$. Ainsi, la condition **(C)** n'est pas satisfaite, d'où le résultat. □

Le théorème 2 est une conséquence directe des lemmes 12 et 13.

Remarque. L'hypothèse de modularité est intervenue dans la démonstration pour établir l'implication **(C)** $\Rightarrow$ **(FS)**.

7. Démonstration du théorème 6

On démontre par récurrence que pour tout $r \geq 1$, tel que $r \leq n+2$, on a l'inclusion

$$\mathbb{Q}(\mu_{2^r})^+ \subseteq K. \tag{7-1}$$

Cela établira le résultat, car $\mathbb{Q}(\mu_{2^{n+2}})^+$ et K sont de même degré 2^n sur $\mathbb{Q}$. L'inclusion (7-1) est vraie si $r=1$ et $r=2$. Soit r un entier tel que $2 \leq r < n+2$ et que (7-1) soit vraie. Il s'agit de vérifier que $\mathbb{Q}(\mu_{2^{r+1}})^+$ est contenu dans K. Posons

$$L = K\mathbb{Q}(\mu_{2^{r+1}})^+.$$

L'extension L/K est non ramifiée en dehors des idéaux premiers de O_K au-dessus de 2, y compris aux places à l'infini. Par suite, son conducteur est une puissance de $\mathscr{L}$. Plus précisément :

Lemme 14. *Le conducteur de L/K divise $4O_K$.*

Démonstration. Soit ζ une racine primitive 2^{r+1}-ième de l'unité. Il existe $x \in O_K$ tel que x appartienne à $\mathscr{L}$ et pas à $\mathscr{L}^2$. On a $r-1 \leq n$. Posons

$$a = x^{2^{n-(r-1)}} \quad \text{et} \quad u = \left(\frac{\zeta+\zeta^{-1}}{a}\right)^2.$$

On a $(\zeta+\zeta^{-1})^2 = \zeta^2+\zeta^{-2}+2$. Parce que ζ^2 est une racine primitive 2^r-ième de l'unité, on déduit de (7-1) que u appartient à K. Par ailleurs, on a

$$[\mathbb{Q}(\mu_{2^{r+1}})^+ : \mathbb{Q}(\mu_{2^r})^+] = 2 \quad \text{et} \quad \mathbb{Q}(\mu_{2^{r+1}})^+ = \mathbb{Q}(\zeta+\zeta^{-1}).$$

Il en résulte que l'on a $[L:K] \leq 2$, puis l'égalité

$$L = K(\sqrt{u}).$$

On a

$$v_{\mathscr{L}}(u) = 0. \tag{7-2}$$

En effet, v étant la valuation de $\overline{\mathbb{Q}_2}$ normalisée par $v(2)=1$, on a

$$v(x) = \frac{1}{2^n}, \quad v(a) = \frac{1}{2^{r-1}} \quad \text{et} \quad v(\zeta+\zeta^{-1}) = \frac{1}{2^{r-1}},$$

d'où (7-2). Ainsi, le discriminant de L/K divise $4O_K$. On obtient le résultat car le conducteur et le discriminant de L/K sont égaux [Cassels et Fröhlich 1967, p. 160, si $[L:K]=2$]. □

Soit K^{4O_K} le corps de classes de rayon modulo $4O_K$ sur K. D'après le lemme précédent, L est contenu dans K^{4O_K}. Par hypothèse, 2 est totalement ramifié dans K et on a $h_K^+ = 1$. D'après l'assertion 1 du théorème 39 de l'Appendice, on a donc $K^{4O_K} = K$. On obtient $L = K$, ce qui montre que $\mathbb{Q}(\mu_{2^{r+1}})^+$ est contenu dans K, d'où le théorème.

Partie III. La méthode modulaire

Les démonstrations du théorème 4 et des résultats annoncés dans le paragraphe 5, reposent sur la méthode modulaire, analogue à celle utilisée par Wiles pour établir le théorème de Fermat sur $\mathbb{Q}$. On peut trouver dans [Freitas et Siksek 2015a] un exposé détaillé de cette méthode. Le principe général consiste à procéder par l'absurde en supposant qu'il existe un point non trivial dans $F_p(K)$. On lui associe ensuite une courbe elliptique définie sur K et en étudiant le module galoisien de ses points de p-torsion, on essaye d'obtenir une contradiction.

Décrivons la mise en œuvre de cette méthode dans notre situation. Soit K un corps de nombres totalement réel, de degré d sur $\mathbb{Q}$, satisfaisant les deux conditions suivantes :

(1) 2 est totalement ramifié dans K.

(2) On a $h_K^+ = 1$.

8. La courbe elliptique E_0/K

Considérons un point $(a, b, c) \in F_p(K)$ tel que $abc \neq 0$. On peut supposer que l'on a

$$a, b, c \in O_K. \tag{8-1}$$

On a $h_K^+ = 1$, en particulier O_K est principal. On supposera désormais, cela n'est pas restrictif, que l'on a

$$aO_K + bO_K + cO_K = O_K. \tag{8-2}$$

Soit E_0/K la cubique, appelée souvent courbe de Frey, d'équation

$$y^2 = x(x - a^p)(x + b^p). \tag{8-3}$$

Les invariants standard qui lui sont associés sont

$$c_4(E_0) = 16\big(a^{2p} + (ab)^p + b^{2p}\big), \quad c_6(E_0) = -32(a^p - b^p)(b^p - c^p)(c^p - a^p), \tag{8-4}$$

$$\Delta(E_0) = 16(abc)^{2p}. \tag{8-5}$$

En particulier, E_0 est une courbe elliptique définie sur K.

8.1. *Réduction de E_0/K.*

Lemme 15. *Soit $\mathfrak{q}$ un idéal premier de O_K distinct de $\mathscr{L}$.*

1) *L'équation (8-3) est minimale en $\mathfrak{q}$.*

2) *Si $\mathfrak{q}$ ne divise pas abc, E_0 a bonne réduction en $\mathfrak{q}$.*

3) *Si* $\mathfrak{q}$ *divise* abc, E_0 *a réduction de type multiplicatif en* $\mathfrak{q}$.

Démonstration. C'est une conséquence de la condition (8-1) ainsi que des formules (8-2), (8-4) et (8-5). □

Pour tout idéal premier $\mathfrak{q}$ de O_K, notons $\Delta_\mathfrak{q}$ un discriminant local minimal de E_0 en $\mathfrak{q}$.

Proposition 16. *Supposons* $p > 4d$. *Quitte à multiplier* (a, b, c) *par une unité convenable de* O_K, *les deux conditions suivantes sont satisfaites :*

(1) E_0 *a réduction de type multiplicatif en* $\mathscr{L}$.

(2) *On a*

$$v_\mathscr{L}(\Delta_\mathscr{L}) = 2pv_\mathscr{L}(abc) - 8d.$$

En particulier, avec une telle normalisation, E_0/K *est semi-stable.*

Démonstration. Le nombre premier 2 étant totalement ramifié dans K, on a $O_K/\mathscr{L} = \mathbb{F}_2$. L'un des entiers a, b, c est donc divisible par $\mathscr{L}$. On peut supposer que $\mathscr{L}$ divise b et que $\mathscr{L}$ ne divise pas ac. On a de plus $h_K^+ = 1$, donc le corps de classes de rayon modulo $4O_K$ sur K est égal à K (théorème 39). Soit U_K le groupe des unités de O_K. Le morphisme naturel $U_K \to (O_K/4O_K)^*$ est surjectif (lemme 41). Il existe ainsi $\varepsilon \in U_K$ tel que l'on ait

$$\varepsilon^{-1} \equiv -a \pmod 4.$$

Posons

$$a' = \varepsilon a, \quad b' = \varepsilon b, \quad c' = \varepsilon c.$$

Considérons alors la courbe elliptique E_0'/K d'équation

$$y^2 = x(x - a'^p)(x + b'^p). \tag{8-6}$$

En effectuant le changement de variables

$$x = 4X \quad \text{et} \quad y = 8Y + 4X,$$

on obtient comme nouveau modèle

$$(W) : \qquad Y^2 + XY = X^3 + \left(\frac{b'^p - a'^p - 1}{4}\right)X^2 - \frac{(a'b')^p}{16}X.$$

On a

$$a'^p + 1 \equiv 0 \pmod 4,$$

et d'après l'hypothèse faite sur p,

$$v_\mathscr{L}(b'^p) = pv_\mathscr{L}(b') \geq p > 4d = 4v_\mathscr{L}(2).$$

Par suite, (W) est un modèle entier. En notant $c_4(W)$ et $\Delta(W)$ les invariants standard qui lui sont associés, on a

$$c_4(E_0') = 2^4 c_4(W) \quad \text{et} \quad \Delta(E_0') = 2^{12}\Delta(W).$$

D'après les formules (8-4) et (8-5), utilisées avec l'équation (8-6), on obtient

$$v_{\mathscr{L}}(c_4(W)) = 0 \quad \text{et} \quad v_{\mathscr{L}}(\Delta(W)) = 2pv_{\mathscr{L}}(a'b'c') - 8d > 0.$$

Ainsi, (W) est un modèle minimal de E_0'/K, qui a donc réduction de type multiplicatif en $\mathscr{L}$. Parce que ε est une unité de O_K, et compte tenu du lemme 15, cela entraîne le résultat. □

Dans le cas où $p > 4d$, on supposera, dans toute la suite, que le triplet $(a, b, c) \in F_p(K)$ est normalisé de sorte que les deux conditions de la proposition 16 soient satisfaites.

8.2. ***Modularité de E_0/K.*** D'après le corollaire 2.1 de [Freitas et Siksek 2015a] :

Théorème 17. *La courbe elliptique E_0/K est modulaire si p est plus grand qu'une constante qui ne dépend que de K.*

D'après la remarque qui suit le corollaire 2.1 de [loc. cit.], ce résultat n'est pas effectif en général. Cependant l'énoncé suivant permet parfois de démontrer qu'une elliptique semi-stable définie sur K est modulaire [Freitas et al. 2015, Theorem 7] :

Théorème 18. *Posons $\ell = 5$ ou $\ell = 7$. Supposons qu'il existe un idéal premier de O_K au-dessus de ℓ en lequel l'extension $K/\mathbb{Q}$ soit non ramifiée. Soit E/K une courbe elliptique semi-stable sur K. Si $E(\overline{K})$ n'a pas de sous-groupe d'ordre ℓ stable par* $\mathrm{Gal}(\overline{K}/K)$, *alors E/K est modulaire.*

9. La représentation $\rho_{E_0,p}$

Notons

$$\rho_{E_0,p} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(E_0[p]) \simeq \mathrm{GL}_2(\mathbb{F}_p)$$

la représentation donnant l'action de $\mathrm{Gal}(\overline{K}/K)$ sur le groupe des points de p-torsion de E_0.

9.1. ***Le conducteur de $\rho_{E_0,p}$.*** Notons N_{E_0} le conducteur de E_0/K. Posons

$$M_p = \prod_{\substack{\mathfrak{q} \mid N_{E_0} \\ p \mid v_{\mathfrak{q}}(\Delta_{\mathfrak{q}})}} \mathfrak{q} \quad \text{et} \quad N_p = \frac{N_{E_0}}{M_p}.$$

Lemme 19. *Supposons $p > 4d$. On a $N_p = \mathscr{L}$.*

Démonstration. D'après le lemme 15 et la formule (8-5), pour tout idéal premier $\mathfrak{q}$ de O_K, distinct de $\mathscr{L}$, on a

$$v_{\mathfrak{q}}(\Delta_{\mathfrak{q}}) \equiv 0 \pmod{p}.$$

La seconde condition de la proposition 16 entraîne alors le résultat. □

Remarque. La terminologie adoptée dans ce paragraphe se justifie par le fait que si l'on a $p > 4d$, on peut démontrer que $\mathscr{L}$ le conducteur de Serre de $\rho_{E_0,p}$ (cf. [Serre 1987] pour $K = \mathbb{Q}$).

9.2. ***Irréductibilité de $\rho_{E_0,p}$.*** Le corps K étant totalement réel, il ne contient pas le corps de classes de Hilbert d'un corps quadratique imaginaire. D'après la proposition de l'Appendice B de [Kraus 2007], on a ainsi l'énoncé suivant :

Théorème 20. *La représentation $\rho_{E_0,p}$ est irréductible si p est plus grand qu'une constante que ne dépend que de K.*

En ce qui concerne l'effectivité de cet énoncé, considérons plus généralement dans la suite de ce paragraphe une courbe elliptique E/K semi-stable. Notons $\rho_{E,p}$ la représentation donnant l'action $\mathrm{Gal}(\overline{K}/K)$ sur son groupe des points de p-torsion. Rappelons un critère permettant souvent d'établir de manière effective que $\rho_{E,p}$ est irréductible (cf. [Kraus 2007]).

Soit p_0 le plus grand nombre premier pour lequel il existe une courbe elliptique définie sur K ayant un point d'ordre p_0 rationnel sur K. Il est borné par une fonction de d [Merel 1996] ; plus précisément, on a (voir [Oesterlé 1996; Derickx 2016])

$$p_0 \leq (1+3^{d/2})^2.$$

Notons $\left[\frac{d}{2}\right]$ la partie entière de $\frac{d}{2}$. Soit U_K^+ le groupe des unités totalement positives de O_K. Pour tout $u \in U_K^+$ et tout entier n tel que $1 \leq n \leq \left[\frac{d}{2}\right]$, on définit le polynôme $H_n^{(u)} \in \mathbb{Z}[X]$ comme suit. Soient H le polynôme minimal de u sur $\mathbb{Q}$ et t son degré. On pose

$$H_1^{(u)} = H \quad \text{et} \quad G = X^t H\left(\frac{Y}{X}\right) \in \mathbb{Z}[Y][X]. \tag{9-1}$$

Pour tout $n \geq 2$, $H_n^{(u)}$ est le polynôme de $\mathbb{Z}[X]$ obtenu en substituant Y par X dans

$$\mathrm{Res}_X\big(H_{n-1}^{(u)}, G\big) \in \mathbb{Z}[Y], \tag{9-2}$$

le résultant par rapport à X de $H_{n-1}^{(u)}$ et G. Il est unitaire de degré t^n et ses racines sont les produits de n racines de H comptées avec multiplicités. Posons

$$A_n = \mathrm{pgcd}_{u\in U_K^+} H_n^{(u)}(1) \quad \text{et} \quad R_K = \prod_{n=1}^{[d/2]} A_n. \tag{9-3}$$

L'énoncé qui suit est une reformulation du théorème 1 de [Kraus 2007] dans le cas où $h_K^+ = 1$ (voir aussi la proposition 4 de [loc. cit.] pour $d = 3$). Seule la condition $h_K^+ = 1$ intervient ici. On n'utilise pas l'hypothèse que 2 est totalement ramifié dans K.

Théorème 21. *Soit p un nombre premier ne divisant pas $D_K R_K$. Si $\rho_{E,p}$ est réductible, alors E/K, ou bien une courbe elliptique sur K liée à E par une K-isogénie de degré p, possède un point d'ordre p rationnel sur K. En particulier, si $p > p_0$ alors $\rho_{E,p}$ est irréductible.*

Démonstration. Rappelons les principaux arguments. Supposons $\rho_{E,p}$ réductible. Il existe des caractères $\varphi, \varphi' : \mathrm{Gal}(\overline{K}/K) \to \mathbb{F}_p^*$ tels que $\rho_{E,p}$ soit représentable sous la forme $\left(\begin{smallmatrix}\varphi & * \\ 0 & \varphi'\end{smallmatrix}\right)$.

Soit $\mathcal{A}_p$ l'ensemble des idéaux premiers de O_K au-dessus de p. Les caractères φ et φ' sont non ramifiés en tout idéal premier qui n'est pas dans $\mathcal{A}_p$. De plus, pour tout $\mathfrak{p} \in \mathcal{A}_p$ l'un des caractères φ et φ' est non ramifié en $\mathfrak{p}$. Par suite, il existe un sous-ensemble $\mathcal{A}$ de $\mathcal{A}_p$ tel que l'un des caractères φ et φ' soit non ramifié en dehors de $\mathcal{A}$ et que pour tout $\mathfrak{p} \in \mathcal{A}$ sa restriction à un sous-groupe d'inertie en $\mathfrak{p}$ soit le caractère cyclotomique.

Supposons $\mathcal{A}$ vide. Alors, φ ou φ' est partout non ramifié aux places finies. Parce que $h_K^+ = 1$, φ ou φ' est donc trivial. Si $\varphi = 1$, E a un point d'ordre p rationnel sur K. Si $\varphi' = 1$, E est liée par une K-isogénie de degré p a une courbe elliptique sur K ayant un point d'ordre p sur K.

Si $\mathcal{A}$ n'est pas vide, alors p divise $D_K R_K$ (voir la fin de la preuve du Theorem 1 de [Kraus 2007], p. 619, alinéa (2)), d'où le résultat. □

Remarque. Si R_K n'est pas nul, on obtient ainsi une constante explicite c_K, telle que pour tout $p > c_K$ et toute courbe elliptique E/K semi-stable sur K, la représentation $\rho_{E,p}$ soit irréductible. Dans ce cas, on obtient une version effective du théorème 20. Par exemple, R_K n'est pas nul si $d \in \{1, 2, 3, 5, 7\}$ [loc. cit., Theorem 2].

10. Le théorème d'abaissement du niveau

Il s'agit de l'analogue du théorème de Ribet intervenant dans la démonstration du théorème de Fermat sur $\mathbb{Q}$ [Ribet 1990]. Dans notre situation, si on a $p > 4d$, il s'énonce comme suit ([Freitas et Siksek 2015a, Theorem 7], les lemmes 15, 19 et l'égalité (8-5)) :

Théorème 22. *Supposons que les conditions suivantes soient satisfaites :*

(1) *On a $p > 4d$.*

(2) *L'indice de ramification de tout idéal premier de O_K au-dessus de p est strictement plus petit que $p-1$ et le corps $\mathbb{Q}(\mu_p)^+$ n'est pas contenu dans K.*

(3) *La courbe elliptique E_0/K est modulaire.*

(4) *La représentation $\rho_{E_0,p}$ est irréductible.*

Alors, il existe $\mathfrak{f} \in \mathcal{H}$ et un idéal premier $\mathfrak{p}$ de l'anneau d'entiers $O_{\mathbb{Q}_\mathfrak{f}}$ de $\mathbb{Q}_\mathfrak{f}$ au-dessus de p, tels que, en notant

$$\rho_{\mathfrak{f},\mathfrak{p}} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}_2(O_{\mathbb{Q}_\mathfrak{f}}/\mathfrak{p})$$

la représentation galoisienne associée à $\mathfrak{f}$ et $\mathfrak{p}$, on ait

$$\rho_{E_0,p} \simeq \rho_{\mathfrak{f},\mathfrak{p}}. \tag{10-1}$$

Proposition 23. *Les hypothèses faites dans l'énoncé du théorème 22 sont satisfaites si p est plus grand qu'une constante qui ne dépend que de K.*

Démonstration. La seconde condition est réalisée si p est non ramifié dans K. Les théorèmes 17 et 20 entraînent alors le résultat. □

Les représentations $\rho_{E_0,p}$ et $\rho_{\mathfrak{f},\mathfrak{p}}$ sont non ramifiées en dehors de $\mathscr{L}$ et des idéaux premiers de O_K au-dessus de p. Parce que $\rho_{E_0,p}$ est irréductible, l'isomorphisme (10-1) se traduit par les conditions suivantes : pour tout idéal premier $\mathfrak{q}$ de O_K, distinct de $\mathscr{L}$, qui n'est pas au-dessus de p, on a

$$a_{\mathfrak{q}}(\mathfrak{f}) \equiv a_{\mathfrak{q}}(E_0) \pmod{\mathfrak{p}} \quad \text{si} \quad E_0 \text{ a bonne réduction en } \mathfrak{q}, \tag{10-2}$$

$$a_{\mathfrak{q}}(\mathfrak{f}) \equiv \pm(\mathrm{Norm}(\mathfrak{q})+1) \pmod{\mathfrak{p}} \quad \text{si} \quad E_0 \text{ a réduction de type multiplicatif en } \mathfrak{q}. \tag{10-3}$$

On en déduit l'énoncé ci-dessous permettant parfois d'obtenir une contradiction à l'existence de $(a,b,c) \in F_p(K)$ (cf. [Freitas et Siksek 2015b, lemme 7.1]). Pour tout idéal premier $\mathfrak{q}$ de O_K, distinct de $\mathscr{L}$, posons

$$A_{\mathfrak{q}} = \bigl\{t \in \mathbb{Z} \mid |t| \le 2\sqrt{\mathrm{Norm}(\mathfrak{q})} \quad \text{et} \quad \mathrm{Norm}(\mathfrak{q})+1 \equiv t \pmod 4\bigr\}, \tag{10-4}$$

$$B_{\mathfrak{f},\mathfrak{q}} = \mathrm{Norm}(\mathfrak{q})\bigl((\mathrm{Norm}(\mathfrak{q})+1)^2 - a_{\mathfrak{q}}(\mathfrak{f})^2\bigr) \prod_{t \in A_{\mathfrak{q}}} (t - a_{\mathfrak{q}}(\mathfrak{f})). \tag{10-5}$$

Proposition 24. *Supposons les quatre conditions du théorème 22 satisfaites. Soient $\mathfrak{f} \in \mathcal{H}$ et $\mathfrak{p}$ un idéal premier de $O_{\mathbb{Q}_\mathfrak{f}}$ au-dessus de p tels que $\rho_{E_0,p} \simeq \rho_{\mathfrak{f},\mathfrak{p}}$. Soit $\mathfrak{q}$ un idéal premier de O_K distinct de $\mathscr{L}$. Alors, p divise la norme de $\mathbb{Q}_\mathfrak{f}$ sur $\mathbb{Q}$ de $B_{\mathfrak{f},\mathfrak{q}}$.*

Démonstration. Si $\mathfrak{q}$ divise p, alors p divise $\mathrm{Norm}(\mathfrak{q})$, en particulier p divise la norme de $\mathbb{Q}_\mathfrak{f}$ sur $\mathbb{Q}$ de $B_{\mathfrak{f},\mathfrak{q}}$. Supposons que $\mathfrak{q}$ ne divise pas p. La courbe elliptique E_0 a bonne réduction ou réduction de type multiplicatif en $\mathfrak{q}$.

Supposons que E_0 ait bonne réduction en $\mathfrak{q}$. Parce que E_0 a tous ses points d'ordre 2 rationnels que K et que $\mathfrak{q}$ est distinct de $\mathscr{L}$, le nombre de points de la courbe elliptique déduite de E_0 par réduction est multiple de 4. Par ailleurs, on a $|a_{\mathfrak{q}}(E_0)| \le 2\sqrt{\mathrm{Norm}(\mathfrak{q})}$ (borne de Weil), donc $a_{\mathfrak{q}}(E_0)$ appartient à $A_{\mathfrak{q}}$. La condition (10-2) implique alors notre assertion dans ce cas.

Si E_0 a réduction de type multiplicatif en $\mathfrak{q}$, la condition (10-3) est satisfaite, d'où le résultat. □

11. Démonstration du théorème 4

Compte tenu des propositions 23 et 24, le théorème 4 résulte de l'énoncé suivant :

Proposition 25. *Pour tout $\mathfrak{f} \in \mathcal{H}$, il existe un idéal premier $\mathfrak{q}$ de O_K, distinct de $\mathscr{L}$, tel que l'on ait $B_{\mathfrak{f},\mathfrak{q}} \neq 0$.*

Démonstration. Soit $\mathfrak{f}$ un élément de $\mathcal{H}$.

Supposons $\mathbb{Q}_\mathfrak{f} \neq \mathbb{Q}$. Il existe alors un idéal premier $\mathfrak{q}$ de O_K, distinct de $\mathscr{L}$, tel que $a_{\mathfrak{q}}(\mathfrak{f})$ ne soit pas dans $\mathbb{Z}$ (cf. [Cremona et Dembélé 2014, Theorem 9]), d'où $B_{\mathfrak{f},\mathfrak{q}} \neq 0$.

Supposons $\mathbb{Q}_\mathfrak{f} = \mathbb{Q}$. D'après la condition (C), il existe un idéal premier $\mathfrak{q}$ de O_K, distinct de $\mathscr{L}$, tel que l'on ait $a_{\mathfrak{q}}(\mathfrak{f}) \not\equiv \mathrm{Norm}(\mathfrak{q})+1 \pmod 4$. En particulier, $a_{\mathfrak{q}}(\mathfrak{f})$ n'est pas dans $A_{\mathfrak{q}}$. De plus, on a $(\mathrm{Norm}(\mathfrak{q})+1)^2 \neq a_{\mathfrak{q}}(\mathfrak{f})^2$: dans le cas contraire, on aurait $a_{\mathfrak{q}}(\mathfrak{f}) = -(\mathrm{Norm}(\mathfrak{q})+1)$. Or $\mathfrak{q}$ étant distinct de $\mathscr{L}$, on a $2(\mathrm{Norm}(\mathfrak{q})+1) \equiv 0 \pmod 4$, ce qui conduit à une contradiction. Par suite, on a $B_{\mathfrak{f},\mathfrak{q}} \neq 0$, d'où l'assertion. □

Partie IV. Les théorèmes 8, 9, 10 et 11

Dans toute cette partie, on suppose qu'il existe un point $(a, b, c) \in F_p(K)$ tel que $abc \neq 0$. Rappelons que pour $p > 4d$, on suppose implicitement qu'il est normalisé comme indiqué dans l'énoncé de la proposition 16.

12. Sur l'irréductibilité de $\rho_{E_0,p}$

Dans le cas où p est ramifié dans K, le théorème 21 ne permet pas d'établir que la représentation $\rho_{E_0,p}$ est irréductible (si tel est le cas). On dispose néanmoins du résultat suivant permettant parfois de conclure, qui vaut sans hypothèse de ramification en p. Pour tout cycle $\mathfrak{m}$ of K, notons $K^{\mathfrak{m}}$ le corps de classes de rayon modulo $\mathfrak{m}$ sur K.

Lemme 26. *Soit $\mathfrak{p}$ un idéal premier de O_K au-dessus de p. Notons $\mathfrak{m}_\infty$ le produit des places archimédiennes de K. Soit $\varphi : \mathrm{Gal}(\overline{K}/K) \to \mathbb{F}_p^*$ un caractère non ramifié en dehors $\mathfrak{m}_\infty\mathfrak{p}$. Alors, le corps laissé fixe par le noyau de φ est contenu dans $K^{\mathfrak{m}_\infty\mathfrak{p}}$.*

Démonstration. Soit $n \geq 1$ un entier. Il suffit de montrer que

$$\mathrm{Gal}\big(K^{\mathfrak{m}_\infty\mathfrak{p}^n}/K^{\mathfrak{m}_\infty\mathfrak{p}}\big) \quad \text{est un } p\text{-groupe.} \tag{12-1}$$

En effet, d'après l'hypothèse faite, il existe $j \geq 1$ tel que le corps laissé fixe par le noyau de φ soit contenu dans $K^{\mathfrak{m}_\infty\mathfrak{p}^j}$. D'après l'assertion (12-1), le groupe $\mathrm{Gal}\big(K^{\mathfrak{m}_\infty\mathfrak{p}^j}/K^{\mathfrak{m}_\infty\mathfrak{p}}\big)$ est contenu dans le noyau φ, ce qui implique alors le résultat.

Démontrons (12-1). Posons $\mathfrak{m} = m_\infty\mathfrak{p}^{n+1}$ et $\mathfrak{n} = \mathfrak{m}_\infty\mathfrak{p}^n$. Notons $U_{\mathfrak{m},1}$ le groupe des unités de O_K congrues à 1 modulo $\mathfrak{m}$ et $U_{\mathfrak{n},1}$ l'analogue de $U_{\mathfrak{m},1}$ en ce qui concerne le cycle $\mathfrak{n}$. Le corollaire 3.2.4 of [Cohen 2000] entraîne l'égalité

$$[K^{\mathfrak{m}} : K^{\mathfrak{n}}](U_{\mathfrak{n},1} : U_{\mathfrak{m},1}) = \mathrm{Norm}(\mathfrak{p}). \tag{12-2}$$

Par ailleurs, pour tout $x \in U_{\mathfrak{n},1}$, on a $x^p \in U_{\mathfrak{m},1}$. Ainsi, $U_{\mathfrak{n},1}/U_{\mathfrak{m},1}$ est un p-groupe. D'après l'égalité (12-2), $[K^{\mathfrak{m}} : K^{\mathfrak{n}}]$ est donc une puissance de p, ce qui entraîne l'assertion (12-1). $\square$

On utilisera ce résultat de la façon suivante. Supposons $p > 4d$ et $\rho_{E_0,p}$ réductible. Soient φ et φ' ses caractères d'isogénie. Ils sont non ramifiés en dehors de $\mathfrak{m}_\infty$ et des idéaux premiers de O_K au-dessus de p. Supposons qu'il existe un idéal premier $\mathfrak{p}$ de O_K au-dessus de p tel que φ ou φ' soit non ramifié en dehors de $\mathfrak{m}_\infty\mathfrak{p}$ et que de plus on ait $[K^{\mathfrak{m}_\infty\mathfrak{p}} : K] \leq 2$. On déduit alors du lemme 26, l'existence d'une courbe elliptique sur K ayant un point d'ordre p rationnel sur K, ce qui, si p est assez grand par rapport à d, conduit à une contradiction.

13. Corps cubiques et modularité

On va démontrer ici un critère permettant parfois d'établir que toute courbe elliptique semi-stable définie sur un corps cubique réel est modulaire. Rappelons que l'entier R_K est défini par la seconde formule de (9-3).

Théorème 27. *Soit K un corps cubique réel satisfaisant les conditions suivantes :*

(1) *On a* $h_K^+ = 1$.

(2) 5 *et* 7 *ne divisent pas* $D_K R_K$.

(3) 3 *n'est pas inerte dans* K.

Alors, toute courbe elliptique semi-stable définie sur K est modulaire.

13.1. ***Courbes elliptiques et points de 35-torsion.*** Commençons par établir l'énoncé qui suit, qui est une conséquence d'un résultat de Bruin et Najman [2016].

Proposition 28. *Soit K un corps cubique tel que* 3 *ne soit pas inerte dans K. Alors, il n'existe pas de courbes elliptiques définies sur K ayant un point d'ordre* 35 *rationnel sur K.*

Démonstration. On utilise le théorème 1 de [Bruin et Najman 2016], ainsi que la remarque 4 à la fin du paragraphe 2 de [loc. cit.] qui est très utile dans son application. Avec les notations de ce théorème, on prend

$$A = \mathbb{Z}/35\mathbb{Z}, \quad L = \mathbb{Q}, \quad m = 1, \quad n = 35, \quad X = X' = X_1(35), \quad \pi = \mathrm{id} \quad \text{et} \quad p = \mathfrak{p}_0 = 3.$$

Il s'agit de vérifier que les six conditions i)-vi) de cet énoncé sont satisfaites. Parce que 3 ne divise pas n, on a $A' = \mathbb{Z}/35\mathbb{Z}$ et $h = 1$. Pour toute pointe Z de $X_1(35)$, l'ensemble $L(Z)$ est le corps de rationalité de Z. C'est donc l'un des corps

$$\mathbb{Q}, \quad \mathbb{Q}(\mu_5), \quad \mathbb{Q}(\mu_7) \quad \text{et} \quad \mathbb{Q}(\mu_{35})^+.$$

Par hypothèse, 3 n'est pas inerte dans K, on a donc

$$S_{K,\mathfrak{p}_0} = \{1, 2\}.$$

La gonalité de $X_1(35)$ vaut 12 et sa Jacobienne est de rang 0 sur $\mathbb{Q}$ [Derickx 2016, p. 19 et lemme 1, p. 30]. Pour tout idéal premier $\mathfrak{p}$ de O_K au-dessus de 3, il n'existe pas de courbes elliptiques définies sur $k(\mathfrak{p})$ ayant un point rationnel d'ordre 35. Par ailleurs, 3 est inerte dans $\mathbb{Q}(\mu_5)$, $\mathbb{Q}(\mu_7)$ et $\mathbb{Q}(\mu_{35})^+$. Les six conditions considérées sont donc satisfaites, d'où le résultat. □

Remarque. Il existe des courbes elliptiques définies sur $\mathbb{F}_{27}$ ayant un point rationnel d'ordre 35, ce qui explique l'hypothèse que 3 n'est pas inerte dans K dans l'énoncé de la proposition (cf. [Waterhouse 1969, Theorem 4.1]).

13.2. ***Démonstration du théorème 27.*** Soit E/K une courbe elliptique semi-stable sur K. On utilise le théorème 18. Par hypothèse, 5 et 7 sont non ramifiés dans K. Il s'agit ainsi de montrer que l'une au moins des représentations $\rho_{E,5}$ et $\rho_{E,7}$ est irréductible. Supposons le contraire, i.e., que $\rho_{E,5}$ et $\rho_{E,7}$ soient réductibles. Parce que l'on a $h_K^+ = 1$ et que 5 et 7 ne divisent pas $D_K R_K$, quitte à remplacer E par une courbe elliptique sur K qui lui est liée par une K-isogénie de degré 1, 5, 7 ou 35, on peut supposer que E a un point d'ordre 5 et un point d'ordre 7 rationnels sur K (théorème 21). Elle possède donc un point d'ordre 35 rationnel sur K, ce qui conduit à une contradiction (proposition 28), d'où le résultat.

14. Corps cubiques et irréductibilité de $\rho_{E_0,13}$

On utilisera dans la démonstration du théorème 8 le résultat suivant.

Théorème 29. *Soit K un corps cubique satisfaisant les conditions suivantes :*

(1) *On a $h_K^+ = 1$.*

(2) *13 ne divise pas $D_K R_K$.*

(3) *3 n'est pas inerte dans K.*

Alors, pour toute courbe elliptique semi-stable E/K, ayant un point d'ordre 2 rationnel sur K, la représentation $\rho_{E,13}$ est irréductible.

Démonstration. Elle est analogue à celle du théorème 27. Soit E/K une courbe elliptique semi-stable ayant un point d'ordre 2 rationnel sur K. Supposons $\rho_{E,13}$ réductible. Parce que $h_K^+ = 1$ et que 13 ne divise pas $D_K R_K$, la courbe elliptique E, ou bien une courbe elliptique sur K liée à E par une K-isogénie de degré 13, possède un point d'ordre 13 rationnel K (théorème 21). Il existe donc une courbe elliptique sur K ayant un point d'ordre 26 rationnel sur K.

Avec les notations du théorème 1 de [Bruin et Najman 2016], on prend $A = \mathbb{Z}/26\mathbb{Z}$, $m = 1$, $n = 26$, $\mathfrak{p}_0 = 3$, $X = X' = X_1(26)$ et π est l'identité de X. Parce que 3 ne divise pas n, on a $A' = \mathbb{Z}/26\mathbb{Z}$ et $h = 1$. Le corps de rationalité des pointes de $X_1(26)$ est $\mathbb{Q}$ ou $\mathbb{Q}(\mu_{13})^+$. Par hypothèse, 3 n'est pas inerte dans K, donc on a $S_{K,\mathfrak{p}_0} = \{1, 2\}$. La gonalité de $X_1(26)$ vaut 6 et sa Jacobienne est de rang 0 sur $\mathbb{Q}$ [Derickx 2016, p. 19 et lemme 1, p. 30]. Pour tout idéal premier $\mathfrak{p}$ de O_K au-dessus de 3, il n'existe pas de courbes elliptiques définies sur $k(\mathfrak{p})$ ayant un point rationnel d'ordre 26. Par ailleurs, dans l'anneau d'entiers de $\mathbb{Q}(\mu_{13})^+$, l'idéal engendré par 3 est le produit de deux idéaux premiers de degré 3, et 3 n'est pas dans $S_{K,\mathfrak{p}_0}$. Le théorème 1 de [Bruin et Najman 2016] entraîne alors contradiction et le résultat. □

15. Démonstration du théorème 8

Gross et Rohrlich [1978, Theorem 5.1] ont démontré que l'ensemble des points rationnels de F_7 et F_{11} sur tout corps cubique est trivial. Par ailleurs, Klassen et Tzermias [1997, Theorem 1] ont établi qu'il en est de même pour F_5. On supposera donc que l'on a

$$p \geq 13.$$

En particulier, l'inégalité $p > 4d$ est satisfaite. De plus, on a $p_0 = 13$ [Parent 2003].

15.1. ***Cas où $D_K = 148$.*** On a [Voight]

$$K = \mathbb{Q}(\alpha) \quad \text{où} \quad \alpha^3 - \alpha^2 - 3\alpha + 1 = 0.$$

Le nombre premier 3 étant inerte dans K, les théorèmes 27 et 29 ne s'appliquent pas.

Lemme 30. *La courbe elliptique E_0/K est modulaire.*

Démonstration. On utilise le théorème 18 avec $\ell = 5$, qui ne divise pas D_K. Supposons que E_0 possède un sous-groupe d'ordre 5 stable par $\mathrm{Gal}(\overline{K}/K)$. Parce que E_0 a tous ses points d'ordre 2 rationnels sur K, il en résulte que E_0 est liée par une K-isogénie de degré au plus 2 à une courbe elliptique sur K ayant un sous-groupe cyclique d'ordre 20 stable par $\mathrm{Gal}(\overline{K}/K)$ (cf. par exemple [Anni et Siksek 2016, p. 1163]). La courbe modulaire $X_0(20)$ est la courbe elliptique, de conducteur 20, numérotée 20A1 dans les tables de Cremona [1997], d'équation

$$y^2 = x^3 + x^2 + 4x + 4.$$

Elle possède six pointes, toutes rationnelles sur $\mathbb{Q}$. Le groupe $X_0(20)(\mathbb{Q})$ est d'ordre 6, et on vérifie avec `Magma`[2] qu'il en est de même du groupe $X_0(20)(K)$. Cela montre que $Y_0(20)(K)$ est vide, d'où une contradiction et notre assertion. □

Lemme 31. *La représentation $\rho_{E_0,p}$ est irréductible.*

Démonstration. Posons $u = \alpha^2$. C'est une unité totalement positive de K. Son polynôme minimal est $H = X^3 - 7X^2 + 11X - 1$. On a $H(1) = 4$, donc R_K divise 4 (formules (9-1) et (9-3)). Par ailleurs, on a $D_K = 4.37$. Cela entraîne le résultat si $p \neq 13, 37$ (théorème 21).

Supposons $\rho_{E_0,13}$ réductible. Dans ce cas, E_0 est liée par une K-isogénie de degré au plus 2 à une courbe elliptique sur K ayant un sous-groupe cyclique d'ordre 52 stable par $\mathrm{Gal}(\overline{K}/K)$. Il existe un morphisme défini sur $\mathbb{Q}$, de degré 3, de la courbe modulaire $X_0(52)$ sur la courbe elliptique $F/\mathbb{Q}$, numérotée 52A1 dans les tables de Cremona [1997, p. 363], d'équation

$$y^2 = x^3 + x - 10.$$

La courbe $X_0(52)$ possède six pointes, toutes rationnelles sur $\mathbb{Q}$. Avec `Magma`, on constate que l'on a $F(K) = F(\mathbb{Q})$, qui est d'ordre 2. On en déduit que $Y_0(52)(K)$ est vide, d'où une contradiction et le fait que $\rho_{E_0,13}$ soit irréductible.

Supposons $\rho_{E_0,37}$ réductible. Soient φ et φ' ses caractères d'isogénie. On a $37O_K = \mathfrak{p}_1^2\mathfrak{p}_2$, où $\mathfrak{p}_i$ est un idéal premier de O_K. L'idéal $\mathfrak{p}_2$ est non ramifié. Par ailleurs, E_0 a en $\mathfrak{p}_2$ réduction semi-stable. D'après l'hypothèse faite sur $\rho_{E_0,37}$, si E_0 a bonne réduction en $\mathfrak{p}_2$, cette réduction est nécessairement de hauteur 1

2. À de nombreuses reprises dans cet article, comme dans la démonstration du lemme 30, on est amené à déterminer le rang sur des corps de nombres de certaines courbes elliptiques définies sur $\mathbb{Q}$. Pour cela, on utilise directement le programme relatif à l'instruction `MordellWeilGroup` du logiciel `Magma`. De plus, cette instruction indique si les résultats obtenus sont inconditionnels. Tel est le cas de tous ceux intervenant dans la suite.

(cf. [Serre 1972, proposition 12]). On en déduit que l'un des caractères φ et φ' est non ramifié en $\mathfrak{p}_2$ [loc. cit., corollaire p. 274 et corollaire p. 277]. Quitte à remplacer E_0 par une courbe elliptique qui lui est liée par une K-isogénie de degré 37, on peut supposer que c'est φ. Par suite, φ est non ramifié en dehors de $\mathfrak{p}_1$ et des places archimédiennes. D'après le lemme 26, le corps laissé fixe par le noyau de φ est donc contenu dans le corps de rayon $K^{\mathfrak{m}_\infty \mathfrak{p}_1}$. On vérifie que l'on a $[K^{\mathfrak{m}_\infty \mathfrak{p}_1} : K] = 2$ [PARI 2015]. Ainsi, φ est d'ordre au plus 2. On a $\varphi \neq 1$, car E_0 n'a pas de point d'ordre 37 rationnel sur K. Le caractère φ est donc d'ordre 2 et la courbe elliptique déduite de E_0 par torsion quadratique par φ a donc un point d'ordre 37 sur K, d'où une contradiction et le résultat. □

Les quatre conditions du théorème d'abaissement de niveau sont donc satisfaites. Par ailleurs, on a $|\mathcal{H}| = 0$, i.e., il n'existe pas de newforms modulaires paraboliques de Hilbert sur K de poids parallèle 2 et de niveau $\mathscr{L}$ [LMFDB 2013]. On obtient ainsi une contradiction à l'existence de (a, b, c), d'où le théorème dans ce cas.

15.2. ***Cas où*** $\boldsymbol{D_K = 404}$**.** On a

$$K = \mathbb{Q}(\alpha) \quad \text{où} \quad \alpha^3 - \alpha^2 - 5\alpha - 1 = 0.$$

On a $3O_K = \wp_1\wp_2$, où $\wp_1$ est un idéal premier de degré 1 et où $\wp_2$ est de degré 2. En particulier, 3 n'est pas inerte dans K.

Le polynôme minimal de $\alpha^2 \in U_K^+$ est $H = X^3 - 11X^2 + 23X - 1$ et on a $H(1) = 12$. On a $D_K = 2^2 \cdot 101$, donc 5, 7 et 13 ne divisent pas $D_K R_K$.

Il résulte alors du théorème 27 que E_0/K est modulaire. Pour $p \neq 101$, les théorèmes 21 et 29 entraînent que $\rho_{E_0,p}$ est irréductible. La décomposition de $101O_K$ en produit d'idéaux premiers est de la forme $\mathfrak{p}_1^2\mathfrak{p}_2$ et on a $[K^{\mathfrak{m}_\infty \mathfrak{p}_1} : K] = 2$. On en déduit, comme dans la démonstration du lemme 31, que $\rho_{E_0,101}$ est irréductible.

Par ailleurs, on a $|\mathcal{H}| = 1$, i.e., il existe une unique newform modulaire parabolique de Hilbert $\mathfrak{f}$ sur K de poids parallèle 2 et de niveau $\mathscr{L}$ [LMFDB 2013]. En particulier, on a $\mathbb{Q}_\mathfrak{f} = \mathbb{Q}$. Soit $\mathfrak{q}$ l'idéal premier de O_K au-dessus de 7 de degré 1. On a $a_\mathfrak{q}(\mathfrak{f}) = -2$. D'après les égalités (10-4) et (10-5), on a

$$A_\mathfrak{q} = \{-4, 0, 4\} \quad \text{et} \quad B_{\mathfrak{f},\mathfrak{q}} = -2^5 \cdot 3^2 \cdot 5 \cdot 7.$$

Ainsi, $B_{\mathfrak{f},\mathfrak{q}}$ n'est pas divisible par p, d'où la conclusion dans ce cas (proposition 24).

15.3. ***Cas où*** $\boldsymbol{D_K = 564}$**.** On a

$$K = \mathbb{Q}(\alpha) \quad \text{où} \quad \alpha^3 - \alpha^2 - 5\alpha + 3 = 0.$$

On a $3O_K = \wp_1^2\wp_2$, où $\wp_i$ est un idéal premier de degré 1.

Le polynôme minimal de $(\alpha+2)^2 \in U_K^+$ est $H = X^3 - 27X^2 + 135X - 1$ et on a $H(1) = 2^2 \cdot 3^3$.

On a $D_K = 2^2 \cdot 3 \cdot 47$. On en déduit que E_0/K est modulaire (théorème 27) et que $\rho_{E_0,p}$ est irréductible pour $p \neq 47$ (théorème 21 et théorème 29). On a $47O_K = \mathfrak{p}_1^2\mathfrak{p}_2$ et $[K^{\mathfrak{m}_\infty \mathfrak{p}_1} : K] = 2$, il en est donc de même de $\rho_{E_0,47}$.

L'ensemble $\mathcal{H}$, qui est de cardinal 2, est formé d'une newform $\mathfrak{f}$ et de sa conjuguée galoisienne telle que $\mathbb{Q}_\mathfrak{f} = \mathbb{Q}(\beta)$ où $\beta^2 + 3\beta - 1 = 0$ [LMFDB 2013]. Soit $\mathfrak{q}$ l'idéal premier de O_K au-dessus de 3 tel que $a_\mathfrak{q}(\mathfrak{f}) = \beta$. Il est de degré 1. On a $A_\mathfrak{q} = \{0\}$ et $B_{\mathfrak{f},\mathfrak{q}} = -3\beta(16-\beta^2)$. Sa norme sur $\mathbb{Q}$ étant -3^6, on obtient la conclusion cherchée.

Cela termine la démonstration du théorème 8.

16. Démonstration du théorème 9

L'ensemble des points rationnels de F_{11} sur tout corps de degré 5 sur $\mathbb{Q}$ est trivial [Gross et Rohrlich 1978, Theorem 5.1]. Tzermias [1998, Theorem 1] a démontré qu'il en est de même pour F_7. On supposera donc que l'on a

$$p \geq 23.$$

En particulier, on a $p > 4d$. Par ailleurs, on a $D_K = 2^4 \cdot 7877$.

Lemme 32. *La courbe elliptique E_0/K est modulaire.*

Démonstration. On utilise le théorème 18 avec $\ell = 7$. Supposons que E_0 possède un sous-groupe d'ordre 7 stable par $\mathrm{Gal}(\overline{K}/K)$. Dans ce cas, E_0 possède un sous-groupe cyclique d'ordre 14 stable par $\mathrm{Gal}(\overline{K}/K)$. La courbe modulaire $X_0(14)$ est la courbe elliptique, de conducteur 14, numérotée 14A1 dans les tables de Cremona [1997], d'équation

$$y^2 + xy + y = x^3 + 4x - 6.$$

Elle possède quatre pointes, qui sont rationnelles sur $\mathbb{Q}$. On vérifie avec `Magma` que l'on a $X_0(14)(K) = X_0(14)(\mathbb{Q})$ qui est d'ordre 6. Par ailleurs, les points non cuspidaux de $X_0(14)$ correspondent à deux classes d'isomorphisme de courbes elliptiques sur $\mathbb{Q}$ d'invariants modulaires entiers (-15^3 et 255^3). Parce que celui de E_0 n'est pas entier en $\mathscr{L}$, on obtient une contradiction et le résultat. □

Lemme 33. *La représentation $\rho_{E_0,p}$ est irréductible.*

Démonstration. L'entier α étant défini par l'égalité (5-1), posons

$$u_1 = (\alpha - 1)^2 \quad \text{et} \quad u_2 = (\alpha^2 + \alpha - 1)^2.$$

Ce sont des unités totalement positives de O_K. On vérifie que l'on a (formules (9-1) et (9-2))

$$H_1^{(u_1)}(1) = -12 \quad \text{et} \quad \mathrm{pgcd}\big(H_2^{(u_1)}(1), H_2^{(u_2)}(1)\big) = 2^{12} \cdot 3 \cdot 5^2.$$

Il en résulte que R_K n'est pas divisible pas un nombre premier plus grand que 7. On a $p_0 = 19$ [Derickx 2016, Chapter III, Theorem 1.1], d'où l'assertion si $p \neq 7877$. Par ailleurs, on a $7877 O_K = \mathfrak{p}_1^2 \mathfrak{p}_2$, où $\mathfrak{p}_1$ est un idéal premier de degré 1 et $\mathfrak{p}_2$ un idéal premier de degré 3. On vérifie que l'on a $[K^{\mathfrak{m}_\infty \mathfrak{p}_1} : K] = 2$, ce qui entraîne le résultat pour $p = 7877$. □

Par ailleurs, on a $|\mathcal{H}| = 2$. Plus précisément, $\mathcal{H}$ est formé d'une newform $\mathfrak{f}$ et de sa conjuguée galoisienne, et on a $\mathbb{Q}_\mathfrak{f} = \mathbb{Q}(\beta)$ où $\beta^2 + \beta - 3 = 0$ [LMFDB 2013]. Soit $\mathfrak{q}$ l'idéal premier de O_K au-dessus de 3 de degré 1. On a $a_{\mathfrak{q}}(\mathfrak{f}) = \beta$. D'après les égalités (10-4) et (10-5), on a $A_{\mathfrak{q}} = \{0\}$ et $B_{\mathfrak{f},\mathfrak{q}} = -3\beta(16-\beta^2)$. La norme sur $\mathbb{Q}$ de $B_{\mathfrak{f},\mathfrak{q}}$ est $-3^5 \cdot 17$, qui n'est pas divisible par p. La proposition 24 implique alors le théorème.

17. Démonstration du théorème 10

On a $D_K = 2^{11} \cdot 37^2$. On vérifie, comme dans la démonstration du lemme 32, que l'on a $X_0(14)(\mathbb{Q}) = X_0(14)(K)$, d'où l'on déduit que E_0/K est modulaire.

Démontrons que $\rho_{E_0,p}$ est irréductible. Par hypothèse, on a $p > 4d = 24$ et $p \neq 37$. L'entier α étant défini par l'égalité (5-2), on considère les unités totalement positives

$$u_1 = \frac{1}{58^2}\big(4\alpha^5 + 19\alpha^4 - 28\alpha^3 - 170\alpha^2 - 16\alpha + 41\big)^2,$$
$$u_2 = \frac{1}{58^2}\big(14\alpha^5 + 23\alpha^4 - 156\alpha^3 - 160\alpha^2 + 176\alpha + 13\big)^2.$$

On vérifie que l'on a

$$H_1^{(u_1)}(1) = 16, \quad H_2^{(u_1)}(1) = 2^{32} \cdot 5^4 \quad \text{et} \quad H_3^{(u_2)}(1) = 2^{216} \cdot 7^{54}.$$

Par suite, R_K n'est pas divisible par un nombre premier plus grand que 7. Il n'existe pas de courbes elliptiques sur K ayant un point d'ordre p rationnel sur K [Derickx 2016, Chapter III, Theorem 1.1)], d'où l'assertion.

Par ailleurs, on vérifie avec `Magma` que l'on a $|\mathcal{H}| = 2$ et que $\mathcal{H}$ est formé d'une newform $\mathfrak{f}$ et de sa conjuguée galoisienne, dont le corps de rationalité est $\mathbb{Q}_{\mathfrak{f}} = \mathbb{Q}(\beta)$ où $\beta^2 - \beta - 21 = 0$. En utilisant la proposition 24, on conclut alors en considérant l'idéal premier $\mathfrak{q}_1$ de O_K de degré 1 au-dessus de 17, tel que $a_{\mathfrak{q}_1}(\mathfrak{f}) = \beta$ et l'idéal premier $\mathfrak{q}_2$ de degré 1 au-dessus de 23, tel que $a_{\mathfrak{q}_2}(\mathfrak{f}) = \beta - 2$.

18. Démonstration du théorème 11

Posons $K_n = \mathbb{Q}(\mu_{2^{n+2}})^+$.

Lemme 34. *Toute courbe elliptique définie sur K_n est modulaire.*

Démonstration. Le corps K_n est le n-ième étage de la $\mathbb{Z}_2$-extension cyclotomique de $\mathbb{Q}$, d'où l'assertion [Thorne 2015, Theorem 1]. □

18.1. ***L'assertion* (1).** Elle est déjà connue si $p = 7$ [Tzermias 1998, Theorem 1] et si $p = 11$ [Gross et Rohrlich 1978]. Le fait que $F_5(K_2)$ soit trivial est une conséquence directe du théorème 2 de [Kraus 2018]. On supposera donc

$$p \geq 13.$$

On est amené à distinguer deux cas suivant que $p = 13$ ou $p \geq 17$ (car $4d = 16$).

18.1.1. *Cas où $p \geq 17$.*

Lemme 35. *Pour tout $p \geq 17$, la représentation $\rho_{E_0,p}$ est irréductible.*

Démonstration. On a $D_{K_2} = 2^{11}$ et $K_2 = \mathbb{Q}(\alpha)$ où

$$\alpha^4 - 4\alpha^2 + 2 = 0.$$

Posons $u = (\alpha+1)^2$. On a $u \in U_{K_2}^+$. Son polynôme minimal est $H_1^{(u)} = X^4 - 12X^3 + 34X^2 - 20X + 1$ et on a $H_1^{(u)}(1) = 4$. On vérifie que $H_2^{(u)}(1) = 2^{16} \cdot 17$. Par ailleurs, on a $p_0 = 17$ [Derickx 2016, Theorem 1.1], d'où le résultat si $p \neq 17$ (théorème 21).

Supposons $p = 17$. La courbe modulaire $X_0(17)$ est $\mathbb{Q}$-isomorphe à la courbe elliptique de conducteur 17, notée $17A1$ dans les tables de Cremona [1997], d'équation

$$y^2 + xy + y = x^3 - x^2 - x - 14.$$

Elle possède deux pointes, qui sont rationnelles sur $\mathbb{Q}$. Avec `Magma`, on constate que l'on a $X_0(17)(K) = X_0(17)(\mathbb{Q})$, qui est cyclique d'ordre 4. Les points non cuspidaux de $X_0(17)(\mathbb{Q})$ correspondent à deux classes d'isomorphisme de courbes elliptiques sur $\mathbb{Q}$ d'invariants modulaires (voir par exemple [Dahmen 2008, p. 30, Table 2.1])

$$j_1 = -\frac{17 \cdot 373^3}{2^{17}} \quad \text{et} \quad j_2 = -\frac{17^2 \cdot 101^3}{2}.$$

Ils sont distincts de $j(E_0)$. En effet, on peut supposer $v_{\mathscr{L}}(b) > 0$ et $v_{\mathscr{L}}(ac) = 0$; on a

$$j(E_0) = 2^8\left(\frac{(a^{34} + (ab)^{17} + b^{34})^3}{(abc)^{34}}\right),$$

d'où $v_{\mathscr{L}}(j(E_0)) = 32 - 34 v_{\mathscr{L}}(abc)$. Par ailleurs, on a $v_{\mathscr{L}}(j_1) = -68$ et $v_{\mathscr{L}}(j_2) = -4$, ce qui entraîne l'assertion et le lemme. □

On constate dans les tables de [LMFDB 2013] que l'on a $|\mathcal{H}| = 0$, d'où le résultat dans ce cas.

18.1.2. *Cas où $p = 13$.* A priori, on ne peut plus normaliser $(a, b, c) \in F_{13}(K_2)$ de sorte que la courbe de Frey E_0/K_2 soit semi-stable. On va donc utiliser le théorème d'abaissement du niveau dans le cas général [Freitas et Siksek 2015a, Theorem 7].

Considérons un point $(a, b, c) \in F_{13}(K_2)$ tel que $abc \neq 0$ et que a, b, c soient premiers entre eux dans O_{K_2}. On a $v_{\mathscr{L}}(abc) \geq 1$. Soit E_0/K_2 la courbe elliptique d'équation

$$y^2 = x(x - a^{13})(x + b^{13}). \tag{18-1}$$

Rappelons que l'on a

$$c_4(E_0) = 2^4\left(a^{26} + (ab)^{13} + b^{26}\right) \quad \text{et} \quad \Delta(E_0) = 2^4(abc)^{26}. \tag{18-2}$$

Lemme 36. *La représentation $\rho_{E_0,13}$ est irréductible.*

Démonstration. Supposons $\rho_{E_0,13}$ réductible. Dans ce cas, E_0/K_2 possède un sous-groupe d'ordre 26 stable par $\mathrm{Gal}(\overline{K_2}/K_2)$. La courbe modulaire $X_0(26)$ possède donc un point rationnel sur K_2 qui n'est pas l'une de ses quatre pointes. Soit $\mathcal{C}/\mathbb{Q}$ la courbe hyperelliptique d'équation

$$y^2 = x^6 - 8x^5 + 8x^4 - 18x^3 + 8x^2 - 8x + 1.$$

Il existe un unique $\mathbb{Q}$-isomorphisme de $X_0(26)$ sur $\mathcal{C}$ appliquant les quatre pointes de $X_0(26)$ sur $(0, 1)$, $(0, -1)$ et les deux points à l'infini de $\mathcal{C}$ (cf. [Mazur et Vélu 1972]). Par suite, il existe un point $P = (x_0, y_0) \in \mathcal{C}(K_2)$ tel que $x_0 \neq 0$.

Soit $F/\mathbb{Q}$ la courbe elliptique d'équation, numérotée 26B1 dans les tables de Cremona, d'équation

$$Y^2 + XY + Y = X^3 - X^2 - 3X + 3.$$

Les formules

$$X = -\frac{(x+1)^2}{(x-1)^2} \quad \text{et} \quad Y = \frac{2(x(x-1)-y)}{(x-1)^3},$$

définissent un morphisme $\varphi : \mathcal{C} \to F$ de degré 2 [loc. cit., 2.3] On vérifie alors avec `Magma` que l'on a $F(K_2) = F(\mathbb{Q})$ qui est cyclique d'ordre 7. On en déduit que l'on a

$$F(K_2) = \big\{0, (-1, 2), (3, -6), (1, 0), (1, -2), (3, 2), (-1, -2)\big\}.$$

On a $x_0 \neq 1$ car sinon $y_0^2 = -16$, or -1 n'est pas un carré dans K_2. Parce que x_0 n'est pas nul, on a ainsi

$$\varphi(P) \in \big\{(3, -6), (1, 0), (1, -2), (3, 2)\big\}.$$

Cela conduit à une contradiction car -3 et -1 ne sont pas des carrés dans K_2, d'où le lemme. □

Remarque. Le même argument que celui utilisé dans démonstration du lemme 31, pour $p = 13$, ne convient pas ici pour conclure. En effet, on peut vérifier avec `Magma`, seulement conditionnellement, que le groupe des points K_2-rationnels de la courbe elliptique numérotée 52A1 dans les tables de Cremona est d'ordre 2.

Supposons $v_{\mathscr{L}}(abc) \geq 2$. Dans ce cas, par les mêmes arguments que ceux utilisés dans la démonstration de la proposition 16, on peut encore normaliser (a, b, c) de sorte que E_0 ait réduction de type multiplicatif en $\mathscr{L}$ et que E_0 soit semi-stable. On peut alors conclure comme dans l'alinéa précédent.

Supposons donc désormais que l'on a

$$v_{\mathscr{L}}(abc) = 1. \tag{18-3}$$

Lemme 37. *Quitte à multiplier (a, b, c) par une unité convenable de O_{K_2}, on a*

$$v_{\mathscr{L}}(N_{E_0}) \in \{5, 6, 8\}.$$

Démonstration. D'après (18-2) et (18-3), on a

$$v_{\mathscr{L}}(c_4(E_0)) = 16, \quad v_{\mathscr{L}}(c_6(E_0)) = 24, \quad v_{\mathscr{L}}(\Delta(E_0)) = 42. \tag{18-4}$$

On a $v_{\mathscr{L}}(j_{E_0}) = 6$, donc E_0 a potentiellement bonne réduction en $\mathscr{L}$.

On peut supposer que l'on a $v_{\mathscr{L}}(b) = 1$, auquel cas $v_{\mathscr{L}}(ac) = 0$. Par ailleurs, il existe une unité ε de O_{K_2} telle que l'on ait $\varepsilon a + 1 \equiv 0 \pmod 4$ (Appendice, théorème 39 et lemme 41). En remplaçant (a, b, c) par $(\varepsilon a, \varepsilon b, \varepsilon c)$, on se ramène au cas où l'on a

$$a^{13} + 1 \equiv 0 \pmod 4. \tag{18-5}$$

Le modèle (18-1) n'est pas minimal en $\mathscr{L}$. En effet, posons

$$x = \alpha^4 X \quad \text{et} \quad y = \alpha^6 Y + \alpha^4 X.$$

On obtient comme nouveau modèle

$$(W) : \qquad Y^2 + \frac{2}{\alpha^2} XY = X^3 + \left(\frac{b^{13} - a^{13} - 1}{\alpha^4}\right) X^2 - \frac{(ab)^{13}}{\alpha^8} X.$$

D'après (18-5) et le fait que 2 soit associé à α^4, c'est un modèle entier. On a

$$c_4(E_0) = \alpha^8 c_4(W), \quad c_6(E_0) = \alpha^{12} c_6(W), \quad \Delta(E_0) = \alpha^{24} \Delta(W),$$

d'où (formules (18-4))

$$v_{\mathscr{L}}(c_4(W)) = 8, \quad v_{\mathscr{L}}(c_6(W)) = 12, \quad v_{\mathscr{L}}(\Delta(W)) = 18.$$

On vérifie avec les tables de [Papadopoulos 1993] que le type de Néron en $\mathscr{L}$ de (W) est I_6^* ou bien que (W) n'est pas minimal en $\mathscr{L}$. Si le type de Néron est I_6^*, on a $v_{\mathscr{L}}(N_{E_0}) = 8$. Si (W) n'est pas minimal, le triplet de valuations de ses invariants minimaux en $\mathscr{L}$ est $(4, 6, 6)$. On constate alors que son type de Néron est II, auquel cas $v_{\mathscr{L}}(N_{E_0}) = 6$, ou bien que son type de Néron est III et on a $v_{\mathscr{L}}(N_{E_0}) = 5$, d'où le lemme. □

Supposons (a, b, c) normalisé comme dans l'énoncé du lemme précédent. Notons $\mathcal{S}_2^+(\mathscr{L}^r)$ le $\mathbb{C}$-espace vectoriel engendré par les newforms modulaires paraboliques de Hilbert sur K_2, de poids parallèle 2 et de niveau $\mathscr{L}^r$. On a [LMFDB 2013]

$$\dim \mathcal{S}_2^+(\mathscr{L}^5) = 1, \quad \dim \mathcal{S}_2^+(\mathscr{L}^6) = 3 \quad \text{et} \quad \dim \mathcal{S}_2^+(\mathscr{L}^8) = 8.$$

Compte tenu des lemmes 34 et 36, les conditions du théorème d'abaissement du niveau sont satisfaites [Freitas et Siksek 2015a, Theorem 7]. Il existe donc $\mathfrak{f} \in \mathcal{S}_2^+(\mathscr{L}^r)$ avec $r \in \{5, 6, 8\}$ et un idéal premier $\mathfrak{p}$ au-dessus de 13 dans $O_{\mathbb{Q}_\mathfrak{f}}$, tels que

$$\rho_{\mathfrak{f},\mathfrak{p}} \simeq \rho_{E_0,13}.$$

Considérons alors le nombre premier 79. Il est totalement décomposé dans K_2. Soit $\mathfrak{q}$ un idéal premier de O_{K_2} au-dessus de 79. On constate que l'on a

$$a_\mathfrak{q}(\mathfrak{f}) \in \{-8, 0, 8\}.$$

Si E_0 a réduction multiplicative en $\mathfrak{q}$, on a

$$a_{\mathfrak{q}}(\mathfrak{f}) \equiv \pm 2 \pmod{\mathfrak{p}},$$

ce qui n'est pas. Ainsi E_0 a bonne réduction en $\mathfrak{q}$. On a donc la congruence

$$a_{\mathfrak{q}}(E_0) \equiv a_{\mathfrak{q}}(\mathfrak{f}) \pmod{\mathfrak{p}}.$$

On a $v_{\mathfrak{q}}(abc) = 0$, donc a^{13}, b^{13}, c^{13} sont des racines 6-ièmes de l'unité modulo $\mathfrak{q}$. On a ainsi

$$a^{13}, b^{13}, c^{13} \equiv 1, 23, 24, 55, 56, 78 \pmod{\mathfrak{q}}.$$

L'égalité $a^{13} + b^{13} + c^{13} = 0$ implique

$$(a^{13}, b^{13}) \equiv (1, 23), (1, 55), (23, 1), (23, 55), (24, 56), (24, 78), (55, 1), (55, 23), (56, 24),\\ (56, 78), (78, 24), (78, 56) \pmod{\mathfrak{q}}.$$

Dans tous les cas, on obtient

$$a_{\mathfrak{q}}(E_0) = \pm 4,$$

d'où la contradiction cherchée et le résultat pour $p = 13$.

18.2. ***L'assertion* (2).**

Lemme 38. *Pour tout* $p > 6724$, *la représentation* $\rho_{E_0,p}$ *est irréductible.*

Démonstration. On a $D_{K_3} = 2^{31}$ et $K_3 = \mathbb{Q}(\alpha)$ où

$$\alpha^8 - 8\alpha^6 + 20\alpha^4 - 16\alpha^2 + 2 = 0.$$

Posons

$$u_1 = \left(-\alpha^6 - 2\alpha^5 + 5\alpha^4 + 10\alpha^3 - 4\alpha^2 - 9\alpha - 1\right)^2, \quad u_2 = \left(\alpha^7 + \alpha^6 - 6\alpha^5 - 5\alpha^4 + 9\alpha^3 + 5\alpha^2 - 3\alpha - 1\right)^2,$$
$$u_3 = \left(-2\alpha^7 - 2\alpha^6 + 11\alpha^5 + 10\alpha^4 - 13\alpha^3 - 9\alpha^2 + \alpha + 1\right)^2.$$

Ce sont des unité totalement positives de O_{K_3}. En utilisant le théorème 21 avec les u_i, on vérifie que on a l'implication

$$R_{K_3} \equiv 0 \pmod{p} \;\Rightarrow\; p \leq 607.$$

Par ailleurs, on a $p_0 < 6724$ (voir [Oesterlé 1996; Derickx 2016]), d'où l'assertion. □

Les conditions du théorème 22 sont satisfaites. On constate avec `Magma` que l'on a $|\mathcal{H}| = 40$. À conjugaison près, $\mathcal{H}$ est formé de quatre newforms $\mathfrak{f}$ telles que $[\mathbb{Q}_{\mathfrak{f}} : \mathbb{Q}] = 4$ et d'une newform dont le corps de rationalité est de degré 24 sur $\mathbb{Q}$.

Les nombres premiers 31 et 97 sont totalement décomposés dans K_3. En utilisant la proposition 24, et en prenant pour $\mathfrak{q}$ un idéal premier de O_{K_3} au-dessus de 31, puis un idéal premier au-dessus de 97, on obtient alors le résultat.

Partie V. Appendice

Soit K un corps de nombres totalement réel. Notons K^{4O_K} le corps de classes de rayon modulo $4O_K$ sur K. On établit ici l'énoncé suivant, que l'on utilise, tout au moins sa première assertion, dans les démonstrations du théorème 6, de la proposition 16 et du lemme 37.

Théorème 39. (1) *Supposons 2 totalement ramifié dans K et $h_K^+ = 1$. Alors, on a $K = K^{4O_K}$.*

(2) *Supposons $K = K^{4O_K}$. Alors, on a $h_K^+ = 1$.*

En particulier :

Corollaire 40. *Supposons 2 totalement ramifié dans K. On a $h_K^+ = 1$ si et seulement si $K = K^{4O_K}$.*

18.3. ***Résultats préliminaires.*** Notons U_K le groupe des unités de O_K et h_K le nombre de classes de K. Rappelons que d désigne le degré de K sur $\mathbb{Q}$. Posons

$$G = (O_K/4O_K)^*.$$

Soit $\varphi : U_K \to G$ le morphisme qui à $u \in U_K$ associe $u + 4O_K$.

Lemme 41. *On a $K = K^{4O_K}$ si et seulement si $h_K = 1$ et φ est une surjection sur G.*

Démonstration. C'est conséquence directe de [Cohen 2000, Proposition 3.2.3]. □

Lemme 42. *Les deux conditions suivantes sont équivalentes :*

(1) *On a $h_K^+ = 1$.*

(2) *On a $h_K = 1$ et toute unité totalement positive est un carré dans K.*

Démonstration. On a l'égalité (cf. [Cohen 2000, Proposition 3.2.3], avec pour $\mathfrak{m}$ le produit des places à l'infini)

$$h_K^+ = \frac{2^d}{[U_K : U_K^+]} h_K.$$

Supposons $h_K^+ = 1$. On a alors $h_K = 1$, puis $[U_K : U_K^+] = 2^d$. D'après le théorème de Dirichlet, on a $[U_K : U_K^2] = 2^d$, d'où $U_K^+ = U_K^2$. Inversement, si $h_K = 1$ et $U_K^+ = U_K^2$, la formule ci-dessus implique $h_K^+ = 1$. □

Lemme 43. *Supposons 2 totalement ramifié dans K. On a*

$$|G/G^2| = 2^d.$$

Démonstration. Soit a un élément de G. Soit $\mathscr{L}$ l'idéal premier de O_K au-dessus de 2. On a $O_K/\mathscr{L} = \mathbb{F}_2$, donc il existe $x \in \mathscr{L}$ tel que $a = 1 + x + 4O_K$. On a $a^2 = 1 + x(2 + x) + 4O_K$, d'où il résulte que l'on a

$$a^2 = 1 \Leftrightarrow x \in 2O_K.$$

Posons $G[2] = \{z \in G \mid z^2 = 1\}$. On en déduit une application $f : G[2] \to 2O_K/4O_K$ définie pour tout $a \in G[2]$ par l'égalité

$$f(a) = x + 4O_K \quad \text{où} \quad a = 1 + x + 4O_K.$$

C'est un isomorphisme de groupes. Le groupe $2O_K/4O_K$ est isomorphe à $O_K/2O_K$ qui est d'ordre 2^d. Par ailleurs, les ordres de $G[2]$ et G/G^2 sont égaux, d'où le lemme. □

Lemme 44. *Supposons* 2 *totalement ramifié dans* K. *Les deux conditions suivantes sont équivalentes* :

(1) *On a* $K = K^{4O_K}$.

(2) *On a* $h_K = 1$ *et toute unité congrue à un carré modulo* 4 *est un carré dans* K.

Démonstration. Notons $\psi : U_K \to G \to G/G^2$ le morphisme naturel déduit de φ.

Supposons $K = K^{4O_K}$. Le morphisme ψ est une surjection (lemme 41). Les ordres de U_K/U_K^2 et G/G^2 sont égaux (lemme 43). Par suite, U_K^2 est le noyau de ψ, donc toute unité congrue à un carré modulo 4 est un carré.

Inversement, supposons la seconde condition satisfaite. Démontrons que le morphisme φ est surjectif, ce qui, d'après le lemme 41, impliquera la première assertion. D'après l'hypothèse faite, le noyau de ψ est U_K^2. Les ordres de U_K/U_K^2 et G/G^2 étant égaux, on en déduit que ψ est une surjection. Ainsi, l'image de $\varphi(U_K)$ dans G/G^2 est G/G^2. Parce que 2 est totalement ramifié dans K, G est un 2-groupe [Cohen 2000, p. 137]. Il en résulte que $\varphi(U_K) = G$, d'où le résultat. (Si p est premier, le sous-groupe de Frattini d'un p-groupe abélien fini A est A^p, voir par exemple [Rotman 1995, Theorem 5.48]. Si B est un sous-groupe de A tel que $BA^p = A$, on a donc $A = B$.) □

18.4. ***Fin de la démonstration du théorème 39.*** (1) Supposons 2 totalement ramifié dans K et $h_K^+ = 1$. Soit u une unité de O_K congrue à un carré modulo 4. L'extension $K(\sqrt{u})/K$ est alors partout non ramifiée aux places finies de K [Cox 1989, Lemma 5.32, p. 114]. On a $h_K^+ = 1$, donc u est un carré dans K. Vu que $h_K = 1$, on a donc $K = K^{4O_K}$ (lemme 44).

(2) Supposons $K = K^{4O_K}$. On a $h_K = 1$. Soit u une unité totalement positive de O_K. D'après le lemme 42, il s'agit de montrer que u est un carré dans K. L'extension $K(\sqrt{u})/K$ est non ramifiée aux places à l'infini et en dehors des idéaux premiers de O_K au-dessus de 2. Le conducteur de l'extension $K(\sqrt{u})/K$ est égal à son discriminant [Cassels et Fröhlich 1967, p. 160], qui divise $4O_K$. Par suite, $K(\sqrt{u})$ est contenu dans K^{4O_K}, d'où $K(\sqrt{u}) = K$ et notre assertion.

Remerciements

J'ai bénéficié de conversations avec J. Oesterlé concernant l'Appendice de cet article et son application au théorème 6. Je l'en remercie vivement. Je remercie également R. Barbaud pour l'aide qu'il m'a apportée dans la réalisation de certains programmes intervenant dans le logiciel `Magma`, ainsi que D. Bernardi pour les commentaires dont il m'a fait part. Par ailleurs, je remercie le referee de cet article pour sa lecture détaillée et pour toutes ses remarques.

Bibliographie

[Anni et Siksek 2016] S. Anni et S. Siksek, "Modular elliptic curves over real abelian fields and the generalized Fermat equation $x^{2\ell}+y^{2m}=z^p$", *Algebra Number Theory* **10**:6 (2016), 1147–1172. MR Zbl

[Bosma et al. 1997] W. Bosma, J. Cannon et C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl

[Breuil et al. 2001] C. Breuil, B. Conrad, F. Diamond et R. Taylor, "On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises", *J. Amer. Math. Soc.* **14**:4 (2001), 843–939. MR Zbl

[Browkin 2006] J. Browkin, "The *abc*-conjecture for algebraic numbers", *Acta Math. Sin.* (*Engl. Ser.*) **22**:1 (2006), 211–222. MR Zbl

[Bruin et Najman 2016] P. Bruin et F. Najman, "A criterion to rule out torsion groups for elliptic curves over number fields", *Res. Number Theory* **2** (2016), art. id. 3. MR Zbl

[Cassels et Fröhlich 1967] J. W. S. Cassels et A. Fröhlich (éditeurs), *Algebraic number theory* (Brighton, 1965), Academic Press, London, 1967. MR

[Cohen 2000] H. Cohen, *Advanced topics in computational number theory*, Graduate Texts in Math. **193**, Springer, 2000. MR Zbl

[Cox 1989] D. A. Cox, *Primes of the form x^2+ny^2*, Wiley, New York, 1989. MR

[Cremona 1997] J. E. Cremona, *Algorithms for modular elliptic curves*, 2nd éd., Cambridge Univ. Press, 1997. MR Zbl

[Cremona et Dembélé 2014] J. E. Cremona et L. Dembélé, "Modular forms over number fields", preprint, 2014, https :// tinyurl.com/cremdem.

[Dahmen 2008] S. R. Dahmen, *Classical and modular methods applied to Diophantine equations*, Ph.D. thesis, Utrecht University, 2008, https ://tinyurl.com/srdahmen.

[Dembélé et Voight 2013] L. Dembélé et J. Voight, "Explicit methods for Hilbert modular forms", pp. 135–198 dans *Elliptic curves, Hilbert modular forms and Galois deformations*, édité par H. Darmon et al., Birkhäuser, Basel, 2013. MR Zbl

[Derickx 2016] M. Derickx, *Torsion points on elliptic curves over number fields of small degree*, Ph.D. thesis, Leiden University, 2016, http ://hdl.handle.net/1887/43186.

[Freitas et Siksek 2015a] N. Freitas et S. Siksek, "The asymptotic Fermat's last theorem for five-sixths of real quadratic fields", *Compos. Math.* **151**:8 (2015), 1395–1415. MR Zbl

[Freitas et Siksek 2015b] N. Freitas et S. Siksek, "Fermat's last theorem over some small real quadratic fields", *Algebra Number Theory* **9**:4 (2015), 875–895. MR Zbl

[Freitas et Siksek 2018] N. Freitas et S. Siksek, "On the asymptotic Fermat's last theorem", preprint, 2018. arXiv

[Freitas et al. 2015] N. Freitas, B. V. Le Hung et S. Siksek, "Elliptic curves over real quadratic fields are modular", *Invent. Math.* **201**:1 (2015), 159–206. MR Zbl

[Fukuda et Komatsu 2011] T. Fukuda et K. Komatsu, "Weber's class number problem in the cyclotomic $\mathbb{Z}_2$-extension of $\mathbb{Q}$, III", *Int. J. Number Theory* **7**:6 (2011), 1627–1635. MR Zbl

[Gross et Rohrlich 1978] B. H. Gross et D. E. Rohrlich, "Some results on the Mordell–Weil group of the Jacobian of the Fermat curve", *Invent. Math.* **44**:3 (1978), 201–224. MR Zbl

[Jarvis et Meekin 2004] F. Jarvis et P. Meekin, "The Fermat equation over $\mathbb{Q}(\sqrt{2})$", *J. Number Theory* **109**:1 (2004), 182–196. MR Zbl

[Klassen et Tzermias 1997] M. Klassen et P. Tzermias, "Algebraic points of low degree on the Fermat quintic", *Acta Arith.* **82**:4 (1997), 393–401. MR Zbl

[Kraus 2007] A. Kraus, "Courbes elliptiques semi-stables sur les corps de nombres", *Int. J. Number Theory* **3**:4 (2007), 611–633. MR Zbl

[Kraus 2018] A. Kraus, "Quartic points on the Fermat quintic", *Ann. Math. Blaise Pascal* **25**:1 (2018), 199–205. MR Zbl

[LMFDB 2013] LMFDB Collaboration, "The L-functions and modular forms database", 2013, http ://www.lmfdb.org.

[Mazur et Vélu 1972] B. Mazur et J. Vélu, "Courbes de Weil de conducteur 26", *C. R. Acad. Sci. Paris Sér. A-B* **275** (1972), 743–745. MR Zbl

[Merel 1996] L. Merel, "Bornes pour la torsion des courbes elliptiques sur les corps de nombres", *Invent. Math.* **124**:1-3 (1996), 437–449. MR Zbl

[Oesterlé 1996] J. Oesterlé, note non publiée, 1996.

[Papadopoulos 1993] I. Papadopoulos, "Sur la classification de Néron des courbes elliptiques en caractéristique résiduelle 2 et 3", *J. Number Theory* **44**:2 (1993), 119–152. MR Zbl

[Parent 2003] P. Parent, "No 17-torsion on elliptic curves over cubic number fields", *J. Théor. Nombres Bordeaux* **15**:3 (2003), 831–838. MR Zbl

[PARI 2015] PARI Group, PARI/GP version 2.7.3, 2015, http ://pari.math.u-bordeaux.fr.

[Ribet 1990] K. A. Ribet, "On modular representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms", *Invent. Math.* **100**:2 (1990), 431–476. MR Zbl

[Rotman 1995] J. J. Rotman, *An introduction to the theory of groups*, 4th éd., Graduate Texts in Math. **148**, Springer, 1995. MR Zbl

[Şengün et Siksek 2018] M. H. Şengün et S. Siksek, "On the asymptotic Fermat's last theorem over number fields", *Comment. Math. Helv.* **93**:2 (2018), 359–375. MR Zbl

[Serre 1972] J.-P. Serre, "Propriétés galoisiennes des points d'ordre fini des courbes elliptiques", *Invent. Math.* **15**:4 (1972), 259–331. MR Zbl

[Serre 1987] J.-P. Serre, "Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$", *Duke Math. J.* **54**:1 (1987), 179–230. MR Zbl

[Siegel 1929] C. L. Siegel, "Über einige Anwendungen diophantischer Approximationen", *Abh. Preuß. Akad. Wiss., Phys.-Math. Kl.* **1929**:1 (1929), 1–41. JFM

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd éd., Graduate Texts in Math. **106**, Springer, 2009. MR Zbl

[Smart 1998] N. P. Smart, *The algorithmic resolution of Diophantine equations*, London Math. Soc. Student Texts **41**, Cambridge Univ. Press, 1998. MR Zbl

[Taylor et Wiles 1995] R. Taylor et A. Wiles, "Ring-theoretic properties of certain Hecke algebras", *Ann. of Math.* (2) **141**:3 (1995), 553–572. MR Zbl

[Thorne 2015] J. A. Thorne, "Elliptic curves over $\mathbb{Q}_\infty$ are modular", 2015. À paraître dans *J. Eur. Math. Soc.* arXiv

[Tzermias 1998] P. Tzermias, "Algebraic points of low degree on the Fermat curve of degree seven", *Manuscripta Math.* **97**:4 (1998), 483–488. MR Zbl

[Voight] J. Voight, "Tables of totally real number fields", https ://math.dartmouth.edu/~jvoight/nf-tables.

[Waterhouse 1969] W. C. Waterhouse, "Abelian varieties over finite fields", *Ann. Sci. École Norm. Sup.* (4) **2**:4 (1969), 521–560. MR Zbl

[Wiles 1995] A. Wiles, "Modular elliptic curves and Fermat's last theorem", *Ann. of Math.* (2) **141**:3 (1995), 443–551. MR Zbl

Communicated by Christopher Skinner
Received 2017-09-24 Revised 2018-04-13 Accepted 2018-09-23

alain.kraus@imj-prg.fr *Institut de Mathématiques de Jussieu - Paris Rive Gauche, Sorbonne Université, UMR 7586 CNRS - Paris Diderot, 75005 Paris, France*

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.333

G-valued local deformation rings and global lifts

Rebecca Bellovin and Toby Gee

We study G-valued Galois deformation rings with prescribed properties, where G is an arbitrary (not necessarily connected) reductive group over an extension of $\mathbb{Z}_l$ for some prime l. In particular, for the Galois groups of p-adic local fields (with p possibly equal to l) we prove that these rings are generically regular, compute their dimensions, and show that functorial operations on Galois representations give rise to well-defined maps between the sets of irreducible components of the corresponding deformation rings. We use these local results to prove lower bounds on the dimension of global deformation rings with prescribed local properties. Applying our results to unitary groups, we improve results in the literature on the existence of lifts of mod l Galois representations, and on the weight part of Serre's conjecture.

1. Introduction

The study of Galois deformation rings was initiated in [Mazur 1989], and was crucial to the proof of Fermat's last theorem in [Wiles 1995], and in particular to the modularity lifting theorems proved in [Wiles 1995; Taylor and Wiles 1995]. Many generalisations of these modularity lifting theorems have been proved over the last 25 years, and it has become increasingly important to consider Galois representations valued in reductive groups other than GL_n. From the point of view of the Langlands program, it is particularly important to be able to use disconnected groups, as the L-groups of nonsplit groups are always disconnected. In particular, it is important to study the structure of local deformation rings for general reductive groups, and to prove lifting results for global deformation rings. We briefly review the history of such results in Section 1.1, but we firstly explain the main theorems of this paper.

Gee was supported in part by a Leverhulme Prize, EPSRC grant EP/L025485/1, Marie Curie Career Integration Grant 303605, by ERC Starting Grant 306326, and by a Royal Society Wolfson Research Merit Award.
MSC2010: primary 11F80; secondary 11F85.
Keywords: Galois deformations.

We begin with a result about local deformation rings. Let $K/\mathbb{Q}_p$ be a finite extension, let $\mathcal{O}$ be the ring of integers in a finite extension E of $\mathbb{Q}_l$ with residue field $\mathbb{F}$, where l is possibly equal to p, and let G be a (not necessarily connected) reductive group over $\mathcal{O}$. Given a representation $\bar{\rho} : \mathrm{Gal}_K \to G(\mathbb{F})$, we consider liftings of $\bar{\rho}$ of some inertial type τ, and in the case $l = p$, some p-adic Hodge type $\boldsymbol{v}$. There is a corresponding universal framed deformation ring $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}$, and we prove the following result (as well as a variant for "fixed determinant ψ" deformations).

Theorem A (Theorem 3.3.2). *Fix an inertial type τ, and if $l = p$ then fix a p-adic Hodge type $\boldsymbol{v}$. Then $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]$ is generically regular. In addition, $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}$ is equidimensional of dimension*

$$1 + \dim_E G + \delta_{l=p} \dim_E (\mathrm{Res}_{E\otimes K/E}\, G)/P_{\boldsymbol{v}},$$

and $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v},\psi}$ is equidimensional of dimension

$$1 + \dim_E G^{\mathrm{der}} + \delta_{l=p} \dim_E (\mathrm{Res}_{E\otimes K/E}\, G)/P_{\boldsymbol{v}}.$$

(We are abusing notation here; $P_{\boldsymbol{v}}$ is a $(\mathrm{Res}_{E\otimes K/E}\, G)^{\circ}_{\bar{E}}$-conjugacy class of parabolic subgroups of $\mathrm{Res}_{E\otimes K/E}\, G$, and we choose a representative defined over E to compute the dimension of the quotient.) We are also able to describe the regular locus of $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]$ precisely in terms of the corresponding Weil–Deligne representations; see Corollary 3.3.4. In the case that $G = \mathrm{GL}_n$ and $l = p$ this is a theorem of Kisin [2008], and results for general groups (but with more restrictive hypotheses than those of Theorem A) were previously proved by Balaji [2013] and Bellovin [2016].

Combining Theorem A with results of [Balaji 2013], we obtain the following result (see Section 4 for any unfamiliar notation or terminology — in particular, $\mathfrak{g}^0_{\mathbb{F}}$ denotes the $\mathbb{F}$-points of the Lie algebra of the derived subgroup of G); in the case of potentially crystalline representations, this is the main result of [loc. cit.].

Theorem B (Proposition 4.2.6). *Let F be totally real, assume that $l > 2$, let S be a finite set of places of F containing all places dividing $l\infty$, and let $\bar{\rho} : \mathrm{Gal}_{F,S} \to G(\overline{\mathbb{F}}_l)$ be a representation admitting a universal deformation ring. Fix inertial types at all places $v \in S$, and Hodge types at all places $v \mid l$, in such a way that the corresponding local deformation rings are nonzero, and let R^{univ} denote the corresponding fixed determinant universal deformation ring for $\bar{\rho}$.*

Assume that $\bar{\rho}$ is odd, and that $H^0(\mathrm{Gal}_{F,S}, (\mathfrak{g}^0_{\mathbb{F}})^*(1)) = 0$. Suppose also that for each place $v \mid l$ the corresponding Hodge type is regular. Then R^{univ} has Krull dimension at least 1.

We use this result to improve on some results about automorphic forms on unitary groups proved using the methods of [Barnet-Lamb et al. 2014]. Beginning with [Clozel et al. 2008], Galois deformations were considered for representations valued in a certain disconnected group $\mathcal{G}_n$, whose connected component is $\mathrm{GL}_n \times \mathrm{GL}_1$ (this group is related to the L-group of a unitary group, see [Buzzard and Gee 2014, §8]). In the case that $G = \mathcal{G}_n$, Theorem B generalises [Barnet-Lamb et al. 2014, Proposition 1.5.1], removing restrictions on the places in S (which were chosen to split in the splitting field of the corresponding unitary group, in order to reduce the local deformation theory to the GL_n case).

We deduce corresponding improvements to a number of results proved using the methods of [loc. cit.], such as the following general result about Serre weights for rank-2 unitary groups, which removes a "split ramification" hypothesis on the ramification of $\bar{r}$ at places away from l.

Theorem C (Theorem 5.2.2). *Let F be an imaginary CM field with maximal totally real subfield F^+, and suppose that F/F^+ is unramified at all finite places, that each place of F^+ above l splits in F, and that $[F^+ : \mathbb{Q}]$ is even. Suppose that l is odd, that $\bar{r} : G_{F^+} \to \mathcal{G}_2(\overline{\mathbb{F}}_l)$ is irreducible and modular, and that $\bar{r}(G_{F(\zeta_l)})$ is adequate.*

Then the set of Serre weights for which $\bar{r}$ is modular is exactly the set of weights given by the sets $W(\bar{r}|_{G_{F_v}})$, $v \mid l$.

(See Remark 5.2.3 for a discussion of further improvements to this result that could be made by techniques orthogonal to those of this paper.) These results are also crucially applied in [Calegari et al. 2018], where they are used to construct lifts of representations valued in $\mathcal{G}_n$ which have prescribed ramification at certain inert places.

1.1. ***A brief historical overview.*** We now give a very brief overview of some of the developments in the deformation theory of Galois representations, which was introduced for representations valued in GL_n in [Mazur 1989]; we apologise for the many important papers that we do not discuss here for reasons of space. The abstract parts of this deformation theory were generalised to arbitrary reductive groups in [Tilouine 1996]. However, for applications to the Langlands program (and in particular to proving automorphy lifting theorems), one needs to study conditions on Galois deformations coming from p-adic Hodge theory.

This was initially done in a somewhat ad hoc fashion, mostly for the group GL_2 and mostly for conditions coming from p-divisible groups, culminating in [Breuil et al. 2001], which used a detailed study of some particular such deformation rings to complete the proof of the Taniyama–Shimura–Weil conjecture. This situation changed with [Kisin 2008], which proved the existence of local deformation rings for GL_n corresponding to general p-adic Hodge theoretic conditions (namely being potentially crystalline or semistable of a given inertial type), and determined the structure of their generic fibres, in particular showing that they are generically regular, and computing their dimensions.

The results of [Kisin 2008] were generalised in [Balaji 2013] to the case of general reductive groups G under the hypothesis of being potentially crystalline, and in [Bellovin 2016] to the case that G is connected, and the inertial type is totally ramified. In the potentially crystalline case the generic fibres of the deformation rings can easily be shown to be regular, whereas in the potentially semistable case, one has to gain some control of the singularities, which is why there are additional restrictions in the theorems of [loc. cit.]. Our Theorem A is a common generalisation of these results to the case that G is possibly disconnected, and the representation is potentially semistable with no condition on the inertial type. (We also simultaneously handle the case that $p \neq l$.)

Another important application of Galois deformation theory to the Langlands program is to prove results showing that mod l representations of the Galois groups of number fields admit lifts to characteristic 0 with prescribed local properties; for example, such results were an important part of Khare and Wintenberger's proof of Serre's conjecture. The first such results were proved in [Ramakrishna 2002] for GL_2, and this method has now been generalised to a wide class of reductive groups; see in particular [Patrikis 2016; Booher 2019a; 2019b]. However, it has two disadvantages: it loses control of the local properties at a finite set of places, and it only applies in cases where formally smooth deformation rings exist.

A different approach was found in [Khare and Wintenberger 2009], which observed that in conjunction with the theory of potential modularity, such lifting results can be deduced from a lower bound on the Krull dimension of a global deformation ring, which was provided by the results of [Böckle 1999]. Kisin [2007] improved on the results of [Böckle 1999], proving a result about presentations of global deformation rings over local ones for GL_n, and deducing a lower bound on the dimensions of global deformation rings. These results were generalised to general reductive groups by Balaji [2013], and given our Theorem A, results such as Theorem B are essentially immediate from Balaji's.

Finally, [Booher and Patrikis 2017] (independently and contemporaneously) proved similar results to those of this paper in the case $l \neq p$ by a related but different method; rather than constructing a large enough supply of unobstructed points, as in this paper, they instead show that all points can be path connected to unobstructed points. We refer to the introduction to [loc. cit.] for a fuller discussion of the difference between the approaches.

1.2. ***Some details.*** We now explain our local results (and their proofs) in more detail. Theorem A is a generalisation of [Kisin 2008, Theorem 3.3.4], which proves the result in the case $l = p$ and $G = \mathrm{GL}_n$. It was previously adapted to the (much easier) case $G = \mathrm{GL}_n$ and $l \neq p$ in [Gee 2011] by using Weil–Deligne representations in place of the filtered (φ, N)-modules employed in [Kisin 2008]. It was also generalised in [Bellovin 2016] to the case that G is connected, $l = p$, and τ is totally ramified. Our approach is in some sense a synthesis of the approaches of [Gee 2011; Bellovin 2016], in that we treat the cases $l \neq p$ and $l = p$ essentially simultaneously, by using Weil–Deligne representations.

We briefly explain our approach, which in broad outline follows that of [Kisin 2008]. It is relatively straightforward (by passing from Galois representations to Weil–Deligne representations using Fontaine's constructions in the case $l = p$, and Grothendieck's monodromy theorem if $l \neq p$) to reduce Theorem A to analogous statements about moduli spaces of Weil–Deligne representations over l-adic fields. These moduli spaces admit an explicit tangent-obstruction theory given by an analogue of Herr's complex computing Galois cohomology in terms of (φ, Γ)-modules, and the key problem is to prove that the H^2 of this complex generically vanishes. We can think of this H^2 as a coherent sheaf over the moduli space, so by considering its support, we can reduce to the problem of exhibiting sufficiently many points at which the H^2 vanishes (which turn out to be precisely the regular points, which in a standard abuse of terminology we refer to as "smooth points").

Our approach to exhibiting these points is related to that taken in [Bellovin 2016], in that it makes use of the theory of associated cocharacters (see Section 2.3), but it is more streamlined and conceptual (for example, we do not need to consider the case $N = 0$ separately, as was done in [loc. cit.]). Surprisingly (at least to us), it is possible to construct all the smooth points that we need by considering the single Weil–Deligne representation $W_K \to \mathrm{SL}_2(\overline{\mathbb{Q}}_l)$ which is trivial on I_K, takes an arithmetic Frobenius element of W_K to

$$\begin{pmatrix} q^{1/2} & 0 \\ 0 & q^{-1/2} \end{pmatrix},$$

where q is the order of the residue field of K, and has

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

It is easy to check that this gives a smooth point of the moduli space of Weil–Deligne representations (while the point with the same representation of W_K but with $N = 0$ is not smooth).

Returning to the case of general G, suppose that the inertial type τ is trivial. If we consider a nilpotent element $N \in \operatorname{Lie} G$, the theory of associated cocharacters allows us to construct a particular homomorphism $\mathrm{SL}_2 \to G$ taking $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ to N, and an elementary calculation using the representation theory of $\mathfrak{sl}_2$ shows that the composition of our fixed representation $W_K \to \mathrm{SL}_2(\overline{\mathbb{Q}}_l)$ with this homomorphism defines a smooth point. We obtain further smooth points by multiplication by elements of $G(\overline{\mathbb{Q}}_l)$ of finite order, and this turns out to give us all the smooth points we need (even when G is not connected). (See Remark 2.3.10 for an interpretation of this construction in terms of the SL_2 version of the Weil–Deligne group.)

In the case of general τ we reduce to the same situation by replacing G by the normaliser in G of τ, which is also a reductive group. This use of Weil–Deligne representations is what allows us to remove the assumption made in [Bellovin 2016] that the inertial type is totally ramified, which was used in order to choose coordinates so that the inertial type τ was invariant under Frobenius. (Similarly, it clarifies the calculations made for GL_n in [Kisin 2008], as the semilinear algebra becomes linear algebra.) Under this assumption, when studying the structure of the moduli space of G-valued (φ, N, τ)-modules one could exploit the fact that Φ was in the centraliser $Z_G(\tau)$ and N was in $\operatorname{Lie} Z_G(\tau)$. Passing to Weil–Deligne representations r lets us argue similarly for general τ: a generator Φ of the unramified quotient of the Weil group normalises the inertial type and N is centralised by the inertial type. Since $Z_G(r|_{I_{L/K}})$ has finite index in the normaliser $N_G(r|_{I_{L/K}})$, we see that N is again in the Lie algebra of the algebraic group containing Φ.

In view of the functorial nature of our construction of smooth points, we are able to produce points on each irreducible component of the generic fibre of the deformation ring which are furthermore "very smooth" in the sense that they give rise to smooth points after restriction to any finite extension K'/K (these points were called "robustly smooth" in [Barnet-Lamb et al. 2014] when $p \neq l$). In particular, the images of such points on the corresponding deformation rings for $\mathrm{Gal}_{K'}$ lie on only one irreducible component, so that we obtain a well-defined "base change" map between irreducible components. We prove a similar result for the maps between deformation rings induced by morphisms of algebraic groups $G \to G'$ (see Section 3.5 for this, and for the case of base change). In particular, this allows one to talk about taking tensor products of components of deformation rings, which is frequently convenient when applying the Harris tensor product trick; see for example [Calegari et al. 2018].

We end this introduction by explaining the structure of the paper. In Section 2, we prove our main results about the structure of the moduli spaces of Weil–Deligne representations; we explain the tangent-obstruction theory and exhibit smooth points, and study the relationship with Galois representations. In doing so we remove the connectedness hypothesis on G made in [Bellovin 2016], by studying exact tensor-filtrations on fibre functors for disconnected reductive groups. We do this via a functor of points approach, using the dynamic approach to parabolic subgroups discussed in [Conrad et al. 2010, §I.2.1]. In Section 3 we deduce our results on the local structure of Galois deformation rings, which we then combine with the results of [Balaji 2013] to prove our lower bound on the dimension of a global deformation ring in Section 4. Finally, in Section 5 we specialise these results to the case of unitary groups.

1.3. ***Notation and conventions.*** All representations considered in this paper are assumed to be continuous with respect to the natural topologies, and we will never draw attention to this.

If K is a field then we write $\mathrm{Gal}_K := \mathrm{Gal}(\overline{K}/K)$ for its absolute Galois group, where $\overline{K}$ is a fixed choice of algebraic closure; we will regard all algebraic extensions of K as subfields of $\overline{K}$ without further comment, so that in particular we can take the compositum of any two such extensions. If L/K is a Galois extension then we write $\mathrm{Gal}_{L/K} := \mathrm{Gal}(L/K)$, a quotient of Gal_K. If K is a number field and v is a place of K then we fix an embedding $\overline{K} \hookrightarrow \overline{K}_v$, so that we have a homomorphism $\mathrm{Gal}_{K_v} \to \mathrm{Gal}_K$. If S is a finite set of places of a number field K, then we let $K(S)$ be the maximal extension of K (inside $\overline{K}$) which is unramified outside S, and write $\mathrm{Gal}_{K,S} := \mathrm{Gal}(K(S)/K)$.

If $K/\mathbb{Q}_p$ is a finite extension for some prime p then we write I_K for the inertia subgroup of Gal_K, W_K for the Weil group, and f_K for the inertial degree of $K/\mathbb{Q}_p$. We let φ denote the arithmetic Frobenius on $\overline{\mathbb{F}}_p$, so that we have an exact sequence

$$1 \to I_K \to W_K \to \langle \varphi^{f_K} \rangle \to 1,$$

and we let $v : W_K \to \mathbb{Z}$ be the function such that $v(g) = i$ if the image of g modulo I_K is $\varphi^{i f_K}$. Recall that a Weil–Deligne representation of W_K is a pair (r, N) consisting of a finite-dimensional representation $r : W_K \to \mathrm{End}(V)$ and a (necessarily nilpotent) endomorphism $N \in \mathrm{End}(V)$ satisfying

$$\rho(g)N = p^{v(g) f_K} N \rho(g)$$

for all $g \in W_K$.

1.3.1. *Parabolic subgroups.* If G is a finite-type affine group scheme over A, and $\lambda : \mathbb{G}_m \to G$ is a cocharacter of G, then there is a subgroup $P_G(\lambda)$ of G associated to λ as follows. Following [Conrad et al. 2010, §I.2.1], for any A-algebra A' we define the functors

$$P_G(\lambda)(A') = \{g \in G(A') \mid \lim_{t\to 0} \lambda(t) g \lambda(t)^{-1} \text{ exists}\},$$
$$U_G(\lambda)(A') = \{g \in P_G(\lambda)(A') \mid \lim_{t\to 0} \lambda(t) g \lambda(t)^{-1} = 1\}.$$

We also let $Z_G(\lambda)$ denote the scheme-theoretic centraliser of λ. All of these functors are representable by subgroup schemes of G, and they are smooth if G is smooth. By construction, the formation of $P_G(\lambda)$, $U_G(\lambda)$, and $Z_G(\lambda)$ commutes with base change on A.

The cocharacter λ induces a grading on the Lie algebra $\mathfrak{g} := \mathrm{Lie}\, G$. Let $\mathfrak{g}_n := \{v \in \mathfrak{g} \mid \mathrm{Ad}(\lambda(t))(v) = t^n v\}$ and let $\mathfrak{g}_{\geq 0} := \bigoplus_{n \geq 0} \mathfrak{g}_n$. Then $\mathrm{Lie}\, P_G(\lambda) = \mathfrak{g}_{\geq 0}$, $\mathrm{Lie}\, U_G(\lambda) = \mathfrak{g}_{\geq 1}$, and $\mathrm{Lie}\, Z_G(\lambda) = \mathfrak{g}_0$.

The multiplication map $Z_G(\lambda) \ltimes U_G(\lambda) \to P_G(\lambda)$ is an isomorphism. Furthermore, the fibres of $U_G(\lambda)$ are unipotent and connected. If the morphism $G \to \operatorname{Spec} A$ has connected reductive fibres, then $P_G(\lambda)$ is a parabolic subgroup scheme with connected fibres, $U_G(\lambda)$ is its unipotent radical, and $Z_G(\lambda)$ is connected and reductive.

1.3.2. *Deformation rings.* Let l be prime, and let $\mathcal{O}$ be the ring of integers in a finite extension $E/\mathbb{Q}_l$ with residue field $\mathbb{F}$. Write $\mathrm{CNL}_{\mathcal{O}}$ for the category of complete local noetherian $\mathcal{O}$-algebras with residue field $\mathbb{F}$.

Let Γ be either the absolute Galois group Gal_K of a finite extension K of $\mathbb{Q}_l$ for some p (possibly equal to l), or a group $\mathrm{Gal}_{K,S}$ where S is a finite set of places of a number field K.

Let G be a smooth affine group scheme over $\mathcal{O}$ whose geometric fibres are reductive (but not necessarily connected), and fix a homomorphism $\bar{\rho} : \Gamma \to G(\mathbb{F})$. A *framed deformation* of $\bar{\rho}$ to a ring $A \in \mathrm{CNL}_{\mathcal{O}}$ is a homomorphism $\rho : \Gamma \to G(A)$ whose reduction modulo $\mathfrak{m}_A$ is equal to $\bar{\rho}$. The functor of framed deformations is represented by the universal framed deformation $\mathcal{O}$-algebra $R^{\square}_{\bar{\rho}}$, an object of $\mathrm{CNL}_{\mathcal{O}}$ [Balaji 2013, Theorem 1.2.2].

Suppose from now on for the rest of the paper that the centre Z_G of G is smooth over $\mathcal{O}$. Write $\mathfrak{g}_{\mathbb{F}}$ and $\mathfrak{z}_{\mathbb{F}}$ for the $\mathbb{F}$-points of the Lie algebras of G and Z_G respectively; Γ acts on $\mathfrak{g}_{\mathbb{F}}$ via the adjoint action composed with $\bar{\rho}$. A *deformation* of $\bar{\rho}$ to A is a $(\ker(G(A) \to G(\mathbb{F})))$-conjugacy class of framed deformations of $\bar{\rho}$ to A. If $H^0(\Gamma, \mathfrak{g}_{\mathbb{F}}) = \mathfrak{z}_{\mathbb{F}}$, then the functor of deformations is represented by the universal framed deformation $\mathcal{O}$-algebra $R_{\bar{\rho}}$, an object of $\mathrm{CNL}_{\mathcal{O}}$; see [Balaji 2013, Theorem 1.2.2] or [Tilouine 1996, Theorem 3.3], together with Comment (2) following [loc. cit., Theorem 3.3].

We will also consider "fixed determinant" versions of these (framed) deformations rings. Let G^{ab} and G^{der} respectively denote the abelianisation and derived subgroup of G, and write $\mathrm{ab} : G \to G^{\mathrm{ab}}$ for the natural map. Write $\mathfrak{g}^0_{\mathbb{F}}$ for the $\mathbb{F}$-points of the Lie algebra of G^{der}. Fix a homomorphism $\psi : \Gamma \to G^{\mathrm{ab}}(\mathcal{O})$ such that $\mathrm{ab} \circ \bar{\rho} = \bar{\psi}$. We let $R^{\square,\psi}_{\bar{\rho}}$ denote the quotient of $R^{\square}_{\bar{\rho}}$ corresponding to deformations ρ with $\mathrm{ab} \circ \rho = \psi$ and $R^{\psi}_{\bar{\rho}}$ denote the quotient of $R_{\bar{\rho}}$ corresponding to framed deformations ρ with $\mathrm{ab} \circ \rho = \psi$.

We write G° for the connected component of G containing the identity. We will always consider representations up to G°-conjugacy, rather than G-conjugacy; note that this is compatible with our definition of deformations, as an element of $(\ker(G(A) \to G(\mathbb{F})))$ is necessarily contained in $G^{\circ}(A)$.

We for the most part allow any coefficient field E, although for some constructions in p-adic Hodge theory we need to allow it to be sufficiently large; we will comment when we do this. The effect of replacing E with a finite extension E' with ring of integers $\mathcal{O}'$ is simply to replace $R^{\square}_{\bar{\rho}}$ and $R_{\bar{\rho}}$ with $R^{\square}_{\bar{\rho}} \otimes_{\mathcal{O}} \mathcal{O}'$ and $R_{\bar{\rho}} \otimes_{\mathcal{O}} \mathcal{O}'$ respectively.

2. Moduli of Weil–Deligne representations

Let $K/\mathbb{Q}_p$ be a finite extension, and let l be a prime, possibly equal to p. In this section we prove analogues for l-adic Weil–Deligne representations of some results on moduli spaces of weakly admissible modules from [Kisin 2008; Bellovin 2016], and remove some hypotheses imposed in those papers; in particular, we allow our groups to be disconnected, and we work with arbitrary inertial types (rather than totally ramified types). In the case that $l = p$ we relate our moduli spaces to those for weakly admissible modules. In Section 3 we will use these results to study the generic fibres of deformation rings in both the case $l = p$ and the case $l \neq p$.

2.1. ***Moduli of Weil–Deligne representations.*** Let $K/\mathbb{Q}_p$ be a finite extension, and let L/K be a finite Galois extension. As in Section 1.3, we let $E/\mathbb{Q}_l$ be a finite extension for some prime l, with ring of integers $\mathcal{O}$. We also continue to let G be a (not necessarily connected) reductive group over $\mathcal{O}$; in fact, throughout this section we will be working with l inverted, and we will write G for G_E without further comment. We write $\mathfrak{g}_E$ for the Lie algebra of G.

A morphism of G-torsors $f : D \to D'$ over an E-scheme X is a morphism of the underlying X-schemes which is equivariant for the action of G_X. Such a morphism is necessarily an isomorphism. The G-equivariant automorphisms of D, which we denote by $\mathrm{Aut}_G(D)$, form a group, and it makes sense to talk about homomorphisms $r : W_K \to \mathrm{Aut}_G(D)$. We also define a sheaf of automorphism groups $\underline{\mathrm{Aut}}_G(D)$ over X; if X' is an X-scheme, its X'-points are given by $\underline{\mathrm{Aut}}_G(D)(X') := \mathrm{Aut}_G(D_{X'})$. This is a representable functor, since $\underline{\mathrm{Aut}}_G(D)$ is étale-locally isomorphic to G_X, which is affine. We abuse notation by writing $\underline{\mathrm{Aut}}_G(D)$ for the group scheme, as well.

Definition 2.1.1. Let $G\text{-}\mathrm{WD}_E(L/K)$ be the category cofibred in groupoids over E-Alg whose fibre over an E-algebra A is a G-torsor D over A together with a pair (r, N), where now $r : W_K \to \mathrm{Aut}_G(D)$ is a representation of the Weil group such that $r|_{I_L}$ is trivial, $N \in \mathrm{Lie}\,\underline{\mathrm{Aut}}_G(D)$, and $N = p^{-v(g) f_K}\,\mathrm{Ad}(r(g))(N)$ for all $g \in W_K$.

Requiring D to be a trivial G-torsor equipped with a trivialising section lets us define a representable functor covering $G\text{-}\mathrm{WD}_E(L/K)$, as follows. The exact sequence

$$0 \to I_K \to W_K \to \langle \varphi^{f_K} \rangle \cong \mathbb{Z} \to 0$$

is noncanonically split, and choosing a splitting is the same as choosing a lift $g_0 \in W_K$ of φ^{f_K}. Thus, to specify a representation $r : W_K \to \mathrm{Aut}_G(D)$, it suffices to specify $r|_{I_K}$ and $r(g_0)$ (which we denote by Φ). Since we are interested in representations which are trivial on I_L, we may replace $r|_{I_K}$ with $r|_{I_{L/K}}$. For an E-algebra A, we let $\mathrm{Rep}_A\, I_{L/K}$ denote the set of A-linear representations of $I_{L/K}$ on $G(A)$.

Definition 2.1.2. Choose $g_0 \in W_K$ lifting φ^{f_K}. We let $Y_{L/K,\varphi,\mathcal{N}}$ be the functor on the category of E-algebras whose A-points are triples

$$(\Phi, N, \tau) \in G(A) \times \mathfrak{g}_E(A) \times \mathrm{Rep}_A\, I_{L/K}$$

which satisfy

- $N = p^{-f_K}\,\mathrm{Ad}(\Phi)(N)$,
- $\Phi \circ \tau(g) \circ \Phi^{-1} = \tau(g_0 g g_0^{-1})$ for all $g \in I_{L/K}$, and
- $N = \mathrm{Ad}(\tau(g))(N)$ for all $g \in I_{L/K}$.

To go from $Y_{L/K,\varphi,\mathcal{N}}$ to $G\text{-}\mathrm{WD}_E(L/K)$, we need to forget the trivialising section and also forget g_0; the representation associated to (Φ, N, τ) is given by

$$r(g_0^n h) = \Phi^n \tau(h),$$

where $n \in \mathbb{Z}$ and $h \in I_K$.

The functor $Y_{L/K,\varphi,\mathcal{N}}$ is visibly represented by a finite-type affine scheme over E, and there is an action of G on $Y_{L/K,\varphi,\mathcal{N}}$ given by changing the trivialising section; explicitly,

$$a \cdot (\Phi, N, \{\tau(g)\}_{g \in I_{L/K}}) := (a\Phi a^{-1}, \mathrm{Ad}(a)(N), \{a\tau(g)a^{-1}\}_{g \in I_{L/K}}).$$

Recall that if Z is an E-scheme equipped with a left-action of an algebraic group H over E, then for any E-scheme S, the groupoid $[Z/H](S)$ over S is the category

$$[Z/H](S) := \{\text{Left } H\text{-bundle } D \to S \text{ and } H\text{-equivariant morphism } D \to Z\}.$$

A morphism $f : D \to D'$ in this fibre category is a morphism of H-torsors over S.

Lemma 2.1.3. *The quotient stack $[Y_{L/K,\varphi,\mathcal{N}}/G]$ is equivalent to the groupoid $G\text{-}\mathrm{WD}_E(L/K)$.*

Proof. We choose $g_0 \in W_K$ lifting φ^{f_K}. Given an A-valued point of $G\text{-}\mathrm{WD}_E(L/K)$ with underlying G-torsor D, the base change $D \times_A D \to D$ (which is projection on the first factor) is a trivial G-torsor (with G acting on the second factor). The identity morphism $D \xrightarrow{\sim} D$ induces a canonical trivialising section $D \to D \times_A D$, namely the diagonal. Pulling back r and N to $D \times_A D$, writing them in coordinates (with respect to the trivialising section), and writing $\tau := r|_{I_{L/K}}$ and $\Phi := r(g_0)$ gives us a morphism $D \to Y_{L/K,\varphi,\mathcal{N}}$.

We need to check that the morphism $D \to Y_{L/K,\varphi,\mathcal{N}}$ is G-equivariant. If A' is an A-algebra, the morphism $D \to Y_{L/K,\varphi,\mathcal{N}}$ carries $x \in D(A')$ to the fibre of (Φ, N, τ) over x. The fibre of $D \times_A D \to D$ over x is a copy of $D_{A'}$, together with a section (defined by taking the fibre of the diagonal over x). If $g \in G(A')$, the fibre of $D \times_A D \to D$ over $g \cdot x$ is also a copy of $D_{A'}$, but the section has been multiplied by g. Thus, our "change-of-basis" formula for triples (Φ, N, τ) implies that the morphism $D \to Y_{L/K,\varphi,\mathcal{N}}$ is G-equivariant, as required. □

Similarly, we let $Y_{L/K,\mathcal{N}}$ denote the functor on the category of E-algebras parametrising pairs

$$(N, \tau) \in \mathfrak{g}_E(A) \times \mathrm{Rep}_A\, I_{L/K}$$

such that $N = \mathrm{Ad}(\tau(g))(N)$ for all $g \in I_{L/K}$; and we let $Y_{L/K}$ be the functor on the category of E-algebras, whose A-points are $\mathrm{Rep}_A\, I_{L/K}$.

Let K'/K be a finite extension, and write L'/K' for the compositum of K' and L. Then L'/K' is Galois, with Galois group $\mathrm{Gal}_{L'/K'} \subset \mathrm{Gal}_{L/K}$. There are versions of the above functors for L'/K' which we write $Y_{L'/K',\varphi,\mathcal{N}}$, $Y_{L'/K',\mathcal{N}}$, and $Y_{L'/K'}$. Restriction of Weil–Deligne representations from W_K to $W_{K'}$ induces morphisms $Y_{L/K,\varphi,\mathcal{N}} \to Y_{L'/K',\varphi,\mathcal{N}}$, $Y_{L/K,\mathcal{N}} \to Y_{L'/K',\mathcal{N}}$ and $Y_{L/K} \to Y_{L'/K'}$.

2.2. ***A tangent-obstruction theory for G-$\mathrm{WD}_E(L/K)$.*** Choose an object $D_A \in G\text{-}\mathrm{WD}_E(L/K)$ with coefficients in an E-algebra A, and let $\mathrm{ad}\, D_A$ denote the Weil–Deligne module induced on $\mathrm{Lie}\, \underline{\mathrm{Aut}}_G D_A$. Choose $g_0 \in W_K$ which lifts φ^{f_K} and write $\Phi := r(g_0)$, let $\mathrm{Ad}(\Phi)$ denote the action on $\mathrm{ad}\, D_A$ given by differentiating the homomorphism $\underline{\mathrm{Aut}}_G D_A \to \underline{\mathrm{Aut}}_G D_A$ given by $g \mapsto \Phi g \Phi^{-1}$, and let ad_N act by $x \mapsto [N, x]$. If $G = \mathrm{GL}_n$ and D_A is the trivial torsor, these actions become $x \mapsto \Phi \circ x \circ \Phi^{-1}$ and $x \mapsto N \circ x - x \circ N$, respectively. Then we have an anticommutative diagram:

$$\begin{tikzcd}
(\mathrm{ad}\, D_A)^{I_{L/K}} \arrow[r, "1-\mathrm{Ad}(\Phi)"] \arrow[d, "\mathrm{ad}_N"] & (\mathrm{ad}\, D_A)^{I_{L/K}} \arrow[d, "\mathrm{ad}_N"] \\
(\mathrm{ad}\, D_A)^{I_{L/K}} \arrow[r, "p^{-f_K}\,\mathrm{Ad}(\Phi)-1"] & (\mathrm{ad}\, D_A)^{I_{L/K}}
\end{tikzcd}$$

Here $g \in I_{L/K}$ acts on $\mathrm{ad}\, D_A$ via $\mathrm{Ad}(\tau(g))$; note that the minus sign in p^{-f_K} arises because g_0 is a lift of arithmetic Frobenius. This diagram does not depend on our choice of g_0, because any two lifts of φ^{f_K} differ by an element of $I_{L/K}$, which acts trivially on $(\mathrm{ad}\, D_A)^{I_{L/K}}$.

The total complex $C^\bullet(D_A)$ of this double complex controls the deformation theory of objects of G-$\mathrm{WD}_E(L/K)$. We write $H^i(\mathrm{ad}\, D_A)$ for the cohomology groups of $C^\bullet(D_A)$. The following result will be proved in a very similar way to [Kisin 2008, Proposition 3.1.2], which is an analogous result for semilinear representations in the case $G = \mathrm{GL}_n$.

Proposition 2.2.1. *Let A be a local E-algebra with maximal ideal $\mathfrak{m}_A$ and let $I \subset A$ be an ideal with $I\mathfrak{m}_A = (0)$. Let $D_{A/I}$ be an object of G-$\mathrm{WD}_E(L/K)$ with coefficients in A/I, with Weil–Deligne representation $(\bar{r}, \bar{N})$. Then*:

(1) *If $H^2(\mathrm{ad}\, D_{A/\mathfrak{m}_A}) = 0$, then there exists an object D_A in G-$\mathrm{WD}_E(L/K)$ with coefficients in A, such that $(A/I) \otimes_A D_A \cong D_{A/I}$.*

(2) *The set of isomorphism classes of liftings of $D_{A/I}$ to D_A is either empty or a torsor under $I \otimes_{A/\mathfrak{m}_A} H^1(\mathrm{ad}\, D_{A/\mathfrak{m}_A})$.*

We begin by proving a preliminary lemma.

Lemma 2.2.2. *Let D_A be a G-torsor over A, and suppose there is a representation $\bar{r} : W_K \to \mathrm{Aut}_G(D_{A/I})$ such that $\bar{r}|_{I_L}$ is trivial. Then there is a representation $r : W_K \to \mathrm{Aut}_G(D_A)$ such that $r|_{I_L}$ is trivial and r lifts $\bar{r}$. Moreover, the set of infinitesimal automorphisms of r (as a lift of $\bar{r}$) is a torsor under*

$$H^0(W_K/I_L, I \otimes_{A/\mathfrak{m}_A} \mathrm{ad}\, D_{A/\mathfrak{m}_A}^{I_L}) = I \otimes_{A/\mathfrak{m}_A} \mathrm{ad}\, D_{A/\mathfrak{m}_A}^{W_K},$$

and the set of lifts of $\bar{r}$ is a torsor under

$$H^1(W_K/I_K, I \otimes_{A/\mathfrak{m}_A} \mathrm{ad}\, D_{A/\mathfrak{m}_A}^{I_{L/K}}).$$

Proof. An isomorphism $\bar{f} : D_{A/I} \to D_{A/I}$ lifts to an isomorphism $f : D_A \to D_A$, and the set of such lifts is a torsor under either a left- or right-action of $H^0(A, I \otimes_{A/\mathfrak{m}_A} D_{A/\mathfrak{m}_A})$ by [Bellovin 2016, Lemma 3.5]. Thus, for each $g \in W_K$, we can lift the map $\bar{r}(g) : D_{A/I} \to D_{A/I}$ to an isomorphism $r(g) : D_A \to D_A$.

The assignment

$$(g_1, g_2) \mapsto r(g_1)r(g_2)r(g_1g_2)^{-1}$$

is a 2-cocycle of W_K/I_L valued in $I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}$. Since we are in characteristic 0, and $I_{L/K}$ is a finite group, the Hochschild–Serre spectral sequence implies that for each $i > 0$, we have an isomorphism

$$H^i(W_K/I_K, I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}^{I_{L/K}}) \xrightarrow{\sim} H^i(W_K/I_L, I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}).$$

In particular,

$$H^2(W_K/I_L, I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}) \cong H^2(\widehat{\mathbb{Z}}, I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}^{I_{L/K}}) = 0,$$

so $\bar{r}$ lifts to a representation $r : W_K \to \operatorname{Aut}_G(D_A)$ with $r|_{I_L} = 0$, as claimed.

An isomorphism $f : D_A \to D_A$ is an infinitesimal automorphism of r if and only if it is the identity modulo I and $r(g) \circ f = f \circ r(g)$ for all $g \in W_K$. Equivalently, f is an element of $I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}$ fixed by W_K, and since I is a vector space over $A/\mathfrak{m}_A$, this is equivalent to $f \in I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}^{W_K}$, as desired.

Finally, if $r' : W_K \to \operatorname{Aut}_G(D)$ is another such lift, then $g \mapsto r'(g)r(g)^{-1}$ is a 1-cocycle of W_K/I_L valued in $I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}$. But

$$H^1(W_K/I_L, I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}) \cong H^1(W_K/I_K, I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}^{I_{L/K}}),$$

so we are done. □

Proof of Proposition 2.2.1. By [Bellovin 2016, Lemma 3.4], the underlying G-torsor $D_{A/I}$ lifts to a G-torsor D_A over $\operatorname{Spec} A$, and D_A is unique up to isomorphism, and by Lemma 2.2.2, $\bar{r}$ lifts to a representation $r : W_K \to \operatorname{Aut}_G(D_A)$. Moreover, by [loc. cit., Lemma 3.7], $\overline{N} \in \operatorname{ad} D_{A/I}$ lifts to some $N \in \operatorname{ad} D_A$ such that $\operatorname{Ad}(r(g))(N) = N$ for all $g \in I_{L/K}$, and any two lifts differ by an element of $I \otimes_{A/\mathfrak{m}_A} (\operatorname{ad} D_{A/\mathfrak{m}_A})^{I_{L/K}}$.

Now D_A, together with r and N, is an object of $G\text{-}\mathrm{WD}_E(L/K)$ if and only if $N = p^{-f_K} \operatorname{Ad}(\Phi)(N)$, where $\Phi := r(\varphi^{f_K})$. We define

$$h := N - p^{-f_K} \operatorname{Ad}(\Phi)(N) \in I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}^{I_{L/K}}.$$

If $H^2(\operatorname{ad} D_{A/\mathfrak{m}_A}) = 0$, then by definition there exist $f, g \in I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}^{I_{L/K}}$ such that $h = \operatorname{ad}_{\overline{N}}(f) + (p^{-f_K} \operatorname{Ad}(\overline{\Phi}) - 1)(g)$. We can view f and g either as elements of $\operatorname{Aut}_G(D_A)$ (congruent to the identity modulo I) or as elements of its tangent space. Thus we claim that if we define $\widetilde{N} := N + g$ and $\widetilde{\Phi} := f^{-1} \circ \Phi$, then $\widetilde{N} = p^{-f_K} \operatorname{Ad}(\widetilde{\Phi})(\widetilde{N})$. Indeed,

$$\begin{aligned}
&\widetilde{N} - p^{-f_K} \operatorname{Ad}(\widetilde{\Phi})(\widetilde{N}) \\
&\qquad = N + g - p^{-f_K}(\operatorname{Ad}(1-f) \circ \operatorname{Ad}(\Phi))(N+g) \\
&\qquad = N + g - p^{-f_K} \operatorname{Ad}(\Phi)(N) - p^{-f_K} \operatorname{Ad}(\Phi)(g) + p^{-f_K}[f, \operatorname{Ad}(\Phi)(N)] + p^{-f_K}[f, \operatorname{Ad}(\Phi)(g)] \\
&\qquad = \operatorname{ad}_{\overline{N}}(f) + p^{-f_K}[f, \operatorname{Ad}(\Phi)(N)] \\
&\qquad = [h, f] = 0.
\end{aligned}$$

Here we have used that $f, g, h \in I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}^{\mathrm{Gal}_{L/K}}$ and $I \cdot I \subset I\mathfrak{m}_A = 0$, so the Lie brackets $[f, \mathrm{Ad}(\Phi)(g)]$ and $[h, f]$ vanish. This proves part (1).

Now suppose that $\widetilde{N} = p^{-f_K} \mathrm{Ad}(\widetilde{\Phi})(\widetilde{N})$, and let $f, g \in I \otimes_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A}^{I_{L/K}}$. Define $\widetilde{N}' := N + g$ and $\widetilde{\Phi}' := f^{-1} \circ \widetilde{\Phi}$. Then

$$\begin{aligned}
&\widetilde{N}' - p^{-f_K} \mathrm{Ad}(\widetilde{\Phi}')(\widetilde{N}') \\
&\qquad = \widetilde{N} + g - p^{-f_K} \mathrm{Ad}(\widetilde{\Phi})(\widetilde{N}) - p^{-f_K} \mathrm{Ad}(\widetilde{\Phi})(g) + p^{-f_K}[f, \mathrm{Ad}(\widetilde{\Phi})(\widetilde{N})] + p^{-f_K}[f, \mathrm{Ad}(\widetilde{\Phi})(g)] \\
&\qquad = (1 - p^{-f_K} \mathrm{Ad}(\widetilde{\Phi}))(g) + [f, \widetilde{N}] \\
&\qquad = -(p^{-f_K} \mathrm{Ad}(\Phi) - 1)(g) - \mathrm{ad}_N(f).
\end{aligned}$$

Thus, $\widetilde{\Phi}', \widetilde{N}'$ give another lift if and only if $(f, g) \in \ker(d^1)$.

Moreover, if $(\widetilde{\Phi}', \widetilde{N}')$ is another lift, it is isomorphic to $(\widetilde{\Phi}, \widetilde{N})$ if and only if there is some $j \in I \otimes_{A/\mathfrak{m}_A} D_{A/\mathfrak{m}_A}^{I_{L/K}}$ such that

$$\widetilde{N}' = \mathrm{Ad}(1 + j)(\widetilde{N}) \quad \text{and} \quad (1 + j)\widetilde{\Phi} = \widetilde{\Phi}'(1 + j).$$

This is equivalent to

$$\widetilde{N} - \widetilde{N}' = \mathrm{ad}_N(j) \quad \text{and} \quad \widetilde{\Phi}(\widetilde{\Phi}')^{-1} = 1 - (1 - \mathrm{Ad}(\Phi))(j).$$

In other words, $(\widetilde{\Phi}, \widetilde{N})$ and $(\widetilde{\Phi}', \widetilde{N}')$ differ by an element of $\mathrm{im}(d^0)$, as required. □

2.3. ***Construction of smooth points.*** We wish to show that "most" points of $Y_{L/K,\varphi,\mathcal{N}}$ are smooth, and so are their images in $Y_{L'/K',\varphi,\mathcal{N}}$ for any finite extension K'/K. In this section we will consider a single fixed extension K'/K, and in Section 2.4 below we will deduce a result for all extensions K'/K simultaneously.

We begin by fixing an inertial type $\tau : I_{L/K} \to G(E)$. This amounts to considering the fibre of $Y_{L/K,\varphi,\mathcal{N}} \to Y_{L/K}$ over the point corresponding to τ. Next, we observe that if we can find $r : W_K \to G(E)$ such that $r|_{I_K} = \tau$, then $\Phi := r(g_0)$ is an element of the algebraic group defined over E

$$N_G(\tau) := \{h \in G \mid hr(g)h^{-1} \in r(I_{L/K}) \text{ for all } g \in I_{L/K}\}.$$

Note that Φ is not necessarily an element of the centraliser

$$Z_G(\tau) := \{h \in G \mid hr(g)h^{-1} = r(g) \text{ for all } g \in I_{L/K}\}.$$

However, since $I_{L/K}$ is finite (and in particular has only finitely many automorphisms), $Z_G(\tau) \subset N_G(\tau)$ has finite index; so we have $Z_G(\tau)^\circ = N_G(\tau)^\circ$ and $\mathrm{Lie}\, Z_G(\tau) = \mathrm{Lie}\, N_G(\tau)$. In particular, this implies that $N_G(\tau)$ and $Z_G(\tau)$ are reductive:

Theorem 2.3.1. *The normaliser $N_G(\tau) := \{h \in G \mid hr(g)h^{-1} \in r(I_{L/K}) \text{ for all } g \in I_{L/K}\}$ of $\tau(I_{L/K})$ is a reductive group.*

Proof. Since we are working over a field of characteristic 0, it is enough to prove that the connected component of the identity $N_G(\tau)^\circ = Z_G(\tau)^\circ = Z_{G^\circ}(\tau)^\circ$ is reductive. But reductivity for the latter group follows from [Prasad and Yu 2002, Theorem 2.1], which states that when a finite group acts on a connected reductive group, the connected component of the identity of the fixed points is reductive. □

Remark 2.3.2. Prasad and Yu prove their result under the assumption that the characteristic of the ground field does not divide the order of the group. Conrad, Gabber, and Prasad prove a more general result [Conrad et al. 2010, Proposition A.8.12], assuming only that the algebraic group acting is geometrically linearly reductive.

Our hypotheses imply that $N \in \operatorname{Lie} Z_G(\tau)$ and $\Phi \in N_G(\tau)$. However, if (r, N) exists and has the correct inertial type, the set of $\Phi \in G(E)$ compatible with $r|_{I_{L/K}}$ and N is a torsor under $Z_G(\tau) \cap Z_G(N)$.

We now briefly recall the theory of associated cocharacters over a field of characteristic 0; we refer the reader to [Jantzen 2004] (in particular Section 5) for further details and proofs. We will not draw attention to the assumption that our ground field has characteristic 0 below (but we will frequently use it); on the other hand, we do explain why the results that we are recalling hold over arbitrary fields of characteristic 0.

If $N \in \mathfrak{g}$ is nilpotent, a cocharacter $\lambda : \mathbb{G}_m \to G$ is said to be *associated* to N if

- $\operatorname{Ad}(\lambda(t))(N) = t^2 N$, and
- λ takes values in the derived subgroup of a Levi subgroup $L \subset G$ for which $N \in \mathfrak{l} := \operatorname{Lie} L$ is distinguished (that is, every torus contained in $Z_L(N)$ is contained in the centre of L).

By [McNinch 2004, Theorem 26], for any N there exists a cocharacter associated to N which is defined over the same field as N. Any two cocharacters associated to N are conjugate under the action of $Z_G(N)^\circ$.

An $\mathfrak{sl}_2$*-triple* is as usual a nonzero triple (X, H, Y) of elements of $\mathfrak{g}$ such that $[H, X] = 2X$, $[H, Y] = -2Y$, and $[X, Y] = H$. The Jacobson–Morozov theorem [Bourbaki 2005, Chapter VIII, §11, Proposition 2] states that for a nonzero nilpotent element N in a semisimple Lie algebra, an $\mathfrak{sl}_2$-triple (N, H, Y) always exists, and any two such triples (N, H, Y) and (N, H', Y') are conjugate under the action of $Z_G(N)^\circ$ [loc. cit., Chapter VIII, §11, Proposition 1]. Given a pair (N, H) such that $[H, N] = 2N$ and $H \in [N, \mathfrak{g}]$, it is possible to construct an $\mathfrak{sl}_2$-triple (N, H, Y) [loc. cit., Chapter VIII, §11, Lemme 6] (or the zero triple if $N = H = 0$). Since SL_2 is simply connected, this implies that there is a homomorphism $\mathrm{SL}_2 \to G$ which sends the "standard" basis for $\mathfrak{sl}_2$ to (N, H, Y).

If we let $\lambda : \mathbb{G}_m \to \mathrm{SL}_2 \to G$ be the composition of the cocharacter $t \mapsto \left(\begin{smallmatrix} t & 0 \\ 0 & t^{-1} \end{smallmatrix}\right)$ with this homomorphism $\mathrm{SL}_2 \to G$, then λ is associated to N. Moreover, the association $\lambda \mapsto d\lambda(1)$ sends cocharacters associated to N to elements H such that $[H, N] = 2N$ and $H \in [N, \mathfrak{g}]$, and this is an injective map [Jantzen 2004, Proposition 5.5] (this reference assumes that the ground field is algebraically closed, but this hypothesis is not used). Thus (in characteristic 0) associated cocharacters are a group-theoretic analogue of the Jacobson–Morozov theorem.

We use the following properties of associated cocharacters; the given reference assumes the ground field is algebraically closed, but these statements can all be checked after extension of the ground field.

Proposition 2.3.3 [Jantzen 2004, 5.9–11]. *Let G be a connected reductive group, let $N \in \mathfrak{g}$ be a nilpotent element, and let $\lambda : \mathbb{G}_m \to G$ be an associated cocharacter for N. Then*:

(1) *The associated parabolic $P_G(\lambda)$ depends only on N, not on the choice of associated cocharacter.*

(2) *We have $Z_G(N) \subset P_G(\lambda)$. In particular, $Z_G(N) = Z_{P_G(\lambda)}(N)$.*

(3) $Z_G(N) = (U_G(\lambda) \cap Z_G(N)) \rtimes (Z_G(\lambda) \cap Z_G(N))$.

(4) $Z_G(\lambda) \cap Z_G(N)$ *is reductive.*

In particular, by Proposition 2.3.3(3), the disconnectedness of $Z_G(N)$ is entirely accounted for by the disconnectedness of $Z_G(\lambda) \cap Z_G(N)$. The connectedness assumption on G for that part is removed in [Bellovin 2016, Proposition 4.9], so we may apply it to groups such as $Z_G(\tau)$ (which is reductive but not necessarily connected).

We will use the following lemma in the proof of Theorem 2.3.6 below.

Lemma 2.3.4. *If λ is an associated cocharacter of N, then the weight-2 part of $\mathfrak{g}$ for the adjoint action of λ is in the image of* ad_N.

Proof. If $N = 0$, then λ is the constant cocharacter and the corresponding weight-2 subspace is trivial. Otherwise, we may find an $\mathfrak{sl}_2$-triple of the form $(N, d\lambda(1), Y)$ and view $\mathfrak{g}$ as a representation of $\mathfrak{sl}_2$. Then the result follows by the representation theory of $\mathfrak{sl}_2$: if $T \in \mathfrak{g}$ is in the weight-2 part, then $\frac{1}{2}[Y, T]$ is in the weight-0 part and

$$\big[N, \tfrac{1}{2}[Y, T]\big] = \tfrac{1}{2}[[N, Y], T] = \tfrac{1}{2}[d\lambda(1), T] = T,$$

so T is in the image of ad_N. □

Let $f : G \to G'$ be a morphism of reductive groups over E, inducing a morphism $\mathfrak{g} \to \mathfrak{g}'$ on Lie algebras, which we also denote by f. We use the following lemma in the proof of Theorem 2.3.8 below.

Lemma 2.3.5. *If λ is an associated cocharacter for $N \in \mathfrak{g}$, then $f \circ \lambda$ is an associated cocharacter for $f(N)$.*

Proof. It is clear that $d\lambda(1)$ is semisimple. Then there exists some $Y \in \mathfrak{g}$ such that $(N, d\lambda(1), Y)$ is an $\mathfrak{sl}_2$-triple, and therefore there is a homomorphism $\mathrm{SL}_2 \to G$ such that the precomposition with the diagonal is λ. The composition $\mathbb{G}_m \to \mathrm{SL}_2 \to G \to G'$ is $f \circ \lambda$. Moreover, if we consider the composition $\mathrm{SL}_2 \to G \to G'$ and differentiate, we get a map $\mathfrak{sl}_2 \to \mathfrak{g}'$ sending the "standard" basis of $\mathfrak{sl}_2$ to $(f(N), f(d\lambda(1)), f(Y))$. This shows that $[f(d\lambda(1)), f(N)] = 2f(N)$ and $f(d\lambda(1))$ is in the image of $\mathrm{ad}_{f(N)}$. Since $f(d\lambda(1)) = d(f \circ \lambda)(1)$, this shows that $f \circ \lambda$ is associated to $f(N)$, by [Jantzen 2004, Proposition 5.5]. □

If K'/K is a finite extension, we write $H^2_{L'/K'}$ for the coherent sheaf on $Y_{L/K,\varphi,\mathcal{N}}$ given by the cokernel of

$$(\mathrm{ad}\,\mathcal{D})^{I_{L'/K'}} \oplus (\mathrm{ad}\,\mathcal{D})^{I_{L'/K'}} \xrightarrow{\mathrm{ad}_{N_{L'}} - (p^{-f_{K'}} \mathrm{Ad}(\Phi^{f_{K'}/f_K}) - 1)} (\mathrm{ad}\,\mathcal{D})^{I_{L'/K'}},$$

where $(\mathcal{D}, \Phi, N, \tau)$ is the universal object over $Y_{L/K,\varphi,\mathcal{N}}$, the operator $\mathrm{ad}_{N_{L'}}$ acts on the first factor and $(p^{-f_{K'}} \mathrm{Ad}(\Phi^{f_{K'}/f_K}) - 1)$ acts on the second factor. Then the fibre of $H^2_{L'/K'}$ at a closed point of $Y_{L/K,\varphi,\mathcal{N}}$ controls the obstruction theory of the restriction to $W_{K'}$ of the corresponding Weil–Deligne representation.

Theorem 2.3.6. *Let K'/K be a finite extension. Then there is a dense open subscheme $U \subset Y_{L/K,\varphi,\mathcal{N}}$ (possibly depending on K') such that $H^2_{L'/K'}|_U = 0$.*

Proof. Since the support of $H^2_{L'/K'}$ is closed, it suffices to show that if we consider the map $Y_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\mathcal{N}}$, then each component of the fibre over some point $N \in Y_{L/K,\mathcal{N}}$ contains a point (Φ, N) whose corresponding H^2 vanishes (when viewed as a point of $Y_{L'/K',\varphi,\mathcal{N}}$).

To do this, we consider a new moduli problem $\widetilde{Y}_{L/K,\varphi,\mathcal{N}}$, which by definition is the functor on the category of E-algebras whose A-points are triples

$$(\Phi, N, \tau) \in N_G(\tau) \times \operatorname{Lie} Z_G(\tau) \times \operatorname{Rep}_A I_{L/K}$$

which satisfy $N = p^{-f_K} \operatorname{Ad}(\Phi)(N)$.

This is representable by an affine scheme which we also write as $\widetilde{Y}_{L/K,\varphi,\mathcal{N}}$, and there is a natural morphism $\widetilde{Y}_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\mathcal{N}}$. Indeed, the map $Y_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\mathcal{N}}$ factors through the natural inclusion $Y_{L/K,\varphi,\mathcal{N}} \hookrightarrow \widetilde{Y}_{L/K,\varphi,\mathcal{N}}$, and the fibres of $Y_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\mathcal{N}}$ are closed and open in the fibres of $\widetilde{Y}_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\mathcal{N}}$. Thus, it suffices to study the fibres of the map $\widetilde{Y}_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\mathcal{N}}$. (Note that the tangent-obstruction complex for objects of $G\text{-}\mathrm{WD}_E(L/K)$ makes sense over $\widetilde{Y}_{L/K,\varphi,\mathcal{N}}$ as well.)

Choose an associated cocharacter $\lambda : \mathbb{G}_m \to Z_G(\tau)^\circ$ for N, so that in particular $\operatorname{Ad}(\lambda(t))(N) = t^2 N$, and let $\Phi := \lambda(p^{f_K/2})$. Then (Φ, N, τ) is a point of $\widetilde{Y}_{L/K,\varphi,\mathcal{N}}$, and we wish to study the restriction $(\Phi^{f_{K'}/f_K}, N_{L'}, \tau|_{I_{L'/K'}})$.

If D denotes the underlying G-torsor for (Φ, N, τ), and $\operatorname{ad} D$ denotes its pushout via the adjoint representation, then $\operatorname{Ad}(\Phi)$ and $\operatorname{Ad}(\Phi^{f_{K'}/f_K})$ are semisimple operators on $(\operatorname{ad} D)^{I_{L/K}}$ and $(\operatorname{ad} D)^{I_{L'/K'}}$, respectively. Therefore, $p^{-f_K} \operatorname{Ad}(\Phi) - 1$ and $p^{f_{K'}} \operatorname{Ad}(\Phi^{f_{K'}/f_K}) - 1$ are semisimple as well (since they are the difference of commuting semisimple operators in characteristic 0).

Thus, to compute the cokernel of $p^{-f_{K'}} \operatorname{Ad}(\Phi^{f_{K'}/f_K}) - 1$, it suffices to compute its kernel. Now $(\operatorname{ad} D)^{I_{L'/K'}}$ is graded by the adjoint action of $\lambda : \mathbb{G}_m \to Z_G(\tau) \subset Z_G(\tau|_{I_{L'/K'}})$, and if $(\operatorname{ad} D)_k^{I_{L'/K'}}$ denotes the weight-k subspace, then $p^{-f_{K'}} \operatorname{Ad}(\Phi^{f_{K'}/f_K}) - 1$ preserves it, so it suffices to compute

$$\ker(p^{-f_{K'}} \operatorname{Ad}(\Phi^{f_{K'}/f_K}) - 1)|_{(\operatorname{ad} D)_k^{I_{L'/K'}}}$$

for each k. But $p^{-f_{K'}} \operatorname{Ad}(\Phi^{f_{K'}/f_K}) - 1$ acts invertibly unless $k = 2$ (in which case it acts by 0), so the cokernel of $p^{-f_{K'}} \operatorname{Ad}(\Phi^{f_{K'}/f_K}) - 1$ is exactly $(\operatorname{ad} D)_2^{I_{L'/K'}}$. By Lemma 2.3.4, the weight-2 part of $\mathfrak{g}^{I_{L'/K'}}$ is in the image of ad_N, so we conclude that $H^2_{L'/K'}$ vanishes at (Φ, N), and at its image in $Y_{L'/K',\varphi,\mathcal{N}}$.

We need to find similar points on every connected component of the fibre of $\widetilde{Y}_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\mathcal{N}}$ over $N \in Y_{L/K,\mathcal{N}}$. This fibre is a torsor under $N_G(\tau) \cap Z_G(N)$, and the disconnectedness of $N_G(\tau) \cap Z_G(N)$ is entirely accounted for by the disconnectedness of $N_G(\tau) \cap Z_G(\lambda) \cap Z_G(N)$, by [Bellovin 2016, Proposition 4.9] (applied with $G' = N_G(\tau)$). On each component of $N_G(\tau) \cap Z_G(N)$, we may therefore by [loc. cit., Lemma 5.3] choose a finite-order element $c \in N_G(\tau) \cap Z_G(\lambda) \cap Z_G(N)$. (Note that $N_G(\tau) \cap Z_G(\lambda) \cap Z_G(N) = Z_{N_G(\tau)}(N) \cap Z_{N_G(\tau)}(\lambda)$ is reductive by Proposition 2.3.3.)

We now check that $H^2_{L/K}$ and $H^2_{L'/K'}$ vanish at the points of $\widetilde{Y}_{L/K,\varphi,N}$ and $\widetilde{Y}_{L'/K',\varphi,N}$, respectively, corresponding to $(\Phi\cdot c, N)$.

Firstly, we claim that $p^{-f_{K'}}\operatorname{Ad}((\Phi\cdot c)^{f_{K'}/f_K})-1$ is semisimple, or equivalently, that $\operatorname{Ad}((\Phi\cdot c)^{f_{K'}/f_K})$ is semisimple. For this, it suffices to check that some iterate of $\operatorname{Ad}((\Phi\cdot c)^{f_{K'}/f_K})$ is semisimple (since we are in characteristic 0). Let n be the order of c. Since c and $\Phi=\lambda(p^{f_K/2})$ commute,

$$\operatorname{Ad}(\Phi^{f_{K'}/f_K}\cdot c)^n=\operatorname{Ad}(\Phi^{nf_{K'}/f_K}\cdot c^n)=\operatorname{Ad}(\Phi^{nf_{K'}/f_K}).$$

But since $\operatorname{Ad}(\Phi)$ is semisimple by construction, so is $\operatorname{Ad}(\Phi^{nf_{K'}/f_K})$, as claimed.

Thus, to compute the cokernel of $p^{-f_{K'}}\operatorname{Ad}((\Phi\cdot c)^{f_{K'}/f_K})-1$, it suffices to compute its kernel, which is contained in the kernel of $p^{-nf_{K'}}\operatorname{Ad}(\Phi^{nf_{K'}/f_K})-1$. Since $p^{-nf_{K'}}\operatorname{Ad}(\Phi^{nf_{K'}/f_K})-1$ acts invertibly on each weight space $(\operatorname{ad} D)_k^{I_{L/K}}$ unless $k=2$, the cokernel of $p^{-f_{K'}}\operatorname{Ad}(\Phi^{f_{K'}/f_K}\cdot c)-1$ is contained in $(\operatorname{ad} D)_2^{I_{L/K}}$. Since $(\operatorname{ad} D)_2^{I_{L/K}}$ is again in the image of ad_N by Lemma 2.3.4, we are done. □

Corollary 2.3.7. *The stack $G\text{-}\mathrm{WD}_E(L/K)$ is generically smooth, and is equidimensional of dimension* 0*; equivalently, the scheme $Y_{L/K,\varphi,\mathcal{N}}$ is generically smooth, and is equidimensional of dimension* $\dim G$*. The nonsmooth locus is precisely the locus of Weil–Deligne representations D with $H^2(\operatorname{ad} D)\neq 0$. Moreover, $Y_{L/K,\varphi,\mathcal{N}}$ is locally a complete intersection and reduced.*

Proof. It is enough to prove the statement for $Y_{L/K,\varphi,\mathcal{N}}$. Let $U\subset Y_{L/K,\varphi,\mathcal{N}}$ be the dense open subscheme provided by Theorem 2.3.6 (with $K'=K$). Then at each closed point x of U, it follows from Lemma 2.2.2 and Proposition 2.2.1 that $Y_{L/K,\varphi,\mathcal{N}}$ is formally smooth at x. Furthermore, for any closed point x of $Y_{L/K,\varphi,\mathcal{N}}$ with corresponding Weil–Deligne representation D_x, the dimension of the tangent space at x is $\dim G-\dim H^0(D_x)+\dim H^1(D_x)$. Since the Euler characteristic of $C^\bullet(D_x)$ is 0, this is equal to $\dim G+\dim H^2(\operatorname{ad} D_x)=\dim G$, and the claim about $H^2(\operatorname{ad} D)$ follows immediately.

To see that $Y_{L/K,\varphi,\mathcal{N}}$ is reduced and locally a complete intersection, we proceed as in the proof of [Bellovin 2016, Corollary 5.4]. We have morphisms $Y_{L/K,\varphi,\mathcal{N}}\to Y_{L/K,\mathcal{N}}\to Y_{L/K}$, and the fibre above a point $\tau\in Y_{L/K}$ is defined by the relation $N=p^{-f_K}\operatorname{Ad}(\Phi)(N)$, where $\Phi\in Z_G(\tau)$ and $N\in\operatorname{Lie} Z_G(\tau)$. In other words, the fibre $Y_{L/K,\varphi,\mathcal{N}}|_\tau$ is cut out of the smooth $(2\dim Z_G(\tau))$-dimensional space $Z_G(\tau)\times\operatorname{Lie} Z_G(\tau)$ by $\dim Z_G(\tau)$ equations.

The quotient map $G\to G/Z_G(\tau)\cong Y_{L/K}$ admits sections étale locally. Thus, there is an étale neighborhood $U\to Y_{L/K}$ of τ such that the U-pullback $Y_{L/K,\varphi,\mathcal{N}}\times_{Y_{L/K}}U$ is isomorphic to $U\times Y_{L/K,\varphi,\mathcal{N}}|_\tau$. Since $Y_{L/K,\varphi,\mathcal{N}}\times_{Y_{L/K}}U$ is étale over $Y_{L/K,\varphi,\mathcal{N}}$, it is equidimensional of dimension $\dim G$. On the other hand, it is cut out of the smooth $(\dim U+2\dim Z_G(\tau))$-dimensional space $U\times Z_G(\tau)\times\operatorname{Lie} Z_G(\tau)$ by $\dim Z_G(\tau)$ equations.

Since $\dim U=\dim Y_{L/K}=\dim G-\dim Z_G(\tau)$ and being locally a complete intersection can be checked étale locally, it follows that $Y_{L/K,\varphi,\mathcal{N}}$ is locally a complete intersection. Moreover, schemes which are local complete intersections are Cohen–Macaulay, by [Matsumura 1989, Theorem 21.3], and Cohen–Macaulay schemes which are generically reduced are reduced everywhere, by [loc. cit., Theorem 17.3], so we are done. □

If $G \to G'$ is a morphism of reductive groups over E, then for any family of G-torsors D over Spec A, we can push out to a family D' of G'-torsors. Therefore, the moduli space $Y_{L/K,\varphi,\mathcal{N}}$ of (framed) G-valued Weil–Deligne representations carries a family D' of G'-torsors, and ad $D' := \operatorname{Lie} \operatorname{Aut}_{G'}(D')$ is a coherent sheaf on $Y_{L/K,\varphi,\mathcal{N}}$. Since D is a trivial G-torsor, D' is a trivial G'-torsor. Since pushing out G-torsors to G'-torsors is functorial, D' is a family of G'-valued Weil–Deligne representations and we can construct the complex $C^\bullet(D')$. We let $H^2_{G'}$ denote its cohomology in degree 2.

Theorem 2.3.8. *Let $f : G \to G'$ be a morphism of reductive groups over E. Then there is a dense open subset $U \subset Y_{L/K,\varphi,\mathcal{N}}$ (possibly depending on G') such that $H^2_{G'}|_U = 0$.*

Proof. As in the proof of Theorem 2.3.6, it suffices to construct a point on each connected component of each fibre of the map $Y_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\mathcal{N}}$ where $H^2_{G'}$ vanishes. In fact, the same points work: by Lemma 2.3.5 the composition $f \circ \lambda$ is an associated cocharacter for $f_*(N)$. Therefore, $H^2_{G'}$ vanishes at the point corresponding to $(\lambda(p^{f_K/2}), N)$. Similarly, if $c \in N_G(\tau) \cap Z_G(\lambda) \cap Z_G(N)$ is a finite-order point, then $H^2_{G'}$ vanishes at the point corresponding to $(\lambda(p^{f_K/2}) \cdot c, N)$. □

Remark 2.3.9. The proofs of Theorems 2.3.6 and 2.3.8 justify the claim we made in the Introduction, that all of the smooth points that we explicitly construct arise from pushing out a single "standard" smooth point for SL_2. Indeed, as discussed above, given an associated cocharacter λ for N, the map $\lambda \mapsto d\lambda(1)$ allows us to determine a homomorphism $\mathrm{SL}_2 \to G$, and we see that the choice of Φ, N made in the proof of Theorem 2.3.6 is the image under this homomorphism of the elements Φ, N for SL_2 discussed in the Introduction.

Remark 2.3.10. The Jacobson–Morozov theorem allows one to think of semisimple Weil–Deligne representations as representations of $W_K \times \mathrm{SL}_2$; see [Gross and Reeder 2010, Proposition 2.2] for a precise statement. From this perspective, our construction of smooth points from associated cocharacters can be summarised as follows: given a nilpotent $N \in \operatorname{Lie} G$, we obtain a map $\mathrm{SL}_2 \to G$, and the corresponding Weil–Deligne representation is obtained by composing with the map

$$W_K \times \mathrm{SL}_2 \to \mathrm{SL}_2$$

which on the first factor is unramified and takes an arithmetic Frobenius to the matrix

$$\begin{pmatrix} p^{f_K} & 0 \\ 0 & p^{-f_K} \end{pmatrix},$$

and is the identity on the second factor.

2.4. ***Tate local duality for Weil–Deligne representations.*** If D is a G-valued Weil–Deligne representation over a field E, we can also prove an analogue of Tate local duality for the complex $C^\bullet(D)$. In addition to allowing us to compute with either kernels or cokernels, this pairing allows us to give an explicit characterisation of the smooth locus (see Corollary 2.4.2). Since we only need the pairing between H^0 and H^2, we have not worked out the details of the pairing on H^1s, which for reasons of space we leave to the interested reader.

To construct pairings $H^i((\mathrm{ad}\,D)^*(1)) \times H^{2-i}(\mathrm{ad}\,D) \to E(1)$, we use the evaluation pairing

$$\mathrm{ev} : (\mathrm{ad}\,D)^* \times \mathrm{ad}\,D \to E.$$

Here the "(1)" means that we multiply the action of $\mathrm{Ad}(\Phi)$ by p^{f_K}; since $(\mathrm{ad}\,D)^*$ and $(\mathrm{ad}\,D)^*(1)$ have the same underlying vector space (as do E and $E(1)$), we have an induced pairing $\mathrm{ev}(1) : (\mathrm{ad}\,D)^*(1) \times \mathrm{ad}\,D \to E(1)$. Note that if $X \in (\mathrm{ad}\,D)^*$, $Y \in \mathrm{ad}\,D$, then

$$\mathrm{ev}(\mathrm{Ad}(\Phi)(X), \mathrm{Ad}(\Phi)(Y)) = \mathrm{ev}(X, Y),$$

and if $X \in (\mathrm{ad}\,D)^*(1)$, $Y \in \mathrm{ad}\,D$, then

$$\mathrm{ev}(1)(\mathrm{Ad}(\Phi)(X), \mathrm{Ad}(\Phi)(Y)) = \mathrm{ev}(p^{f_K}\,\mathrm{Ad}(\Phi)(X), \mathrm{Ad}(\Phi)(Y)) = p^{f_K}\,\mathrm{ev}(X, Y) = \mathrm{Ad}(\Phi)(\mathrm{ev}(1)(X, Y)).$$

Proposition 2.4.1. *Let D be as above. Then the evaluation pairing induces a perfect pairing*

$$H^0((\mathrm{ad}\,D)^*(1)) \times H^2(\mathrm{ad}\,D) \to E(1).$$

Proof. We first check that the pairing $\mathrm{ev}(1) : (\mathrm{ad}\,D)^*(1) \times \mathrm{ad}\,D \to E(1)$ descends to a well-defined pairing $H^0((\mathrm{ad}\,D)^*(1)) \times H^2(\mathrm{ad}\,D) \to E(1)$. If $X \in (\mathrm{ad}\,D)^*(1)^{I_{L/K}}$ is in the kernel of ad_N and the kernel of $1 - \mathrm{Ad}(\Phi)$, and $Y \in (\mathrm{ad}\,D)^{I_{L/K}}$, then

$$\begin{aligned} \mathrm{ev}(1)(X, Y + \mathrm{ad}_N(Z)) &= \mathrm{ev}(1)(X, Y) + \mathrm{ev}(1)(X, \mathrm{ad}_N(Z)) \\ &= \mathrm{ev}(1)(X, Y) - \mathrm{ev}(1)(\mathrm{ad}_N(X), Z) \\ &= \mathrm{ev}(1)(X, Y), \end{aligned}$$

and

$$\begin{aligned} \mathrm{ev}(1)(X, Y + (p^{-f_K}\,\mathrm{Ad}(\Phi) - 1)(Z)) &= \mathrm{ev}(1)(X, Y) + \mathrm{ev}(1)(X, p^{-f_K}\,\mathrm{Ad}(\Phi)(Z)) - \mathrm{ev}(1)(X, Z) \\ &= \mathrm{ev}(1)(X, Y) + p^{-f_K}\,\mathrm{ev}(1)(\mathrm{Ad}(\Phi)(X), \mathrm{Ad}(\Phi)(Z)) - \mathrm{ev}(1)(X, Z) \\ &= \mathrm{ev}(1)(X, Y) + \mathrm{ev}(1)(X, Z) - \mathrm{ev}(1)(X, Z) \\ &= \mathrm{ev}(1)(X, Y), \end{aligned}$$

so the pairing is indeed well-defined.

Next, we need to check that this pairing is perfect. Suppose $X \in H^0((\mathrm{ad}\,D)^*(1))$ and $\mathrm{ev}(1)(X, Y) = 0$ for all $Y \in H^2(\mathrm{ad}\,D)$. Then $\mathrm{ev}(1)(X, Y) = 0$ for all $Y \in (\mathrm{ad}\,D)^{I_{L/K}}$, so $X = 0$. This implies that the natural map $H^0((\mathrm{ad}\,D)^*(1)) \to (H^2(\mathrm{ad}\,D)^*)(1)$ is injective.

On the other hand, let $f : H^2(\mathrm{ad}\,D) \to E(1)$ be an element of $(H^2(\mathrm{ad}\,D)^*)(1)$. By composition, we have a linear functional

$$f : (\mathrm{ad}\,D)^{I_{L/K}} \to H^2(\mathrm{ad}\,D) \to E(1).$$

This is an element of $((\mathrm{ad}\,D)^{I_{L/K}})^*(1)$; we need to show that $\mathrm{ad}_N(f) = (1 - \mathrm{Ad}(\Phi))(f) = 0$. But for any $Y \in (\mathrm{ad}\,D)^{I_{L/K}}$,

$$\mathrm{ev}(1)(\mathrm{ad}_N(f), Y) = \mathrm{ev}(f, -\mathrm{ad}_N(Y)) = 0$$

since f factors through $H^2(\operatorname{ad} D)$. Similarly, for any $Y \in (\operatorname{ad} D)^{I_{L/K}}$,

$$\begin{aligned}\operatorname{ev}(1)((1-\operatorname{Ad}(\Phi))(f), Y) &= \operatorname{ev}(1)(f, Y) - \operatorname{ev}(1)(\operatorname{Ad}(\Phi)(f), Y)\\ &= \operatorname{ev}(1)(f, Y) - \operatorname{ev}(1)(f, p^{-f_K}\operatorname{Ad}(\Phi)^{-1}(Y))\\ &= \operatorname{ev}(1)(f, (1-p^{-f_K}\operatorname{Ad}(\Phi)^{-1})(Y))\\ &= \operatorname{ev}(1)\big(f, (p^{f_K}\operatorname{Ad}(\Phi)-1)(p^{-f_K}\operatorname{Ad}(\Phi^{-1})(Y))\big) = 0\end{aligned}$$

Since $\operatorname{Ad}(\Phi) : (\operatorname{ad} D)^{I_{L/K}} \to (\operatorname{ad} D)^{I_{L/K}}$ is an isomorphism, this suffices. □

Corollary 2.4.2. *The nonsmooth locus of the stack G-$\mathrm{WD}_E(L/K)$ is precisely the locus of Weil–Deligne representations D with $H^0((\operatorname{ad} D)^*(1)) \neq 0$.*

Proof. This is immediate from Corollary 2.3.7 and Proposition 2.4.1. □

We now use Corollary 2.4.2 to deduce that there is a dense set of points of $Y_{L/K,\varphi,\mathcal{N}}$ which give smooth points for every finite extension K'/K.

Definition 2.4.3. A point $x \in Y_{L/K,\varphi,\mathcal{N}}$ is *very smooth* if its image in $Y_{L'/K',\varphi,\mathcal{N}}$ is smooth for every finite extension K'/K.

Lemma 2.4.4. *Fix a finite extension E'/E. There is a finite extension K'/K (which depends only on E') such that $H^2_{L'/K'}$ vanishes at $x \in Y_{L/K,\varphi,\mathcal{N}}(E')$ if and only if x is very smooth.*

Proof. Suppose (D, Φ, N, τ) corresponds to a point of $Y_{L/K,\varphi,\mathcal{N}}$ such that $H^2_{L''/K''}$ does not vanish at its image in $Y_{L''/K'',\varphi,\mathcal{N}}$. By Corollary 2.4.2, this holds if and only if $H^0((\operatorname{ad} D)^*(1))$ does not vanish.

Thus, it suffices to consider the injectivity of

$$1 - p^{f_{K''}}\operatorname{Ad}(\Phi^{f_{K''}/f_K})^* : (\operatorname{ad} D)^{I_{L''/K''}} \to (\operatorname{ad} D)^{I_{L''/K''}}$$

on $\ker(\operatorname{ad}_N)$, where $\operatorname{Ad}(\Phi^{f_{K''}/f_K})^*$denotes the dual of $\operatorname{Ad}(\Phi^{f_{K''}/f_K})$. If this map is not injective, this implies that $p^{f_K}\operatorname{Ad}(\Phi)^*$ has a generalised eigenvalue λ satisfying $\lambda^{f_{K''}/f_K} = 1$. But the characteristic polynomial of $\operatorname{Ad}(\Phi)$ acting on $\operatorname{ad} D$ has degree $\dim \operatorname{ad} D = \dim G$ and there are only finitely many roots of unity with minimal polynomial of bounded degree over E'. It follows that there are only a finite number of possibilities for λ.

In other words, to check whether $1 - p^{f_{K''}}\operatorname{Ad}(\Phi^{f_{K''}/f_K})^*$ has a nontrivial kernel for any finite extension K''/K, it suffices to consider some fixed K' such that $f_{K'}/f_K$ is divisible by all n such that $\phi(n) \le \dim G$ and such that $\tau|_{I_{L'/K'}}$ is trivial (where $\phi(n)$ denotes Euler's totient function), as required. □

Corollary 2.4.5. *The set of closed points of G-$\mathrm{WD}_E(L/K)$ which are very smooth is Zariski dense.*

Proof. Let E'/E be a finite extension such that $Y_{L/K,\varphi,\mathcal{N}}(E')$ is Zariski dense in $Y_{L/K,\varphi,\mathcal{N}}$. By Lemma 2.4.4, there is a finite extension K'/K such that $x \in Y_{L/K,\varphi,\mathcal{N}}(E')$ is very smooth if $H^2_{L'/K'}$ vanishes at x. By Theorem 2.3.6, there is a Zariski dense open subscheme $U \subset Y_{L/K,\varphi,\mathcal{N}}$ such that $H^2_{L'/K'}|_U = 0$. But then the intersection $U \cap Y_{L/K,\varphi,\mathcal{N}}(E')$ is a Zariski dense subset of $Y_{L/K,\varphi,\mathcal{N}}$ consisting of very smooth points, so we are done. □

2.5. ***l-adic Hodge theory.*** We suppose in this subsection that $l \neq p$. We briefly recall some results from [Fontaine 1994], which will allow us to relate l-adic representations of Gal_K to Weil–Deligne representations.

Recall that by a theorem of Grothendieck, a continuous representation $\rho : \mathrm{Gal}_K \to \mathrm{GL}_d(E)$ is automatically potentially semistable, in the sense that there is a finite extension L/K such that $\rho|_{I_L}$ is unipotent. After making a choice of a compatible system of l-power roots of unity in $\overline{K}$, we see from [loc. cit., Propositions 1.3.3, 2.3.4] that there is an equivalence of Tannakian categories between the category of E-linear representations of Gal_K which become semistable over L, and the full subcategory of Weil–Deligne representations (r, N) of W_K over E with the properties that $r|_{I_L}$ is trivial and the roots of the characteristic polynomial of any arithmetic Frobenius element of W_L are l-adic units (such an equivalence is given by the functor $\underline{\widehat{WD}}_{pst}$ of [loc. cit., §2.3.7]).

2.6. ***The case $l = p$: (φ, N)-modules.*** In this section we let $l = p$, and we explain the relationship between Weil–Deligne representations and (φ, N)-modules. Let K_0, L_0 be the maximal unramified subfields of K, L respectively, of respective degrees f_K, f_L over $\mathbb{Q}_p$. Let $E/\mathbb{Q}_p$ be a finite extension, which is large enough that it contains the image of all embeddings $L_0 \hookrightarrow E$, so that we may identify $E \otimes_{\mathbb{Q}_p} L_0$ with $\bigoplus_{L_0 \hookrightarrow E} E$. Let φ denote the arithmetic Frobenius.

If D is a $\mathrm{Res}_{E\otimes_{\mathbb{Q}_p} L_0/E}\, G$-torsor over $\mathrm{Spec}\, A$, we may also view D as a G-torsor over $A \otimes_{\mathbb{Q}_p} L_0$. Then any automorphism $g : L_0 \to L_0$ extends to an automorphism of $A \otimes_{\mathbb{Q}_p} L_0$, and we may pull D back to a G-torsor g^*D over $A \otimes_{\mathbb{Q}_p} L_0$. Then we may view g^*D as a $\mathrm{Res}_{E\otimes_{\mathbb{Q}_p} L_0/E}\, G$-torsor over $\mathrm{Spec}\, A$, which we also denote by g^*D. In particular, we may pull D back by Frobenius and obtain another $\mathrm{Res}_{E\otimes_{\mathbb{Q}_p} L_0/E}\, G$-torsor φ^*D over $\mathrm{Spec}\, A$.

This motivates us to define the following groupoid on E-algebras.

Definition 2.6.1. The category of G-valued $(\varphi, N, \mathrm{Gal}_{L/K})$-modules, which we denote by $G\text{-}\mathrm{Mod}_{L/K,\varphi,N}$, is the groupoid whose fibre over an E-algebra A consists of a $\mathrm{Res}_{E\otimes L_0/E}\, G$-torsor D over A, equipped with

- an isomorphism $\Phi : \varphi^*D \xrightarrow{\sim} D$,
- a nilpotent element $N \in \mathrm{Lie}\,\mathrm{Aut}_G\, D$, and
- for each $g \in \mathrm{Gal}_{L/K}$, an isomorphism $\tau(g) : g^*D \xrightarrow{\sim} D$.

These are required to satisfy the following compatibilities:

(1) $\underline{\mathrm{Ad}}\Phi(N) = \frac{1}{p}N$.

(2) $\underline{\mathrm{Ad}}\tau(g)(N) = N$ for all $g \in \mathrm{Gal}_{L/K}$.

(3) $\tau(g_1 g_2) = \tau(g_1) \circ g_1^*\tau(g_2)$ for all $g_1, g_2 \in \mathrm{Gal}_{L/K}$.

(4) $\tau(g) \circ g^*\Phi = \Phi \circ \varphi^*\tau(g)$ for all $g \in \mathrm{Gal}_{L/K}$.

Here $\underline{\mathrm{Ad}}\Phi$ and $\underline{\mathrm{Ad}}\tau(g)$ are "twisted adjoint" actions on $\mathrm{Lie}\,\mathrm{Aut}_G\, D$; after pushing out Y by a representation $\sigma \in \mathrm{Rep}_E(G)$, they are given by $M \mapsto \Phi_\sigma \circ M \circ \Phi_\sigma^{-1}$ and $M \mapsto \tau(g)_\sigma \circ M \circ \tau(g)_\sigma^{-1}$, respectively.

Note that the action of $\mathrm{Gal}_{L/K}$ on scalars factors through the abelian quotient $\langle \varphi^{f_K} \rangle$, which also commutes with φ, so $(g_1 g_2)^* = g_1^* \circ g_2^*$ and $g^* \varphi^* = \varphi^* g^*$.

Requiring D to be a trivial $\mathrm{Res}_{E\otimes L_0/E}$-torsor equipped with a trivialising section lets us define a representable functor which covers $G\text{-Mod}_{L/K,\varphi,N,\tau}$, as follows.

Definition 2.6.2. Let $X_{L/K,\varphi,\mathcal{N}}$ denote the functor on the category of E-algebras whose A-points are triples

$$(\Phi, N, \tau) \in (\mathrm{Res}_{E\otimes L_0/E}\, G)(A) \times (\mathrm{Res}_{E\otimes L_0/E}\, \mathfrak{g}_E)(A) \times \mathrm{Rep}_{A\otimes L_0} \mathrm{Gal}_{L/K}$$

which satisfy

- $N = p\underline{\mathrm{Ad}}(\Phi)(N)$,
- $\tau(g) \circ \Phi = \Phi \circ \tau(g)$, and
- $\underline{\mathrm{Ad}}(\tau(g))(N) = N$ for all $g \in \mathrm{Gal}_{L/K}$.

This functor is visibly representable by a finite-type affine scheme over E, which we also denote by $X_{L/K,\varphi,\mathcal{N}}$. Moreover, there is a left action of $\mathrm{Res}_{E\otimes L_0/E}\, G$ on $X_{L/K,\varphi,\mathcal{N}}$ coming from changing the choice of trivialising section. Explicitly,

$$a \cdot (\Phi, N, \{\tau(g)\}_{g\in \mathrm{Gal}_{L/K}}) = (a\Phi\varphi(a)^{-1}, \mathrm{Ad}(a)(N), \{a\tau(g)(g\cdot a)^{-1}\}_{g\in\mathrm{Gal}_{L/K}}).$$

As in Lemma 2.1.3, we have the following:

Lemma 2.6.3. *The stack quotient $[X_{L/K,\varphi,\mathcal{N}}/\,\mathrm{Res}_{E\otimes L_0/E}\, G]$ is isomorphic to $G\text{-Mod}_{L/K,\varphi,N}$.*

Proof. The proof follows as in Lemma 2.1.3. □

Given a $(\varphi, N, \mathrm{Gal}_{L/K})$-module, there is a standard recipe due to Fontaine for constructing a Weil–Deligne representation, and there is an analogous construction for $\mathrm{Res}_{E\otimes L_0/E}\, G$-torsors. Indeed, let A be an E-algebra. Given a $\mathrm{Res}_{E\otimes L_0/E}\, G$-torsor D over A, and an embedding $\sigma : L_0 \hookrightarrow E$, the σ-isotypic part is a G-torsor over A which we denote by D_σ. Moreover, if N_σ denotes the σ-isotypic component of N, then $N_\sigma \in \mathrm{Lie}\,\mathrm{Aut}_G(D_\sigma)$ is nilpotent.

Given an isomorphism $\Phi : \varphi^* D \xrightarrow{\sim} D$, the composition $\Phi^{f_L} := \Phi \circ \varphi^*(\Phi) \circ \cdots \circ (\varphi^{f_L - 1})^*(\Phi)$ restricts to an isomorphism $D_\sigma \to D_\sigma$ for each σ.

Lemma 2.6.4. *For any σ and any E-algebra A, the association $(D, \Phi) \rightsquigarrow (D_\sigma, \Phi^{f_L})$ defines an equivalence of categories between $\mathrm{Res}_{E\otimes L_0/E}\, G$-torsors D over A equipped with an isomorphism $\Phi : \varphi^* D \xrightarrow{\sim} D$, and G-torsors D_σ over A equipped with an isomorphism $\Phi'_\sigma : D_\sigma \xrightarrow{\sim} D_\sigma$.*

Proof. Write the embeddings $\sigma_i : L_0 \hookrightarrow E$, $i \in \mathbb{Z}/f_L\mathbb{Z}$, with the numbering chosen so that $\sigma_1 = \sigma$, and Φ induces isomorphisms $\sigma_i : D_{i+1} \xrightarrow{\sim} D_i$ for each i (where we write D_i for D_{σ_i}).

Let $A \to A'$ be an fpqc cover trivialising D, so that $D_{A'}$ is a trivial torsor and we may choose a section. Then we can write $\Phi = (\Phi_1, \ldots, \Phi_{f_L})$.

We define

$$\underline{a} := (1, (\Phi_2 \cdots \Phi_{f_L})^{-1}, (\Phi_3 \cdots \Phi_{f_L})^{-1}, \ldots, \Phi_{f_L}^{-1}).$$

Then if we multiply our choice of trivialising section by $\underline{a}$, we replace Φ by

$$\underline{a}\Phi\varphi(\underline{a})^{-1} = (\Phi_1 \cdots \Phi_{f_L}, 1, \dots, 1).$$

Thus, we can recover $(D_{A'}, \Phi)$ from $((D_\sigma)_{A'}, \Phi^{f_L})$.

Furthermore, $D_{A'}$ is equipped with a descent datum, since it is the base change of D. Therefore, $(D_i)_{A'}$ has a descent datum, and since $(D_i)_{A'} \to \operatorname{Spec} A'$ is affine, it is effective.

Now suppose that $f = (f_1, \dots, f_{f_L}) : D \xrightarrow{\sim} D'$ is an isomorphism of $\operatorname{Res}_{E\otimes L_0/E} G$-torsors equipped with isomorphisms $\Phi : \varphi^* D \xrightarrow{\sim} D$ and $\Phi' : \varphi^* D' \xrightarrow{\sim} D'$. We obtain a corresponding isomorphism $f_{A'} : D_{A'} \xrightarrow{\sim} D'_{A'}$, together with a covering datum. Then each $f_i : D_i \xrightarrow{\sim} D'_i$ is an isomorphism of G-torsors, and we have

$$f_i \circ \Phi_i = \Phi'_i \circ f_{i+1} : D_{i+1} \to D'_i.$$

Multiplying the trivialising section of $D_{A'}$ by $\underline{a}$ and multiplying the trivialising section of $D_{A'}$ by $\underline{a}'$ has the effect of replacing $\underline{f}$ with $\underline{a}' \circ \underline{f} \circ \underline{a}^{-1}$. Then if we let $\underline{a}$ and $\underline{a}'$ be as above, $\underline{f}$ becomes $(f_1, \dots, f_1)$. Thus, we can also recover morphisms of pairs $(D, \Phi) \to (D', \Phi')$ from the associated morphisms of pairs $(D_i, \Phi^{f_L}) \to (D'_i, (\Phi')^{f_L})$, as required. □

Now suppose that D is a $\operatorname{Res}_{E\otimes L_0/E} G$-torsor equipped with an isomorphism $\Phi : \varphi^* D \xrightarrow{\sim} D$, and suppose in addition that D is equipped with a semilinear action τ of $\operatorname{Gal}_{L/K}$, compatible with Φ in the sense that $\Phi \circ \varphi^* \tau(g) = \tau(g) \circ g^*(\Phi)$ for all $g \in \operatorname{Gal}_{L/K}$. For each σ, we will construct a Weil–Deligne representation on D_σ which is trivial on I_L.

There is a surjective map $W_K \twoheadrightarrow \operatorname{Gal}_{L/K}$ which restricts to a surjection $I_K \twoheadrightarrow I_{L/K}$. If $g \in W_K$, we write $\bar{g}$ for its image in $\operatorname{Gal}_{L/K}$. For $g \in W_K$, we have an isomorphism

$$\tau(\bar{g}) : g^* D \xrightarrow{\sim} D$$

and we have an isomorphism

$$\Phi^{-v(g) f_K} := D \xrightarrow{\Phi^{-1}} \varphi^* D \xrightarrow{\varphi^*\Phi^{-1}} \cdots \xrightarrow{(g\varphi^{-1})^*\Phi^{-1}} g^* D.$$

Accordingly, we define $r(g) : D_\sigma \xrightarrow{\sim} D_\sigma$ to be the restriction of

$$r(g) := \tau(\bar{g}) \circ \Phi^{-v(g) f_K} : D \xrightarrow{\sim} D.$$

Note that $r|_{I_L}$ is trivial.

Lemma 2.6.5. *Let D be a G-torsor and let $r : W_K \to \operatorname{Aut}_G(D)$ be a homomorphism such that $r|_{I_L}$ is trivial. Then $r(W_L)$ centralises $r(W_K)$.*

Proof. Let $g \in W_K$ and let $h \in W_L$. Then $v(ghg^{-1}h^{-1}) = 0$, so $ghg^{-1}h^{-1} \in I_K$. Moreover, $W_L \subset W_K$ is a normal subgroup, so that $ghg^{-1}h^{-1} \in W_L$. But $I_K \cap W_L = I_L$, so $r(ghg^{-1}h^{-1}) = 1$, as required. □

We now prove the equivalence between Weil–Deligne representations and (φ, N)-modules. In the case that $G = \mathrm{GL}_n$, the following lemma is [Breuil and Schneider 2007, Proposition 4.1].

Lemma 2.6.6. *The map $r : W_K \to \mathrm{Aut}_G(D_\sigma)$ is a homomorphism, and $(D, \Phi, N, \tau) \rightsquigarrow (D_\sigma, r, N_\sigma)$ is an equivalence of categories between G-$\mathrm{Mod}_{L/K,\varphi,N}$ and G-$\mathrm{WD}_E(L/K)$.*

Proof. Since $\tau(\bar{g}) \circ g^*(\Phi) = \Phi \circ \varphi^*(\tau(\bar{g}))$, we have $\Phi^{-1} \circ \tau(\bar{g}) = \varphi^*(\tau(\bar{g})) \circ g^*(\Phi^{-1})$ as isomorphisms $g^*D \xrightarrow{\sim} \varphi^*D$. It follows that

$$\begin{aligned} r(g_1)r(g_2) &= (\tau(\bar{g}_1) \circ \Phi^{-v(g_1)f_K}) \circ (\tau(\bar{g}_2) \circ \Phi^{-v(g_2)f_K}) \\ &= \tau(\bar{g}_1) \circ (\varphi^{v(g_1)f_K})^*(\tau(\bar{g}_2) \circ \Phi^{-v(g_1g_2)f_K}) \\ &= \tau(\overline{g_1g_2}) \circ \Phi^{-v(g_1g_2)f_K} = r(g_1g_2) \end{aligned}$$

and r is a homomorphism. Another short computation shows that

$$N_\sigma = p^{-v(g)f_K}\,\mathrm{Ad}(r(g))(N_\sigma),$$

so that (E_σ, r, N_σ) is a G-valued Weil–Deligne representation.

The association $(D, \Phi, N, \tau) \rightsquigarrow (D_\sigma, r, N_\sigma)$ is clearly functorial. Moreover, if $f : D \to D'$ is a morphism of G-valued $(\varphi, N, \mathrm{Gal}_{L/K})$-modules, then $\Phi' \circ \varphi^*(f) = f \circ \Phi$. This implies that f is determined by its restriction $f|_{D_\sigma}$ to the σ-isotypic piece, and therefore, the functor is fully faithful.

We need to check that this functor is essentially surjective. In other words, we need to check that we can construct (D, Φ, N, τ) from (D_σ, r, N_σ). To do so, we number the embeddings as σ_i, as in the proof of Lemma 2.6.4. For each element $h \in I_{L/K}$, we fix a lift to an element $\tilde{h} \in I_K$; note that since $r|_{I_L}$ is trivial, $r(\tilde{h})$ is independent of the choice of $\tilde{h}$.

To construct $\Phi^{f_L}|_{D_i}$ from r, we observe that if $g_0 \in W_K$ lifts φ^{f_K} and (D_i, r, N_i) is in the essential image of our functor, then

$$r(g_0^{f_L/f_K}) = \tau(\bar{g}_0^{f_L/f_K})\Phi^{-f_L}.$$

But $\bar{g}_0^{f_L/f_K} \in I_{L/K}$, so we can define $\Phi^{f_L}|_{D_i} := r(g_0^{f_L/f_K})^{-1} r(\widetilde{\bar{g}_0^{f_L/f_K}})$.

We need to check that $\Phi^{f_L}|_{D_i}$ does not depend on our choice of g_0. Indeed, if $h \in I_K$, then

$$(g_0h)^{f_L/f_K} = h_1 \cdots h_{f_L/f_K-1} g_0^{f_L/f_K},$$

where $h_i := g_0^i h g_0^{-i} \in I_K$, so we may write $(g_0h)^{f_L/f_K} = h' g_0^{f_L/f_K}$ for some $h' \in I_K$. Then $r(\tilde{\bar{h}}') = r(h')$, so

$$r((g_0h)^{f_L/f_K})^{-1} r(\widetilde{\overline{g_0h}^{f_L/f_K}}) = r(g_0^{f_L/f_K})^{-1} r(h')^{-1} r(\tilde{\bar{h}}') r(\widetilde{\bar{g}_0^{f_L/f_K}}) = r(g_0^{f_L/f_K})^{-1} r(\widetilde{\bar{g}_0^{f_L/f_K}}),$$

as required.

Lemma 2.6.4 now implies that we can construct (D, Φ) from $(D_i, \Phi^{f_L}|_{D_i})$. Since $W_K \to \mathrm{Gal}_{L/K}$ is surjective, we define for $g \in \mathrm{Gal}_{L/K}$

$$\tau(g) := r(\tilde{g}) \circ \Phi^{v(\tilde{g})f_K} = r(\tilde{g}) \circ (\Phi \circ \cdots \circ (\varphi^{-1})^* g^*\Phi)$$

as a map $D_{i+v(g)f_K} \to D_i$. We need to check that this is well-defined. Note that the kernel of $W_K \to \mathrm{Gal}_{L/K}$ is W_L, and if $h \in W_L$, then $v(h) = (f_L/f_K) \cdot i$ for some $i \in \mathbb{Z}$. Thus, for any $h \in W_L$,

$$r(\tilde{g}h) \circ \Phi^{v(\tilde{g}h)f_K} = r(\tilde{g})r(h) \circ \Phi^{i \cdot f_L} \circ \Phi^{v(\tilde{g})f_K},$$

so it suffices to show that $r(h)\circ\Phi^{i\cdot f_L}=1$. Since $r|_{I_L}$ is trivial, it suffices to consider the case $i=1$, i.e., h generates the unramified quotient of W_L. But then

$$r(h)\circ\Phi^{f_L}=r(h)r(g_0^{f_L/f_K})^{-1}r(\widetilde{\bar{g}_0^{f_L/f_K}});$$

on the one hand $hg_0^{-f_L/f_K}\in I_K$ and $\widetilde{\bar{g}_0^{f_L/f_K}}\in I_K$, and on the other hand $g_0^{-f_L/f_K}\widetilde{\bar{g}_0^{f_L/f_K}}\in W_L$. It follows that

$$hg_0^{-f_L/f_K}\widetilde{\bar{g}_0^{f_L/f_K}}\in I_K\cap W_L=I_L$$

and the result follows.

We can also construct $\tau(g):D_{j+v(\tilde{g})f_K}\xrightarrow{\sim}D_j$ for the remaining σ_j-isotypic factors. Indeed, the desired compatibility between Φ and τ forces us to set

$$\varphi^*\tau(g):=\Phi^{-1}\circ\tau(g)\circ g^*\Phi:D_{i+v(\tilde{g})f_K+1}\xrightarrow{\sim}D_{i+1} \tag{2-6-1}$$

(and we proceed inductively).

We need to check that this is well-defined. More precisely, we need to check that $(\varphi^{f_L})^*\tau(g)=\tau(g)$ for all $g\in\mathrm{Gal}_{L/K}$. In other words, we need to check that

$$\tau(g)\circ(g^*\Phi\circ\varphi^*g^*\Phi\circ\cdots\circ(\varphi^{f_L-1})^*g^*\Phi)=(\Phi\circ\varphi^*\Phi\circ\cdots\circ(\varphi^{f_L-1})^*\Phi)\circ\tau(g)$$

as isomorphisms $D_{i+v(\tilde{g})f_K}\xrightarrow{\sim}D_i$, or equivalently that

$$\tau(g)\circ g^*\Phi^{f_L}=\Phi^{f_L}\circ\tau(g).$$

But

$$\begin{aligned}\tau(g)\circ g^*\Phi^{f_L}&=(r(\tilde{g})\circ\Phi^{v(\tilde{g})f_K})\circ g^*(\Phi^{f_L})\\&=r(\tilde{g})\circ\Phi^{f_L}\circ\Phi^{v(\tilde{g})f_K}\\&=r(\tilde{g})\cdot r(g_0^{-f_L/f_K}\widetilde{\bar{g}_0^{f_L/f_K}})\circ\Phi^{v(\tilde{g})f_K}\\&=r(g_0^{-f_L/f_K}\widetilde{\bar{g}_0^{f_L/f_K}})\cdot r(\tilde{g})\circ\Phi^{v(\tilde{g})f_K}\\&=\Phi^{f_L}\circ\tau(g).\end{aligned}$$

Here we used Lemma 2.6.5 and the fact that $g_0^{-f_L/f_K}\widetilde{\bar{g}_0^{f_L/f_K}}\in W_L$.

It remains to show that τ is a semilinear representation, or more precisely, that $\tau(g_1g_2)=\tau(g_1)\circ g_1^*\tau(g_2)$ for all $g_1,g_2\in\mathrm{Gal}_{L/K}$. Now by (2-6-1) we see that

$$\begin{aligned}\tau(g_1)\circ g_1^*\tau(g_2)&=\tau(g_1)\circ\big(((g_1\varphi^{-1})^*\Phi^{-1}\circ\cdots\circ\Phi^{-1})\circ\tau(g_2)\circ(g_2^*\Phi\circ\cdots\circ(g_1\varphi^{-1})^*g_2^*\Phi)\big)\\&=\tau(g_1)\circ((g_1\varphi^{-1})^*\Phi^{-1}\circ\cdots\circ\Phi^{-1})\circ\tau(g_2)\circ g_2^*(\Phi\circ\cdots\circ(g_1\varphi^{-1})^*\Phi)\\&=r(\tilde{g}_1)\circ r(\tilde{g}_2)\circ\Phi^{v(\tilde{g}_2)f_K}\circ g_2^*\Phi^{v(\tilde{g}_1)f_K}\\&=r(\tilde{g}_1)r(\tilde{g}_2)\circ\Phi^{v(\tilde{g}_1\tilde{g}_2)f_K}\\&=\tau(g_1g_2),\end{aligned}$$

as required.

Finally, we construct N. We have N_i, and we use the desired relation $N = p\underline{\mathrm{Ad}}(\Phi)(N)$ to construct the Frobenius-conjugates of N_i. It then follows that for any $g \in \mathrm{Gal}_{L/K}$

$$\begin{aligned}\underline{\mathrm{Ad}}(\tau(g))(N) &= \underline{\mathrm{Ad}}(r(\tilde{g}) \circ \Phi^{v(g) f_K})(N) \\ &= \mathrm{Ad}(r(\tilde{g}) \circ \Phi^{v(g) f_K})(p^{-v(g) f_K} \mathrm{Ad}(\Phi^{-v(g) f_K})(N)) \\ &= \mathrm{Ad}(r(\tilde{g}))(N) = N\end{aligned}$$

so we are done.

The assignment $(D_i, r, N_i) \rightsquigarrow (D, \Phi, N, \tau)$ is clearly functorial and quasi-inverse to $(D, \Phi, N, \tau) \rightsquigarrow (D_i, r, N_i)$. $\square$

2.7. ***Exact ⊗-filtrations for disconnected groups.*** In this section we prove some results on tensor filtrations that we will apply to the Hodge filtration in p-adic Hodge theory.

Let G be an affine group scheme over a field k of characteristic 0, let A be a k-algebra, and let η be a fibre functor from $\mathrm{Rep}_k(G)$ to Proj_A. More precisely, $\mathrm{Rep}_k(G)$ is the category of k-linear finite-dimensional representations of G, Proj_A is the category of finite projective A-modules (which we will also think of as being vector bundles on $\mathrm{Spec}\, A$), and by a "fibre functor" we mean that:

(1) η is k-linear, exact, and faithful.

(2) η is a tensor functor; that is, $\eta(V_1 \otimes_k V_2) = \eta(V_1) \otimes_A \eta(V_2)$.

(3) If $\mathbf{1}$ denotes the trivial representation of G, then $\eta(\mathbf{1})$ is the trivial A-module of rank 1.

Given a fibre functor $\eta : \mathrm{Rep}_k(G) \to \mathrm{Proj}_A$ and an A-algebra A', there is a natural fibre functor $\eta' : \mathrm{Rep}_k(G) \to \mathrm{Proj}_{A'}$ given by composing η with the natural base extension functor $\iota_{A'} : \mathrm{Proj}_A \to \mathrm{Proj}_{A'}$ sending M to $M \otimes_A A'$.

Definition 2.7.1. Let $\omega, \eta : \mathrm{Rep}_k(G) \rightrightarrows \mathrm{Proj}_A$ be fibre functors. Then $\underline{\mathrm{Hom}}^{\otimes}(\omega, \eta)$ is the functor on A-algebras given by

$$\underline{\mathrm{Hom}}^{\otimes}(\omega, \eta)(A') := \mathrm{Hom}^{\otimes}(\iota_{A'} \circ \omega, \iota_{A'} \circ \eta).$$

Here $\underline{\mathrm{Hom}}^{\otimes}$ refers to natural transformations of functors which preserve tensor products.

Theorem 2.7.2 [Deligne and Milne 1982, Theorem 3.2]. *Let $\omega : \mathrm{Rep}_k(G) \to \mathrm{Vec}_k$ be the natural forgetful functor*:

(1) *For any fibre functor $\eta : \mathrm{Rep}_k(G) \to \mathrm{Proj}_A$, the functor $\underline{\mathrm{Hom}}^{\otimes}(\iota_A \circ \omega, \eta)$ is representable by an affine scheme faithfully flat over* $\mathrm{Spec}\, A$; *it is therefore a G-torsor.*

(2) *The functor $\eta \rightsquigarrow \underline{\mathrm{Hom}}^{\otimes}(\iota_A \circ \omega, \eta)$ is an equivalence between the category of fibre functors η : $\mathrm{Rep}_k(G) \to \mathrm{Proj}_A$ and the category of G-torsors over* $\mathrm{Spec}\, A$. *The quasi-inverse assigns to any G-torsor X over A the functor η sending any $\rho : G \to \mathrm{GL}(V)$ to the $M \in \mathrm{Proj}_A$ associated to the push-out of X over A.*

Corollary 2.7.3. *Let $\eta : \mathrm{Rep}_k(G) \to \mathrm{Proj}_A$ be a fibre functor, corresponding to a G-torsor $X \to$* $\mathrm{Spec}\, A$. *Then the functor $\underline{\mathrm{Aut}}^{\otimes}(\eta)$ is representable by the A-group scheme $\mathrm{Aut}_G(X)$. This is a form of G_A.*

We now assume that η is equipped with an exact $\otimes$-filtration; i.e., for each $V \in \mathrm{Rep}_k(G)$, we have a decreasing filtration $\mathcal{F}^\bullet(\eta(V))$ of vector sub-bundles on each $\eta(V)$ such that:

(1) The specified filtrations are functorial in V.

(2) The specified filtrations are tensor-compatible, in the sense that

$$\mathcal{F}^n\eta(V \otimes_k V') = \sum_{p+q=n} \mathcal{F}^p\eta(V) \otimes_A \mathcal{F}^q\eta(V') \subset \eta(V) \otimes_A \eta(V').$$

(3) $\mathcal{F}^n(\eta(\mathbf{1})) = \eta(\mathbf{1})$ if $n \leq 0$ and $\mathcal{F}^n(\eta(\mathbf{1})) = 0$ if $n \geq 1$.

(4) The associated functor from $\mathrm{Rep}_k(G)$ to the category of graded projective A-modules is exact.

Equivalently, an exact $\otimes$-filtration of η is the same as a factorisation of η through the category of filtered vector bundles over $\mathrm{Spec}\, A$.

We define two auxiliary subfunctors of $\underline{\mathrm{Aut}}^{\otimes}(\eta)$:

- $P_{\mathcal{F}} = \underline{\mathrm{Aut}}^{\otimes}_{\mathcal{F}}(\eta)$ is the functor on A-algebras such that

$$P_{\mathcal{F}}(A') = \{\lambda \in \underline{\mathrm{Aut}}^{\otimes}(\eta)(A') \mid \lambda(\mathcal{F}^n\eta(V)) \subset \mathcal{F}^n\eta(V) \text{ for all } V \in \mathrm{Rep}_k(G) \text{ and } n \in \mathbb{Z}\}.$$

- $U_{\mathcal{F}} = \underline{\mathrm{Aut}}^{\otimes !}_{\mathcal{F}}(\eta)$ is the functor on A-algebras such that

$$U_{\mathcal{F}}(A') = \{\lambda \in \underline{\mathrm{Aut}}^{\otimes}(\eta)(A') \mid (\lambda - \mathrm{id})(\mathcal{F}^n\eta(V)) \subset \mathcal{F}^{n+1}\eta(V) \text{ for all } V \in \mathrm{Rep}_k(G) \text{ and } n \in \mathbb{Z}\}.$$

By [Saavedra Rivano 1972, Chapter IV, 2.1.4.1], these functors are both representable by closed subgroup schemes of $\mathrm{Aut}_G(X)$, and they are smooth if G is. This holds for any affine group G over k (since it is automatically flat); there is no need for reductivity or connectedness hypotheses. Furthermore, $\mathrm{Lie}\, P_{\mathcal{F}} = \mathcal{F}^0(\mathrm{Lie}\, \underline{\mathrm{Aut}}^{\otimes}(\eta))$ and $\mathrm{Lie}\, U_{\mathcal{F}} = \mathcal{F}^1(\mathrm{Lie}\, \underline{\mathrm{Aut}}^{\otimes}(\eta))$, by the same result.

We also have a notion of a $\otimes$-grading on η: a $\otimes$-*grading* of η is the specification of a grading $\eta(V) = \bigoplus_{n\in\mathbb{Z}} \eta(V)_n$ of vector bundles on each $\eta(V)$ such that:

(1) The specified gradings are functorial in V.

(2) The specified grading are tensor-compatible, in the sense that

$$\eta(V \otimes_k V')_n = \bigoplus_{p+q=n} (\eta(V)_p \otimes_A \eta(V')_q).$$

(3) $\eta(\mathbf{1})_0 = \eta(\mathbf{1})$.

Equivalently, a $\otimes$-grading of η is a factorisation of η through the category of graded vector bundles on $\mathrm{Spec}\, A$. A $\otimes$-grading induces a homomorphism of A-group schemes $\mathbb{G}_m \to \underline{\mathrm{Aut}}^{\otimes}(\eta)$.

Given a $\otimes$-grading of η, we may construct a $\otimes$-filtration of η, by setting

$$\mathcal{F}^n\eta(V) = \bigoplus_{n' \geq n} \eta(V)_{n'}.$$

We say that a $\otimes$-filtration $\mathcal{F}^\bullet$ is *splittable* if it arises in this way, and we say that $\mathcal{F}^\bullet$ is *locally splittable* if fpqc-locally on $\mathrm{Spec}\, A$ it arises in this way. A *splitting* of $\mathcal{F}^\bullet$ is a $\otimes$-grading on η giving rise to $\mathcal{F}^\bullet$.

Given an exact $\otimes$-filtration $\mathcal{F}^\bullet$ on η, we may define a fibre functor $\mathrm{gr}(\eta)$ equipped with a $\otimes$-grading by setting

$$\mathrm{gr}(\eta)(V)_n := \mathcal{F}^n(V)/\mathcal{F}^{n+1}(V).$$

Thus, a splitting of $\mathcal{F}^\bullet$ is equivalent to an isomorphism of filtered fibre functors $\mathrm{gr}(\eta) \cong \eta$.

In fact, by a theorem of Deligne (proved in [Saavedra Rivano 1972, Chapter IV, 2.4]), every $\otimes$-filtration is locally splittable (in fact, splittable Zariski-locally on $\operatorname{Spec} A$), because G is smooth and A has characteristic 0 (this result also holds under various other sets of hypotheses on G and A). Again, this does not require G to be reductive or connected. If $\lambda : \mathbb{G}_m \to \underline{\mathrm{Aut}}^{\otimes}(\eta)$ is a cocharacter splitting the filtration, then $P_{\mathcal{F}} = U_{\mathcal{F}} \rtimes Z_G(\lambda)$, by [loc. cit., Chapter IV, 2.1.5.1]. In particular, λ factors through $P_{\mathcal{F}}$.

If $\mathcal{F}^\bullet$ is a splittable filtration on η, we may consider the functor $\underline{\mathrm{Scin}}(\eta, \mathcal{F}^\bullet)$ of splittings. Then $\underline{\mathrm{Scin}}(\eta, \mathcal{F}^\bullet)$ is the same as the functor $\underline{\mathrm{Isom}}^{\otimes !}_{\mathcal{F}}(\mathrm{gr}_{\mathcal{F}}(\eta), \eta)$, which is the subset of $\underline{\mathrm{Isom}}^{\otimes}_{\mathcal{F}}(\mathrm{gr}_{\mathcal{F}}(\eta), \eta)$ inducing the identity $\mathrm{gr}_{\mathcal{F}}(\eta) \to \mathrm{gr}_{\mathcal{F}}(\eta)$. Thus, $\underline{\mathrm{Scin}}(\eta, \mathcal{F}^\bullet)$ is a left torsor under $U_{\mathcal{F}}$. It follows that the composition $\lambda : \mathbb{G}_m \to P_{\mathcal{F}} \to P_{\mathcal{F}}/U_{\mathcal{F}}$ is independent of the choice of splitting.

In other words, $P_{\mathcal{F}}$ and $U_{\mathcal{F}}$ depend only on the filtration, and if it is locally splittable, there is a homomorphism $\bar{\lambda} : \mathbb{G}_m \to P_{\mathcal{F}}/U_{\mathcal{F}}$ which also only depends on the filtration. If the filtration is actually splittable, a choice of splitting lets us lift $\bar{\lambda}$ to a cocharacter $\lambda : \mathbb{G}_m \to P_{\mathcal{F}}$. In that case, since both $\underline{\mathrm{Scin}}(\eta, \mathcal{F})$ and the set of lifts of cocharacters from $P_{\mathcal{F}}/U_{\mathcal{F}}$ to $P_{\mathcal{F}}$ are torsors under $U_{\mathcal{F}}$ (in the latter case, $U_{\mathcal{F}}$ acts by conjugation), they are isomorphic. In particular, any two cocharacters $\lambda, \lambda' : \mathbb{G}_m \rightrightarrows P_{\mathcal{F}}$ splitting the $\otimes$-filtration $\mathcal{F}$ are conjugate by $U_{\mathcal{F}}$.

Let $\mathcal{G} := \underline{\mathrm{Aut}}^{\otimes}(\eta)$, so that the geometric fibres of $\mathcal{G}$ are isomorphic to $G_{\bar{k}}$. Then for any geometric point $x \in \operatorname{Spec} A$, the $G^\circ(\kappa(x))$-conjugacy class of $\mathcal{F}^\bullet_x$ induces a unique $G^\circ(\kappa(x))$-conjugacy class of cocharacters, and this conjugacy class is Zariski-locally constant on $\operatorname{Spec} A$.

Recall that when $\lambda : \mathbb{G}_m \to \mathcal{G}$ is a cocharacter, we defined subgroups $U_{\mathcal{G}}(\lambda) \subset P_{\mathcal{G}}(\lambda) \subset \mathcal{G}$ in Section 1.3.

Proposition 2.7.4. *Suppose that G is a (possibly disconnected) algebraic group. Let $\eta : \mathrm{Rep}_k(G) \to \mathrm{Proj}_A$ be a fibre functor equipped with a splittable exact $\otimes$-filtration $\mathcal{F}^\bullet$, and let $\lambda : \mathbb{G}_m \to \underline{\mathrm{Aut}}^{\otimes}(\eta)$ be a splitting. Let $\mathcal{G}$ denote the group scheme representing $\underline{\mathrm{Aut}}^{\otimes}(\eta)$. Then $P_{\mathcal{F}} = P_{\mathcal{G}}(\lambda)$, $U_{\mathcal{F}} = U_{\mathcal{G}}(\lambda)$, and the fibres of $U_{\mathcal{F}}$ are connected.*

Proof. We consider the map $\mu : \mathbb{G}_m \times P_{\mathcal{F}} \to \underline{\mathrm{Aut}}^{\otimes}(\eta)$ defined by $\mu(t, g) := \lambda(t) g \lambda(t^{-1})$, and for $g \in P_{\mathcal{F}}(A')$, we let $\mu_g : (\mathbb{G}_m)_{A'} \to (\underline{\mathrm{Aut}}^{\otimes}(\eta))_{A'}$ be the restriction $\mu|_{\mathbb{G}_m \times \{g\}}$. Let $\sigma : G \to \mathrm{GL}(V)$ be a representation of G. Then the pushout $\eta(V)$ is a filtered vector bundle, and if $g \in P_{\mathcal{F}}(A')$, the action of g preserves the filtration on $\eta(V)$. The choice of a splitting in particular specifies an isomorphism $\mathrm{gr}^\bullet(\eta(V)) \xrightarrow{\sim} \eta(V)$, and $t \in \mathbb{G}_m(A')$ acts via t^n on $(\eta(V))_n$.

Let $\sigma_*(\lambda)$ denote the corresponding cocharacter $\sigma_*(\lambda) : \mathbb{G}_m \to \mathrm{Aut}_{\mathrm{GL}(V)}(\eta(V))$. Since this cocharacter induces the filtration on $\eta(V)$, we see that the morphism

$$\sigma_*(\mu_g) := \sigma_*(\lambda)(t) g \sigma_*(\lambda)(t^{-1}) : \mathbb{G}_m \to P_{\mathrm{Aut}_{\mathrm{GL}(V)}(\eta(V))}(\sigma_*(\lambda))$$

extends uniquely to a morphism

$$\widetilde{\sigma_*(\mu_g)} : \mathbb{A}^1 \to P_{\mathrm{Aut}_{\mathrm{GL}(V)}(\eta(V))}(\sigma_*(\lambda)).$$

We claim that the collection $\{\widetilde{\sigma_*(\mu_g)}\}_\sigma$ is functorial in σ and tensor-compatible. Indeed, since the collection $\{\widetilde{\sigma_*(\mu_g)}|_{\mathbb{G}_m}\}_\sigma$ is functorial in σ and tensor-compatible, and the extensions to $\mathbb{A}^1$ are unique, it follows that $\{\widetilde{\sigma_*(\mu_g)}\}_\sigma$ is functorial in σ and tensor-compatible. Thus, there is a morphism $\tilde{\mu}_g : \mathbb{A}^1 \to \underline{\mathrm{Aut}}^{\otimes}_{\mathcal{F}}(\eta)$ whose restriction to $\mathbb{G}_m$ is μ_g. It follows that $g \in P_{\mathcal{G}}(\lambda)(A')$.

Suppose in addition that $g \in U_{\mathcal{F}}(A')$. Then for every representation $\sigma : G \to \mathrm{GL}(V)$, g induces the identity map from $\mathrm{gr}^\bullet(\sigma(\mathcal{F}^\bullet))$ to itself. It follows that $\widetilde{\sigma_*(\mu_g)}(0) = \mathbf{1}$ for all σ, and therefore $\tilde{\mu}_g(0) = \mathbf{1}$.

On the other hand, if $g \in P_{\mathcal{G}}(\lambda)(A')$, then the morphism $\mu_g : (\mathbb{G}_m)_{A'} \to \underline{\mathrm{Aut}}^{\otimes}(\eta)_{A'}$ defined by $t \mapsto \lambda(t) g \lambda(t^{-1})$ extends to a morphism $\tilde{\mu}_g : (\mathbb{A}^1)_{A'} \to \underline{\mathrm{Aut}}^{\otimes}(\eta)_{\mathbb{A}'}$. It therefore induces a family of morphisms

$$\sigma_*(\tilde{\mu}_g) : (\mathbb{A}^1)_{A'} \to \mathrm{GL}(V)_{A'}$$

and so $\sigma_*(g) \in P_{\mathrm{Aut}_{\mathrm{GL}(V)}(\eta(V))}(\sigma_*(\lambda))$. But then $\sigma_*(g)$ preserves the filtration on $\eta(V)$ induced by $\sigma_*(\lambda)$; since this holds for all $V \in \mathrm{Rep}_k(G)$, we have $g \in P_{\mathcal{F}}(A')$. A similar argument shows that if $g \in U_{\mathcal{G}}(\lambda)(A')$, then $g \in U_{\mathcal{F}}(A')$.

Finally, since $\tilde{\mu}_g : \mathbb{A}^1 \to \underline{\mathrm{Aut}}^{\otimes}(\eta)$ is a morphism from a connected scheme such that $\tilde{\mu}_g(0) = \mathbf{1}$ and $\tilde{\mu}_g(1) = g$, we see that g is in the connected component of the identity for all $g \in U_{\mathcal{F}}(A')$. □

Lemma 2.7.5. *Let $\mathcal{F}^\bullet$ be a locally splittable exact $\otimes$-filtration on η. Then the geometric fibres of $P_{\mathcal{F}}$ are parabolic subgroups of $G_{\bar{k}}$.*

Proof. We may work locally on Spec A and assume that we have a cocharacter $\lambda : \mathbb{G}_m \to \mathcal{G}_A$ splitting the exact $\otimes$-filtration. Then $P_{\mathcal{F}} \cong P_{\mathcal{G}}(\lambda)$. Since the formation of $P_{\mathcal{G}}(\lambda)$ commutes with base change on A, we may assume that $A = k = \bar{k}$ and $\mathcal{G} = G = G_{\bar{k}}$. Then $P_{G^\circ}(\lambda) \subset G^\circ$ is a parabolic subgroup, so $G^\circ / P_{G^\circ}(\lambda)$ is proper. There is a sequence of maps

$$G^\circ / P_{G^\circ}(\lambda) \to G / P_{G^\circ}(\lambda) \twoheadrightarrow G / P_G(\lambda).$$

Since $G^\circ \subset G$ has finite index, the properness of $G^\circ / P_{G^\circ}(\lambda)$ implies the properness of $G / P_{G^\circ}(\lambda)$. This implies that $G / P_G(\lambda)$ is proper, so $P_G(\lambda) \subset G$ is a parabolic subgroup. □

We will also need the following result:

Theorem 2.7.6 [SGA 3_{II} 1970, Exposé IX, Théorème 3.6]. *Let S be an affine scheme, S_0 a subscheme defined by a nilpotent ideal J, H a group of multiplicative type over S, G a smooth group scheme over S, and $\mu_0 : H \times_S S_0 \to G \times_S S_0$ a homomorphism of S_0-groups.*

Then there exists a homomorphism $\mu : H \to G$ of S-groups which lift μ_0, and any two such lifts are conjugate by an element of $G(S)$ which reduces to the identity modulo J.

Corollary 2.7.7. *Let A be an artin local k-algebra with maximal ideal $\mathfrak{m}_A$, and let $I \subset A$ be an ideal such that $I\mathfrak{m}_A = (0)$. Then if D_A is a G-torsor over A such that the reduction $D_{A/I} := D_A \otimes_A A/I$ is* *equipped with an exact $\otimes$-filtration $\mathcal{F}^\bullet_{A/I}$, then the set of lifts of $\mathcal{F}^\bullet_{A/I}$ to an exact $\otimes$-filtration on D_A is nonempty, and is a torsor under $I \otimes_{A/\mathfrak{m}_A} (\operatorname{ad} D_{A/\mathfrak{m}_A}/\mathcal{F}^0_{A/\mathfrak{m}_A}(\operatorname{ad} D_{A/\mathfrak{m}_A}))$.*

Proof. Suppose that $D_{A/I}$ is a G-torsor over $\operatorname{Spec} A/I$, equipped with an exact $\otimes$-filtration $\mathcal{F}^\bullet_{A/I}$. Since A/I is local, $\mathcal{F}^\bullet_{A/I}$ is split, so it is induced by a cocharacter $\lambda_{A/I} : \mathbb{G}_m \to \operatorname{Aut}_G(D_{A/I})$. By Theorem 2.7.6, $\lambda_{A/I}$ lifts to a cocharacter $\lambda_A : \mathbb{G}_m \to \operatorname{Aut}_G(D_A)$. Then λ_A induces an exact $\otimes$-filtration $\mathcal{F}^\bullet_A$ on D_A which lifts that on $D_{A/I}$.

Suppose there are two exact $\otimes$-filtrations, $\mathcal{F}^\bullet_A$ and $\mathcal{F}'^\bullet_A$ on D_A lifting $\mathcal{F}^\bullet_{A/I}$, induced by cocharacters λ_A and λ'_A, respectively, which lift $\lambda_{A/I}$. Then λ_A and λ'_A are conjugate by an element of $\operatorname{Aut}_G(D_A)$ which is the identity modulo I. In other words, there is some $j \in \operatorname{ad} D_{A/\mathfrak{m}_A} \otimes_{A/\mathfrak{m}_A} I$ such that $\lambda'_A = (1+j)\lambda_A(1-j)$. This implies that $\mathcal{F}^\bullet_A$ and $\mathcal{F}'^\bullet_A$ are conjugate.

On the other hand, conjugation by $1+j$ preserves $\mathcal{F}^\bullet_A$ if and only if $1+j \in P_{\mathcal{F}_A}(\operatorname{Aut}_G(D_A))$. This holds if and only if $j \in \mathcal{F}^0_{A/\mathfrak{m}_A} \operatorname{Lie} \operatorname{Aut}_G(D_{A/\mathfrak{m}_A}) \otimes_{A/\mathfrak{m}_A} I = \mathcal{F}^0_{A/\mathfrak{m}_A} \operatorname{ad} D_{A/\mathfrak{m}_A} \otimes_{A/\mathfrak{m}_A} I$. □

2.8. *p-adic Hodge theory.* Our goal is to study deformations of potentially semistable Galois representations. That is, we wish to consider deformations of representations $\rho : \operatorname{Gal}_K \to G(E)$ such that $\rho|_{\operatorname{Gal}_L}$ is semistable. Such representations can be described by linear algebra. Briefly, for every representation $\sigma : G \to \operatorname{GL}_d$, $\sigma \circ \rho$ is a potentially semistable representation, and $D^L_{\mathrm{st}}(\sigma \circ \rho)$ is a weakly admissible filtered $(\varphi, N, \operatorname{Gal}_{L/K})$-module. The formation of $D^L_{\mathrm{st}}(\sigma \circ \rho)$ is exact and tensor-compatible in σ, and if $\mathbf{1}$ denotes the trivial representation of G, then $D^L_{\mathrm{st}}(\mathbf{1} \circ \rho)$ is the trivial filtered $(\varphi, N, \operatorname{Gal}_{L/K})$-module with coefficients in E.

Therefore, as in [Bellovin 2016, §A.2.8–9], $\sigma \mapsto D^L_{\mathrm{st}}(\sigma \circ \rho)$ is a fibre functor $\eta : \operatorname{Rep}_E(G) \to \operatorname{Proj}_{E \otimes_{\mathbb{Q}_p}} L_0$, and we obtain from ρ a G-torsor $D = D^L_{\mathrm{st}}(\rho)$ over $E \otimes L_0$ equipped with

- an isomorphism $\Phi : \varphi^* D \xrightarrow{\sim} D$,
- a nilpotent element $N \in \operatorname{Lie} \operatorname{Aut}_G D$,
- for each $g \in \operatorname{Gal}_{L/K}$, an isomorphism $\tau(g) : g^* D \xrightarrow{\sim} D$,
- a $\operatorname{Gal}_{L/K}$-stable exact $\otimes$-filtration on D_L, or equivalently (by Galois descent), an exact $\otimes$-filtration on the $\operatorname{Res}_{E \otimes K/E} G$-torsor $D_L^{\operatorname{Gal}_{L/K}}$ over K.

These satisfy the requisite compatibilities such that forgetting the filtration on $D^L_{\mathrm{st}}(\rho)$ gives us an object of $G\text{-}\operatorname{Mod}_{L/K,\varphi,N}$.

Definition 2.8.1. The category of G-valued filtered $(\varphi, N, \operatorname{Gal}_{L/K})$-modules, which we denote by $G\text{-}\operatorname{Mod}_{L/K,\varphi,N,\mathrm{Fil}}$, is the category cofibred in groupoids over E-Alg whose fibre over an E-algebra A consists of a $\operatorname{Res}_{E \otimes L_0/E} G$-torsor D over A, equipped with

- an isomorphism $\Phi : \varphi^* D \xrightarrow{\sim} D$,
- a nilpotent element $N \in \operatorname{Lie} \operatorname{Aut}_G D$,
- for each $g \in \operatorname{Gal}_{L/K}$, an isomorphism $\tau(g) : g^* D \xrightarrow{\sim} D$,

- a $\mathrm{Gal}_{L/K}$-stable exact $\otimes$-filtration on D_L, or equivalently, an exact $\otimes$-filtration on the $\mathrm{Res}_{E\otimes K/E}\, G$-torsor $D_L^{\mathrm{Gal}_{L/K}}$ over A.

The $\mathrm{Res}_{E\otimes L_0/E}\, G$-torsor D, together with Φ, N, and $\{\tau(g)\}_{g\in\mathrm{Gal}_{L/K}}$, is required to be an object of $G\text{-}\mathrm{Mod}_{L/K,\varphi,N}$.

Definition 2.8.2. Suppose that $\rho:\mathrm{Gal}_K\to G(E)$ is a potentially semistable Galois representation which becomes semistable when restricted to Gal_L. The *p-adic Hodge type* $\boldsymbol{v}$ of ρ is the $(\mathrm{Res}_{E\otimes K/E}\, G)^\circ(\bar{E})$-conjugacy class of cocharacters $\lambda:\mathbb{G}_m\to(\mathrm{Res}_{E\otimes K/E}\, G)_{\bar{E}}$ which split the $\otimes$-filtration on $D_{\mathrm{st}}^L(\rho)_L^{\mathrm{Gal}_{L/K}}$. We let $P_{\boldsymbol{v}}$ denote the $(\mathrm{Res}_{E\otimes K/E}\, G)^\circ(\bar{E})$-conjugacy class of $P_{\mathrm{Res}_{E\otimes K/E}\, G}(\lambda)$ for $\lambda\in\boldsymbol{v}$.

While we do not need it, for completeness we record the following definition and result, which control the deformation theory of filtered $(\varphi, N, \mathrm{Gal}_{L/K})$-modules. Given an object $D_A\in G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}}$, we consider the diagram

$$\begin{array}{ccccc}
(\mathrm{ad}\, D_A)^{\mathrm{Gal}_{L/K}} & \longrightarrow & (\mathrm{ad}\, D_A)^{\mathrm{Gal}_{L/K}}\oplus(\mathrm{ad}\, D_A)^{\mathrm{Gal}_{L/K}} & \longrightarrow & (\mathrm{ad}\, D_A)^{\mathrm{Gal}_{L/K}} \\
\downarrow & & & & \\
(\mathrm{ad}\, D_{A,L}/\mathrm{Fil}^0\mathrm{ad}\, D_{A,L})^{\mathrm{Gal}_{L/K}} & & & &
\end{array}$$

where the top line is the total complex of

$$\begin{array}{ccc}
(\mathrm{ad}\, D_A)^{\mathrm{Gal}_{L/K}} & \xrightarrow{1-\underline{\mathrm{Ad}}(\Phi)} & (\mathrm{ad}\, D_A)^{\mathrm{Gal}_{L/K}} \\
\downarrow{\scriptstyle \mathrm{ad}_N} & & \downarrow{\scriptstyle \mathrm{ad}_N} \\
(\mathrm{ad}\, D_A)^{\mathrm{Gal}_{L/K}} & \xrightarrow{p\underline{\mathrm{Ad}}(\Phi)-1} & (\mathrm{ad}\, D_A)^{\mathrm{Gal}_{L/K}}
\end{array}$$

and the vertical map is the natural quotient map. We let $C^\bullet_{\mathrm{Fil}}$ denote its total complex. Then $C^\bullet_{\mathrm{Fil}}$ controls the deformation theory of D_A:

Proposition 2.8.3. *Let A be an artin local E algebra with maximal ideal $\mathfrak{m}_A$ and let $I\subset A$ be an ideal such that $I\mathfrak{m}_A=(0)$. Let $D_{A/I}$ be an object of $G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}}(A/I)$ and set $D_{A/\mathfrak{m}_A}:=D_{A/I}\otimes_{A/I}A/\mathfrak{m}_A$:*

(1) *If $H^2_{\mathrm{Fil}}(D_{A/I})=0$, then there exists an object $D_A\in G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}}(A)$ lifting $D_{A/I}$.*

(2) *The set of isomorphism classes of lifts of $D_{A/I}$ to $D_A\in G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}}(A)$ is either empty or a torsor under $H^1_{\mathrm{Fil}}(D_{A/\mathfrak{m}_A})\otimes_{A/\mathfrak{m}_A}I$.*

Proof. This follows by combining [Bellovin 2016, Proposition 3.2] and Corollary 2.7.7. □

3. Local deformation rings

As in Section 1.3.2, we let $K/\mathbb{Q}_p$ be a finite extension for some prime p, possibly equal to l, and let $\bar{\rho}:\mathrm{Gal}_K\to G(\mathbb{F})$ be a continuous representation. We have a universal framed deformation $\mathcal{O}$-algebra $R^\square_{\bar{\rho}}$, and if we fix a homomorphism $\psi:\Gamma\to G^{\mathrm{ab}}(\mathcal{O})$ such that $\mathrm{ab}\circ\bar{\rho}=\bar{\psi}$, we also have the quotient $R^{\square,\psi}_{\bar{\rho}}$ corresponding to framed deformations ρ with $\mathrm{ab}\circ\rho=\psi$. When we define quotients of $R^\square_{\bar{\rho}}$, there are corresponding quotients of $R_{\bar\rho}^{\square,\psi}$, which we will not explicitly define, but will denote by a superscript ψ. An inertial type is by definition a $G^\circ(\bar E)$-conjugacy class of representations $\tau : I_K \to G(\bar E)$ with open kernel which admit extensions to Gal_K; any such τ is defined over some finite extension of E. We choose a finite Galois extension L/K for which $\tau|_{I_L}$ is trivial. If E'/E is a finite extension, and $\rho : \mathrm{Gal}_K \to G(E')$ is a representation, which we assume to be potentially semistable if $l = p$, then we say that ρ has type τ if the restriction to I_K (forgetting N) of the corresponding Weil–Deligne representation $\mathrm{WD}(\rho)$ is equivalent to τ.

3.1. ***The case $l \neq p$.*** Suppose firstly that $l \neq p$. The proof of [Balaji 2013, Proposition 3.0.12] shows that for each τ we may define a $\mathbb{Z}_l$-flat quotient $R_{\bar\rho}^{\square,\tau}$ of $R_{\bar\rho}^{\square}$ whose characteristic-0 points correspond to representations of type τ. The usual construction of the Weil–Deligne representation associated to a Galois representation makes sense over $R_{\bar\rho}^{\square}[1/l]$, so we have a natural morphism

$$\operatorname{Spec} R_{\bar\rho}^{\square,\tau}[1/l] \to G\text{-}\mathrm{WD}_E(L/K).$$

3.2. ***The case $l = p$.*** Now suppose that $l = p$. If we fix a p-adic Hodge type $\boldsymbol{v}$ in the sense of Definition 2.8.2 (that is, a $(\mathrm{Res}_{E\otimes K/E}\, G)^\circ(\bar E)$-conjugacy class of cocharacters $\lambda : \mathbb{G}_m \to (\mathrm{Res}_{E\otimes K/E}\, G)_{\bar E}$), and an inertial type τ, then by [Balaji 2013, Proposition 3.0.12] there is a unique $\mathbb{Z}_l$-flat quotient $R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}$ of $R_{\bar\rho}^{\square}$ with the property that if B is a finite local E-algebra, then a morphism $R_{\bar\rho}^{\square} \to B$ factors through $R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}$ if and only if the corresponding representation $\rho : \mathrm{Gal}_K \to G(B)$ is potentially semistable with Hodge type $\boldsymbol{v}$ and inertial type τ. For each finite-dimensional representation V of G, we may compose with the representation $\mathrm{Gal}_K \to G(R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}[1/p])$ to obtain a representation $\mathrm{Gal}_K \to \mathrm{GL}(V)(R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}[1/p])$. Then exactly as in [Kisin 2008, Theorem 2.5.5] we obtain a corresponding ($\mathrm{GL}(V)$-valued) filtered $(\varphi, N, \mathrm{Gal}_{L/K})$-module over $R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}[1/p]$ (note that we have been working with covariant functors in this paper, while Kisin uses contravariant functors; it is necessary to dualise the construction in [loc. cit., §2.4]). As these filtered $(\varphi, N, \mathrm{Gal}_{L/K})$-modules are exact and tensor-compatible, we obtain a G-valued filtered $(\varphi, N, \mathrm{Gal}_{L/K})$-module over $R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}[1/p]$. By Lemma 2.6.6, we again have a natural morphism

$$\operatorname{Spec} R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}[1/l] \to G\text{-}\mathrm{WD}_E(L/K).$$

3.3. ***Denseness of very smooth points.*** We continue to fix an inertial type τ and (if $p = l$) a p-adic Hodge type $\boldsymbol{v}$. For convenience, if $l \neq p$ then for the rest of this section we write $R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}$ for $R_{\bar\rho}^{\square,\tau}$; this notational convention allows us to treat the cases $l \neq p$ and $l = p$ simultaneously. We study the generic fibre $R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}[1/l]$ via the morphism

$$\operatorname{Spec} R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}[1/l] \to G\text{-}\mathrm{WD}_E(L/K). \tag{3-3-1}$$

In a standard abuse of terminology, we say that a closed point $x \in \operatorname{Spec} R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}[1/l]$ is *smooth* if the (completed) local ring at x is regular. We will see in the proof of Theorem 3.3.2 that these are the points whose images in $G\text{-}\mathrm{WD}_E(L/K)$ are smooth points, which perhaps justifies this terminology. Similarly, we say that x is *very smooth* if for any finite extension K'/K, the image of x in (with obvious notation) $\operatorname{Spec} R^{\square,\tau|_{I_{K'}},\boldsymbol{v}_{K'}}_{\bar{\rho}|_{G_{K'}}}[1/l]$ is smooth.

As in [Kisin 2009, Proposition 2.3.5], if $x \in \operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ is a closed point corresponding to a representation ρ_x, then the completed local ring A_x at x pro-represents framed deformations of ρ_x which are potentially semistable of p-adic Hodge type $\boldsymbol{v}$ (if $l = p$), and have inertial type τ.

Proposition 3.3.1. (1) *If x is a closed point of the Jacobson scheme* $\operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$, *then the completion at x of the morphism (3-3-1) is formally smooth.*

(2) *The morphism (3-3-1) is flat.*

Proof. The formal smoothness follows from the proofs of [Kisin 2008, Lemma 3.2.1, Proposition 3.3.1], which carries over verbatim to our setting (since the morphism of groupoids from framed deformations to unframed deformations is formally smooth). Part (2) then follows from the fact that formally smooth morphisms between locally noetherian schemes are flat, which in turn follows from [EGA IV$_1$ 1964, §0 Théorème 19.7.1]. □

Theorem 3.3.2. *Assume that $R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}} \neq 0$. There is a dense open subscheme $U \subset \operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ which is regular, and there is a Zariski dense subset of $\operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ consisting of very smooth points. Furthermore, $\operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ is equidimensional of dimension $\dim G + \delta_{l=p} \dim \operatorname{Res}_{E\otimes K/E} G/P_{\boldsymbol{v}}$, locally a complete intersection, and reduced.*

Similarly, $\operatorname{Spec} R^{\square,\tau,\boldsymbol{v},\psi}_{\bar{\rho}}[1/l]$ contains a regular dense open subscheme and a Zariski dense subset of very smooth points, and is equidimensional of dimension $\dim G^{\mathrm{der}} + \delta_{l=p} \dim(\operatorname{Res}_{E\otimes K/E} G)/P_{\boldsymbol{v}}$.

Remark 3.3.3. In contrast to previous work (in particular [Kisin 2008; Gee 2011; Bellovin 2016]), we only claim that U is regular, not formally smooth over $\mathbb{Q}_p$. We are grateful to Jeremy Booher and Stefan Patrikis [2017] for drawing our attention to this.

Proof. Since the formation of scheme-theoretic images is compatible with flat base change, the existence of a dense open subscheme U consisting of smooth points follows from Corollary 2.3.7 and Proposition 3.3.1. The existence of a Zariski dense subset of very smooth points follows from Corollary 2.4.5. We claim that if $x \in \operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ is a closed point in U, then the completion A_x of $R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ at x is a formally smooth $\mathbb{Q}_p$-algebra, and is in particular regular. Indeed, if $\mathfrak{m}_x$ is the maximal ideal of A_x, then $\operatorname{Spec} A_x/\mathfrak{m}_x^n \subset U$ for all $n \geq 1$ (since U is open). Let B be a local $\mathbb{Q}_p$-algebra with maximal ideal $\mathfrak{m}_B$ and let $I \subset B$ be an ideal such that $I\mathfrak{m}_B = (0)$. If there is a local homomorphism $A_x \to B/I$, let $D_{B/I}$ be the induced object of $G\text{-}\mathrm{WD}_E(L/K)(B/I)$. Then $H^2(\operatorname{ad} D_{B/I}) = 0$, since the homomorphism $A_x \to B/I$ factors through $A/\mathfrak{m}_x^n$ for some n. It follows that $D_{B/I}$ lifts to $D_B \in G\text{-}\mathrm{WD}_E(L/K)(B)$. Since $\operatorname{Spf} A_x \to G\text{-}\mathrm{WD}_E(L/K)$ is formally smooth, D_B is induced from a map $A_x \to B$ lifting $A \to B/I$. Since $R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ is Noetherian, it follows that the localisation of $R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ at x is regular [Stacks 2005–, Tag 07NY], so U is regular by [loc. cit., Tag 02IT], as claimed.

Thus, to compute the dimension of $\operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$, it is enough to compute the dimension of the tangent spaces at closed points in U. Let x be such a closed point, let E' be its residue field, and write

A_x for the completion of $R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ at x. Since the morphism $\operatorname{Spf} A_x \to G\text{-}\mathrm{WD}_E(L/K)$ is formally smooth by Proposition 3.3.1, it is versal at x. More precisely, in the case that $l \neq p$ we see (by the equivalence between Galois representations and Weil–Deligne representations recalled in Section 2.5) that the induced map $\operatorname{Spf} A_x \to G\text{-}\mathrm{WD}_E(L/K)^\wedge_x$ (with the right-hand side denoting the completion of the target at x) is a $\widehat{G}$-torsor, where $\widehat{G}$ is the completion of G_E along the closed subgroup given by the centraliser of the representation corresponding to x, in the sense that there is an evident isomorphism

$$\operatorname{Spf} A_x \times \widehat{G} \xrightarrow{\sim} \operatorname{Spf} A_x \times_{G\text{-}\mathrm{WD}_E(L/K)^\wedge_x} \operatorname{Spf} A_x.$$

In particular, we have $\dim A_x \times_{G\text{-}\mathrm{WD}_E(L/K)^\wedge_x} A_x = \dim A_x + \dim \widehat{G}$, and the claim about the dimension then follows from [Emerton and Gee 2017, Lemma 2.40] and Corollary 2.4.5.

If $l = p$, let $D_x := D^L_{\mathrm{st}}(\rho_x)$; it is equipped with a filtration $\mathcal{F}^\bullet_x$. We consider the set $(\operatorname{Spf} A_x)(E'[\varepsilon])$. Forgetting the framing on liftings is a formally smooth morphism of groupoids and makes the tangent space at x into a Lie G-torsor over the groupoid of unframed deformations. But since $E'[\varepsilon]$ is an artin local E-algebra, by [Bellovin 2016, Proposition 2.4] the category of (unframed) potentially semistable representations of Gal_K over $E'[\varepsilon]$ deforming ρ_x is equivalent to the subcategory of $G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}}(E'[\varepsilon])$ deforming $D^L_{\mathrm{st}}(\rho_x)$.

There is a natural morphism of groupoids

$$G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}} \to G\text{-}\mathrm{Mod}_{L/K,\varphi,N}$$

and therefore a commutative diagram:

$$\begin{tikzcd}
G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}}(E'[\varepsilon]) \arrow[r] \arrow[d] & G\text{-}\mathrm{Mod}_{L/K,\varphi,N}(E'[\varepsilon]) \arrow[d] \\
G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}}(E') \arrow[r] & G\text{-}\mathrm{Mod}_{L/K,\varphi,N}(E')
\end{tikzcd}$$

By Corollary 2.7.7, the fibres of

$$G\text{-}\mathrm{Mod}_{L/K,\varphi,N,\mathrm{Fil}}(E'[\varepsilon]) \to G\text{-}\mathrm{Mod}_{L/K,\varphi,N}(E'[\varepsilon])$$

over the filtered G-torsor D_x are torsors under $(\operatorname{ad} D_x/\mathcal{F}^0(\operatorname{ad} D_x))^{\mathrm{Gal}_{L/K}}$. Since $G\text{-}\mathrm{Mod}_{L/K,\varphi,N} \cong G\text{-}\mathrm{WD}_E(L/K)$ is equidimensional of dimension 0 and $x \in \operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ is a smooth point, we conclude that

$$\begin{aligned}
\dim A_x &= \dim \operatorname{Lie} G + \dim(\operatorname{ad} D_x/\mathcal{F}^0(\operatorname{ad} D_x))^{\mathrm{Gal}_{L/K}} \\
&= \dim G + \dim \operatorname{Res}_{E\otimes K/E} G/P_{\boldsymbol{v}}
\end{aligned}$$

as desired.

To prove that $R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$ is reduced and locally a complete intersection, we consider the fibre product $\operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l] \times_{G\text{-}\mathrm{WD}_E(L/K)} Y_{L/K,\varphi,\mathcal{N}}$. This is a G-torsor, hence smooth, over $\operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l]$, so it suffices to prove that this fibre product is reduced and locally a complete intersection. But by Proposition 3.3.1, the natural morphism $\operatorname{Spec} R^{\square,\tau,\boldsymbol{v}}_{\bar{\rho}}[1/l] \times_{G\text{-}\mathrm{WD}_E(L/K)} Y_{L/K,\varphi,\mathcal{N}} \to Y_{L/K,\varphi,\mathcal{N}}$ is formally smooth, so completed local rings at points of Spec $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l] \times_{G\text{-}\mathrm{WD}_E(L/K)} Y_{L/K,\varphi,\mathcal{N}}$ are power series rings over completed local rings of $Y_{L/K,\varphi,\mathcal{N}}$. Since the latter are reduced and complete intersections (by Corollary 2.4.5), the same holds for the former.

The corresponding statements for $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v},\psi}$ can be proved in the same way; we leave the details to the reader. □

The following is a generalisation of [Allen 2016a, Theorem D] (which treats the case that $l = p$ and $G = \mathrm{GL}_n$). We let x be a closed point of $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]$ with residue field E_x (a finite extension of E), and write $\rho_x : \mathrm{Gal}_K \to G(E_x)$ for the corresponding representation.

Corollary 3.3.4. *The point x is a formally smooth point of $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]$ if and only if*

$$H^0((\mathrm{ad}\,\mathrm{WD}(\rho_x))^*(1)) = 0.$$

Proof. Corollary 2.4.2 implies that the formally smooth points of $G\text{-}\mathrm{WD}_E(L/K)$ are precisely those points x for which $H^0((\mathrm{ad}\,D_x)^*(1))$. Thus, we need to show that $x \in \mathrm{Spec}\, R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]$ is formally smooth if and only if its image in $G\text{-}\mathrm{WD}_E(L/K)$ is formally smooth.

We have a morphism

$$\mathrm{Spec}\, R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]_x^\wedge \to G\text{-}\mathrm{WD}_E(L/K)_x^\wedge,$$

which is formally smooth by Proposition 3.3.1. But this implies that for any $\mathbb{Q}_p$-finite artin local ring B, the map

$$\mathrm{Spec}\, R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]_x^\wedge(B) \to G\text{-}\mathrm{WD}_E(L/K)_x^\wedge(B)$$

is surjective. Hence, Spec $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]_x^\wedge$ is formally smooth if and only if $G\text{-}\mathrm{WD}_E(L/K)_x^\wedge$ is formally smooth. □

Remark 3.3.5. If G is the L-group of a quasisplit reductive group over K, then it seems plausible that the condition of Corollary 3.3.4 could be equivalent to the condition that the (conjectural) L-packet of representations associated to the Frobenius semisimplification of $\mathrm{WD}(\rho_x)$ contains a generic element. In the case that $G = \mathrm{GL}_n$ (where the L-packets are singletons) and $\mathrm{WD}(\rho_x)$ is Frobenius semisimple, this is proved in [Allen 2016a, §1], and in the general case it is closely related to [Gross and Prasad 1992, Conjecture 2.6] (which relates genericity to poles at $s = 1$ of the adjoint L-function).

Remark 3.3.6. In the case that $l \neq p$, the equivalence between Galois representations and Weil–Deligne representations means that we can rewrite the condition in Corollary 3.3.4 as $H^0(\mathrm{Gal}_K, \mathrm{ad}\,\rho_x^*(1)) = 0$.

We can also consider the quotient $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v},N=0}$, corresponding to the union of the irreducible components of $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]$ for which the monodromy operator N vanishes identically (if $l = p$, this is the locus of potentially crystalline representations, and if $l \neq p$, it is the locus of potentially unramified representations).

Theorem 3.3.7. *Fix an inertial type τ, and if $l = p$ then fix a p-adic Hodge type $\boldsymbol{v}$. Assume that $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v},N=0} \neq 0$. Then $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v},N=0}[1/l]$ is regular, and is equidimensional of dimension*

$$\dim_E G + \delta_{l=p} \dim_E(\mathrm{Res}_{E\otimes K/E}\, G)/P_{\boldsymbol{v}}.$$

Similarly $R_{\bar\rho}^{\square,\tau,\boldsymbol{v},N=0,\psi}[1/l]$ *is regular and equidimensional of dimension*

$$\dim_E G^{\mathrm{der}} + \delta_{l=p} \dim_E(\mathrm{Res}_{E\otimes K/E}\, G)/P_{\boldsymbol{v}}.$$

Proof. This can be proved in exactly the same way as Theorem 3.3.2, replacing the use of the three term complex $\mathcal{C}^\bullet(D)$ considered in Proposition 2.2.1 with the two term complex

$$(\mathrm{ad}\, D_A)^{I_{L/K}} \xrightarrow{1-\mathrm{Ad}(\Phi)} (\mathrm{ad}\, D_A)^{I_{L/K}}$$

concentrated in degrees 0 and 1; see [Kisin 2008, Theorem 3.3.8] for more details in the case that $l = p$ and $G = \mathrm{GL}_n$. $\square$

3.4. *Components of deformation rings.* We now prove the following reassuring lemma, which shows that the components of universal deformation rings are invariant under $G(\mathcal{O})$-conjugacy. It is a generalisation of [Barnet-Lamb et al. 2014, Lemma 1.2.2], which treats the case $G = \mathrm{GL}_n$; the proof there is by an explicit homotopy, while we use the theory of reductive group schemes over $\mathcal{O}$ to construct less explicit homotopies.

Lemma 3.4.1. *Let $h \in G(\mathcal{O}')$ be an element which reduces to the identity modulo the maximal ideal, where $\mathcal{O}'$ is the ring of integers in a finite extension of E. Then conjugation by h induces a map* $\mathrm{Spec}(R_{\bar\rho}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}')[1/l] \to \mathrm{Spec}(R_{\bar\rho}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}')[1/l]$, *and it fixes each irreducible component.*

Before we prove it, we record a preliminary lemma on irreducible components of the generic fibre of $R_{\bar\rho}^{\square,\tau,\boldsymbol{v}}$:

Lemma 3.4.2. *Let $A := \mathcal{O}[\![X_1, \dots, X_n]\!]/I$ be the quotient of a power series ring. If $x, x' \in (\mathrm{Spf}\, A)^{\mathrm{rig}}$ lie on the same irreducible component, then they lie on the same irreducible component of* $\mathrm{Spec}\, A[1/l]$.

Proof. If $x = x'$ as points of $(\mathrm{Spf}\, A)^{\mathrm{rig}}$, then by [de Jong 1995, Lemma 7.1.9], $x = x'$ as points of $\mathrm{Spec}\, A[1/l]$. Thus, we may assume that $x \neq x'$. Let $A \to \tilde{A}$ denote the normalisation of A. Then by [Conrad 1999, Theorem 2.1.3], $(\mathrm{Spf}\, \tilde{A})^{\mathrm{rig}} \to (\mathrm{Spf}\, A)^{\mathrm{rig}}$ is a normalisation of the rigid space $(\mathrm{Spf}\, A)^{\mathrm{rig}}$, and x, x' lift to points $\tilde{x}, \tilde{x}' \in (\mathrm{Spf}\, \tilde{A})^{\mathrm{rig}}$ on the same connected component. By [de Jong 1995, Lemma 7.1.9], $\tilde{x}$ and $\tilde{x}'$ correspond to distinct closed points of $\mathrm{Spec}\, \tilde{A}[1/l]$.

If $\tilde{x}$ and $\tilde{x}'$ lie on distinct connected components of $\mathrm{Spec}\, \tilde{A}[1/l]$, there are idempotents $e_x, e_{x'} \in \tilde{A}[1/l]$ such that e_x is 1 at $\tilde{x}$ and 0 at $\tilde{x}'$ and $e_{x'}$ is 1 at $\tilde{x}'$ and 0 at $\tilde{x}$. Again by [loc. cit., Lemma 7.1.9], the natural map $(\mathrm{Spf}\, \tilde{A})^{\mathrm{rig}} \to \mathrm{Spec}\, \tilde{A}[1/l]$ induces isomorphisms on residue fields of closed points. It follows that the pullbacks of e_x and $e_{x'}$ to $(\mathrm{Spf}\, \tilde{A})^{\mathrm{rig}}$ are again idempotents (in the global sections of the structure sheaf of $(\mathrm{Spf}\, \tilde{A})^{\mathrm{rig}}$) such that e_x is 1 at $\tilde{x}$ and 0 at $\tilde{x}'$ and $e_{x'}$ is 1 at $\tilde{x}'$ and 0 at $\tilde{x}$. But this would contradict the fact that $\tilde{x}$ and $\tilde{x}'$ lie on the same connected component of $(\mathrm{Spf}\, \tilde{A})^{\mathrm{rig}}$, so they must actually lie on the same connected component of $\mathrm{Spec}\, \tilde{A}[1/l]$. This in turn implies that they lie on the same irreducible component of $\mathrm{Spec}\, A[1/l]$. $\square$

Proof of Lemma 3.4.1. Let $R_{\bar\rho}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'' \to \mathcal{O}''$ be a homomorphism corresponding to a lift $\rho : \mathrm{Gal}_K \to G(\mathcal{O}'')$, where $\mathcal{O}''$ is the ring of integers in a finite extension of E and contains $\mathcal{O}'$. We continue to write

h for the image of h in $G(\mathcal{O}'')$. There is a finite surjective morphism

$$\operatorname{Spec}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'')[1/l] \to \operatorname{Spec}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}')[1/l],$$

so to show that conjugation by h preserves irreducible components of $\operatorname{Spec}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}')[1/l]$, it suffices to show that conjugation by h preserves irreducible components of $\operatorname{Spec}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'')[1/l]$. Moreover, by Lemma 3.4.2, it suffices to work with the rigid analytic generic fibre $\operatorname{Spf}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'')^{\mathrm{rig}}$ of $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}''$.

After possibly extending $\mathcal{O}''$, we may assume that G splits over $\mathcal{O}''$. Since h is residually the identity element of G, it is a point of G°. After possibly further increasing $\mathcal{O}''$, there is some Borel subgroup $B_{\mathcal{O}''[1/l]} \subset G^\circ_{\mathcal{O}''[1/l]}$ containing the image of h; it extends to a Borel subgroup $B \subset G^\circ_{\mathcal{O}''}$ which contains h. Since $\mathcal{O}''$ is local, by [Conrad 2014, Proposition 5.2.3] there is a cocharacter $\lambda : (\mathbb{G}_m)_{\mathcal{O}''} \to G^\circ_{\mathcal{O}''}$ such that $B = P_{G^\circ}(\lambda) = U_{G^\circ}(\lambda) \rtimes Z_{G^\circ}(\lambda)$. Write h_z for the projection of h to $Z_{G^\circ}(\lambda)$ and h_u for the projection to $U_{G^\circ}(\lambda)$. Since this decomposition is unique, both h_z and h_u reduce to the identity modulo ϖ (where ϖ is a uniformiser of $\mathcal{O}''$).

Since $Z_{G^\circ}(\lambda)$ is a split torus, there is a map $z_t : (\mathbb{G}_m)_{\mathcal{O}''} \to G^\circ_{\mathcal{O}''}$ which specialises to both h_z and the identity. After analytifying this map, h_z and the identity lie in the same residue disk. Choosing coordinates on this residue disk, and rescaling them if necessary, we obtain a Galois representation $\tilde{\rho} : \mathrm{Gal}_K \to G(\mathcal{O}''[\![T]\!])$ by considering the conjugation map $z_t \rho z_t^{-1} : \mathrm{Gal}_K \to G(\mathcal{O}''[T])$. This induces a homomorphism $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'' \to \mathcal{O}''[\![T]\!]$, which in turn induces a morphism of rigid spaces $\operatorname{Spf}(\mathcal{O}''[\![T]\!])^{\mathrm{rig}} \to \operatorname{Spf}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'')^{\mathrm{rig}}$. Since the source is irreducible and its image contains points corresponding to both ρ and $h_z \rho h_z^{-1}$, they lie on the same irreducible component of $\operatorname{Spf}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'')^{\mathrm{rig}}$.

Thus, we may assume that $h \in U_{G^\circ}(\lambda)$. By definition, if A is an $\mathcal{O}'$-algebra,

$$U_{G^\circ}(\lambda)(A) = \{g \in G^\circ(A) \mid \lim_{t\to 0} \lambda(t) g \lambda(t)^{-1} = 1\},$$

so conjugating h by λ induces a map $u_t : \mathbb{A}^1_{\mathcal{O}''} \to G_{\mathcal{O}''}$ with $u_1 = h$ and $u_0 = 1$. We therefore obtain a Galois representation $\tilde{\rho}' : \mathrm{Gal}_K \to G(\mathcal{O}''\langle T\rangle)$ by l-adically completing the map $u_t \rho u_t^{-1} : \mathrm{Gal}_K \to G(\mathcal{O}''[T])$. Since u_t is the identity modulo ϖ, $\tilde{\rho}'$ in fact lands in $G(\mathcal{O}''\langle \varpi T\rangle)$, and therefore in $G(\mathcal{O}''[\![\varpi T]\!])$. This induces a map $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'' \to \mathcal{O}''[\![\varpi T]\!]$, and therefore a morphism of rigid spaces $\operatorname{Spf}(\mathcal{O}''[\![\varpi T]\!])^{\mathrm{rig}} \to \operatorname{Spf}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'')^{\mathrm{rig}}$. Since the source is irreducible and its image contains points corresponding to both ρ and $h_u \rho h_u^{-1}$, they lie on the same irreducible component of $\operatorname{Spf}(R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}} \otimes_{\mathcal{O}} \mathcal{O}'')^{\mathrm{rig}}$, as required. □

3.5. ***Tensor products of components, and base change.*** By a "component for $\bar{\rho}$" we mean a choice of τ and $\boldsymbol{v}$ (in the case $l = p$) such that $R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l] \neq 0$, and a choice of an irreducible component of $\operatorname{Spec} R_{\bar{\rho}}^{\square,\tau,\boldsymbol{v}}[1/l]$.

Let $\bar{r} : \mathrm{Gal}_K \to \mathrm{GL}_n(\mathbb{F})$ and $\bar{s} : \mathrm{Gal}_K \to \mathrm{GL}_m(\mathbb{F})$ be representations, let C be a component for $\bar{r}$ and let D be a component for $\bar{s}$. Let K'/K be a finite extension. The following lemma will be useful in Section 5.

Lemma 3.5.1. *There is a unique component $C \otimes D$ for $\bar{r} \otimes \bar{s}$ with the property that, if $r : \mathrm{Gal}_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_l)$ and $s : \mathrm{Gal}_K \to \mathrm{GL}_m(\overline{\mathbb{Q}}_l)$ correspond to closed points of C and D respectively, then $r \otimes s$ corresponds* *to a closed point of $C \otimes D$. Similarly, there is a unique component $C|_{K'}$ for $\bar{r}|_{\mathrm{Gal}_{K'}}$ such that for all r, $r|_{\mathrm{Gal}_{K'}}$ corresponds to a closed point of $C|_{K'}$.*

Proof. If a point of $\operatorname{Spec} R_{\bar{r}}^{\square,\tau,\boldsymbol{v}}[1/l]$ or a point of $\operatorname{Spec} R_{\bar{r}\otimes\bar{s}}^{\square,\tau,\boldsymbol{v}}[1/l]$ is smooth, then it lies on a unique irreducible component. Then the first part follows as in the proof of Theorem 3.3.2, replacing the appeal to Corollary 2.4.5 with one to Theorem 2.3.8, applied to the tensor product map

$$\mathrm{GL}_n \times \mathrm{GL}_m \to \mathrm{GL}_{nm}\,.$$

The second part follows from Theorem 3.3.2 (more precisely, from the existence of very smooth points on each irreducible component). $\square$

In the setting of the previous lemma, we will sometimes say that the component $C \otimes D$ is the tensor product of the components C and D, and that $C|_{K'}$ is the base change to K' of the component C.

4. Global deformation rings

4.1. ***A result of Balaji.*** In this section we recall one of the main results of [Balaji 2013], which we will then combine with the results of Section 3 to prove Proposition 4.2.6, which gives a lower bound for the dimension of certain global deformation rings. In [loc. cit., §4.2] the group G is assumed to be connected, but this is unnecessary. Indeed, the assumption is only made in order to use the results of [Tilouine 1996, §5], where it is also assumed that G is connected; however, this assumption is never used in any of the arguments of [loc. cit., §5], which apply unchanged to general G. Accordingly, we will freely use the results of [Balaji 2013, §4.2] without assuming that G is connected. We assume in this section that E is taken large enough that G_E is quasisplit.

Let F be a number field, and let S be a finite set of places of F containing all of the places dividing $l\infty$. We work in the fixed determinant setting, and accordingly we fix homomorphisms $\bar{\rho} : \mathrm{Gal}_{F,S} \to G(\mathbb{F})$ and $\psi : \mathrm{Gal}_{F,S} \to G^{\mathrm{ab}}(\mathcal{O})$ such that $\mathrm{ab} \circ \bar{\rho} = \bar{\psi}$.

Write $R_{F,S}^{\square,\psi} \in \mathrm{CNL}_{\mathcal{O}}$ for the universal fixed determinant framed deformation $\mathcal{O}$-algebra of $\bar{\rho}$. Let $\Sigma \subset S$ be a subset containing all of the places lying over l. For each $v \in \Sigma$, we let $R_v^{\square,\psi}$ denote the universal fixed determinant framed deformation $\mathcal{O}$-algebra of $\bar{\rho}|_{\mathrm{Gal}_{F_v}}$, and we set

$$R_{\Sigma}^{\square,\psi} := \widehat{\bigotimes}_{v\in\Sigma,\mathcal{O}} R_v^{\square,\psi}.$$

The following result is a special case of [Balaji 2013, Proposition 4.2.5].

Proposition 4.1.1. *Suppose that $H^0(\mathrm{Gal}_{F,S}, (\mathfrak{g}_{\mathbb{F}}^0)^*(1)) = 0$, and let*

$$s := (|\Sigma| - 1)\dim_{\mathbb{F}} \mathfrak{g}_{\mathbb{F}}^0 + \sum_{v\,|\,\infty,\, v\notin\Sigma} \dim_{\mathbb{F}} H^0(\mathrm{Gal}_{F_v}, \mathfrak{g}_{\mathbb{F}}^0).$$

Then for some $r \geq 0$ there is a presentation

$$R_{F,S}^{\square,\psi} \xrightarrow{\sim} R_{\Sigma}^{\square,\psi}[\![x_1, \ldots, x_r]\!]/(f_1, \ldots, f_{r+s}).$$

4.2. ***Global deformation rings of fixed type.*** We now combine our local results with Proposition 4.1.1 to prove a lower bound for the Krull dimension of a global deformation ring, following Balaji. This lower bound will only be nontrivial in the following setting.

Definition 4.2.1. If $l > 2$ then we say that $\bar{\rho}$ is *discrete series and odd* if F is totally real, and if for all places $v \mid \infty$ of F we have $\dim_{\mathbb{F}} H^0(\mathrm{Gal}_{F_v}, \mathfrak{g}^0_{\mathbb{F}}) = \dim_E G - \dim_E B$, where B is a Borel subgroup of G.

Remark 4.2.2. Recall that we chose E to be large enough that G_E is quasisplit, so this definition makes sense. The condition that $\bar{\rho}$ is discrete series and odd is needed to make the usual Taylor–Wiles method work; see the introduction to [Clozel et al. 2008]. If G is the L-group of a simply connected group then one can check that this condition is equivalent to F being totally real and $\bar{\rho}$ being odd in the sense of [Gross 2007] (cf. [Balaji 2013, Lemma 4.3.1]). We use the term “discrete series” because the (conjectural) Galois representations associated to tempered automorphic representations which are discrete series at infinite places are expected to satisfy this property; see Section 5 for an example of this, and [Gross 2007] for a more general discussion.

Definition 4.2.3. We say that a p-adic Hodge type $\boldsymbol{v}$ is *regular* if the conjugacy class $P_{\boldsymbol{v}}$ consists of parabolic subgroups of $\mathrm{Res}_{E\otimes K/E}\, G$ whose connected components are Borel subgroups of $(\mathrm{Res}_{E\otimes K/E}\, G)^\circ$.

Remark 4.2.4. If $G = \mathrm{GL}_n$ then Definition 4.2.3 is equivalent to the usual definition, that for each embedding $K \hookrightarrow E$ the Hodge–Tate weights are pairwise distinct.

Remark 4.2.5. If E'/E is a field extension, then

$$(\mathrm{Res}_{E\otimes K/E}\, G)_{E'} \cong \mathrm{Res}_{E'\otimes K/E'}\, G.$$

Furthermore, the formation of $P_{\mathrm{Res}_{E\otimes K/E} G}(\lambda)$ is compatible with extension of scalars from E to E'. Thus, if $\boldsymbol{v}$ is regular after extending scalars, it was regular over E (and $\mathrm{Res}_{E\otimes K/E}\, G$ is automatically quasisplit).

Write S^∞ for the set of finite places in S. For each place $v \in S^\infty$, we fix an inertial type τ_v, and if $v \mid l$ then we fix a Hodge type $\boldsymbol{v}_v$. If $v \nmid l$ (resp. if $v \mid l$), we let $\bar{R}_v$ be a quotient of the corresponding fixed determinant framed deformation ring $R^{\square,\tau_v,\psi}_{\bar{\rho}|_{\mathrm{Gal}_{F_v}}}$ (resp. $R^{\square,\tau_v,\boldsymbol{v}_v,\psi}_{\bar{\rho}|_{\mathrm{Gal}_{F_v}}}$) corresponding to a nonempty union of irreducible components of the generic fibre. Set

$$R^{\square,\mathrm{univ}} := R^{\square,\psi}_{F,S} \otimes_{R^{\square,\psi}_{\Sigma},\mathcal{O}} \widehat{\bigotimes}_{v\in S^\infty} \bar{R}_v;$$

this is nonzero, because we are assuming that each $\bar{R}_v$ is nonzero.

Assume that $H^0(\mathrm{Gal}_{F,S}, \mathfrak{g}_{\mathbb{F}}) = \mathfrak{z}_{\mathbb{F}}$, so that $\bar{\rho}$ admits a universal fixed determinant deformation $\mathcal{O}$-algebra $R^{\psi}_{F,S} \in \mathrm{CNL}_{\mathcal{O}}$, and write R^{univ} for the quotient of $R_{F,S}$ corresponding to $R^{\square,\mathrm{univ}}$ (as in the discussion preceding [Barnet-Lamb et al. 2014, Lemma 1.3.3], this quotient exists by Lemma 3.4.1). In the case that we fix potentially crystalline types at the places $v \mid l$, and do not fix types at places away from l, the following result is [Balaji 2013, Theorem 4.3.2]; the general case follows from the same arguments as those of Balaji, given the input of our local results.

Proposition 4.2.6. *Assume that $l > 2$, that $\bar{\rho}$ is a discrete series and odd (so that in particular F is totally real), and that $H^0(\mathrm{Gal}_{F,S}, (\mathfrak{g}^0_{\mathbb{F}})^*(1)) = 0$. Maintain our assumption that the local deformation rings $\bar{R}_v$ are nonzero.*

Suppose that for each place $v \mid l$ the Hodge type $\boldsymbol{v}_v$ is regular. Then R^{univ} has Krull dimension at least 1.

Proof. By Proposition 4.1.1 (taking $\Sigma = S^\infty$) we see that for some $r \geq \dim_{\mathbb{F}} \mathfrak{g}^0_{\mathbb{F}}$ we have a presentation

$$R^{\square,\mathrm{univ}} \xrightarrow{\sim} \Big(\widehat{\bigotimes}_{v\in S^\infty} \bar{R}_v\Big)[\![x_1, \dots, x_r]\!]/(f_1, \dots, f_{r+s}),$$

where

$$s = (|S^\infty| - 1)\dim_{\mathbb{F}} \mathfrak{g}^0_{\mathbb{F}} + \sum_{v \mid \infty} \dim_{\mathbb{F}} H^0(\mathrm{Gal}_{F_v}, \mathfrak{g}^0_{\mathbb{F}}).$$

Since $R^{\square,\mathrm{univ}}$ is formally smooth over R^{univ} of relative dimension $\dim_{\mathbb{F}} \mathfrak{g}^0_{\mathbb{F}}$, it follows that the Krull dimension of R^{univ} is at least

$$\dim \widehat{\bigotimes}_{v\in S^\infty,\mathcal{O}} \bar{R}_v - |S^\infty| \dim_{\mathbb{F}} \mathfrak{g}^0_{\mathbb{F}} - \sum_{v \mid \infty} \dim_{\mathbb{F}} H^0(\mathrm{Gal}_{F_v}, \mathfrak{g}^0_{\mathbb{F}}),$$

which by Theorem 3.3.2, and our assumption that each Hodge type $\boldsymbol{v}_v$ is regular, is equal to

$$1 + \sum_{v \mid p} [F_v : \mathbb{Q}_p] \dim_E G/B - \sum_{v \mid \infty} \dim_{\mathbb{F}} H^0(\mathrm{Gal}_{F_v}, \mathfrak{g}^0_{\mathbb{F}}),$$

which in turn (by the assumption that $\bar{\rho}$ is discrete series and odd) equals 1, as required. □

5. Unitary groups

5.1. ***The group $\mathcal{G}_n$.*** Let F be a CM field with maximal totally real subfield F^+. In this section we generalise some results of [Barnet-Lamb et al. 2014] on the deformation theory of Galois representations associated to polarised representations of Gal_F, by allowing ramification at primes of F^+ which are inert or ramified in F. This allows us to make cleaner statements, and is also useful in applications; for example, in Theorem 5.2.2 we remove a "split ramification" condition in the proof of the weight part of Serre's conjecture for rank-2 unitary groups. Our results are also needed in [Calegari et al. 2018], where they are used to construct lifts with specified ramification at certain places of F^+ which are inert in F.

Recall from [Clozel et al. 2008] the reductive group $\mathcal{G}_n$ over $\mathbb{Z}$ given by the semidirect product of $\mathcal{G}^0_n = \mathrm{GL}_n \times \mathrm{GL}_1$ by the group $\{1, \jmath\}$, where

$$\jmath(g, a)\jmath^{-1} = (a(g^t)^{-1}, a).$$

We let $\nu : \mathcal{G}_n \to \mathrm{GL}_1$ be the character which sends (g, a) to a and sends $\jmath$ to -1. Our results in this section are for the most part a straightforward application of the results of the earlier sections to the particular case $G = \mathcal{G}_n$, but we need to begin by comparing our definitions to those of [loc. cit.]; we will follow the notation of that paper where possible.

Fix a place $v \mid \infty$. By [Clozel et al. 2008, Lemma 2.1.1], for any ring R there is a natural bijection between the set of homomorphisms $\rho : \mathrm{Gal}_{F^+} \to \mathcal{G}_n(R)$ inducing an isomorphism $\mathrm{Gal}_{F^+} / \mathrm{Gal}_F \xrightarrow{\sim} \mathcal{G}_n/\mathcal{G}_n^0$, and the set of triples $(r, \mu, \langle\cdot,\cdot\rangle)$ where $r : \mathrm{Gal}_F \to \mathrm{GL}_n(R)$, $\mu : \mathrm{Gal}_{F^+} \to R^\times$, and $\langle\cdot,\cdot\rangle : R^n \times R^n \to R$ is a perfect R-linear pairing such that $\langle x, y\rangle = -\mu(c_v)\langle y, x\rangle$, and $\langle r(\delta)x, r^{c_v}(\delta)y\rangle = \mu(\delta)\langle x, y\rangle$ for all $\delta \in \mathrm{Gal}_F$. We refer to such a triple as a μ-polarised representation of Gal_F, and we will sometimes denote it as a pair (r, μ), the pairing being implicit.

This bijection is given by setting $r := \rho|_{\mathrm{Gal}_F}$ (more precisely, the projection of $\rho|_{\mathrm{Gal}_F}$ to $\mathrm{GL}_n(R)$), $\mu := \nu \circ \rho$, and $\langle x, y\rangle = x^t A^{-1} y$, where $\rho(c_v) = (A, -\mu(c_v))\jmath$. If v is a finite place of F^+ which is inert or ramified in F, then we have an induced bijection between representations $\mathrm{Gal}_{F_v^+} \to \mathcal{G}_n(R)$ and μ-polarised representations $\mathrm{Gal}_{F_v} \to \mathrm{GL}_n(R)$.

There is an isomorphism $\mathrm{GL}_1 \to Z_{\mathcal{G}_n}$ given by $g \mapsto (g, g^2) \in \mathrm{GL}_1 \to \mathrm{GL}_1 \subset \mathrm{GL}_n \times \mathrm{GL}_1$, and we have $\mathcal{G}_n^{\mathrm{der}} = \mathrm{GL}_n \times 1$, and $\mathcal{G}_n^{\mathrm{ab}} = \mathrm{GL}_1 \times \{1, \jmath\}$. (It is easy to check by direct calculation that $\mathcal{G}_n^{\mathrm{der}} \subset \mathcal{G}_n^\circ$, and indeed $\mathcal{G}_n^{\mathrm{der}} \subset \mathrm{GL}_n \times 1$. Since $\mathrm{GL}_n^{\mathrm{der}} = \mathrm{SL}_n$, we have $\mathrm{SL}_n \times 1 \subset \mathcal{G}_n^{\mathrm{der}}$, and since $\jmath(1, a)\jmath^{-1}(1, a^{-1}) = (a, 1)$, we also have $\mathrm{GL}_1 \times 1 \subset \mathcal{G}_n^{\mathrm{der}}$, whence $\mathrm{GL}_n \times 1 \subset \mathcal{G}_n^{\mathrm{der}}$. Similarly, one checks easily that $Z_{\mathcal{G}_n} \subset \mathcal{G}_n^\circ$, so that $Z_{\mathcal{G}_n} \subset \mathrm{GL}_1 \times \mathrm{GL}_1$. If $(g, a) \in \mathrm{GL}_1 \times \mathrm{GL}_1$ then $\jmath(g, a)\jmath^{-1} = (ag^{-1}, a)$, so we see that $(g, a) \in Z_{\mathcal{G}_n}$ if and only if $a = g^2$, as required.)

We fix a prime $l > 2$ and a representation $\bar\rho : \mathrm{Gal}_{F^+} \to \mathcal{G}_n(\mathbb{F})$ with $\bar\rho^{-1}(\mathcal{G}_n^0(\mathbb{F})) = \mathrm{Gal}_F$. We fix a character $\mu : \mathrm{Gal}_{F^+} \to \mathcal{O}^\times$ with $\nu \circ \bar\rho = \bar\mu$. Write $\psi : \mathrm{Gal}_{F^+} \to \mathcal{G}_n^{\mathrm{ab}}(\mathcal{O})$ for the character taking $g \in \mathrm{Gal}_F$ to $(\mu(g), 1)$ and $g \in \mathrm{Gal}_{F^+} \setminus \mathrm{Gal}_F$ to $(-\mu(g), \jmath)$.

Note that if $R \in \mathrm{CNL}_{\mathcal{O}}$ then a deformation $\rho : \mathrm{Gal}_{F^+} \to \mathcal{G}_n(R)$ of $\bar\rho$ has $\mathrm{ab} \circ \rho = \psi$ if and only if $\nu \circ \rho = \mu$, in which case we say that it is μ-polarised. By [Allen 2016b, Proposition 2.2.3], restriction to Gal_F gives an equivalence between the μ-polarised (framed) deformations of $\bar\rho$ and the μ-polarised (framed) deformations r of $\bar r := \bar\rho|_{\mathrm{Gal}_F} : \mathrm{Gal}_F \to \mathrm{GL}_n(\mathbb{F})$, the latter by definition being those r which satisfy $r^c \cong r^\vee \mu$ (where we are writing c for c_v, as r^c is independent of the choice of $v \mid \infty$).

The same equivalence pertains to deformations of $\bar\rho|_{\mathrm{Gal}_{F_v^+}}$, where v is inert or ramified in F. On the other hand, if v splits as $\tilde v \tilde v^c$ in F, then restriction to $\mathrm{Gal}_{F_{\tilde v}}$ gives an equivalence between μ-polarised (framed) deformations of $\bar\rho|_{\mathrm{Gal}_{F_v^+}}$ and (framed) deformations of $\bar r|_{\mathrm{Gal}_{F_{\tilde v}}}$; thus at such places the deformation theory of representations valued in $\mathcal{G}_n$ is reduced to the case of GL_n. It is for this reason that [Clozel et al. 2008] and its sequels only permit ramification at places which split in F.

By [loc. cit., Lemma 2.1.3], $\bar\rho$ is discrete series and odd in the sense of Definition 4.2.1 if and only if for each place $v \mid \infty$ of F^+ with corresponding complex conjugation $c_v \in \mathrm{Gal}_{F^+}$ we have $\bar\mu(c_v) = -1$. This is by definition equivalent to the corresponding polarised representation $(\bar\rho|_{\mathrm{Gal}_F}, \bar\mu)$ being totally odd in the sense of [Barnet-Lamb et al. 2014, §2.1].

Let S be a finite set of places of F^+, including all the places where $\bar r$ or μ are ramified, all the infinite places, and all the places dividing l. The following is a generalisation of [loc. cit., Proposition 1.5.1] (which is the case that every finite place in S splits in F, and is actually proved in [Clozel et al. 2008]); note that the assumption that $\bar\rho|_{\mathrm{Gal}_{F(\zeta_l)}}$ is absolutely irreducible is missing from the statement of [Barnet-Lamb et al. 2014, Proposition 1.5.1], but should have been included there. Note also that this assumption implies that $\bar\rho$

admits a universal deformation ring; indeed, we have $H^0(\mathrm{Gal}_{F^+}, \mathfrak{g}_{\mathbb{F}}) = H^0(\mathrm{Gal}_{F^+}, \mathfrak{gl}_{n,\mathbb{F}} \times \mathfrak{gl}_{1,\mathbb{F}}) = \mathfrak{gl}_{1,\mathbb{F}}$ by Schur's lemma (note that $\mathrm{Gal}(F/F^+)$ acts by -1 on the scalar matrices in $\mathfrak{gl}_{n,\mathbb{F}}$).

Corollary 5.1.1. *Let $l > 2$ be prime, and let $\bar{\rho} : \mathrm{Gal}_{F^+} \to \mathcal{G}_n(\mathbb{F})$ be such that $\bar{\rho}|_{\mathrm{Gal}_{F(\zeta_l)}}$ is absolutely irreducible. Assume that $\bar{\rho}$ is discrete series and odd.*

Let μ be a de Rham lift of $\bar{\mu}$, and let S be a finite set of places of F^+ including all the places at which either $\bar{r}$ or μ is ramified, and all the places dividing $l\infty$. For each finite place $v \in S$, fix an inertial type τ_v, and if $v \mid l$, fix a regular Hodge type $\boldsymbol{v}_v$. Fix quotients of the corresponding local μ-polarised framed deformation rings which correspond to a (nonempty) union of irreducible components of the generic fibre.

Let R^{univ} be the universal deformation ring for μ-polarised deformations of $\bar{\rho}$ which are unramified outside S, and lie on the given union of irreducible components for each finite place $v \in S$. Then R^{univ} has Krull dimension at least 1.

Proof. By Proposition 4.2.6, we need only check that $H^0(\mathrm{Gal}_{F^+,S}, (\mathfrak{gl}_{n,\mathbb{F}})^*(1))$ vanishes, where $\mathfrak{gl}_{n,\mathbb{F}}$ is the Lie algebra of $\mathcal{G}_n^{\mathrm{der}}$. By inflation-restriction this group injects into

$$H^0(\mathrm{Gal}_{F(\zeta_l)}, (\mathfrak{gl}_{n,\mathbb{F}})^*(1))^{\mathrm{Gal}(F(\zeta_l)/F^+)} = H^0(\mathrm{Gal}_{F(\zeta_l)}, (\mathfrak{gl}_{n,\mathbb{F}}))^{\mathrm{Gal}(F(\zeta_l)/F^+)}.$$

Since $\bar{\rho}|_{\mathrm{Gal}_{F(\zeta_l)}}$ is absolutely irreducible by assumption, this group vanishes by Schur's lemma (noting again that $\mathrm{Gal}(F/F^+)$ acts by -1 on the scalar matrices in $\mathfrak{gl}_{n,\mathbb{F}}$). □

5.2. ***Existence of lifts and the weight part of Serre's conjecture.*** We now prove a strengthening of [Barnet-Lamb et al. 2013, Theorem A.4.1], removing the condition that the places at which our Galois representations are ramified are split in F. We refer the reader to [Barnet-Lamb et al. 2014] for any unfamiliar terminology; in particular, potential diagonalisability is defined in [loc. cit., §1.4], while adequacy and the notion of a polarised Galois representation being potentially diagonalisably automorphic are defined in [loc. cit., §2.1].

Theorem 5.2.1. *Let l be an odd prime not dividing n, and suppose that $\zeta_l \notin F$. Let $\bar{\rho} : \mathrm{Gal}_{F^+} \to \mathcal{G}_n(\mathbb{F})$ be such that $\bar{\rho}|_{\mathrm{Gal}_{F(\zeta_l)}}$ is absolutely irreducible. Assume that $\bar{\rho}$ is discrete series and odd. Let S be a finite set of places of F^+, including all places dividing $l\infty$.*

Let μ be a de Rham lift of $\bar{\mu}$, and let S be a finite set of places of F^+ including all the places at which either $\bar{r}$ or μ is ramified, and all the places dividing $l\infty$. For each finite place $v \in S$, fix an inertial type τ_v, and if $v \mid l$, fix a regular Hodge type $\boldsymbol{v}_v$. Fix quotients of the corresponding local μ-polarised framed deformation rings which correspond to an irreducible component of the generic fibre; if $v \mid l$, assume also that this component is potentially diagonalisable

Assume further that there is a finite extension of CM fields F'/F such that F' does not contain ζ_l, all finite places of $(F')^+$ above S split in F, and $\bar{\rho}(\mathrm{Gal}_{F'(\zeta_l)})$ is adequate, and assume that there exists a lift $\rho' : \mathrm{Gal}_{F^+,S} \to \mathcal{G}_n(\mathcal{O})$ of $\bar{\rho}|_{\mathrm{Gal}_{(F')^+,S}}$ with $\nu \circ \rho' = \mu|_{\mathrm{Gal}_{F^+,S}}$, with the further property that ρ' is potentially diagonalisably automorphic.

Then there is a lift

$$\rho : \mathrm{Gal}_{F^+,S} \to \mathcal{G}_n(\mathcal{O})$$

of $\bar{\rho}$ *such that*:

(1) $\nu \circ \rho = \mu$.

(2) *If* $v \in S$ *is a finite place, then* $\rho|_{G_{F_v^+}}$ *corresponds to a point on our chosen component of the local deformation ring.*

(3) $\rho|_{\mathrm{Gal}_{(F')^+,S}}$ *is potentially diagonalisably automorphic.*

Proof. Let R^{univ} be the universal deformation ring for μ-polarised deformations of $\bar{\rho}$ which are unramified outside S, and lie on the given irreducible component for each finite place $v \in S$. Then R^{univ} has Krull dimension at least 1 by Corollary 5.1.1. We claim that R^{univ} is a finite $\mathcal{O}$-algebra. Admitting this claim, we can choose a homomorphism $R^{\mathrm{univ}} \to E$, and let ρ be the corresponding representation. This satisfies properties (1) and (2) by construction.

Let $R^{\mathrm{univ}}_{F'}$ be the universal deformation ring for $\mu|_{G_{(F')^+,S}}$-polarised deformations of $\bar{r}|_{G_{F',S}}$ which lie on the base changes of our chosen components. By [Barnet-Lamb et al. 2014, Lemma 1.2.3(1)], R^{univ} is a finite $R^{\mathrm{univ}}_{F'}$-algebra, so in order to prove the claim it is enough to show that $R^{\mathrm{univ}}_{F'}$ is a finite $\mathcal{O}$-algebra.

By [Barnet-Lamb et al. 2013, Theorem A.4.1] (with F there taken to equal F'), there is a representation $\rho'' : G_{(F')^+,S} \to \mathcal{G}_n(\mathcal{O})$ corresponding to an $\mathcal{O}$-point of $R^{\mathrm{univ}}_{F'}$, which is furthermore potentially diagonalisably automorphic. Then $R^{\mathrm{univ}}_{F'}$ is a finite $\mathcal{O}$-algebra by [Barnet-Lamb et al. 2014, Theorem 2.3.2]. as required. Finally, property (3) holds by [loc. cit., Theorem 2.3.2] (applied to ρ'' and $\rho|_{G_{(F')^+,S}}$). □

We now apply this result to the weight part of Serre's conjecture for unitary groups. We restrict ourselves to the case $n = 2$, where the existing results in the literature are strongest; our results should also allow the removal of the hypothesis of "split ramification" from results in the literature for higher-rank unitary groups, such as the results of [Barnet-Lamb et al. 2018]. We recall that if $K/\mathbb{Q}_l$ is a finite extension, there is associated to any representation $\bar{\rho} : \mathrm{Gal}_K \to \mathrm{GL}_2(\mathbb{F})$ a set $W(\bar{\rho})$ of Serre weights. A definition of $W(\bar{\rho})$ was first given in [Buzzard et al. 2010] in the case that $K/\mathbb{Q}_l$ is unramified, and various generalisations and alternative definitions have subsequently been proposed. As a result of the main theorems of [Gee et al. 2015; Calegari et al. 2017], all of these definitions are equivalent; we refer the reader to the introductions to those papers for a discussion of the various definitions.

Suppose that F is an imaginary CM field with maximal totally real subfield F^+ such that F/F^+ is unramified at all finite places, that each place of F^+ above l splits in F, and that $[F^+ : \mathbb{Q}]$ is even. Then as in [Barnet-Lamb et al. 2013] we have a unitary group G/F^+ which is quasisplit at all finite places and compact at all infinite places. If $\bar{r} : \mathrm{Gal}_{F^+} \to \mathcal{G}_2(\bar{\mathbb{F}}_l)$ is irreducible, the notion of $\bar{r}$ being modular of a Serre weight is defined in [loc. cit., Definition 2.1.9]. This definition (implicitly) insists that $\bar{r}$ is only ramified at places which split in F, and we relax it as follows: we change the definition of a good compact open subgroup $U \subset G(\mathbb{A}^{\infty}_{F^+})$ in [loc. cit., Definition 2.1.5] to require only that at all places $v \mid l$ we have $U_v = G(\mathcal{O}_{F_v^+})$, and at all places $v \nmid l$ we have $U_v \subset G(\mathcal{O}_{F_v^+})$. (Consequently, we are now considering automorphic forms of arbitrary level away from l, whereas in [loc. cit.] the level is hyperspecial at all places which do not split in F.)

Having made this change, everything in [loc. cit., §2] goes through unchanged, except that all mentions of "split ramification" can be deleted. The following theorem strengthens [Gee et al. 2014, Theorem A], removing a hypothesis on the ramification away from l (and also a hypothesis on the ramification at l, although that could already have been removed thanks to the results of [Gee et al. 2015]).

Theorem 5.2.2. *Let F be an imaginary CM field with maximal totally real subfield F^+, and suppose that F/F^+ is unramified at all finite places, that each place of F^+ above l splits in F, and that $[F^+ : \mathbb{Q}]$ is even. Suppose that l is odd, that $\bar{r} : G_{F^+} \to \mathcal{G}_2(\overline{\mathbb{F}}_l)$ is irreducible and modular, and that $\bar{r}(G_{F(\zeta_l)})$ is adequate.*

Then the set of Serre weights for which $\bar{r}$ is modular is exactly the set of weights given by the sets $W(\bar{r}|_{G_{F_v}})$, $v \mid l$.

Proof. We begin by observing that the proof of [Barnet-Lamb et al. 2013, Theorem 5.1.3] goes through in our more general context (that is, without assuming "split ramification"). Indeed, we have already observed that the results of [loc. cit., §2] are valid in our context, and chasing back through the references, we see that the only change that needs to be made is to relax the hypotheses in [loc. cit., Theorem 3.1.3] by no longer requiring that the places $v \in S$, $v \nmid l$, split in F. This follows by replacing the citation of [loc. cit., Theorem A.4.1] in the proof of [loc. cit., Theorem 3.1.3] with a reference to Theorem 5.2.1 above (after making a further extension of F' to arrange that all of the places of $(F')^+$ lying over S split in F').

This shows that $\bar{r}$ is modular of every weight given by the $W(\bar{r}|_{G_{F_v}})$, $v \mid l$. For the converse, observe that [loc. cit., Corollary 4.1.8] also holds in our context (again, since the results of [loc. cit., §2] go through); the result then follows immediately from [Gee et al. 2015, Theorem 6.1.8]. □

Remark 5.2.3. It is presumably possible to prove in the same way a further strengthening of Theorem 5.2.2 where we allow our unitary group to be ramified at some finite places (and thus allow $[F^+ : \mathbb{Q}]$ to be odd, and F/F^+ to be ramified at some finite places), but to do so would involve a lengthier discussion of automorphic representations on unitary groups, which would take us too far afield.

Remark 5.2.4. We have assumed that the places of F^+ above l split in F, because the weight part of Serre's conjecture has not been considered in the literature for unitary groups which do not split above l (although if l is unramified in F, and we are in the generic semisimple case, such a conjecture is a special case of the conjectures of [Gee et al. 2018]). However, it seems likely that it is possible to formulate and prove a generalisation of Theorem 5.2.2 which removes this assumption, following the ideas of [Gee and Kisin 2014; Gee and Geraghty 2015] (that is, using the Breuil–Mézard conjecture for potentially Barsotti–Tate representations). Again, this would take us too far afield from the main concerns of this paper, so we do not pursue this; and in any case we understand that this will be carried out in forthcoming work of Koziol and Morra.

Acknowledgements

We would like to thank Matthew Emerton for emphasising the importance of Weil–Deligne representations to us, and for his comments on an earlier draft of this paper. We would also like to thank Jeremy Booher, George Boxer, Stefan Patrikis, and Jacques Tilouine for helpful conversations, and Brian Conrad, Mark Kisin and Daniel Le for their comments on an earlier draft. We would like to thank the referees for their careful reading of the paper and their many helpful comments.

References

[Allen 2016a] P. B. Allen, "Deformations of polarized automorphic Galois representations and adjoint Selmer groups", *Duke Math. J.* **165**:13 (2016), 2407–2460. MR Zbl

[Allen 2016b] P. B. Allen, "On automorphic points in polarized deformation rings", 2016. To appear in *Amer. J. Math.* arXiv

[Balaji 2013] S. Balaji, *G-valued potentially semi-stable deformation rings*, Ph.D. thesis, University of Chicago, 2013, available at https://search.proquest.com/docview/1346024943.

[Barnet-Lamb et al. 2013] T. Barnet-Lamb, T. Gee, and D. Geraghty, "Serre weights for rank two unitary groups", *Math. Ann.* **356**:4 (2013), 1551–1598. MR Zbl

[Barnet-Lamb et al. 2014] T. Barnet-Lamb, T. Gee, D. Geraghty, and R. Taylor, "Potential automorphy and change of weight", *Ann. of Math.* (2) **179**:2 (2014), 501–609. MR Zbl

[Barnet-Lamb et al. 2018] T. Barnet-Lamb, T. Gee, and D. Geraghty, "Serre weights for $U(n)$", *J. Reine Angew. Math.* **735** (2018), 199–224. MR Zbl

[Bellovin 2016] R. Bellovin, "Generic smoothness for G-valued potentially semi-stable deformation rings", *Ann. Inst. Fourier* (*Grenoble*) **66**:6 (2016), 2565–2620. MR Zbl

[Böckle 1999] G. Böckle, "A local-to-global principle for deformations of Galois representations", *J. Reine Angew. Math.* **509** (1999), 199–236. MR Zbl

[Booher 2019a] J. Booher, "Minimally ramified deformations when $l \neq p$", *Compos. Math.* **155**:1 (2019), 1–37. MR Zbl

[Booher 2019b] J. Booher, "Producing geometric deformations of orthogonal and symplectic Galois representations", *J. Number Theory* **195** (2019), 115–158. Zbl

[Booher and Patrikis 2017] J. Booher and S. Patrikis, "G-valued Galois deformation rings when $l \neq p$", 2017. To appear in *Math. Res. Lett.* arXiv

[Bourbaki 2005] N. Bourbaki, *Lie groups and Lie algebras, chapters 7–9*, Springer, 2005. MR Zbl

[Breuil and Schneider 2007] C. Breuil and P. Schneider, "First steps towards p-adic Langlands functoriality", *J. Reine Angew. Math.* **610** (2007), 149–180. MR Zbl

[Breuil et al. 2001] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, "On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises", *J. Amer. Math. Soc.* **14**:4 (2001), 843–939. MR Zbl

[Buzzard and Gee 2014] K. Buzzard and T. Gee, "The conjectural connections between automorphic representations and Galois representations", pp. 135–187 in *Automorphic forms and Galois representations, I* (Durham, 2011), edited by F. Diamond et al., London Math. Soc. Lecture Note Ser. **414**, Cambridge Univ. Press, 2014. MR Zbl

[Buzzard et al. 2010] K. Buzzard, F. Diamond, and F. Jarvis, "On Serre's conjecture for mod l Galois representations over totally real fields", *Duke Math. J.* **155**:1 (2010), 105–161. MR Zbl

[Calegari et al. 2017] F. Calegari, M. Emerton, T. Gee, and L. Mavrides, "Explicit Serre weights for two-dimensional Galois representations", *Compos. Math.* **153**:9 (2017), 1893–1907. MR Zbl

[Calegari et al. 2018] F. Calegari, M. Emerton, and T. Gee, "Globally realizable components of local deformation rings", preprint, 2018. arXiv

[Clozel et al. 2008] L. Clozel, M. Harris, and R. Taylor, "Automorphy for some l-adic lifts of automorphic mod l Galois representations", *Publ. Math. Inst. Hautes Études Sci.* **108** (2008), 1–181. MR Zbl

[Conrad 1999] B. Conrad, "Irreducible components of rigid spaces", *Ann. Inst. Fourier* (*Grenoble*) **49**:2 (1999), 473–541. MR Zbl

[Conrad 2014] B. Conrad, "Reductive group schemes", pp. 93–444 in *Autour des schémas en groupes, I*, Panor. Synthèses **42/43**, Soc. Math. France, Paris, 2014. MR Zbl

[Conrad et al. 2010] B. Conrad, O. Gabber, and G. Prasad, *Pseudo-reductive groups*, New Math. Monographs **17**, Cambridge Univ. Press, 2010. MR Zbl

[Deligne and Milne 1982] P. Deligne and J. S. Milne, "Tannakian categories", pp. ii+414 in *Hodge cycles, motives, and Shimura varieties*, Lecture Notes in Math. **900**, Springer, 1982. MR Zbl

[EGA IV_1 1964] A. Grothendieck and J. A. Dieudonné, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, I", *Inst. Hautes Études Sci. Publ. Math.* **20** (1964), 5–259. MR Zbl

[Emerton and Gee 2017] M. Emerton and T. Gee, "Dimension theory and components of algebraic stacks", preprint, 2017. arXiv

[Fontaine 1994] J.-M. Fontaine, "Représentations l-adiques potentiellement semi-stables", pp. 321–347 in *Périodes p-adiques* (Bures-sur-Yvette, French, 1988), Astérisque **223**, Soc. Math. France, Paris, 1994. MR Zbl

[Gee 2011] T. Gee, "Automorphic lifts of prescribed types", *Math. Ann.* **350**:1 (2011), 107–144. MR Zbl

[Gee and Geraghty 2015] T. Gee and D. Geraghty, "The Breuil–Mézard conjecture for quaternion algebras", *Ann. Inst. Fourier* (*Grenoble*) **65**:4 (2015), 1557–1575. MR Zbl

[Gee and Kisin 2014] T. Gee and M. Kisin, "The Breuil–Mézard conjecture for potentially Barsotti–Tate representations", *Forum Math. Pi* **2** (2014), art. id. e1. MR Zbl

[Gee et al. 2014] T. Gee, T. Liu, and D. Savitt, "The Buzzard–Diamond–Jarvis conjecture for unitary groups", *J. Amer. Math. Soc.* **27**:2 (2014), 389–435. MR Zbl

[Gee et al. 2015] T. Gee, T. Liu, and D. Savitt, "The weight part of Serre's conjecture for GL(2)", *Forum Math. Pi* **3** (2015), art. id. e2. MR Zbl

[Gee et al. 2018] T. Gee, F. Herzig, and D. Savitt, "General Serre weight conjectures", *J. Eur. Math. Soc.* **20**:12 (2018), 2859–2949. MR Zbl

[Gross 2007] B. H. Gross, "Odd Galois representations", preprint, 2007, available at https://tinyurl.com/grossoddg.

[Gross and Prasad 1992] B. H. Gross and D. Prasad, "On the decomposition of a representation of SO_n when restricted to SO_{n-1}", *Canad. J. Math.* **44**:5 (1992), 974–1002. MR Zbl

[Gross and Reeder 2010] B. H. Gross and M. Reeder, "Arithmetic invariants of discrete Langlands parameters", *Duke Math. J.* **154**:3 (2010), 431–508. MR Zbl

[Jantzen 2004] J. C. Jantzen, "Nilpotent orbits in representation theory", pp. 1–211 in *Lie theory*, edited by J.-P. Anker and B. Orsted, Progr. Math. **228**, Birkhäuser, Boston, 2004. MR Zbl

[de Jong 1995] A. J. de Jong, "Crystalline Dieudonné module theory via formal and rigid geometry", *Inst. Hautes Études Sci. Publ. Math.* **82** (1995), 5–96. MR Zbl

[Khare and Wintenberger 2009] C. Khare and J.-P. Wintenberger, "On Serre's conjecture for 2-dimensional mod p representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$", *Ann. of Math.* (2) **169**:1 (2009), 229–253. MR Zbl

[Kisin 2007] M. Kisin, "Modularity of 2-dimensional Galois representations", pp. 191–230 in *Current developments in mathematics*, edited by D. Jerison et al., Int. Press, Somerville, MA, 2007. MR Zbl

[Kisin 2008] M. Kisin, "Potentially semi-stable deformation rings", *J. Amer. Math. Soc.* **21**:2 (2008), 513–546. MR Zbl

[Kisin 2009] M. Kisin, "Moduli of finite flat group schemes, and modularity", *Ann. of Math.* (2) **170**:3 (2009), 1085–1180. MR Zbl

[Matsumura 1989] H. Matsumura, *Commutative ring theory*, 2nd ed., Cambridge Studies in Adv. Math. **8**, Cambridge Univ. Press, 1989. MR Zbl

[Mazur 1989] B. Mazur, "Deforming Galois representations", pp. 385–437 in *Galois groups over* $\mathbb{Q}$ (Berkeley, 1987), edited by Y. Ihara et al., Math. Sci. Res. Inst. Publ. **16**, Springer, 1989. MR Zbl

[McNinch 2004] G. J. McNinch, "Nilpotent orbits over ground fields of good characteristic", *Math. Ann.* **329**:1 (2004), 49–85. MR Zbl

[Patrikis 2016] S. Patrikis, "Deformations of Galois representations and exceptional monodromy", *Invent. Math.* **205**:2 (2016), 269–336. MR Zbl

[Prasad and Yu 2002] G. Prasad and J.-K. Yu, "On finite group actions on reductive groups and buildings", *Invent. Math.* **147**:3 (2002), 545–560. MR Zbl

[Ramakrishna 2002] R. Ramakrishna, "Deforming Galois representations and the conjectures of Serre and Fontaine–Mazur", *Ann. of Math.* (2) **156**:1 (2002), 115–154. MR Zbl

[Saavedra Rivano 1972] N. Saavedra Rivano, *Catégories Tannakiennes*, Lecture Notes in Math. **265**, Springer, 1972. MR Zbl

[SGA 3_{II} 1970] M. Demazure and A. Grothendieck, *Schémas en groupes, Tome II: Groupes de type multiplicatif, et structure des schémas en groupes généraux, Exposés VIII–XVIII* (Séminaire de Géométrie Algébrique du Bois Marie 1962–1964), Lecture Notes in Math. **152**, Springer, 1970. MR Zbl

[Stacks 2005–] Stacks project authors, "The Stacks project", electronic reference, 2005–, available at https://tinyurl.com/on7qtt4.

[Taylor and Wiles 1995] R. Taylor and A. Wiles, "Ring-theoretic properties of certain Hecke algebras", *Ann. of Math.* (2) **141**:3 (1995), 553–572. MR Zbl

[Tilouine 1996] J. Tilouine, *Deformations of Galois representations and Hecke algebras*, Narosa, New Delhi, 1996. MR Zbl

[Wiles 1995] A. Wiles, "Modular elliptic curves and Fermat's last theorem", *Ann. of Math.* (2) **141**:3 (1995), 443–551. MR Zbl

Communicated by Samit Dasgupta
Received 2017-10-06 Revised 2018-11-08 Accepted 2018-12-24

r.bellovin@imperial.ac.uk *Department of Mathematics, Imperial College London, London, United Kingdom*

toby.gee@imperial.ac.uk *Department of Mathematics, Imperial College London, London, United Kingdom*

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.379

Functorial factorization of birational maps for qe schemes in characteristic 0

Dan Abramovich and Michael Temkin

We prove functorial weak factorization of projective birational morphisms of regular quasiexcellent schemes in characteristic 0 broadly based on the existing line of proof for varieties. From this general functorial statement we deduce factorization results for algebraic stacks, formal schemes, complex analytic germs, Berkovich analytic and rigid analytic spaces, answering a present need in nonarchimedean geometry. Techniques developed for this purpose include a method for functorial factorization of toric maps, variation of GIT quotients relative to general noetherian qe schemes, and a GAGA theorem for Stein compacts.

1. Introduction

1.1. The class of qe schemes (originally "quasiexcellent schemes") is the natural class of schemes on which problems around resolution of singularities are of interest. They can also be used as a bridge for studying the same type of problems in other geometric categories; see [Temkin 2008, Section 5]. In this paper we address the problem of functorial factorization of birational morphisms between regular qe schemes of characteristic 0 into blowings up and down of regular schemes along regular centers. We rely on general foundations developed in [Abramovich and Temkin 2017; 2018] and the approach for varieties of [Włodarczyk 2000; Abramovich et al. 2002]. As a consequence of both this generality of qe schemes

This research is supported by BSF grant 2010255.
MSC2010: primary 14E05; secondary 14A20, 14E15, 14L30, 32H04.
Keywords: birational geometry, blowing up, bimeromorphic maps.

and of functoriality, we are able to deduce factorization of birational or bimeromorphic morphisms in other geometric categories of interest.

Theorem 1.3.3 below answers a question in [Simons 2015] in characteristic 0 unconditionally, and in positive and mixed characteristics conditionally on resolution. It provides a complete proof of a key result, Proposition 3 in [Gillet and Soulé 2000], using the argument of Section 3.4 of that paper for general base ring Λ and not relying on results of J. Franke yet unpublished after 25 years. It justifies [Kontsevich and Soibelman 2006, Theorem 9].

1.2. ***Blowings up and weak factorizations.*** We start with a morphism of noetherian qe regular schemes $\phi : X_1 \to X_2$ given as the blowing up of a coherent sheaf of ideals I on the qe scheme X_2. In addition, we provide ϕ with a *boundary* (D_1, D_2), where each D_i is a normal crossings divisor in X_i and $D_1 = \phi^{-1} D_2$. Let $U = X_2 \smallsetminus (D_2 \cup V(I))$ be the maximal open subscheme of X_2 such that I is the unit ideal on U and the boundary is disjoint from U. The restriction of ϕ on U is the trivial blowing up (i.e., the blowing up of the empty center); in particular, we canonically have an isomorphism $\phi^{-1}U \to U$. We often keep the ideal I implicit in the notation, even though it determines ϕ (but see Section 2.1.8 for a construction in the reverse direction). The reader may wish to focus on the following two cases of interest: (i) $D_2 = \varnothing$; (ii) $V(I) \subseteq D_2$.

A *weak factorization* of a blowing up $\phi : X_1 \to X_2$ is a diagram of regular qe schemes

$$X_1 = V_0 \overset{\varphi_1}{\dashrightarrow} V_1 \overset{\varphi_2}{\dashrightarrow} \cdots \overset{\varphi_{l-1}}{\dashrightarrow} V_{l-1} \overset{\varphi_l}{\dashrightarrow} V_l = X_2$$

along with regular schemes Z_i for $i = 1, \dots, l$ and ideal sheaves J_i for $i = 1, \dots, l-1$ satisfying the following conditions:

(1) $\phi = \varphi_l \circ \varphi_{l-1} \circ \cdots \circ \varphi_2 \circ \varphi_1$.

(2) The maps $V_i \dashrightarrow X_2$ are morphisms; these maps as well as φ_i induce isomorphisms on U.

(3) For every $i = 1, \dots, l$ either $\varphi_i : V_{i-1} \dashrightarrow V_i$ or $\varphi_i^{-1} : V_i \dashrightarrow V_{i-1}$ is a morphism given as the blowing up of Z_i, which is respectively a subscheme of V_i or V_{i-1} disjoint from U.

(4) The inverse image $D_{V_i} \subset V_i$ of $D_2 \subset X_2$ is a normal crossings divisor and Z_i has normal crossings with D_{V_i}.

(5) For every $i = 1, \dots, l-1$, the morphism $V_i \to X_2$ is given as the blowing up of the corresponding coherent ideal sheaf J_i on X_2, which is the unit ideal on U.

To include $V_0 \to X_2$, we define $J_0 = I$. The ideals J_i are a convenient way to encode functoriality, especially when we later pass to other geometric categories.

These conditions are the same as (1)–(5) in [Abramovich et al. 2002, Theorem 0.3.1], except that here the centers of blowing up and ideal sheaves are specified. Condition (2) is formulated for convenience; it is a consequence of (3) and (5). Note that here, as in [loc. cit., Theorem 0.3.1], the centers are not assumed irreducible, in contrast with [loc. cit., Theorem 0.1.1]. With these conditions, the most basic form of our main theorem is as follows:

Theorem 1.2.1 (weak factorization). *Every birational blowing up $\phi : X_1 \to X_2$ of a noetherian qe regular $\mathbb{Q}$-scheme has a weak factorization $X_1 = V_0 \dashrightarrow V_1 \dashrightarrow \cdots \dashrightarrow V_{l-1} \dashrightarrow V_l = X_2$.*

The adjective "weak" serves to indicate that blowings up and down may alternate arbitrarily among the maps φ_i, as opposed to a *strong factorization*, where one has a sequence of blowings up followed by a sequence of blowings down. We note that at present strong factorization is not known even for toric threefolds.

Theorem 1.2.1 generalizes [Włodarczyk 2003, 0.0.1; Abramovich et al. 2002, Theorem 0.1.1], where the case of varieties is considered. But we wish to prove a more precise theorem.

1.3. ***Functorial weak factorization.*** The class of data (X_2, I, D_2), namely morphisms $\phi : X_1 \to X_2$ of noetherian qe regular schemes given as blowings up of ideals I, with divisor D_2 as in Section 1.2, can be made into *the regular surjective category of blowings up*, denoted by Bl, by defining arrows as follows:

Definition 1.3.1. An arrow from the blowing up $\phi' : X_1' = \mathrm{Bl}_{I'}(X_2') \to X_2'$ to $\phi : X_1 = \mathrm{Bl}_I(X_2) \to X_2$ is a regular and surjective morphism $g : X_2' \to X_2$ such that $g^*I = I'$ and $g^{-1}D_2 = D_2'$. In particular, g induces a canonical isomorphism $X_1' \to X_1 \times_{X_2} X_2'$ and D_1' is the preimage of D_1 under $X_1' \to X_1$.

Similarly, weak factorizations can be made into the regular surjective category of weak factorizations, denoted by Fact, by defining arrows as follows:

Definition 1.3.2. A morphism in Fact from a weak factorization

$$X_1' = V_0' \dashrightarrow V_1' \dashrightarrow \cdots \dashrightarrow V_{l-1}' \dashrightarrow V_l' = X_2'$$

of $\phi' : X_1' \to X_2'$, with centers Z_i' and ideals J_i', to a weak factorization

$$X_1 = V_0 \dashrightarrow V_1 \dashrightarrow \cdots \dashrightarrow V_{l-1} \dashrightarrow V_l = X_2$$

of $\phi : X_1 \to X_2$, with centers Z_i and ideals J_i, consists of a regular surjective morphism $g : X_2' \to X_2$ such that $g^*I = I'$, $g^*J_i = J_i'$, inducing $g_i : V_i' \to V_i$ such that $Z_i' = g_i^{-1}Z_i$ or $g_{i-1}^{-1}Z_i$ as appropriate. In particular $\varphi_i \circ g_{i-1} = g_i \circ \varphi_i$ and $g_i^{-1}D_{V_i} = D_{V_i'}$.

Note that given a factorization of ϕ, any morphism from a factorization of ϕ' is uniquely determined by $g : X_2' \to X_2$.

If we wish to restrict to schemes in a given characteristic p we denote the categories $\mathrm{Bl}(\mathrm{char} = p)$ and $\mathrm{Fact}(\mathrm{char} = p)$ respectively. If we wish to restrict the dimension we write $\mathrm{Bl}(\mathrm{char} = p, \dim \le d)$ and $\mathrm{Fact}(\mathrm{char} = p, \dim \le d)$.

There is an evident forgetful functor $\mathrm{Fact} \to \mathrm{Bl}$ taking a weak factorization $X_1 = V_0 \dashrightarrow V_1 \dashrightarrow \cdots \dashrightarrow V_{l-1} \dashrightarrow V_l = X_2$ to its composition $\phi : X_1 \to X_2$. The weak factorization theorem provides a section, when strong resolution of singularities holds:

Theorem 1.3.3. (1) *<u>Functorial weak factorization</u>: There is a functor*

$$\mathrm{Bl}(\mathrm{char} = 0) \to \mathrm{Fact}(\mathrm{char} = 0)$$

assigning to a blowing up $\phi : X_1 \to X_2$ *in characteristic* 0 *a weak factorization*

$$X_1 = V_0 \dashrightarrow V_1 \dashrightarrow \cdots \dashrightarrow V_{l-1} \dashrightarrow V_l = X_2,$$

so that the composite $\mathrm{Bl}(\mathrm{char}=0) \to \mathrm{Fact}(\mathrm{char}=0) \to \mathrm{Bl}(\mathrm{char}=0)$ *is the identity.*

(2) *Conditional factorization in positive and mixed characteristics*: *If functorial embedded resolution of singularities applies in characteristic* p *for schemes of dimension* $\leq d+1$, *then there is a functor*

$$\mathrm{Bl}(\mathrm{char}=p, \dim \leq d) \to \mathrm{Fact}(\mathrm{char}=p, \dim \leq d)$$

which is a section of $\mathrm{Fact}(\mathrm{char}=p, \dim \leq d) \to \mathrm{Bl}(\mathrm{char}=p, \dim \leq d)$. *If functorial embedded resolution of singularities applies over* $\mathbb{Z}$ *for schemes of dimension* $\leq d+1$, *then there is a functor*

$$\mathrm{Bl}(\dim \leq d) \to \mathrm{Fact}(\dim \leq d)$$

which is a section of $\mathrm{Fact}(\dim \leq d) \to \mathrm{Bl}(\dim \leq d)$.

This generalizes a theorem for *varieties* in characteristic 0 [Abramovich et al. 2002, Theorem 0.3.1 and Remark (3) thereafter; Włodarczyk 2006, Theorem 1.1, 2009, Theorem 0.0.1], where the factorization is only shown to be functorial for isomorphisms. The precise statements we need for part (2) are spelled out below as Hypothetical Statements 2.2.13 and 2.3.6.

Remark 1.3.4 (preservation of G-normality). Borisov and Libgober [2005, Definition 3.1] introduced G-normal divisors and in Theorem 3.8 of the same paper they showed that this condition can be preserved in the algorithm of [Abramovich et al. 2002]. The same holds true here, using the same argument of [Borisov and Libgober 2005, Theorem 3.8], by performing the sequence of blowings up associated to the barycentric subdivision on the schemes $W_{i\pm}^{\mathrm{res}}$ obtained in Section 5.4. Details are left to the interested reader.

1.4. ***Applications of functoriality.*** We need to justify the somewhat heavy functorial treatment. Of course functoriality may be useful if one wants to make sure the factorization is equivariant under group actions and separable field extensions; this has been of use already in the case of varieties. But it also serves as a tool to transport our factorization result to other geometric spaces.

Blowings up of regular objects is a concept which exists in categories other than schemes, for instance, in Artin stacks, qe formal schemes, complex semianalytic germs (see Appendix B), Berkovich k-analytic spaces, rigid k-analytic spaces. For brevity we denote the full subcategory of *qe noetherian* objects in any of these categories by $\mathfrak{Sp}$. Functoriality, as well as the generality of qe schemes, is crucial in proving the following:

Theorem 1.4.1 (factorization in other categories). *Any blowing up* $X_1 \to X_2$ *of either noetherian qe regular algebraic stacks, or regular objects of* $\mathfrak{Sp}$, *in characteristic* 0 *has a weak factorization* $X_1 = V_0 \dashrightarrow V_1 \dashrightarrow \cdots \dashrightarrow V_{l-1} \dashrightarrow V_l = X_2$. *The same holds in positive and mixed characteristics* (*when relevant*) *if functorial embedded resolution of singularities for qe schemes applies in positive and mixed characteristics.*

See Theorem 6.1.3 for the case of stacks and Theorem 6.4.5 for other categories, where functoriality is also shown; in other words Theorem 1.3.3 applies in each of the categories $\mathfrak{Sp}$. In addition, the argument deducing Theorem 6.1.3 from Theorem 1.3.3 is a formal one based on functoriality, so the same argument can be used to extend Theorem 6.4.5 to stacks in the categories of formal schemes, Berkovich spaces, etc., once an appropriate theory of stacks is constructed; see for instance [Simpson 1996; Noohi 2005; Ulirsch 2015; Yu 2018; Porta and Yu 2016].

1.5. ***The question of stronger functoriality.*** It is natural to replace the category Bl by the category Bl_r with the same objects but where arrows $g : X_2' \to X_2$ as in Definition 1.3.1 are not required to be surjective, only regular. In a similar way one can replace the category Fact by a category Fact_r. As explained in [Temkin 2008, §2.3.3] for resolution of singularities, removing the surjectivity assumption requires imposing an equivalence relation on factorizations in which two factorizations which differ by a step which is the blowing up of the unit ideal are considered equivalent. It is conceivable that the analogue of Theorem 1.3.3 may hold for $\mathrm{Fact}_r \to \mathrm{Bl}_r$.

1.6. ***Factorization of birational and bimeromorphic maps.*** Our results for projective morphisms imply results for birational and bimeromorphic maps. We start with the case of schemes. By a proper birational map $f : X_1 \dashrightarrow X_2$ of reduced schemes we mean an isomorphism $f_0 : U_1 \to U_2$ of dense open subschemes such that the closure $Y \subset X_1 \times X_2$ of the graph of f_0 is proper over each X_i. Assume that X_1 and X_2 are regular. The factorization problem for the birational map f reduces to factorization of the proper morphisms $Y^{\mathrm{res}} \to X_i$, where Y^{res} is a resolution of Y. Assume, now, that $f : X_1 \to X_2$ is a proper birational morphism. By a blow-up version of Chow's lemma (e.g., it follows from the flattening of Raynaud–Gruson) there exists a blowing up $Y = \mathrm{Bl}_I(X_2) \to X_2$ that factors through X_1. Then $Y = \mathrm{Bl}_{f^{-1}I}(X_1)$ and hence the resolution Y^{res}, which is a blowing up of Y, is also a blowing up of both X_i. Thus, factorization of f reduces to the factorization for blowings up, which was dealt with in Theorem 1.3.3.

Now, assume that $\mathfrak{Sp}$ is any geometric category. The definition of a proper bimeromorphic map $f : X_1 \to X_2$ is similar to the definition of a proper birational map with two addenda: in the case of stacks we require that the morphisms $Y \to X_i$ are representable, and in the case of analytic spaces or formal schemes we require that U is open in Y (in particular, $Y \to X_i$ are bimeromorphic). Then the general factorization problem immediately reduces to the case when f is a proper morphism. Furthermore, if objects of $\mathfrak{Sp}$ are compact and if Chow's lemma holds in $\mathfrak{Sp}$ then the problem reduces further to the case when f is a blowing up. For complex analytic spaces, Chow's lemma was proved by Hironaka [1975, Corollary 2]. It extends immediately to the complex analytic germs we consider in this paper, and these are indeed compact. Most probably, it also holds in all other categories $\mathfrak{Sp}$ we mentioned, but this does not seem to be worked out so far.

2. qe schemes and functoriality

2.1. ***Projective morphisms and functorial constructions.*** In our method, it will be important to describe certain morphisms we will obtain as the blowing up of a concrete ideal or an explicitly described projective morphism, since further constructions will depend on these data. Moreover, this should be done functorially with respect to surjective regular morphisms. In the current section we develop a few basic functorial constructions of this type.

There are a few ways to describe a projective morphism: using Proj, using ample sheaves, or using projective fibrations, but each approach involves choices. Neither description is "more natural" than the others, and we will have to switch between them. Similarly to [EGA II 1961] we choose the language of projective fibrations to be the basic one and we will show how other descriptions are canonically reduced to projective fibrations.

2.1.1. *Projective fibrations.* Let X be a scheme. For a coherent $\mathcal{O}_X$-module E consider the projective fibration $\mathbb{P}(E) = \mathbb{P}_X(E) := \operatorname{Proj}_X \operatorname{Sym}^\bullet(E)$ associated with E. It has a canonical twisting sheaf $\mathcal{O}_{\mathbb{P}(E)}(1)$, and $E \to \pi_*\mathcal{O}(1)$ is an isomorphism. This construction is functorial for all morphisms: if $\phi : X' \to X$ is any morphism and $E' = \phi^*E$ then $\mathbb{P}_{X'}(E') = X' \times_X \mathbb{P}_X(E)$, and $\mathcal{O}_{\mathbb{P}(E')}(1)$ is the pullback of $\mathcal{O}_{\mathbb{P}(E)}(1)$.

2.1.2. *Projective morphisms.* By the usual definition [EGA II 1961, 5.5.2], a morphism $f : Y \to X$ is projective if it factors through a closed immersion $i : Y \hookrightarrow \mathbb{P}_X(E)$ for a coherent $\mathcal{O}_X$-module E. In this paper, we will use the convention that by saying "f is projective" we fix E and i. In particular, Y acquires a canonical relatively very ample sheaf $\mathcal{O}_Y(1) = \mathcal{O}_{\mathbb{P}(E)}(1)|_Y$. The base change or *pullback* $f' : Y' = Y \times_X X' \to X'$ of f with respect to a morphism $\phi : X' \to X$ is projective via the embedding $Y' \hookrightarrow \mathbb{P}_{X'}(E')$, where $E' = \phi^*E$. We will use the notation $f' = \phi^*(f)$. Also, we say that f is *projectively the* identity over an open U of X if $E|_U = \mathcal{O}_U$ and $Y|_U = U$.

2.1.3. *Relation to* Proj. For a projective morphism $f : Y \to X$ we also obtain a canonical description of Y as a Proj. Namely, if $I_Y \subseteq \mathcal{O}_{\mathbb{P}(E)}$ denotes the ideal defining Y then $Y = \operatorname{Proj}_X A$, where $A^\bullet = \operatorname{Sym}^\bullet(E)/I_Y$ is a quasicoherent $\mathcal{O}_X$-algebra with coherent graded components, generated over $A^0 = \mathcal{O}_X$ by its degree-1 component A^1. Again this structure is functorial for all morphisms: if $\phi : X' \to X$ is any morphism and $A' = \phi^*A$ then $\operatorname{Proj}_{X'} A' = X' \times_X \operatorname{Proj}_X A$.

Conversely, if a graded $\mathcal{O}_X$-algebra $A^\bullet$ has coherent components and is generated over $A^0 = \mathcal{O}_X$ by A^1 then $\operatorname{Sym}^\bullet(A^1) \twoheadrightarrow A^\bullet$ and we obtain a closed immersion $i : \operatorname{Proj}_X A \hookrightarrow \mathbb{P}_X(A^1)$. Thus, $Y = \operatorname{Proj}_X A$ is projective over X, and the associated graded quasicoherent algebra is A itself. This construction is also functorial for all morphisms.

Remark 2.1.4. We note that the construction of a projective morphism from Proj is right inverse to the construction of Proj from a projective morphism, but they are not inverse: going from a projective morphism to Proj and back to a projective morphism usually changes the projective fibration.

Remark 2.1.5. In this paper we use superscripts to denote degrees of homogeneous components of a graded object, as in $A^i \subset A^\bullet$. When considering weights of a given $\mathbb{G}_m$-action we will use subscripts. We hope this will not cause confusion.

2.1.6. *General* Proj. Consider now a general quasicoherent graded $\mathcal{O}_X$-algebra with coherent graded components, which is only assumed to be generated over $A_0 = \mathcal{O}_X$ in finitely many degrees. Writing

$A^{M\bullet} = \bigoplus_j A^{Mj}$ for a positive integer M, we have a canonical isomorphism $Y = \mathrm{Proj}_X A^\bullet \simeq \mathrm{Proj}_X A^{M\bullet}$. For a suitable M the algebra $A^{M\bullet}$ is generated in degree 1 by A^M. If we take the minimal M_0 such that $A^{M\bullet}$ is generated in degree 1, then L is not functorial for all morphisms. Rather it is functorial for all *flat surjective* morphisms $X' \to X$: if $A^{M\bullet}$ is generated in degree 1 then $(A')^{M\bullet}$ is generated in degree 1, and the opposite is true whenever $X' \to X$ is flat surjective; this follows since surjectivity of $((A')^1)^{\otimes n} \to (A')^n$ implies surjectivity of $(A^1)^{\otimes n} \to A^n$ by flat decent. Combining this construction with the previous one we obtain an interpretation of $Y \to X$ as a projective morphism, and this construction is functorial for all flat surjective morphisms.

Remark 2.1.7. This construction applies to the following situation: assume $f : Y \to X$ is a proper morphism of noetherian schemes and L is an f-ample sheaf. Then $A^\bullet = \mathcal{O}_X \oplus \bigoplus_{k=1}^\infty f_*(L^k)$ is generated in finitely many degrees and $Y = \mathrm{Proj}_X A$. Therefore, L gives rise to an interpretation of f as a projective morphism functorially for all surjective flat morphisms.

2.1.8. *Blowings up.* An important variant is that of blowings up. Consider a coherent ideal sheaf I on X. The Rees algebra $R_X(I) = \bigoplus_{k=0}^\infty I^k$ is generated in degree 1, and we define $\mathrm{Bl}_I(X) = \mathrm{Proj}_X R_X(I)$. In particular, $\mathrm{Bl}_I(X)$ is projective over X with the closed immersion $\mathrm{Bl}_I(X) \hookrightarrow \mathbb{P}_X(I)$, and if I is the unit ideal on an open U of X then $\mathrm{Bl}_I(X) \to X$ is projectively the identity on U. If $\phi : X' \to X$ is a morphism, then $I^k\mathcal{O}_{X'} = (I\mathcal{O}_{X'})^k = (I')^k$ and $\phi^*(I^k) \to I^k\mathcal{O}_{X'}$ is surjective, giving a canonical morphism $\phi' : \mathrm{Bl}_{I'}(X') \to \mathrm{Bl}_I(X)$ over ϕ. Clearly $(\phi')^*L = L'$. So a blowing up is functorially projective. If moreover $X' \to X$ is flat, then $\mathrm{Bl}_{I'}(X') = X' \times_X \mathrm{Bl}_X I$.

We will need an opposite construction, using a variant of [Hartshorne 1977, Theorem II.7.17] for regular schemes. Assume X is regular and $f : Y \to X$ is a proper birational morphism with a relatively ample sheaf L (e.g., if $Y \to X$ is projective we can take $L = \mathcal{O}_Y(1)$). Then after replacing L by a positive power which is functorial for flat surjective morphisms, we have $Y = \mathrm{Proj}_X A^\bullet$, where $A^\bullet$ is generated over $A_0 = \mathcal{O}_X$ by its degree-1 component, and $A^k = f_*L^k$.

Locally on X, write L^k as a fractional ideal on Y, giving it as a fractional ideal $F_{L,k}$ on X since $Y \to X$ is birational. Since $A^\bullet$ is generated in degree 1, we have $F_{L,k} = F_{L,1}^k$; see [loc. cit., Theorem II.7.17, Step 5]. Since X is factorial, there is a unique expression $F_{L,1} = MI$, where M is an invertible fractional ideal and I is an *ideal sheaf* without invertible factors. Explicitly, $F_{L,1}^*$ is invertible, so we can write $I = F_{L,1}^* F_{L,1}$ and $M = F_{L,1}^{**}$. It follows that $F_{L,k} = M^k I^k$. Note that while the construction is local on X and depends on an embedding of L in the fraction field, the ideal sheaf I glues canonically. Locally on X we have a canonical isomorphism $Y \simeq \mathrm{Bl}_I(X)$, which evidently glues canonically. We have obtained that a projective birational morphism $f : Y \to X$ with X regular is a blowing up, functorially for flat surjective morphisms $X' \to X$ of regular schemes. In addition, if f is projectively the identity on $U \subseteq X$ then I is the unit ideal on U.

For future reference we record the following well-known result that follows from the universal property of blowings up.

Lemma 2.1.9. *If X is an integral scheme and a blowing up $Y = \mathrm{Bl}_I(X) \to X$ factors through a proper birational morphism $Z \to X$ then $Y = \mathrm{Bl}_{I\mathcal{O}_Z}(Z)$.*

2.1.10. *Sequences of projective morphisms.* Now assume $Z \xrightarrow{g} Y \xrightarrow{f} X$ is a sequence of projective morphisms of noetherian schemes, say $Z \hookrightarrow \mathbb{P}_Y(F)$ and $Y \hookrightarrow \mathbb{P}_X(E)$ for a coherent $\mathcal{O}_Y$-module F and a coherent $\mathcal{O}_X$-module E. For a large enough n the map $f^* f_*(F(n)) \xrightarrow{\alpha} F(n)$ is surjective; hence $\mathbb{P}_Y(F) = \mathbb{P}_Y(F(n))$ embeds into $\mathbb{P}_X(E \otimes f_*F(n))$ and we obtain a closed immersion $Z \hookrightarrow \mathbb{P}_X(E \otimes f_*F(n))$. Choosing the minimal n such that α is surjective we obtain a construction that realizes composition of projective morphisms as a projective morphism functorially for flat surjective morphisms $X' \to X$.

If X is regular we can combine this with the previous statements, so if $Y_m \to \cdots \to Y_1 \to X$ is a sequence of birational projective morphisms which are projectively the identity over an open $U \subseteq X$, then $Y_m \to X$ is a blowing up of an ideal sheaf which is the unit ideal on U, and this is functorial for flat and surjective morphisms of regular schemes.

Remark 2.1.11. We will not use this, but blowings up can also be composed in terms of ideals. One can show that if X is normal then the composition of $Y = \mathrm{Bl}_I(X) \xrightarrow{f} X$ and $\mathrm{Bl}_J(Y) \to Y$ is of the form $\mathrm{Bl}_{f_*(f^{-1}(I^n)J)}(X) \to X$ for a large enough n.

2.2. ***qe schemes and resolution of pairs.***

2.2.1. *qe schemes.* The class of quasiexcellent schemes was introduced by Grothendieck as the natural class where problems related to resolution of singularities behave well. The name "quasiexcellent" is perhaps not very elegant (it was not introduced by Grothendieck), and we feel it harmless to refer to them as *qe schemes*.

First recall that regular morphisms are a generalization of smooth morphisms in situations of morphisms which are not necessarily of finite type. Following [EGA IV$_2$ 1965, 6.8.1] a morphism of schemes $f : Y \to X$ is said to be *regular* if

- the morphism f is flat and
- all geometric fibers of $f : Y \to X$ are regular.

A locally noetherian scheme X is a *qe scheme* if the following two conditions hold:

- For any scheme Y of finite type over X, the regular locus Y_{reg} is open.
- For any point $x \in X$, the completion morphism $\operatorname{Spec} \widehat{\mathcal{O}}_{X,x} \to \operatorname{Spec} \mathcal{O}_{X,x}$ is regular.

It is a known, but nontrivial fact, that a scheme Y of finite type over a qe scheme is also a qe scheme; see, for example, [Matsumura 1980, 34.A]. A ring A is a qe ring if $\operatorname{Spec} A$ is a qe scheme.

2.2.2. *Resolution of pairs.* Consider a pair (X, Z), where X is a reduced qe scheme and Z is a nowhere dense closed subset of X. By a *resolution* of (X, Z) we mean a birational projective morphism $f : X' \to X$ such that X' is regular, $Z' = f^{-1}(Z)$ is a simple normal crossings divisor, and f is projectively the identity outside of the union of Z and the singular locus X_{sing} of X. Since [EGA IV$_2$ 1965, 7.9.6], it is universally hoped that every qe scheme admits a good resolution of singularities; the same should also hold for pairs; see Remark 2.2.3 below. If X is noetherian of characteristic 0 then (X, Z) can be resolved by [Temkin 2012, Theorem 1.1].

Remark 2.2.3. (i) Usually, resolution of pairs is constructed in two steps:

(1) Resolve X by a projective morphism $f : X' \to X$. Usually, this is achieved by a sequence of blowings up $X_l \to \cdots \to X_0 = X$. One can also achieve that the centers are regular, though this requires an additional effort.

(2) Resolve $Z' = f^{-1}(Z)$ by a further projective morphism $f' : X'' \to X'$. Usually, this is achieved by a sequence of blowings up $X'' = X'_n \to \cdots \to X'_0 = X'$ whose centers are regular and have simple normal crossings with the accumulated exceptional divisor, so that all schemes X'_i remain regular and exceptional divisors E'_i are simple normal crossings. In addition, one achieves a *principalization* of Z' as a subscheme; i.e., $Z' \times_{X'} X'_n$ is a divisor supported on E'_n.

(ii) The best known results for general noetherian qe schemes beyond characteristic 0 are resolution of qe threefolds, see [Cossart and Piltant 2014], and principalization of surfaces in regular qe schemes, see [Cossart et al. 2009]. In particular, a noetherian qe pair (X, Z) can be resolved whenever $\dim(X) \leq 3$.

2.2.4. *Compatibility with morphisms.* By a morphism of pairs $\phi : (Y, T) \to (X, Z)$ we will always mean a morphism $\phi : Y \to X$ such that $T = \phi^{-1}(Z)$. We say that resolutions $f_X : X' \to X$ and $f_Y : Y' \to Y$ of (X, Z) and (Y, T) are *compatible* with ϕ if $f_Y = \phi^*(f_X)$.

Remark 2.2.5. As we mentioned, often resolution of pairs has a natural structure of a composition of blowings up. The definition of compatibility in this case is similar with the only difference that the blowing up sequence of Y is obtained from the pullback of the blowing up sequence of X by removing all blowings up with empty centers. The latter contraction procedure is only needed when f is not surjective.

2.2.6. *Functorial resolution.* Let $\mathcal{C}$ be a class of pairs (X, Z), where X is a reduced noetherian qe scheme and Z is a closed subscheme. Throughout this paper, by a *functorial resolution* on $\mathcal{C}$ we mean a rule that assigns to any pair $(X, Z) \in \mathcal{C}$ a resolution $(X', Z') \to (X, Z)$ in a way compatible with arbitrary *surjective regular* morphisms between pairs in $\mathcal{C}$. In addition, we always make the following assumption on the resolution of *normal crossings pairs*, i.e., pairs (X, Z) with regular X and normal crossings Z (not necessarily simple):

Assumption 2.2.7. *For any normal crossings pair (X, Z) in $\mathcal{C}$ its resolution $X' \to X$ can be functorially represented as a composition of blowings up whose centers are regular and have normal crossings with the union of the preimage of Z and the accumulated exceptional divisor.*

Remark 2.2.8. (i) This definition provides the minimal list of properties we will use. As we remarked earlier, usually one proves finer desingularization results obtaining, in particular, that $Z \times_X X'$ is a divisor and the resolution is functorial for nonsurjective morphisms as well.

(ii) It seems that any reasonable resolution should satisfy the assumption. In fact, most (if not any) algorithms appearing in the literature apply to normal crossings pairs (X, Z) via the following *standard algorithm*: first one blows up the maximal multiplicity locus of Z, then one blows up the maximal multiplicity locus of the strict transform of Z, etc. It is easy to see that the standard algorithm satisfies the assumption.

2.2.9. *Resolution of singularities of qe schemes: characteristic 0.* Functorial resolution of pairs is known in characteristic 0:

Theorem 2.2.10. *There exists a functorial resolution, satisfying Assumption 2.2.7, on the class $\mathcal{C}_{\mathrm{char}=0}$ whose elements are pairs (X, Z) with X a reduced noetherian qe scheme over $\mathbb{Q}$.*

Proof. By [Temkin 2018, Theorem 1.1.7] there exists a blowing up sequence

$$\mathcal{F}_{\mathrm{princ}}(X, Z) : X' \to \cdots \to X$$

whose centers lie over $Z \cup X_{\mathrm{sing}}$ and such that X' is regular and $Z' = f^{-1}(Z)$ is a simple normal crossings divisor. Moreover, this sequence is functorial in regular morphisms. By Section 2.1.10, the morphism $X' \to X$ is a projective morphism functorially in surjective regular (even flat) morphisms. Finally, a direct (but tedious) inspection shows that the algorithm $\mathcal{F}_{\mathrm{princ}}$ of [loc. cit.] resolves normal crossings pairs via the standard algorithm. $\square$

Remark 2.2.11. Functoriality of this resolution implies that one also gets a functorial way to resolve an arbitrary qe pair over $\mathbb{Q}$ (locally noetherian but not necessarily noetherian) by a morphism $f : X' \to X$. In general, there is no natural way to provide f with an appropriate structure, neither as a single blowing up nor a sequence of blowings up. However, f can be realized as an infinite composition whose restrictions onto noetherian open subschemes of X are finite; e.g., the case of $Z = \varnothing$ is worked out in [Temkin 2008, Theorem 5.3.2].

2.2.12. *Positive and mixed characteristics hypothesis.* In Theorem 1.3.3 (3), the precise hypothetical statement we need about resolutions of pairs is the following analogue of Theorem 2.2.10:

Hypothetical Statement 2.2.13. (1) *Functorial resolution: The classes $\mathcal{C}_{\mathrm{char}=p,\dim\le d+1}$ and $\mathcal{C}_{\dim\le d+1}$ of pairs (X, Z), where X is a reduced noetherian qe $\mathbb{F}_p$-scheme or $\mathbb{Z}$-scheme, respectively, of dimension $\le d+1$, each admit a functorial resolution $f_{(X,Z)} : X' \to X$ satisfying Assumption 2.2.7.*

(2) *Equivariance: Moreover, the resolution is compatible with any G-action on (X, Z), where $G = \mathbb{G}_m$ or $G = (\mathbb{G}_a)^d$, in the sense that $a^*(f_{(X,Z)}) = p_X^*(f_{(X,Z)})$, where $a : G \times X \to X$ is the action morphism and $p_X : G \times X \to X$ is the projection.*

In mixed characteristics we will also need:

(3) *Functoriality of toroidal charts: assume that X is a toroidal scheme* (see [Abramovich and Temkin 2017, §2.3.4]) *of dimension at most $d+1$ and $j : X \to Y = \operatorname{Spec} \mathbb{Z}[M]$ is a toroidal chart* (see [loc. cit., §2.3.17]), *T is a toric subscheme of Y and $Z = X \times_Y T$. Then $j^*(f_{(Y,T)}) = f_{(X,Z)}$.*

We note that the equivariance statement (2) in dimension $d+1$ follows from statement (1) in dimension $d+2$, but here we wish to only make assumptions up to dimension $d+1$. It is conceivable that a version of (2) sufficient for our needs follows from (1) by taking slices, but we will not pursue this question.

Let us say that a pair (X, Z) is *locally monoidal* if locally X admits a logarithmic structure making it into a logarithmically regular logarithmic scheme so that the ideal of Z is monoidal. It is expected that there should exist a canonical resolution of such pairs of combinatorial nature, which is, in particular, independent of the characteristics. Our statement (3) asserts such independence in mixed characteristics; in pure characteristics it is a consequence of equivariance. It is analogous to Hypothetical Statement 2.3.6(3) below. Similarly to Hypothetical Statement 2.3.6, proving statements (1)–(3) for locally monoidal pairs is expected to be easier than the general case. For example, it is proved in [Illusie and Temkin 2014, Theorem 3.4.9] for logarithmically regular logarithmic schemes (with a single logarithmic structure), but the known functoriality [loc. cit., Theorem 3.4.15] is not enough to extend it to locally monoidal schemes. In addition, very recently Buonerba [2015] resolved certain locally monoidal varieties.

2.3. ***Principalization of ideal sheaves.*** In addition to resolution of pairs, we will need a version of functorial principalization of coherent ideal sheaves on a qe regular scheme X with a simple normal crossings divisor D, which will often be called the *boundary*. In fact, we will only need a particular case of locally monoidal ideals as introduced below.

2.3.1. *Permissible sequences.* A blowing up sequence $X_n \to \cdots \to X_0 = X$ will be called *permissible* if its centers $V_i \subset X_i$ are regular and have simple normal crossings with $D_i \subset X_i$, which is defined to be the union of the preimage of D and the accumulated exceptional divisor. Note that in such case each X_i is regular and each D_i is a boundary.

2.3.2. *Principalization.* We consider the category of triples (X, D, I), where (X, D) is a noetherian regular qe scheme with a boundary, I is a coherent ideal sheaf, and arrows are regular morphisms $f : X' \to X$ such that $I\mathcal{O}_{X'} = I'$ and $f^{-1}D = D'$. A *principalization* of I is a permissible sequence of blowings up $\phi_{(X,D,I)} : X_n \to \cdots \to X_0 = X$ such that:

(1) Each center V_i lies in the union of D_i with the locus where I is not the unit ideal.

(2) $I_n = I\mathcal{O}_{X_n}$ is a divisorial ideal supported on D_n. In particular, $V(I_n)$ is a divisor with a simple normal crossings reduction.

Principalizations form a category again, and functorial principalization provides a functor from triples (X, I, D) to principalizations $\phi_X : X' \to X$. As we do not require the morphism f to be surjective, we have to use the equivalence relation mentioned in Section 1.5. However, we will only apply the result in the context of surjective morphisms, so this equivalence will not figure in any of our applications.

2.3.3. *Known results.* Functorial principalization of ideal sheaves for *varieties* over a field of characteristic 0 is known; e.g., see [Bierstone and Milman 1997, Sections 11,13] or [Kollár 2007, Theorem 3.26]. The second author is in the process of writing a general functorial principalization of ideal sheaves on noetherian regular qe schemes over $\mathbb{Q}$ with the methods of [Temkin 2018]; we will manage not to use this result. For general qe schemes, the best known result is principalization on threefolds; see [Cossart and Piltant 2014].

Remark 2.3.4. (i) Classically, one only blows up centers over the locus where I is not trivial. On the other hand, usually one works with ordered boundaries $D = \bigcup_{i=0}^{n} D_i$, where D_i are smooth components.

Ordering the boundary restricts functoriality and, in fact, it is not critical. For example, the boundaries in [Cossart et al. 2009] are not ordered.

(ii) Since we allow blowings up that modify the whole D, we can freely use the classical results to resolve (X, D, I): first apply the standard principalization $f : X_n \to \cdots \to X$ to (X, D); then D_n is a simple normal crossings divisor ordered by the history of blowings up, and we can apply a classical algorithm to $(X_n, D_n, I\mathcal{O}_{X_n})$.

2.3.5. *Locally monoidal ideals.* A triple (X, D, I) with X regular, D a boundary and I an ideal sheaf on X is said to be *locally monoidal* if there is an open covering $\coprod U_\alpha \to X$, logarithmically regular structures (U_α, M_α) in the sense of [Kato 1994; Abramovich and Temkin 2017, §2.3.1] such that D is part of the toroidal divisor, and monoid ideals $I_\alpha \subset M_\alpha$ such that I_{U_α} is generated by the image of I_α under $M_\alpha \to \mathcal{O}_{U_\alpha}$.

Hypothetical Statement 2.3.6. (1) *Each locally monoidal $\mathbb{F}_p$-triple or $\mathbb{Z}$-triple (X, D, I) of dimension $\le d$ admits a principalization*

$$\phi_{(X,D,I)} : \widetilde{X} \to \cdots \to X$$

in a manner functorial for regular morphisms $X' \to X$.

(2) *Moreover, if $a : G \times X \to X$ is an action of $G = (\mathbb{G}_a)^d$ such that I and D are equivariant, that is, $a^{-1}I = p_X^{-1}I$ and $a^{-1}D = p_X^{-1}D$, where $p_X : G \times X \to X$ is the projection, then $\widetilde{X} \to X$ is G-equivariant as well.*

Again in mixed characteristics we also need:

(3) *Functoriality of toroidal charts: assume that (X, D, I) is locally monoidal of dimension $\le d$ and $j : (X, D) \to (Y = \operatorname{Spec} \mathbb{Z}[M], D_Y)$ is a toroidal chart such that $I = j^{-1}I_0$ for a toric ideal I_0 on Y. Then the sequence $\phi_{(X,D,I)}$ is the pullback of $\phi_{(Y,D_Y,I_0)}$.*

Remark 2.3.7. (i) In fact, the hypothesis asserts that toric ideals on schemes $\operatorname{Spec} \mathbb{Z}[M]$ can be principalized so canonically that given a locally monoidal triple (X, D, I) any toroidal chart induces the same principalization of I.

(ii) The results of [Illusie and Temkin 2014, Section 3.1.14] suggest that this statement may be within reach: in that paper the local nonfunctorial problem is solved, and the problem reduces to making the process functorial even if one changes the logarithmic structure M_α on U_α.

(iii) Here and below, given a morphism $f : Y \to X$ and ideal $I \subset \mathcal{O}_X$ we use the common notation $f^{-1}I$ for the ideal sheaf more precisely denoted by $(f^*I)\mathcal{O}_Y$, hoping this notation will not confuse the reader. We find the notation $(a^*I)\mathcal{O}_{G\times X}$ too heavy, and writing more simply $I\mathcal{O}_{G\times X}$ would not work in (2) above.

2.3.8. *The characteristic-0 case.* To make our results unconditional in characteristic 0 we should prove that parts (1) and (2) of Hypothetical Statement 2.3.6 hold for schemes over $\mathbb{Q}$. In fact, we will even deal with a larger class of triples using the case of varieties and methods of [Illusie and Temkin 2014, Theorem 2.4.1, p. 95].

A triple (X, D, I) is said to be $\mathbb{Q}$-*absolute* if there exists an open covering $\coprod U_\alpha \to X$, regular $\mathbb{Q}$-varieties Z_α, regular morphisms $f_\alpha : U_\alpha \to Z_\alpha$, ideal sheaves I_α on Z_α and divisors $D_\alpha \subset Z_\alpha$ such that $f_\alpha^{-1} I_\alpha = I|_{U_\alpha}$ and $f_\alpha^{-1} D_\alpha = D|_{U_\alpha}$. The collection of $\mathbb{Q}$-absolute triples forms a full subcategory of the category of triples. Functorial principalization of $\mathbb{Q}$-absolute triples (X, D, I) is a functor from this subcategory to principalizations of the corresponding ideals.

The statement we need is the following:

Proposition 2.3.9. *There exists a functorial principalization* $\phi_X : \widetilde{X} \to X$ *of* $\mathbb{Q}$*-absolute triples* (X, D, I).

Proof. We may replace $\coprod U_\alpha$ by a finite covering, since X is noetherian. We write $U_{\alpha\beta} = U_\alpha \times_X U_\beta$. Now, we will use the ideas from the proof of [Illusie and Temkin 2014, Theorem 2.4.3].

First we construct a principalization. It suffices to construct a principalization of $\coprod (U_\alpha, D|_{U_\alpha}, I|_{U_\alpha})$ whose two pullbacks to the fiber product $W := \coprod U_{\alpha\beta}$ coincide. The triple $(Z, D_Z, I_Z) := \coprod (Z_\alpha, D_\alpha, I_\alpha)$ has a principalization compatible with D_α coming from the principalization functor for $\mathbb{Q}$-varieties. This pulls back to a principalization of $\coprod (U_\alpha, D|_{U_\alpha}, I|_{U_\alpha})$ and we need to show that the two pullbacks to W coincide. We have two regular morphisms $f, g : W \to Z$. By Popescu's theorem [1986], see also [Spivakovsky 1999], f is the limit of smooth morphisms $f_\gamma : W_\gamma \to Z$. By [EGA IV$_3$ 1966, Proposition 8.13.1], g factors through a morphism $g_\gamma : W_\gamma \to Z$ for a large enough γ and then [Illusie and Temkin 2014, Proposition 2.4.3] implies that replacing W_γ by a neighborhood of the image of W we can achieve that g_γ is also smooth. Since the two pullbacks of I_Z and D_Z to W coincide, there is some γ such that the two pullbacks of I_Z and D_Z to W_γ coincide. It follows by functoriality of principalization for varieties that the two principalizations on W_γ coincide, and therefore they coincide on W, as required.

We now demonstrate that this principalization is functorial. Consider a regular surjective morphism $f : (X_1, D_1, I_1) \to (X_2, D_2, I_2)$ with coverings $\coprod U_{1\alpha}$ and $\coprod U_{2\beta}$ and $\mathbb{Q}$ varieties $Z_{1\alpha}$ and $Z_{2\alpha}$. Then composing $U_{2\beta} \to Z_{2\beta}$ with f we get another covering $\coprod f^{-1} U_{2\beta}$ with regular maps to $Z_{2\beta}$, so it is enough to show that the resulting principalizations on X_1 coincide. We now write $W = \coprod U_{1\alpha} \times_{X_1} f^{-1} U_{2\beta}$, which maps to $Z_1 = \coprod Z_{1\alpha}$ and $Z_2 = \coprod Z_{2\beta}$. By the same argument as earlier we have that $W \to Z_1 \times Z_2$ is the limit of a family $W_\gamma \to Z_1 \times Z_2$, where the two maps $W_\gamma \to Z_i$ are smooth. As above we conclude that the ideals and divisors coincide on some W_γ and the two principalizations coincide on W and therefore on X_1. $\square$

3. Functorial toroidal factorization

3.1. ***Statement.*** We follow the treatment of toroidal schemes in [Abramovich and Temkin 2017, Section 2.3], in particular they carry logarithmic structures in the Zariski topology. A toroidal ideal I on a toroidal scheme X with logarithmic structure M is the ideal generated by the image of a monomial ideal in M through $M \to \mathcal{O}_X$. We define a category $\mathrm{TorBl}_{\mathrm{rs}}$ of toroidal blowings up, similar to Bl:

(1) An object is a birational transformation $X_1 \to X_2$ where X_1, X_2 are toroidal and regular, and $X_1 \to X_2$ is given as the normalized blowing up of a toroidal ideal $I \subset \mathcal{O}_{X_2}$.

(2) An arrow from $X'_1 \to X'_2$ to $X_1 \to X_2$ consists of a regular surjective morphism $g : X'_2 \to X_2$ such that $U_{X_2} = g^{-1}U_{X_2}$ and $I' = I\mathcal{O}_{X'_2}$.

We similarly define a *toroidal weak factorization* $X_1 = V_0 \dashrightarrow V_1 \dashrightarrow \cdots \dashrightarrow V_{l-1} \dashrightarrow V_l = X_2$ of a toroidal blowing up $X_1 \to X_2$, where the schemes V_i, ideals J_i and centers Z_i are toroidal. These form *the regular surjective category* $\mathrm{TorFact}_{\mathrm{rs}}$ *of toroidal weak factorizations* in a manner similar to the above.

Proposition 3.1.1. *Let $X_1 \to X_2$ be a toroidal morphism of toroidal schemes obtained by normalized blowing up a toroidal ideal. Then there is a toroidal weak factorization $X_1 = V_0 \dashrightarrow V_1 \dashrightarrow \cdots \dashrightarrow V_{l-1} \dashrightarrow V_l = X_2$ in a functorial manner: there is a section* $\mathrm{TorBl}_{\mathrm{rs}} \to \mathrm{TorFact}_{\mathrm{rs}}$ *of the forgetful functor* $\mathrm{TorFact}_{\mathrm{rs}} \to \mathrm{Bl}$.

Remark 3.1.2. Jarosław Włodarczyk informed us that one can prove a stronger result: a factorization procedure which is functorial for all regular strict morphisms $g : X'_2 \to X_2$, not required to be surjective. His proposed argument involves subtle modifications at the heart of the algorithm in [Włodarczyk 2009, Sections 4 and 5]. The proof we provide at the end of this section shows that *any* procedure for toric factorization gives rise to a functorial procedure.

3.2. *Cone complexes.* Before proving Proposition 3.1.1 we need to discuss a generalization of the polyhedral cone complexes with integral structure of [Kempf et al. 1973] which was introduced in [Abramovich et al. 2015, 2.5] to accommodate any toroidal embedding in the sense of [Kempf et al. 1973], allowing for self-intersections and monodromy. In this paper we only assign polyhedral cone complexes to Zariski toroidal schemes, without self-intersections or monodromy, but the generalized polyhedral cone complexes are used as a combinatorial tool to achieve functoriality.

Fix a toroidal scheme X. Recall that the polyhedral complex of [Kempf et al. 1973] or the equivalent Kato fan of [Kato 1994] assigns a polyhedral cone σ_Z with integral structure to each toroidal stratum $Z \subset X$; each inclusion $Z' \hookrightarrow \bar{Z} \subset X$ gives rise to a linear map $\nu : \sigma_Z \to \sigma_{Z'}$, which identifies σ_Z as a face of $\sigma_{Z'}$ in such a way that the integral structure on σ_Z is the restriction of the integral structure of $\sigma_{Z'}$: this is called a *face map*. The diagram $(\{\sigma_Z\}, \{\nu\})$ is a poset, defining a polyhedral cone complex $\Sigma(X)$ as in [Abramovich et al. 2015, Section 2.2]. Anticipating Section 3.3 we denote it by $\Sigma(X) = \varinjlim(\{\sigma_Z\}, \{\nu\})$, where the colimit is taken in the category of generalized cone complexes of [loc. cit., Section 2.6]. This polyhedral cone complex is similar to the fan of a toric variety, but is not embedded in a space $N_{\mathbb{R}}$ and the intersection of two cones may be the union of faces rather than just one face.

A map of polyhedral cone complexes $\varinjlim(\{\sigma'_i\}, \{\nu'_k\}) \to \varinjlim(\{\sigma_j\}, \{\nu_l\})$ is defined to be a collection of cone maps $\sigma'_i \to \sigma_{j(i)}$ compatible with the face maps ν'_k and ν_k. A toroidal map $X' \to X$ gives rise to a map of cone complexes; here are a few well-known relationships:

(1) A proper birational toroidal morphism gives rise to a subdivision, and there is an equivalence of categories between proper toroidal birational morphisms and subdivisions. Blowings up of ideals correspond to subdivisions determined by piecewise linear continuous integral functions which are convex on each cone; following [Kempf et al. 1973] we call these *projective subdivisions* (in the combinatorial literature they are *coherent subdivisions*).

(2) A regular morphism $g : X'_2 \to X_2$ such that $U_{X'_2} = g^{-1}U_{X_2}$ gives rise to a map of complexes $\Sigma(g) : \Sigma(X') \to \Sigma(X)$ where all the maps $\sigma'_i \to \sigma_{j(i)}$ are face maps — this is called a face map of complexes.

(3) If the map $g : X'_2 \to X_2$ is also surjective then $\Sigma(g)$ is surjective.

(4) The scheme X is regular if and only if all the cones $\sigma_i \subset \Sigma(X)$ are nonsingular in the usual toric sense.

(5) If X is regular then the closure of a stratum is always regular (this would fail if we allowed self-intersections); we call such subschemes *toroidal centers*.

(6) The blowing up $X' \to X$ of an irreducible toroidal center $\bar{Z}$ on a *regular* X corresponds to the star subdivision $\Sigma' \to \Sigma(X)$ at the barycenter of σ_Z. The blowing up $X' \to X$ of any regular toroidal subscheme W corresponds to the simultaneous star subdivision $\Sigma' \to \Sigma(X)$ at the barycenters of all the cones corresponding to the connected components of W.

Thus Proposition 3.1.1 would follow if the projective subdivision $\Sigma(X_1) \to \Sigma(X_2)$ can be factored as a composition of such simultaneous star subdivisions and their inverses, in such a way that the intermediate steps are projective subdivisions of $\Sigma(X_2)$, in a functorial manner with respect to surjective face maps. This will be our Lemma 3.5.1 below.

Morelli's π-desingularization lemma of fan cobordisms [Włodarczyk 2003, Lemma 10.4.3] gives a nonfunctorial result in the case of fans; this was generalized in [Abramovich et al. 1999] to polyhedral cone complexes. In [Abramovich et al. 2002] it is made functorial under *automorphisms,* which is not sufficient for our purposes here.

Consider the category whose objects are projective subdivisions $\Sigma_1 \to \Sigma_2$ of nonsingular cone complexes given by a fixed piecewise linear continuous integral function $f : \Sigma_2 \to \mathbb{R}$ convex on each cone and arrows $(\Sigma'_2, f') \to (\Sigma_2, f)$ induced by surjective face maps $h : \Sigma'_2 \to \Sigma_2$ with $f' = f \circ h$. Functoriality would be easily achieved if the connected component of any object $\Sigma_1 \to \Sigma_2$ in this category had a final object, as we show below in Lemma 3.5.1. Indeed, this would mean that applying Morelli's lemma to the final object would induce a factorization for the whole component, giving the result. Unfortunately final objects usually do not exist in the category of cone complexes. Our next goal is to enlarge this category so that final objects do exist; see Lemma 3.3.1 below.

3.3. ***Generalized cone complexes and existence of final objects.*** A generalized cone complex is given by any finite diagram $(\{\sigma_j\}, \{\nu_l\})$ of cones and face maps. We allow for more than one face map $\sigma_j \to \sigma_l$, including nontrivial self-face maps $\sigma_j \to \sigma_j$. We think of a generalized cone complex Σ as a structure imposed on the topological space $\underline{\Sigma} = \varinjlim(\{\sigma_j\}, \{\nu_l\})$. Thus an arrow of generalized cone complexes $(\{\sigma'_i\}, \{\nu'_k\}) \to (\{\sigma_j\}, \{\nu_l\})$ is given by compatible cone maps as above; an arrow is a face map if it is given by compatible face maps; and an arrow is declared to be an isomorphism if it is a face map inducing a bijection of sets $\varinjlim(\{\sigma'_i\}, \{\nu'_k\}) \to \varinjlim(\{\sigma_j\}, \{\nu_l\})$. See [Abramovich et al. 2015, §2.6].

Cone complexes are a full subcategory of generalized cone complexes. They are distinguished by the property that, for any cones τ, σ of Σ, a face map $\nu : \tau \to \sigma$ in Σ is unique if it exists. Thus Proposition 3.1.1 would again follow if any projective subdivision $\Sigma_1 \to \Sigma_2$ of *generalized* nonsingular cone complexes can be factored as a composition of simultaneous star subdivisions and their inverses, in a functorial manner with respect to surjective cone maps. The advantage of working with generalized cone complexes is the following:

Lemma 3.3.1. *The connected component of the projective subdivision $\Sigma_1 \to \Sigma_2$ of generalized cone complexes in the category induced by surjective face maps $\Sigma'_2 \to \Sigma_2$ has a final object.*

Proof. The projective subdivision $\Sigma_1 \to \Sigma_2$ is induced by an implicit piecewise linear convex integral function $f : \Sigma_2 \to \mathbb{R}$. Write $\Sigma_2 = (\{\sigma_j\}, \{\nu_l\})$. Then $\nu_l : \sigma_i \to \sigma_j$ has the property that $f_{\sigma_i} = f_{\sigma_j} \circ \nu_l$. Let $\{\mu_k\}$ be the collection of all face maps $\mu_k : \sigma_m \to \sigma_n$ with the property that $f_{\sigma_m} = f_{\sigma_n} \circ \mu_k$. Then $\Delta := (\{\sigma_j\}, \{\mu_k\})$ is a generalized cone complex, the maps f_{σ_j} glue to give a piecewise linear integral function $\tilde{f} : \Delta \to \mathbb{R}$, and since $\{\nu_l\} \subset \{\mu_k\}$ we have a map of diagrams $g : \Sigma_2 \to \Delta$ such that $f = \tilde{f} \circ g$.

It is convenient to have another presentation of Δ. Choose one representative $\bar{\sigma}$ from each isomorphism class of cones in Δ. Given two such representatives $\bar{\tau}$ and $\bar{\sigma}$, consider all maps $\bar{\nu}_l : \bar{\tau} \to \bar{\sigma}$ in Δ. Clearly $\bar{\Delta} = (\{\bar{\sigma}\}, \{\bar{\nu}_l\})$ maps as a subdiagram to Δ, and the map is an isomorphism since it is clearly a bijection on set-theoretic limits.

We claim that $(\Delta, \tilde{f})$ is a final object in the component of (Σ_2, f) in the category of generalized cone complexes with piecewise linear integral function. For this it suffices to show that if (Σ'_2, f') is an object and $h : \Sigma_2 \to \Sigma'_2$ is a surjective face map such that $f' \circ h = f$ then $g = g'' \circ h$, where $g'' : \Sigma'_2 \to \Delta$ is a morphism so that $f' = \tilde{f} \circ g''$.

First, if we apply the construction of Δ to Σ'_2 we get a map $g' : \Sigma'_2 \to \Delta'$ which sits in a commutative diagram:

$$\begin{tikzcd}
\Sigma_2 \arrow[r, "g"] \arrow[d, "h"'] \arrow[rrd, "f"] & \Delta \arrow[d, "\tilde{h}"] \arrow[rd, "\tilde{f}"] & \\
 & & \mathbb{R} \\
\Sigma'_2 \arrow[r, "g'"'] \arrow[rru, "f'"] & \Delta' \arrow[ru, "\tilde{f}'"'] &
\end{tikzcd}$$

On the other hand $\bar{\Delta} \simeq \Delta$ and $\bar{\Delta}' \simeq \Delta'$, and the map $\bar{\Delta} \to \bar{\Delta}'$ induced by $\tilde{h}$ is an isomorphism of diagrams: since h is a surjective face map, any cone in Σ'_2 is isomorphic to a cone of Σ_1 via an isomorphism compatible with f and vice versa. So $\tilde{h}$ gives a bijection between the isomorphism classes of cones, and the maps $\bar{\nu}$ between cones are determined by the compatibility of the function $\tilde{f} = \tilde{f}'$ on them. So $\Delta \to \Delta'$ is an isomorphism, giving the requisite map of generalized complexes $g'' = \tilde{h}^{-1} \circ g'$. $\square$

3.4. ***Barycentric subdivisions and factorization for generalized cone complexes.*** We proceed to extend the factorization of subdivisions of cone complexes to generalized cone complexes. We do it by a reduction step using barycentric subdivisions:

Lemma 3.4.1. (1) [Abramovich et al. 2015, 2.5] *The barycentric subdivision $B(\Delta)$ of a generalized cone complex Δ is a projective subdivision obtained by a sequence of simultaneous star subdivisions. If Δ is nonsingular then the star subdivisions are smooth. The generalized cone complex $B(\Delta)$ is in fact a cone complex.*

(2) [Abramovich et al. 1999, Lemma 8.7] *The barycentric subdivision $B(\Delta)$ of a nonsingular cone complex Δ is a projective subdivision obtained by a sequence of simultaneous smooth star subdivisions. The nonsingular cone complex $B(\Delta)$ is in fact isomorphic to a fan.*

Proof. (1) Write $\Delta = (\{\sigma_j\}, \{\mu_k\})$. We need to show that if τ_B, σ_B are cones in $B(\Delta)$, then a face map $\tau_B \to \sigma_B$ in $B(\Delta)$ is unique if it exists. Suppose the minimal cone containing the image of τ_B is τ and the corresponding cone for σ_B is σ. Then it suffices to show that the restriction to τ_B of a face map $\psi : \tau \to \sigma$ in Δ carrying τ_B into σ_B is unique if it exists. We can write $\sigma_B = \langle b(\sigma_{i_1}), \dots b(\sigma_{i_k})\rangle$ uniquely as the cone generated by the barycenters $b(\sigma_{i_r})$ of faces σ_{i_r} of σ of dimensions $i_1 < \cdots < i_k$, and similarly $\tau_B = \langle b(\tau_{j_1}), \dots b(\tau_{j_l})\rangle$. So ψ must carry $b(\tau_{j_s})$ to the barycenter of a cone of σ of dimension j_s; in other words $\psi(b(\tau_{j_s})) = b(\sigma_{j_s})$. Since $\{b(\tau_{j_1}), \dots, b(\tau_{j_l})\}$ spans τ_B this means that the restriction of ψ is unique if it exists.

(2) Consider the vector space $V = \bigoplus_{\sigma\in\Delta} \mathbb{R}_\sigma$ with one basis element for each cone of σ. Assume Δ is a cone complex. In [Abramovich et al. 1999, Lemma 8.7] it is shown that $B(\Delta)$ has a real embedding in V, and the image is the real support of a fan. The embedding is obtained by sending $b(\sigma)$ to the unit vector $e_\sigma \in \mathbb{R}_\sigma \subset V$. Here we assume that Δ is nonsingular, and we need to check that the embedding gives an isomorphism of cone complexes, namely that the integral structures coincide. Note that the lattice in any cone $\langle b(\sigma_{i_1}), \dots, b(\sigma_{i_k})\rangle$ in $B(\Delta)$ is generated by the elements $b(\sigma_{i_1}), \dots, b(\sigma_{i_k})$. The image of this lattice in V is precisely generated by $e(\sigma_{i_1}), \dots, e(\sigma_{i_k})$, and coincides with the intersection of the cone $\langle e(\sigma_{i_1}), \dots, e(\sigma_{i_k})\rangle$ with $\bigoplus_{\sigma\in\Delta} \mathbb{Z}_\sigma$. So the image of $B(\Delta)$ is indeed a fan, as required. ☐

Lemma 3.4.2. *Let Δ be a nonsingular generalized cone complex and $f : \Delta \to \mathbb{R}$ a piecewise linear function, convex and integral on each cone, such that the corresponding subdivision $\Delta_1 \to \Delta$ is nonsingular. Then $\Delta_1 \to \Delta$ admits a factorization into nonsingular star subdivisions and their inverses, with all intermediate steps projective over Δ.*

Proof. By Lemma 3.4.1 we may replace Δ_1 by its second barycentric subdivision, so we may assume Δ_1 is isomorphic to a fan. The common subdivision of $B(B(\Delta_1))$ and $B(B(\Delta))$ is a projective subdivision of $B(B(\Delta_1))$, so there is a sequence of star subdivisions $\Delta_1' \to B(B(\Delta_1))$ such that $\Delta_1' \to \Delta$ factors through a projective subdivision $\Delta_1' \to \Delta' := B(B(\Delta))$. Since Δ' is isomorphic to a fan and Δ_1' is a projective subdivision, Morelli's π-desingularization lemma applies, see [Morelli 1996] or [Włodarczyk 2003, Lemma 10.4.3], giving a factorization by star subdivisions and their inverses, all projective over Δ'.

Combining these transformation, we obtain the desired factorization, with all steps projective over Δ:

$$\begin{tikzcd}
 & & \Delta_1' \arrow[rr, dotted, "\text{factorized}"] \arrow[dl, "\text{star subdivision sequence}"'] & & \Delta' \arrow[dr, equal] & & \\
 & B(B(\Delta_1)) \arrow[dl, "\text{star subdivision sequence}"'] & & & & B(B(\Delta)) \arrow[dr, "\text{star subdivision sequence}"] & \\
\Delta_1 \arrow[rrrrrr, "\text{projective subdivision}"'] & & & & & & \Delta
\end{tikzcd}$$ □

3.5. *Functoriality for generalized cone complexes.*

Lemma 3.5.1. *The factorization in Lemma 3.4.2 can be made functorial for surjective face maps: we can associate to (Δ, f) a factorization so that, given a surjective face map $\phi : \Sigma \to \Delta$, the factorization of $(\Sigma, f \circ \phi)$ is the pullback of the factorization of (Δ, f) along ϕ.*

Proof. For each connected component of the category of pairs (Δ, f) with face maps between them choose a final object $(\tilde{\Delta}, \tilde{f})$. By Lemma 3.4.2 there is a factorization $\tilde{\Delta}_1 \dashrightarrow \cdots \dashrightarrow \tilde{\Delta}$ of $(\tilde{\Delta}, \tilde{f})$. Given an arbitrary (Δ, f) it has a morphism $\psi_\Delta : \Delta \to \tilde{\Delta}$ to the final object $(\tilde{\Delta}, \tilde{f})$, so that $f = f \circ \psi_\Delta$. The pullback $\Delta_1 \dashrightarrow \cdots \dashrightarrow \Delta$ of $\tilde{\Delta}_1 \dashrightarrow \cdots \dashrightarrow \tilde{\Delta}$ along ψ_Δ is a factorization of (Δ, f), and its pullback along ϕ is simply the pullback $\Sigma_1 \dashrightarrow \cdots \dashrightarrow \Sigma$ along $\psi_\Delta \circ \phi = \psi_\Sigma$ of $\tilde{\Delta}_1 \dashrightarrow \cdots \dashrightarrow \tilde{\Delta}$, so the process is functorial. □

3.6. *Functoriality for toroidal factorization.*

Proof of Proposition 3.1.1. The toroidal morphism $X_1 \to X_2$ corresponds to a subdivision $\Sigma(X_1) \to \Sigma(X_2)$ induced by a piecewise linear function $f : \Sigma(X_2) \to \mathbb{R}$ convex and integral on each cone. This is functorial: a surjective regular morphism $X_2' \to X_2$ gives rise to a surjective face map $\phi : \Sigma(X_2)' \to \Sigma(X_2)$ such that $X_1' \to X_2'$ corresponds to $f \circ \phi$.

By Lemma 3.5.1 we have a factorization $\Sigma(X_1) \dashrightarrow \cdots \dashrightarrow \Sigma(X_2)$, functorial for surjective face maps, into nonsingular star subdivisions and their inverses, with all intermediate steps functorially projective over $\Sigma(X_2)$. This gives rise to a toroidal factorization $X_1 \dashrightarrow \cdots \dashrightarrow X_2$ into blowings up and down, which is functorial for surjective regular morphisms, where the terms are functorially projective over X_2. □

4. Birational cobordisms

A key tool in the factorization algorithm is the notion of *birational cobordism*, introduced in [Włodarczyk 2000], where it is motivated by analogy with Morse theory. In this paper we adopt the approach of [Abramovich et al. 2002], which relies on geometric invariant theory and variation of linearizations; see [Brion and Procesi 1990; Thaddeus 1996; Dolgachev and Hu 1998].

4.1. ***Geometric Invariant Theory of $\mathbb{P}(E)$.*** Given a nonzero coherent sheaf E on X_2, the data of a $\mathbb{G}_m$-action $\rho : \mathbb{G}_m \to \operatorname{Aut} E$ on E is equivalent to the data of a $\mathbb{Z}$-grading $E = \bigoplus_{a \in \mathbb{Z}} E_a$, which is necessarily a finite sum: $E = \bigoplus_{a=a_{\min}}^{a_{\max}} E_a$. The homogeneous factor E_a is characterized by

$$\rho(t)v = t^a v \quad \text{for all } v \in E_a.$$

Here and later we use the informal notation $v \in E_a$ to indicate that v is a local section of E_a. Given such data, there is a resulting action of $\mathbb{G}_m$ on $\mathrm{Sym}^\bullet(E)$ and a linearized action on $\mathbb{P}(E) = \mathbb{P}_{X_2}(E)$.

We require the following:

Assumption 4.1.1. *The sheaves $E_{a_{\min}}$ and $E_{a_{\max}}$ are everywhere nonzero, so the morphisms $\mathbb{P}(E_{a_{\min}}) \to X_2$ and $\mathbb{P}(E_{a_{\max}}) \to X_2$ are surjective.*

Given an integer a viewed as a character of $\mathbb{G}_m$, we define a new action of $\mathbb{G}_m$ on E by

$$\rho_a(t)v = t^{-a}\rho(t)(v).$$

This induces an action on $\mathrm{Sym}^\bullet(E)$ and on $(\mathbb{P}(E), \mathcal{O}_{\mathbb{P}(E)}(1))$ which we also denote by ρ_a. Writing $(\mathrm{Sym}^\bullet(E))^{\rho_a}$ for the ring of invariants under this action, we define

$$\mathbb{P}(E)/\!\!/_a\mathbb{G}_m := \mathrm{Proj}_{X_2}(\mathrm{Sym}^\bullet(E))^{\rho_a}.$$

As customary, we unwind this as follows: we define the *unstable locus of ρ_a* to be the closed subscheme

$$\mathbb{P}(E)_a^{\mathrm{un}} := \mathbb{P}\Big(\bigoplus_{b<a} E_b\Big) \coprod \mathbb{P}\Big(\bigoplus_{b>a} E_b\Big), \tag{1}$$

and the *semistable locus* to be the complementary open

$$\mathbb{P}(E)_a^{\mathrm{sst}} := \mathbb{P}(E) \smallsetminus \mathbb{P}(E)_a^{\mathrm{un}}.$$

We have the following well-known facts:

Lemma 4.1.2. (1) *The semistable locus $\mathbb{P}(E)_a^{\mathrm{sst}}$ is nonempty precisely when $a_{\min} \le a \le a_{\max}$.*

(2) *Consider the rational map $q_a : \mathbb{P}(E) \to \mathbb{P}(E)/\!\!/_a\mathbb{G}_m$ induced by the inclusion $(\mathrm{Sym}^\bullet(E))^{\rho_a} \subset (\mathrm{Sym}^\bullet(E))$. Then q_a restricts to an affine $\mathbb{G}_m$-invariant morphism $\mathbb{P}(E)_a^{\mathrm{sst}} \to \mathbb{P}(E)/\!\!/_a\mathbb{G}_m$ which is a submersive universal categorical quotient*; *thus $\mathbb{P}(E)/\!\!/_a\mathbb{G}_m = \mathbb{P}(E)_a^{\mathrm{sst}}/\!\!/\mathbb{G}_m$.*

(3) *For $a_{\min} \le a_1 < a_2 \le a_{\max}$ we have $\mathbb{P}(E)_{a_1}^{\mathrm{sst}} \subset \mathbb{P}(E)_{a_2}^{\mathrm{sst}}$ precisely when $\bigoplus_{a=a_1}^{a_2-1} E_a = 0$, and similarly $\mathbb{P}(E)_{a_1}^{\mathrm{sst}} \supset \mathbb{P}(E)_{a_2}^{\mathrm{sst}}$ precisely when $\bigoplus_{a=a_1+1}^{a_2} E_a = 0$. In particular $\mathbb{P}(E)_{a_1}^{\mathrm{sst}} = \mathbb{P}(E)_{a_2}^{\mathrm{sst}}$ precisely when $\bigoplus_{a=a_1}^{a_2} E_a = 0$.*

(4) *If $a_{\min} \le a_1 < a_2 \le a_{\max}$ and $\bigoplus_{a=a_1}^{a_2-1} E_a = 0$, then the inclusion $\mathbb{P}(E)_{a_1}^{\mathrm{sst}} \subset \mathbb{P}(E)_{a_2}^{\mathrm{sst}}$ induces a projective morphism*

$$\mathbb{P}(E)_{a_1}^{\mathrm{sst}}/\!\!/\mathbb{G}_m \to \mathbb{P}(E)_{a_2}^{\mathrm{sst}}/\!\!/\mathbb{G}_m.$$

Similarly if $\bigoplus_{a=a_1+1}^{a_2} E_a = 0$ we have a projective morphism

$$\mathbb{P}(E)_{a_1}^{\mathrm{sst}}/\!\!/\mathbb{G}_m \leftarrow \mathbb{P}(E)_{a_2}^{\mathrm{sst}}/\!\!/\mathbb{G}_m.$$

Proof. (1) We have $a \le a_{\max}$ if and only if $\mathbb{P}\big(\bigoplus_{b<a} E_b\big) \ne \mathbb{P}(E)$, and $a_{\min} \le a$ if and only if $\mathbb{P}\big(\bigoplus_{b>a} E_b\big) \ne \mathbb{P}(E)$.

(2a) Affine cover of the quotient: The scheme $\mathbb{P}(E)/\!\!/_a\mathbb{G}_m = \mathrm{Proj}_{X_2}(\mathrm{Sym}^\bullet(E))^{\rho_a}$ is covered by principal open sets

$$D_f^0 := (\mathbb{P}(E)/\!\!/_a\mathbb{G}_m) \smallsetminus Z_{\mathbb{P}(E)/\!\!/_a\mathbb{G}_m}(f) \tag{2}$$

associated to nonzero homogeneous invariant elements of the form $f = \prod_{j=1}^s f_j$, where $f_j \in E_{a+\delta_j}$ with $\sum \delta_j = 0$.

(2b) Common zero locus of $\{f\}$: We note that the common zero locus of elements of E_c is $\mathbb{P}(E/E_c) = \mathbb{P}\big(\bigoplus_{b\ne c} E_b\big)$. Now observe that any element $f = \prod_{j=1}^s f_j$ as above has a factor f_j with $\delta_j \ge 0$ and a factor f_j with $\delta_j \le 0$. This means that f vanishes on $\mathbb{P}\big(\bigoplus_{b<a} E_b\big)$ and on $\mathbb{P}\big(\bigoplus_{b>a} E_b\big)$, so f vanishes on $\mathbb{P}(E)_a^{\mathrm{un}}a$.

Conversely if $x \notin \mathbb{P}(E)_a^{\mathrm{un}}a$ then we have some coordinates $f_1 \in E_{a+\delta_1}$, $\delta_1 \le 0$ and $f_2 \in E_{a+\delta_2}$, $\delta_2 \ge 0$, which do not vanish: $f_1(x) \ne 0 \ne f_2(x)$. Taking any positive r, s so that $r\delta_1 + s\delta_2 = 0$ we can form $f = f_1^r f_2^s$, and $f(x) \ne 0$. This implies that the common zero locus of the elements $f = \prod_{j=1}^s f_j$ above in $\mathbb{P}(E)$ is precisely $\mathbb{P}(E)_a^{\mathrm{un}}a$.

(2c) Compatible affine cover of $\mathbb{P}(E)_a^{\mathrm{sst}}$: It follows that $\mathbb{P}(E)_a^{\mathrm{sst}}$ is covered by principal open sets

$$D_f = \mathbb{P}(E) \smallsetminus Z_{\mathbb{P}(E)}(f), \tag{3}$$

the inverse image of the affine open D_f^0 of equation (2) is the affine open D_f of equation (3), and $\mathbb{P}(E)_a^{\mathrm{sst}} \to \mathbb{P}(E)/\!\!/_a\mathbb{G}_m$ is an affine morphism.

(2d) Coordinates and invariants: The coordinate ring of D_f^0 is the degree-0 component of $(\mathrm{Sym}^\bullet(E))^{\rho_a}\big[\frac{1}{f}\big]$, which is the ρ_a-invariant summand of the degree-0 component of $(\mathrm{Sym}^\bullet(E))\big[\frac{1}{f}\big]$. The latter is the coordinate ring of D_f. In particular, $D_f^0 = D_f/\!\!/\mathbb{G}_m$ is a submersive universal categorical quotient; see [Abramovich and Temkin 2018, Lemma 4.2.6 and Corollary 4.2.11]. It follows from the definition, see [Mumford et al. 1994, Remark 5, p. 8], that $\mathbb{P}(E)_a^{\mathrm{sst}} \to \mathbb{P}(E)/\!\!/_a\mathbb{G}_m$ is a submersive universal categorical quotient.

(3) The situation is symmetric, so we only address the first statement. If $\bigoplus_{a=a_1}^{a_2-1} E_a = 0$ then $\mathbb{P}\big(\bigoplus_{b<a_2} E_b\big) = \mathbb{P}\big(\bigoplus_{b<a_1} E_b\big) \subset \mathbb{P}(E)_a^{\mathrm{un}}a_1$ and certainly $\mathbb{P}\big(\bigoplus_{b>a_2} E_b\big) \subset \mathbb{P}\big(\bigoplus_{b>a_1} E_b\big) \subset \mathbb{P}(E)_a^{\mathrm{un}}a_1$, so $\mathbb{P}(E)_a^{\mathrm{un}}a_1 \subset \mathbb{P}(E)_a^{\mathrm{un}}a_2$ as needed.

Conversely, if $v \in \mathbb{P}\big(\bigoplus_{a=a_1}^{a_2-1} E_a\big)$ over $x \in X_2$ and we take $w \in \mathbb{P}(E_{a_{\min}})$ also over x, then either $v \in \mathbb{P}(E_{a_1}) \subset \mathbb{P}(E)_{a_1}^{\mathrm{sst}}$ or else $(v+w) \in \mathbb{P}(E)_{a_1}^{\mathrm{sst}}$. In either case, if $\bigoplus_{a=a_1}^{a_2-1} E_a \ne 0$ we have $\mathbb{P}(E)_{a_1}^{\mathrm{sst}} \not\subset \mathbb{P}(E)_{a_2}^{\mathrm{sst}}$, as needed.

(4) The situation is symmetric, so we only address the first case, where $a_{\min} \le a_1 < a_2 \le a_{\max}$ and $\bigoplus_{a=a_1}^{a_2-1} E_a = 0$, so that $\mathbb{P}(E)_{a_1}^{\mathrm{sst}} \subset \mathbb{P}(E)_{a_2}^{\mathrm{sst}}$ by (3). Since $\mathbb{P}(E)_{a_i}^{\mathrm{sst}} \to \mathbb{P}(E)/\!\!/_{a_i}\mathbb{G}_m$ are categorical quotients, we have a canonical morphism φ_{a_1/a_2} making the following diagram commutative:

$$\begin{tikzcd}
\mathbb{P}(E)^{\mathrm{sst}}_{a_1} \arrow[r, hook] \arrow[d] & \mathbb{P}(E)^{\mathrm{sst}}_{a_2} \arrow[d] \\
\mathbb{P}(E)/\!\!/_{a_1}\mathbb{G}_m \arrow[r, "\varphi_{a_1/a_2}"] & \mathbb{P}(E)/\!\!/_{a_2}\mathbb{G}_m.
\end{tikzcd}$$

But $\mathbb{P}(E)/\!\!/_{a_i}\mathbb{G}_m$ are projective over X_2; hence φ_{a_1/a_2} is projective. □

This lemma gives the familiar "wall and chamber decomposition" of the interval $[a_{\min}, a_{\max}]$ in the character lattice $\mathbb{Z}$ into segments where the quotients $\mathbb{P}(E)^{\mathrm{sst}}_{a_1}/\!\!/_{\mathbb{G}_m}$ are constant.

All the constructions above are compatible with arbitrary morphisms $X_2' \to X_2$, except that the values of $a_{\min}$ and $a_{\max}$ and the ample sheaf for ϕ_{a_1/a_2} are only compatible with surjective morphisms $X_2' \to X_2$.

Remark 4.1.3. One can show that the quotient morphism $\mathbb{P}(E)^{\mathrm{sst}}_a \to \mathbb{P}(E)^{\mathrm{sst}}_a/\!\!/\mathbb{G}_m$ is in fact universally submersive. If in addition $E_a = 0$, it can be shown that the quotient morphism is a universal geometric quotient $\mathbb{P}(E)^{\mathrm{sst}}_a \to \mathbb{P}(E)^{\mathrm{sst}}_a/\mathbb{G}_m$. These facts follow from [Mumford et al. 1994, Theorem 1.1 and Amplification 1.3], which are stated for schemes over a field in characteristic 0 but apply here since $\mathbb{G}_m$ is a linearly reductive group-scheme over $\mathbb{Z}$. Since we do not need these facts, we will not provide a detailed proof, though we will use the notation $\mathbb{P}(E)^{\mathrm{sst}}_a/\mathbb{G}_m$ when $E_a = 0$.

4.2. *Geometric invariant theory of $B \subset \mathbb{P}(E)$.* Continuing the discussion, let $B \subset \mathbb{P}(E)$ be a closed reduced $\mathbb{G}_m$-stable subscheme. It is the zero locus of a homogeneous and $\mathbb{G}_m$-homogeneous ideal $I_B \subset \mathrm{Sym}^\bullet E$. We define $B^{\mathrm{un}}_a a := B \cap \mathbb{P}(E)^{\mathrm{un}}_a a$ and $B^{\mathrm{sst}}_a := B \cap \mathbb{P}(E)^{\mathrm{sst}}_a$. The image of $q_a : B^{\mathrm{sst}}_a \to \mathbb{P}(E)/\!\!/_a\mathbb{G}_m$ is denoted by $B/\!\!/_a\mathbb{G}_m$. We have canonically $B/\!\!/_a\mathbb{G}_m = \mathrm{Proj}_{X_2}((\mathrm{Sym}^\bullet E/I_B)^{\rho_a})$. We write $a_{\min}(B) = \min\{a \mid B \cap \mathbb{P}(E_a) \neq \varnothing\}$ and similarly $a_{\max}(B) = \max\{a \mid B \cap \mathbb{P}(E_a) \neq \varnothing\}$. We deduce the analogous, still well-known, facts, which follow immediately from Lemma 4.1.2:

Lemma 4.2.1. (1) *The semistable locus B^{sst}_a is nonempty precisely when $a_{\min}(B) \le a \le a_{\max}(B)$.*

(2) *The map $q_a : B^{\mathrm{sst}}_a \to \mathbb{P}(E)/\!\!/_a\mathbb{G}_m$ is an affine $\mathbb{G}_m$-invariant morphism, inducing a categorical quotient $B^{\mathrm{sst}}_a \to B^{\mathrm{sst}}_a/\!\!/\mathbb{G}_m = B/\!\!/_a\mathbb{G}_m$.*

(3) *For $a_1 < a_2$ we have $B^{\mathrm{sst}}_{a_1} \subset B^{\mathrm{sst}}_{a_2}$ precisely when $B \cap \mathbb{P}\big(\bigoplus_{a=a_1}^{a_2-1} E_a\big) = \varnothing$, and similarly $B^{\mathrm{sst}}_{a_1} \supset B^{\mathrm{sst}}_{a_2}$ precisely when $B \cap \mathbb{P}\big(\bigoplus_{a=a_1+1}^{a_2} E_a\big) = \varnothing$. In particular $B^{\mathrm{sst}}_{a_1} = B^{\mathrm{sst}}_{a_2}$ precisely when $B \cap \mathbb{P}\big(\bigoplus_{a=a_1}^{a_2} E_a\big) = \varnothing$.*

(4) *If $a_1 < a_2$ and $B \cap \mathbb{P}\big(\bigoplus_{a=a_1}^{a_2-1} E_a\big) = \varnothing$, then the inclusion $B^{\mathrm{sst}}_{a_1} \subset B^{\mathrm{sst}}_{a_2}$ induces a projective morphism $B^{\mathrm{sst}}_{a_1}/\!\!/\mathbb{G}_m \to B^{\mathrm{sst}}_{a_2}/\!\!/\mathbb{G}_m$. Similarly if $B \cap \mathbb{P}\big(\bigoplus_{a=a_1+1}^{a_2} E_a\big) = \varnothing$ we have a projective morphism $B^{\mathrm{sst}}_{a_1}/\!\!/\mathbb{G}_m \leftarrow B^{\mathrm{sst}}_{a_2}/\!\!/\mathbb{G}_m$.*

This time we obtain a "wall and chamber decomposition" of the interval $[a_{\min}(B), a_{\max}(B)]$. We denote the "walls", namely the values of a for which $B \cap \mathbb{P}(E_a) \neq \varnothing$, by $a_{\min}(B) = a_0 < a_1 < \cdots < a_m = a_{\max}(B)$.

By replacing the embedding $B \subset \mathbb{P}(E)$ by the Veronese re-embedding $B \subset \mathbb{P}(\mathrm{Sym}^2 E)$ we may, and will, assume:

Assumption 4.2.2. $$a_i + 1 < a_{i+1}.$$

We set $B^{\rm sst}_{a_i+} = B^{\rm sst}_{a_i+1}$ and $B^{\rm sst}_{a_i-} = B^{\rm sst}_{a_i-1}$, and note that $B^{\rm sst}_{a_i+} = B^{\rm sst}_{a_{i+1}-}$. Assumption 4.2.2 implies that now we always have projective morphisms $\varphi_{a_i\pm}$:

$$\begin{tikzcd}[column sep=small] B^{\rm sst}_{a_i-}/\mathbb{G}_m \arrow[rr, dashed, "\varphi_i"] \arrow[dr, "\varphi_{a_i-}"'] & & B^{\rm sst}_{a_i+}/\mathbb{G}_m = B^{\rm sst}_{a_{i+1}-}/\mathbb{G}_m \arrow[dl, "\varphi_{a_i+}"] \arrow[drr, "\varphi_{a_{i+1}-}"] & & \cdots \\ & B^{\rm sst}_{a_i}/\!\!/\mathbb{G}_m & & & B^{\rm sst}_{a_{i+1}}/\!\!/\mathbb{G}_m \end{tikzcd} \tag{4}$$

Finally, we will assume the following:

Assumption 4.2.3. *Each irreducible component of B meets both $\mathbb{P}(E_{a_{\min}(B)})$ and $\mathbb{P}(E_{a_{\max}(B)})$.*

Under this assumption the quotients $B^{\rm sst}_a/\!\!/\mathbb{G}_m$ are all birational to each other, as long as $a_{\min}(B) < a < a_{\max}(B)$. For the extreme values we have isomorphisms

$$B \cap \mathbb{P}(E_{a_{\min}(B)}) \to B^{\rm sst}_{a_{\min}(B)}/\!\!/\mathbb{G}_m,$$
$$B \cap \mathbb{P}(E_{a_{\max}(B)}) \to B^{\rm sst}_{a_{\max}(B)}/\!\!/\mathbb{G}_m.$$

Remark 4.2.4. As in Remark 4.1.3, it can be shown that $B^{\rm sst}_a \to B^{\rm sst}_a/\!\!/\mathbb{G}_m$ is universally submersive, and if $B \cap \mathbb{P}(E_a) = \varnothing$ we have a universal geometric quotient $B^{\rm sst}_a \to B^{\rm sst}_a/\mathbb{G}_m$.

4.3. *Definition of a birational cobordism.* The notion of a birational cobordism for a blowing up we use in this paper extends the notion of *compactified relatively projective embedded birational cobordism of* [Abramovich et al. 2002, 2.4] by allowing a nonempty boundary. Ignoring the issue of the boundary, it is far more restrictive than the notion introduced in [Włodarczyk 2000].

Let $\phi : X_1 \to X_2$ be an object of the category Bl (Definition 1.3.1). A birational cobordism for ϕ is a scheme B which is the blowing up of a $\mathbb{G}_m$-invariant ideal on $\mathbb{P}^1_{X_2}$, and embedded, in a manner satisfying Assumptions 4.2.2 and 4.2.3, as a $\mathbb{G}_m$-stable subscheme in $\mathbb{P}(E)$ for a $\mathbb{G}_m$-sheaf E on X_2, such that

(1) $X'_1 = B^{\rm sst}_{a_0+}/\mathbb{G}_m = B^{\rm sst}_{a_0}/\!\!/\mathbb{G}_m$ is obtained from X_1 by principalizing D_1,

(2) $X'_2 = B^{\rm sst}_{a_m-}/\mathbb{G}_m = B^{\rm sst}_{a_m}/\!\!/\mathbb{G}_m$ is obtained from X_2 by principalizing D_2, and

(3) the following diagram of rational maps commutes:

$$\begin{tikzcd} B^{\rm sst}_{a_0} \arrow[r, "q_{a_0}"] \arrow[d, dashed, "\alpha"'] & X'_1 \arrow[r] & X_1 \arrow[d, "\phi"] \\ B^{\rm sst}_{a_m} \arrow[r, "q_{a_m}"] & X'_2 \arrow[r] & X_2 \end{tikzcd}$$

where α is the birational map induced by the open dense inclusions

$$B^{\rm sst}_{a_0} \subset B \supset B^{\rm sst}_{a_m}.$$

The birational cobordism is said to respect the open set $U \subset X_2$ if U is contained in the image of $(B^{\rm sst}_{a_0+} \cap B^{\rm sst}_{a_m-})/\mathbb{G}_m$. This happens whenever the ideal on $\mathbb{P}^1_{X_2}$ whose blowing up is B restricts to the unit ideal on $\mathbb{P}^1_U$. We say that a birational cobordism B of ϕ is *regular* if B is regular and the preimage D_B of D_2 is a simple normal crossings divisor.

4.4. ***Construction of regular birational cobordism.*** We claim that one can associate a regular birational cobordism to any blowing up in Bl functorially, and we formalize this claim as follows. There is an evident category $\mathrm{Cob}_{\mathrm{rs}}$ of regular birational cobordisms of blowings up $\phi : X_1 \to X_2$ in Bl, with an evident forgetful functor $\mathrm{Cob}_{\mathrm{rs}} \to \mathrm{Bl}$. A morphism of regular birational cobordisms $B' \to B$ is uniquely determined by a regular surjective morphism $g : X'_2 \to X_2$.

Proposition 4.4.1. *The functor* $\mathrm{Cob}_{\mathrm{rs}} \to \mathrm{Bl}$ *has a section* $\mathrm{Bl} \to \mathrm{Cob}_{\mathrm{rs}}$.

We provide a sketch of proof here, and more detail in Appendix A.

Sketch of proof. Following the construction of [Abramovich et al. 2002, Theorem 2.3.1], consider the blowing up of the ideal $I \otimes \mathcal{O}_{\mathbb{P}^1_{X_2}} + I_{\{0\}}$. This is a birational cobordism B_I for ϕ, but it may be singular. Let $D_{B_I} \subset B_I$ be the preimage of D_2. Applying resolution of pairs to (B_I, D_{B_I}) we obtain a regular birational cobordism (B, D_B) for ϕ. Here we use Theorem 2.2.10 if the characteristic is 0, and parts (1) and (2) with $G = \mathbb{G}_m$ of the Hypothetical Statement 2.2.13 otherwise. $\square$

5. Factoring the map

Throughout this section "functorial" means "functorial in $X_1 \to X_2$ with respect to surjective regular morphisms". By *total transform* of a divisor $D \subset X$ under a (normalized) blowing up $\mathrm{Bl}_J(X) \to X$ we mean the union of the preimage of D and the total transform of J.

5.1. ***Initial factorization.*** Proposition 4.4.1 provides a functorial birational cobordism (B, D_B) of ϕ. Departing slightly from the notation of [Abramovich et al. 2002, Theorem 2.6.2], we write $W_{i\pm} = B^{\mathrm{sst}}_{a_i\pm}/\mathbb{G}_m$, and $W_i = B^{\mathrm{sst}}_{a_i}/\!\!/\mathbb{G}_m$. Since $W_{i+} \simeq W_{(i+1)-}$ we have a functorial factorization

$$\begin{tikzcd}[column sep=small]
 & W_{1-} \arrow[dr, "\varphi_{1-}"] & & W_{2-} \arrow[dl, "\varphi_{1+}"'] \arrow[dr, "\varphi_{2-}"] & & & & W_{m-} \arrow[dl, "\varphi_{(m-1)+}"'] \arrow[dr, equal, "\varphi_{m-}"] & \\
X'_1 = W_0 \arrow[ur, equal, "\varphi_{0+}"] & & W_1 & & {} & \cdots & {} & & W_m = X'_2
\end{tikzcd} \tag{5}$$

with all terms functorially projective over X_2. Since the cobordism is compatible with U, the morphisms $W_{i\pm} \to X_2$ and $W_i \to X_2$ and hence also the morphisms $\varphi_{i\pm}$ are isomorphisms on U. Note that since $W_{m-1} \dashrightarrow W_m$ is a morphism it follows that $\varphi_{(m-1)+}$ is an isomorphism, but this fact does not feature in our arguments. In general the terms W_i and $W_i\pm$ in this factorization are singular, but we will use them to construct a nonsingular factorization.

5.2. ***Blowing up torific ideals.***

5.2.1. *Torific ideals.* Let $D_i \subset W_i$, $D_{i\pm} \subset W_{i\pm}$, $D_{a_i} \subset B^{\mathrm{sst}}_{a_i}$ and $D_{a_i\pm} \subset B^{\mathrm{sst}}_{a_i\pm}$ denote the preimages of D_2. We will show how main results of [Abramovich and Temkin 2017] imply that since (W_i, D_i) is given as a quotient of $(B^{\mathrm{sst}}_{a_i}, D_{a_i})$, it can be made toroidal by a canonical torific blowing up. Since B is regular and

D_B is a simple normal crossings divisor, $(B^{\rm sst}_{a_i}, D_{a_i})$ is a toroidal scheme with a relatively affine $\mathbb{G}_m$-action. In [loc. cit., §5.4.1] one functorially associates to $(B^{\rm sst}_{a_i}, D_{a_i})$ a $\mathbb{G}_m$-equivariant *normalized torific ideals* J_i^B and J_i on $B^{\rm sst}_{a_i}$ and W_i, respectively. By abuse of language, the ideal sheaves $J_{i\pm} = J_i\mathcal{O}_{W_{i\pm}}$ will also be called normalized torific ideals.

Theorem 5.2.2. *For every $1 \le i \le m-1$ the ideal sheaves J_i and $J_{i\pm}$ are functorial and restrict to the unit ideal on U. Furthermore, let $W_i^{\rm tor} = \mathrm{Bl}_{J_i} W_i$ and $W_{i\pm}^{\rm tor} = \mathrm{Bl}_{J_{i\pm}} W_{i\pm}$, and denote by $D_i^{\rm tor} \subset W_i^{\rm tor}$ and $D_{i\pm}^{\rm tor} \subset W_{i\pm}^{\rm tor}$ the total transforms of D_i and $D_{i\pm}$, respectively. Then*

(1) $(W_i^{\rm tor}, D_i^{\rm tor})$ *and* $(W_{i\pm}^{\rm tor}, D_{i\pm}^{\rm tor})$ *are toroidal, and*

(2) *the morphisms $\varphi_{i\pm}$ induce toroidal morphisms*

$$\varphi_{i\pm}^{\rm tor} : (W_{i\pm}^{\rm tor}, D_{i\pm}^{\rm tor}) \to (W_i^{\rm tor}, D_i^{\rm tor})$$

that restrict to isomorphisms on U.

Proof. The ideals J_i are functorial by [Abramovich and Temkin 2017, Theorem 1.1.2(iii)]; hence $J_{i\pm}$ are functorial too. Since the action of $\mathbb{G}_m$ on $B^{\rm sst}_{a_i}$ is already toroidal on $\mathbb{P}^1_U$, we know by [loc. cit., Theorem 1.1.2(iv)] that the J_i restrict to the unit ideal of U.

By [loc. cit., Lemma 4.2.12] $\mathbb{G}_m$ acts in a relatively affine way on $B^{\rm tor}_{a_i} := \mathrm{Bl}_{J_i^B}(B^{\rm sst}_{a_i})$. Let $D^{\rm tor}_{a_i} \subset B^{\rm tor}_{a_i}$ be the total transform of D_{a_i}; then by [loc. cit., Theorem 1.1.2], $(B^{\rm tor}_{a_i}, D^{\rm tor}_{a_i})$ is a toroidal scheme with toroidal action of $\mathbb{G}_m$, and $W_i^{\rm tor} = B^{\rm tor}_{a_i}/\!\!/\mathbb{G}_m$. Note that $D_i^{\rm tor}$ is the image of $D^{\rm tor}_{a_i}$; hence $(W_i^{\rm tor}, D_i^{\rm tor})$ is toroidal by [loc. cit., Theorem 1.1.3(i)].

By [loc. cit., Lemma 5.5.5], $W_{i\pm}^{\rm tor} = (B^{\rm tor}_{a_i})_\pm/\!\!/\mathbb{G}_m$. Set $(D^{\rm tor}_{a_i})_\pm = D^{\rm tor}_{a_i}|_{(B^{\rm tor}_{a_i})_\pm}$; then $\mathbb{G}_m$ acts toroidally on $((B^{\rm tor}_{a_i})_\pm, (D^{\rm tor}_{a_i})_\pm)$ and hence the quotient $(W_{i\pm}^{\rm tor}, D_{i\pm}^{\rm tor})$ is toroidal by [loc. cit., Theorem 1.1.3(i)]. Note also that $\varphi_{i\pm}$ induce toroidal morphisms $\varphi_{i\pm}^{\rm tor}$ by [loc. cit., Proposition 5.5.2]. □

We note that in general $W_{i+}^{\rm tor} \ne W_{(i+1)-}^{\rm tor}$. The steps $W_{i-} \to W_i \leftarrow W_{i+}$ in the factorization (5) now look as follows:

$$\begin{tikzcd}
 & W_{i-}^{\rm tor} \arrow[dl] \arrow[dr, "\varphi_{i-}^{\rm tor}"] & & W_{i+}^{\rm tor} \arrow[dl, "\varphi_{i+}^{\rm tor}"'] \arrow[dr] & \\
W_{(i-1)+} = W_{i-} \arrow[drr, "\varphi_{i-}"] & & W_i^{\rm tor} \arrow[d] & & W_{i+} = W_{(i+1)-} \arrow[dll, "\varphi_{i+}"'] \\
 & & W_i & &
\end{tikzcd} \tag{6}$$

Remark 5.2.3. In [Abramovich et al. 2002, Lemma 3.2.8] it is stated with a sketch of proof that the ideals J_i can be chosen so that $\varphi_{i\pm}^{\rm tor}$ are isomorphisms. We will not use this statement. We note however that this follows from [Thaddeus 1996, Theorem 3.5]: if the l-torific ideal I_l generates all I_{Ml}, $M \ge 1$, and also I_{-l} generates all I_{-Ml}, $M \ge 1$, then once $l, -l \in S_i$, where S_i is the ample set of characters on $B^{\rm sst}_{a_i}$ used to determine J_i^B in [Abramovich and Temkin 2017], we have $\varphi_{i\pm}^{\rm tor}$ are isomorphisms. One can choose such l in a manner functorial for regular surjective morphisms.

5.3. *Resolution and local charts.*

5.3.1. *Canonical resolution.* Extending the notation of [Abramovich et al. 2002, Section 4.2] to qe schemes with a boundary, we write $W_{i\pm}^{\mathrm{res}} \to W_{i\pm}$ for the resolution of the pair $(W_{i\pm}, D_{i\pm})$ and denote the preimage of D_2 in $W_{i\pm}^{\mathrm{res}}$ by $D_{i\pm}^{\mathrm{res}}$. This morphism is functorially projective and is projectively the identity on U. In characteristic 0 we use Theorem 2.2.10, and otherwise we invoke Hypothetical Statement 2.2.13(1). Thus, $W_{i\pm}^{\mathrm{res}}$ is regular and $D_{i\pm}^{\mathrm{res}}$ is a simple normal crossings divisor.

Note that the resolution process is independent of the toroidal structures and hence coincides for $(W_{(i-1)+}, D_{(i-1)+}) = (W_{i-}, D_{i-})$. Thus, $(W_{(i-1)+}^{\mathrm{res}}, D_{(i-1)+}^{\mathrm{res}}) = (W_{i-}^{\mathrm{res}}, D_{i-}^{\mathrm{res}})$ and this provides a bridge between $W_{(i-1)+}^{\mathrm{tor}}$ and W_{i-}^{tor}:

$$W_{(i-1)+}^{\mathrm{tor}} \dashrightarrow W_{(i-1)+}^{\mathrm{res}} = W_{i-}^{\mathrm{res}} \dashleftarrow W_{i-}^{\mathrm{tor}}$$

Remark 5.3.2. Since $W_{1-} = X_1'$ is regular, $X_1'' := W_{1-}^{\mathrm{res}}$ is obtained from X_1' by principalization of D_1' and similarly $X_2'' := W_{m-}^{\mathrm{res}}$ is obtained from X_2' by principalization of D_2'. Both D_1' and D_2' are simple normal crossings divisors, so we could alternatively take $W_{1-}^{\mathrm{res}} = X_1'$ and $W_{m-}^{\mathrm{res}} = X_m'$. Our choice above helps to make notation uniform, though it results in a slightly longer factorization.

Remark 5.3.3. The singularities requiring resolution in this step are far from general: it is shown in the proof of Lemma 5.3.7 below that Zariski locally one can obtain a toroidal scheme from $(W_{i\pm}, D_{i\pm})$ simply by enlarging the divisor $D_{i\pm}$. At least over an algebraically closed field they admit resolution of singularities, see [Włodarczyk 2003, Theorem 8.3.2], and it seems reasonable to expect the same in general, and in a functorial manner.

5.3.4. *Localization.* In Section 5.4 we will connect $W_{i\pm}^{\mathrm{res}}$ and $W_{i\pm}^{\mathrm{tor}}$ by principalizing the ideal $J_{i\pm}^{\mathrm{res}} := J_{i\pm}\mathcal{O}_{W_{i\pm}^{\mathrm{res}}}$, but to use our principalization conjectures in positive and mixed characteristics we should first check that $J_{i\pm}^{\mathrm{res}}$ is locally monoidal, so we start with defining local toroidal charts of all our constructions. We will work locally at a point $x \in W_{i\pm}$, so consider the localization $W_x := \operatorname{Spec} \mathcal{O}_{W_{i\pm},x}$. We set $W_x^{\mathrm{res}} = W_{i\pm}^{\mathrm{res}} \times_{W_{i\pm}} W_x$ and similarly for W_x^{tor} and other $W_{i\pm}$-schemes we will introduce later. For brevity, we also set $B_x = B_{a_i\pm}^{\mathrm{sst}} \times_{W_{i\pm}} W_x$, $D_{B_x} = D_{a_i\pm} \times_{B_{a_i\pm}^{\mathrm{sst}}} W_x$, and $D_x = D_{i\pm} \times_{W_{i\pm}} W_x$. We use the terminology of [Abramovich and Temkin 2018] regarding strictly local actions and strongly equivariant morphisms, and of [Abramovich and Temkin 2017] regarding simple actions and toroidal actions.

5.3.5. *Local toroidal charts.* The action of $\mathbb{G}_m$ on B_x is simple since $\mathbb{G}_m$ is connected and local since $B_x /\!\!/ \mathbb{G}_m = W_x$. Let O be the closed orbit of B_x and $G_O = \operatorname{Spec}(\mathbb{Z}[L_O])$ its stabilizer. Note that O is a torsor under the $k(x)$-group-scheme $\boldsymbol{D}_{K_O} := \operatorname{Spec} k(x)[K_O]$ with $K_O = \operatorname{Ker}(\mathbb{Z} \twoheadrightarrow L_O)$. We have two possibilities: (1) O is a point (i.e., the action is strictly local), $G_O = \mathbb{G}_m$, and $L_O = \mathbb{Z}$, or (2) the orbit is a torus, $G_O = \mu_n$, and $L_O = \mathbb{Z}/n\mathbb{Z}$. For a toric monoid P we will use the notation $\boldsymbol{A}_P = \operatorname{Spec} \mathbb{Z}[P]$ and $\bar{E}_P = \boldsymbol{A}_P \setminus \boldsymbol{A}_{P^{\mathrm{gp}}}$. By [Abramovich and Temkin 2017, Theorem 3.6.11] there exists a strongly equivariant strict morphism $h : (B_x, D_{B_x}) \to (\boldsymbol{A}_P, E_P)$, with a suitable $\mathbb{Z}$-graded toric monoid of the form $P = \overline{M}_O \oplus K_O \oplus \mathbb{N}^{\sigma_O}$ and $E_P = \boldsymbol{A}_P \smallsetminus \boldsymbol{A}_{\overline{M}_O^{\mathrm{gp}} \oplus K_O \oplus \mathbb{N}^{\sigma_O}}$. Note that the action on $(\boldsymbol{A}_P, E_P)$ is not toroidal, but it becomes toroidal if we enlarge the toroidal structure to $\overline{E}_P$.

5.3.6. *The quotient charts.* Let $M = P_0$ be the trivially graded part of P. Then $Y := \boldsymbol{A}_M = \boldsymbol{A}_P /\!/ \mathbb{G}_m$ and we consider the divisor $E = E_P /\!/ \mathbb{G}_m$ on Y, which is a subdivisor of the toroidal divisor $\overline{E}_M = \overline{E}_P /\!/ \mathbb{G}_m$. The $\mathbb{G}_m$-action on $(\boldsymbol{A}_P, E_P)$ gives rise to the normalized torific ideal J_Y on Y, and let $Y^{\mathrm{tor}} \to Y$ be the blowing up along J_Y. By [Abramovich and Temkin 2017, Theorem 1.1.2(iii)], the torifications of Y and W_x are compatible with respect to the quotient morphism $h /\!/ G : W_x \to Y$; namely, $J_{W_x} = J_{i\pm}|_{W_x}$ coincides with $J_Y \mathcal{O}_{W_x}$ and $W_x^{\mathrm{tor}} = W_x \times_Y Y^{\mathrm{tor}}$.

In addition, consider the resolution $Y^{\mathrm{res}} \to Y$ of the pair (Y, E) as defined in Theorem 2.2.10 and Hypothetical Statement 2.2.13(1). Since the resolution is $\boldsymbol{A}_{M^{\mathrm{gp}}}$-equivariant, Y^{res} is a toric scheme too. Recall that the resolution is compatible with toroidal charts: this follows from the functoriality if X is defined over a field, and we use Hypothetical Statement 2.2.13(3) in mixed characteristics. Therefore, $W_x^{\mathrm{res}} = W_x \times_Y Y^{\mathrm{res}}$ and the ideal $J_{i\pm}^{\mathrm{res}} = J_{i\pm} \mathcal{O}_{W_x^{\mathrm{res}}} = J_Y \mathcal{O}_{W_x^{\mathrm{res}}}$ comes from the ideal $J_Y^{\mathrm{res}} = J_Y \mathcal{O}_{Y^{\mathrm{res}}}$ on Y^{res}.

Lemma 5.3.7. *The ideal $J_{i\pm}^{\mathrm{res}}$ is locally monoidal.*

Proof. We will work locally at $x \in W_{i\pm}$. Let $\overline{D}_{B_x} \subset B_x$ and $\overline{D}_x \subset W_x$ be the preimages of $\overline{E}_P$ and $\overline{E}_M$, respectively. Since h is strongly equivariant, the induced morphism $\bar{h} : (B_x, \overline{D}_{B_x}) \to (\boldsymbol{A}_P, \overline{E}_P)$ is a strongly equivariant toroidal chart. The action on the target of $\bar{h}$ is toroidal; hence the action on the source is toroidal by [Abramovich and Temkin 2017, Lemma 3.1.9(iv)] and $\bar{h} /\!/ G : (W_x, \overline{D}_x) \to (Y, \overline{E}_M)$ is a toroidal chart by [loc. cit., Theorem 1.1.3(iii)].

The resolution $Y^{\mathrm{res}} \to Y$ is $\boldsymbol{A}_{M^{\mathrm{gp}}}$-equivariant; hence it is obtained by blowing up a toroidal ideal, and if $\overline{E}^{\mathrm{res}}$ denotes the total transform of $\overline{E}_M$ then the morphism $(Y^{\mathrm{res}}, \overline{E}^{\mathrm{res}}) \to (Y, \overline{E}_M)$ is toroidal. In addition, the pullback of $\bar{h} /\!/ G$ gives rise to a toroidal chart $g : (W_x^{\mathrm{res}}, \overline{D}_x^{\mathrm{res}}) \to (Y^{\mathrm{res}}, \overline{E}^{\mathrm{res}})$ with $D_x^{\mathrm{res}} \subseteq \overline{D}_x^{\mathrm{res}}$. Since the action on $(Y, \overline{E}_M)$ is toroidal, the ideal J_Y is toroidal with respect to $\overline{E}_M$ by [loc. cit., Lemma 4.4.5(i)]. Thus, J_Y^{res} is toroidal with respect to $\overline{E}^{\mathrm{res}}$ and hence its pullback $J_{i\pm}^{\mathrm{res}}$ is toroidal with respect to $\overline{D}_x^{\mathrm{res}}$. The lemma follows. ☐

5.4. *Tying the maps together.*

5.4.1. *Principalization of torific ideals.* Thanks to Lemma 5.3.7 we can define $W_{i\pm}^{\mathrm{can}}$ to be the canonical principalization of $J_{i\pm}^{\mathrm{res}}$ in the sense of Section 2.3. It is obtained by a *functorial sequence of blowings up of nonsingular centers disjoint from U starting from $W_{i\pm}^{\mathrm{res}}$*; see Proposition 2.3.9. In positive and mixed characteristics we require Hypothetical Statement 2.3.6.

By the universal property of blowing up, the maps $W_{i\pm}^{\mathrm{can}} \dashrightarrow W_{i\pm}^{\mathrm{tor}}$ are morphisms. The map $W_{i\pm}^{\mathrm{can}} \to W_i$ is a composition of maps given functorially by blowing up ideals restricting to the unit ideal on U. By Section 2.1.10 the morphism $W_{i\pm}^{\mathrm{can}} \to W_i$ itself is given by blowing up a functorial ideal $\widetilde{J}_{i\pm}^{\mathrm{can}}$ restricting to the unit ideal on U. So, by Lemma 2.1.9 the morphism $W_{i\pm}^{\mathrm{can}} \to W_{i\pm}^{\mathrm{tor}}$ is given by blowing up the functorial ideal $J_{i\pm}^{\mathrm{can}} = \widetilde{J}_{i\pm}^{\mathrm{can}} \mathcal{O}_{W_{i\pm}^{\mathrm{tor}}}$. By $D_{i\pm}^{\mathrm{can}}$ we denote the total transform of $D_{i\pm}^{\mathrm{tor}}$. Diagram (6) now looks as follows:

$$
\begin{tikzcd}[column sep=small]
 & & & W^{\mathrm{can}}_{i-} \arrow[dl] \arrow[d] & & W^{\mathrm{can}}_{i+} \arrow[d] \arrow[dr] & & & \\
W^{\mathrm{res}}_{(i-1)+} \arrow[dd] & = & W^{\mathrm{res}}_{i-} \arrow[dd] & W^{\mathrm{tor}}_{i-} \arrow[ddl] \arrow[dr, "\varphi^{\mathrm{tor}}_{i-}"] & & W^{\mathrm{tor}}_{i+} \arrow[dl, "\varphi^{\mathrm{tor}}_{i+}"'] \arrow[ddr] & W^{\mathrm{res}}_{i+} \arrow[dd] & = & W^{\mathrm{res}}_{(i+1)-} \arrow[dd] \\
 & & & & W^{\mathrm{tor}}_i \arrow[dd] & & & & \\
W_{(i-1)+} & = & W_{i-} \arrow[drr, "\varphi_{i-}"] & & & & W_{i+} \arrow[dll, "\varphi_{i+}"'] & = & W_{(i+1)-} \\
 & & & & W_i & & & &
\end{tikzcd}
\tag{7}
$$

Lemma 5.4.2. *The ideal $J^{\mathrm{can}}_{i\pm}$ is toroidal. Thus, $(W^{\mathrm{can}}_{i\pm}, D^{\mathrm{can}}_{i\pm}) \to (W^{\mathrm{tor}}_{i\pm}, D^{\mathrm{tor}}_{i\pm})$ is a functorial toroidal blowing up.*

Proof. Step 1: reduction to toric case. We will work locally at $x \in W_{i\pm}$. We already used in Section 5.3.6 that torification and resolution are compatible with toroidal charts to show, in the notation introduced there, that $W^{\mathrm{tor}}_x = W_x \times_Y Y^{\mathrm{tor}}$, $W^{\mathrm{res}}_x = W_x \times_Y Y^{\mathrm{res}}$ and $J^{\mathrm{res}}_x = J^{\mathrm{res}}_Y \mathcal{O}_{W^{\mathrm{res}}_x}$. Let $Y^{\mathrm{can}} \to Y^{\mathrm{res}}$ be the principalization of J^{res}_Y. Then by the same functoriality argument $W^{\mathrm{can}}_x = W_x \times_Y Y^{\mathrm{can}}$.

By the universal property of blowings up, $Y^{\mathrm{can}} \to Y$ factors through Y^{tor}. We have $Y^{\mathrm{can}} = \mathrm{Bl}_{\widetilde{J}^{\mathrm{can}}_Y}(Y)$ for a functorial ideal $\widetilde{J}^{\mathrm{can}}$ on Y; hence by Lemma 2.1.9, $Y^{\mathrm{can}} = \mathrm{Bl}_{J^{\mathrm{can}}_Y}(Y^{\mathrm{tor}})$, where $J^{\mathrm{can}}_Y = \widetilde{J}^{\mathrm{can}}_Y \mathcal{O}_{Y^{\mathrm{tor}}}$. Again, the construction of the ideals $J^{\mathrm{can}}_{\pm}$ is compatible with charts. So $J^{\mathrm{can}}_{i\pm}\mathcal{O}_{W^{\mathrm{can}}_x}$ is the pullback of J^{can}_Y. Thus, it suffices to prove that the ideal J^{can}_Y is toroidal.

Step 2: proof in the toric case. It is shown in [Abramovich et al. 2002, Proposition 4.2.1] that $(Y^{\mathrm{can}}, E^{\mathrm{can}}) \to (Y^{\mathrm{tor}}, E^{\mathrm{tor}})$ is toroidal: here we produce this morphism by blowing up the normalized toroidal ideals of [Abramovich and Temkin 2018] instead of the torific ideal of [Abramovich et al. 2002], but these morphisms have the same equivariance properties. In [Abramovich et al. 2002] the ideal blown up is not shown to be toroidal. This can be shown as follows. As in [loc. cit., Proposition 4.2.2] one constructs an action of $\mathbb{G}_a^k$ on (Y, E). One shows that the morphism $Y^{\mathrm{tor}} \to Y$ of charts is equivariant under this action, as well as the normalized torific ideal J_Y; the scheme Y^{tor} is written as a product of $\mathbb{G}_a^k$ with a toric scheme providing its toroidal structure. It suffices to show that the ideal defining the blowing up $Y^{\mathrm{can}} \to Y^{\mathrm{tor}}$ is a $\mathbb{G}_a^k$-equivariant monomial ideal, since then its generating monomials are not divisible by the coordinates of the $\mathbb{G}_a^k$ factor.

Since the blowing up $Y^{\mathrm{res}} \to Y$ is the canonical resolution of singularities of (Y, E), the ideal defining this blowing up on a toric chart is monomial and $\mathbb{G}_a^k$-equivariant. Also the torific ideal on Y^{res} is monomial and $\mathbb{G}_a^k$-equivariant; therefore the same is true for the ideal defining its functorial principalization $Y^{\mathrm{can}} \to Y^{\mathrm{res}}$, as required. Note that in the case of nonzero characteristic we have used $\mathbb{G}_a^k$-equivariance from Hypothetical Statements 2.2.13 and 2.3.6. □

The above lemma implies that the composition $W^{\mathrm{can}}_{i\pm} \to W^{\mathrm{tor}}_i$ is a toroidal morphism given by blowing up a functorial toroidal ideal we denote by $\bar{J}^{\mathrm{can}}_{i\pm}$. Let $W'_i \to W^{\mathrm{tor}}_i$ be the normalized blowing up of the product ideal $\bar{J}^{\mathrm{can}}_{i-}\bar{J}^{\mathrm{can}}_{i+}$, giving rise to toroidal morphisms $W'_i \to W^{\mathrm{can}}_{i\pm}$. By [Illusie and Temkin 2014, Theorem 3.4.9] there is a functorial toroidal resolution of singularities $W^{\mathrm{tor\,res}}_i \to W'_i$. This gives the following:

Lemma 5.4.3. *There is a toroidal nonsingular modification $W^{\mathrm{tor\,res}}_i \to W^{\mathrm{tor}}_i$ obtained by blowing up a functorial ideal such that the maps $W^{\mathrm{tor\,res}}_i \dashrightarrow W^{\mathrm{can}}_{i\pm}$ are both toroidal morphisms.*

Note that these latter maps are again blowings up of the pullbacks of the ideal defining $W^{\mathrm{tor\,res}}_i \to W^{\mathrm{tor}}_i$, which is functorial as well. Since the morphism is toroidal, it induces the identity on U, and the toroidal ideal blown up is the unit ideal on U.

We now have pieces of the diagram above looking as follows:

$$\begin{tikzcd}
 & & & W^{\mathrm{tor\,res}}_i \arrow[dl, "(\mathrm{TorBl_{rs}})"'] \arrow[dr, "(\mathrm{TorBl_{rs}})"] & & & \\
 & & W^{\mathrm{can}}_{i-} \arrow[dl, "(\text{blow-up sequence})"'] \arrow[dr] & & W^{\mathrm{can}}_{i+} \arrow[dl] \arrow[dr, "(\text{blow-up sequence})"] & & \\
W^{\mathrm{res}}_{(i-1)+} \arrow[r, equal] \arrow[d] & W^{\mathrm{res}}_{i-} \arrow[d] & & W^{\mathrm{tor}}_i \arrow[dd] & & W^{\mathrm{res}}_{i+} \arrow[r, equal] \arrow[d] & W^{\mathrm{res}}_{(i+1)-} \arrow[d] \\
W_{(i-1)+} \arrow[r, equal] & W_{i-} \arrow[drr] & & & & W_{i+} \arrow[r, equal] \arrow[dll] & W_{(i+1)-} \\
 & & & W_i & & &
\end{tikzcd}$$

All maps are functorially the blowings up of ideals. The top diamond is at the same time toroidal, with maps given by blowings up of functorial toroidal ideals, so the toroidal structure is functorial in $X_1 \to X_2$. By Proposition 3.1.1, the two top maps $W^{\mathrm{tor\,res}}_i \to W^{\mathrm{can}}_{i\pm}$ have a functorial toroidal weak factorization; since it is toroidal it induces isomorphisms on U. This gives a factorization of the top diamond of the diagram above as follows:

$$\begin{tikzcd}
 & W^{\mathrm{tor\,res}}_i \arrow[dl, "(\mathrm{TorFact_{rs}})"'] \arrow[dr, "(\mathrm{TorFact_{rs}})"] & \\
W^{\mathrm{can}}_{i-} \arrow[dr] & & W^{\mathrm{can}}_{i+} \arrow[dl] \\
 & W^{\mathrm{tor}}_i &
\end{tikzcd}$$

Note that $W^{\mathrm{res}}_{1-} = X''_1$ and $W^{\mathrm{res}}_{m-} = X''_2$ by Remark 5.3.2. By construction, $X''_i \to X'_i$ and $X'_i \to X_i$ are resolutions of normal crossings pairs (X'_i, D'_i) and (X_i, D_i), respectively; hence $X''_i \to X_i$ factor as sequences of blowings up of regular centers compatible with U_i and D_i thanks to Assumption 2.2.7. Putting these together we functorially obtain a diagram:

$$\begin{tikzcd}
X''_1 \arrow[d, "(\mathrm{Fact})"'] \arrow[dr, equal] & & W^{\mathrm{tor\,res}}_1 \arrow[dl, "(\mathrm{Fact})"'] \arrow[dr, "(\mathrm{Fact})"] & & & & W^{\mathrm{tor\,res}}_{m-1} \arrow[dl, "(\mathrm{Fact})"'] \arrow[dr, "(\mathrm{Fact})"] & & X''_2 \arrow[dl, equal] \arrow[d, "(\mathrm{Fact})"] \\
X_1 \arrow[r, dashed] & W^{\mathrm{res}}_{1-} \arrow[rr, dashed, "\varphi_1"'] & & W^{\mathrm{res}}_{2-} \arrow[r, dashed, "\varphi_2"'] & \cdots \arrow[r, dashed, "\varphi_{m-1}"'] & W^{\mathrm{res}}_{(m-1)-} \arrow[rr, dashed, "\varphi_m"'] & & W^{\mathrm{res}}_{m-} \arrow[r, dashed] & X_2.
\end{tikzcd}$$

Note that W_i are given by blowing up of functorial ideals on X_2, and that $W_{i\pm}^{\mathrm{res}}$ are obtained by blowing up functorial ideals on W_i, all restricting to the identity on U. Similarly, the terms appearing in the diagonal arrows are given by blowing up of functorial ideals on $W_{i\pm}^{\mathrm{res}}$. By the result of Section 2.1.10 all terms appearing are obtained by blowing up of functorial ideals on X_2 restricting to the unit ideal on U. In the case $X_i \setminus U$ are normal crossings divisors, we have guarantees that the same holds for $W_{i\pm}^{\mathrm{res}}$. It follows that the same holds for all terms in the sequence forming $W_{i\pm}^{\mathrm{can}} \to W_{i\pm}^{\mathrm{res}}$ by the properties of canonical principalization, and for the terms in a factorization of $W_i^{\mathrm{tor\,res}} \to W_{i\pm}^{\mathrm{can}}$ since these are all nonsingular toroidal schemes. Renaming all these terms $V_i,\ i = 1, \dots, l$, Theorem 1.3.3 follows. $\square$

5.5. ***Summary of resolution steps.*** Results around resolution of singularities were used in several steps in the proof of Theorem 1.3.3. We recall here these steps and what they require. While our main theorem requires the procedures to be functorial, we emphasize the equivariance and functoriality properties necessary for the factorization theorem to hold even without requiring the factorization to be functorial.

The first resolution process appears in the construction of the birational cobordism in Proposition 4.4.1. This appears explicitly in Step 3a in Appendix A, where we resolve the pair (B_I, D_{B_I}), which has dimension $\dim X_2 + 1$. It is crucial that the process be $\mathbb{G}_m$-equivariant.

In Section 5.3.1 we apply resolution of singularities to $W_{i\pm}$, which has dimension $\dim X_2$. The singularities of $W_{i\pm}$ are all locally monomial. Similarly, in Section 5.4 we apply principalization of the ideals $J_{i\pm}^{\mathrm{res}}$, which are locally monoidal ideals. On the other hand these two steps require the resolution and principalization to be equivariant in a strong sense: Lemma 5.4.2 requires the process to be compatible with toric charts, and the process on the toric schemes must be both torus equivariant and $\mathbb{G}_a^k$-equivariant. Finally, Lemma 5.4.3 requires toroidal resolution of singularities, which is as functorial as one could wish.

6. Extending the factorization to other categories

In this section we use the factorization for schemes to construct an analogous factorization for blowings up of formal schemes, complex and nonarchimedean analytic spaces, and stacks. We follow the general outline of the argument in [Temkin 2008, Sections 5.1–5.2], though we decided to elaborate more details related to the relative GAGA issues. In fact, for this construction to work one only needs to have a reasonable comparison theory between algebraic blowing ups and their analytifications, but some of these results do not seem to be covered by the literature, especially in the complex analytic case.

6.1. ***Stacks.*** Once functorial factorization for schemes is established, it extends to stacks straightforwardly.

6.1.1. *Basic notions.* Our terminology concerning stacks follows that of [Temkin 2008, §5.1]. In particular, by a *stack* we mean an Artin stack $\mathfrak{X}$ and $\mathfrak{X}$ is *qe* (respectively, *regular*) if it admits a smooth covering $W \to \mathfrak{X}$ with W a qe (respectively, a regular) scheme. The definition of blowing up along a closed subscheme is compatible with flat morphisms and hence extends to stacks. We define the regular surjective category of blowings up of stacks $\mathrm{Bl}_{\mathrm{rs}}^{\mathfrak{St}}$ and the regular surjective category of weak factorizations of blowings up of stacks $\mathrm{Fact}_{\mathrm{rs}}^{\mathfrak{St}}$ as in Definitions 1.3.1 and 1.3.2.

6.1.2. *Factorization for stacks.* We are now in position to extend the factorization to stacks.

Theorem 6.1.3. *There is a functor* $\mathrm{Bl}^{\mathfrak{St}}_{\mathrm{rs}}(\mathrm{char}=0) \to \mathrm{Fact}^{\mathfrak{St}}_{\mathrm{rs}}(\mathrm{char}=0)$ *from the regular surjective category of blowings up* $f : \mathfrak{X}' \to \mathfrak{X}$ *in characteristic* 0 *to the regular surjective category of factorizations*

$$\mathfrak{X}' = \mathfrak{X}_0 \dashrightarrow \mathfrak{X}_1 \dashrightarrow \cdots \dashrightarrow \mathfrak{X}_{l-1} \dashrightarrow \mathfrak{X}_l = \mathfrak{X}$$

in characteristic 0 *such that the composite*

$$\mathrm{Bl}^{\mathfrak{St}}_{\mathrm{rs}}(\mathrm{char}=0) \to \mathrm{Fact}^{\mathfrak{St}}_{\mathrm{rs}}(\mathrm{char}=0) \to \mathrm{Bl}^{\mathfrak{St}}_{\mathrm{rs}}(\mathrm{char}=0)$$

is the identity. The same holds in positive and mixed characteristics if Hypothetical Statements 2.2.13 and 2.3.6 hold true.

Proof. Choose a smooth covering of $\mathfrak{X}$ by a qe scheme W. Then W and $R = W \times_{\mathfrak{X}} W$ are regular qe schemes and the projections $p_{1,2} : R \rightrightarrows W$ are surjective and smooth. The pullbacks $W' \to W$ and $R' \to R$ of $\mathfrak{X}' \to \mathfrak{X}$ are objects of Bl; hence Theorem 1.3.3 provides their regular factorizations $(W_\bullet)$ and $(R_\bullet)$. By the functoriality, these factorizations are compatible with both p_1 and p_2. Since both pullbacks of the factorization $(W_\bullet)$ to R coincide, flat descent implies that $(W_\bullet)$ comes from a factorization $(\mathfrak{X}_\bullet)$ of $\mathfrak{X}' \to \mathfrak{X}$.

To see that the factorization $(\mathfrak{X}_\bullet)$ is independent of a smooth covering $W \to \mathfrak{X}$, we note that any smooth covering $W' \to \mathfrak{X}$ that factors through W induces the same factorization of $\mathfrak{X}' \to \mathfrak{X}$, as follows from the functoriality of factorization with respect to the morphism $W' \to W$.

Finally, assume that $(\mathfrak{Y}' \to \mathfrak{Y}) \to (\mathfrak{X}' \to \mathfrak{X})$ is a morphism in $\mathrm{Bl}^{\mathfrak{St}}_{\mathrm{rs}}$. Then there exist smooth coverings by qe schemes $W \to \mathfrak{X}$ and $T \to \mathfrak{Y}$ such that the morphism $\mathfrak{Y} \to \mathfrak{X}$ lifts to a regular surjective morphism $T \to W$. It then follows easily from the functoriality of factorization with respect to $T \to W$ that the factorization for stacks we constructed is compatible with $\mathfrak{Y} \to \mathfrak{X}$. Thus, the factorization for stacks is functorial. □

6.2. *Geometric spaces.*

6.2.1. *Categories.* We will work with the geometric spaces of the following four classes, that will simply be called *spaces*:

(1) qe formal schemes as defined in [Temkin 2008, Section 2.4.3].

(2) Semianalytic germs of complex analytic spaces; see Appendix B.

(3) k-analytic spaces of Berkovich for a complete nonarchimedean field k; see [Berkovich 1993, Section 1].

(3′) Rigid k-analytic spaces, where k is as above and nontrivially valued.

To make notation uniform, the category of all such spaces will be denoted by $\mathfrak{Sp}$ in each of the four cases.

Remark 6.2.2. (i) The case (3′) is added for the sake of completeness. It is essentially included in (3) because the category of qcqs (i.e., quasicompact and quasiseparated) rigid spaces is equivalent to the category of compact strictly analytic Berkovich spaces, and all our arguments will be "local enough".

(ii) Probably, there exist other contexts where our methods apply, e.g., semialgebraic geometry. We do not explore this direction here, but we will deal with the above four cases in a uniform way that should make it simpler for the interested reader to extend our results to other possible settings.

6.2.3. *Affinoid spaces.* We say that a space X is *affinoid* if it is of the following type:

(1) $X = \mathrm{Spf}(A)$ is affine.

(2) $(\mathcal{X}, X)$ is an affinoid germ of a complex analytic space; see Section B.6.

(3) $X = \mathcal{M}(A)$ is an affinoid k-analytic space.

(3′) $X = \mathrm{Sp}(A)$ is an affinoid rigid space over k.

6.2.4. *Admissible affinoid coverings.* To simplify the discussion we consider only affinoid coverings $X = \bigcup_{i\in I} X_i$ of a qcqs space by its affinoid domains. Such a covering is called *admissible* if it possesses a finite refinement. Here is the main property of admissible coverings, which may fail for nonadmissible ones (e.g., the covering of a germ $(\mathcal{X}, X)$ by one-pointed subgerms $(\mathcal{X}, x)$ with $x \in X$).

Lemma 6.2.5. *Assume that $X = \bigcup_{i\in I} X_i$ is an admissible covering of an affinoid space. Then for any coherent $\mathcal{O}_X$-module $\mathcal{F}$ the Čech complex*

$$0 \to \mathcal{F}(X) \to \prod_i \mathcal{F}(X_i) \to \prod_{i,j} \mathcal{F}(X_i \cap X_j) \to \cdots$$

is acyclic.

Proof. For formal schemes this is classical, and for nonarchimedean geometry this is Tate's acyclicity theorem and its extension to Berkovich spaces. It remains to deal with complex germs. It suffices to deal with the case of finite coverings, and then we can replace the direct products with direct sums. Choosing a small enough representative $\mathcal{X}$ of X we can assume that $\mathcal{X}$ is Hausdorff. Choose families of Stein domains $V_0 \supset V_1 \cdots$ and $V_{0i} \supset V_{1i} \cdots$ for each $i \in I$ such that $X = \bigcap_{n=0}^{\infty} V_n$ and $X_i = \bigcap_{n=0}^{\infty} V_{ni}$. For each $n \in \mathbb{N}$ the union $\bigcup_{i\in I} V_{ni}$ is a neighborhood of X and hence it contains some V_m. Let $m = m(n)$ be the minimal number for which the latter happens. The intersections $U_{ni} = V_m \cap V_{ni}$ are Stein domains since $\mathcal{X}$ is Hausdorff; hence V_m is covered by Stein domains U_{ni} and we obtain the acyclic Čech complex

$$0 \to \mathcal{F}(V_m) \to \bigoplus_i \mathcal{F}(U_{ni}) \to \bigoplus_{i,j} \mathcal{F}(U_{ni} \cap U_{nj}) \to \cdots .$$

Since $\lim_{n\to\infty} m(n) = \infty$ and $X_i = \bigcap_n U_{ni}$, passing to the limit on n we obtain the sequence from the formulation of the lemma. It remains to use that the filtered colimit is an exact functor. □

6.2.6. *Regular spaces.* Each category of spaces possesses a natural notion of regular spaces; see [Temkin 2008, Section 5.2.2]. In fact, a space X is regular if it possesses an admissible affinoid covering $X=\bigcup_i X_i$ such that the rings $A_i=\mathcal{O}_X(X_i)$ are regular. In particular, it follows from Lemma B.6.1 that a germ of analytic space $(\mathcal{X}, X)$ is regular if and only if $\mathcal{X}$ is smooth in a neighborhood of X.

By $\mathfrak{Sp}_{\mathrm{reg}}$ we denote the full subcategory of $\mathfrak{Sp}$ consisting of quasicompact regular objects, and we do not impose any separatedness assumption.

6.2.7. *Smooth and regular morphisms.* Also, the category $\mathfrak{Sp}$ has a natural notion of smooth morphisms. In cases (1), (2) and (3′) this is the classical notion (with the obvious adjustment in (2)) and in (3) this is the notion of quasismooth morphisms as defined in [Ducros 2018, Section 5].

In cases (2), (3) and (3′) any morphism is of finite type, so we identify the notions of smooth and regular morphisms. Regular morphisms of qe formal schemes were defined in [Temkin 2008, 2.4.12]: a morphism $f: Y\to X$ is called *regular* if it admits an open covering of the form $f_i:\mathrm{Spf}(B_i)\to\mathrm{Spf}(A_i)$ such that the homomorphisms $A_i\to B_i$ are regular.

Lemma 6.2.8. *If $Y\to X$ is a regular morphism of affinoid spaces in $\mathfrak{Sp}$ then the homomorphism $\mathcal{O}_X(X)\to\mathcal{O}_Y(Y)$ is regular.*

Proof. Case (1) is covered by [Temkin 2008, Lemma 2.4.6]. Case (3), and hence also case (3′), follows from [Ducros 2009, Theorem 3.3; 2018, Theorem 5.5.3] and the fact that for any affinoid space $Z=\mathcal{M}(C)$ the map $Z\to\mathrm{Spec}(C)$ is surjective by [Berkovich 1993, Proposition 2.1.1]. Case (2) is dealt with similarly using that if Z is an affinoid germ, $z\in Z$ and $f:Z\to T=\mathrm{Spec}(\mathcal{O}_Z(Z))$ is the natural map then $f(Z)$ is the set of all closed points and the homomorphism $\mathcal{O}_{T,f(z)}\to\mathcal{O}_{Z,z}$ is regular by Lemma B.6.1. □

6.3. ***Relative GAGA.*** Assume that X is an affinoid space, $A=\mathcal{O}_X(X)$ and $\mathcal{X}=\mathrm{Spec}\,A$. Relative GAGA relates the theory of $\mathcal{X}$-schemes and X-spaces.

6.3.1. *Analytification functor.* There exists an analytification/formal completion functor from $\mathcal{X}$-schemes of finite type to X-spaces. For uniformity, we will usually call this functor *analytification* and set $\mathcal{Y}\mapsto Y=\mathcal{Y}^{\mathrm{an}}$. It is constructed as follows:

(i) The analytification of $\mathbb{A}^n_{\mathcal{X}}$ is $\mathbb{A}^n_X$.

(ii) If $\mathcal{Y}$ is $\mathcal{X}$-affine, say $\mathcal{Y}=\mathrm{Spec}\,B$ with $B=A[t_1,\dots,t_n]/(f_1,\dots,f_m)$, then $\mathcal{Y}^{\mathrm{an}}$ is the vanishing locus of $f_1,\dots,f_m$ in $\mathbb{A}^n_X$. It is easily seen to be independent of the A-presentation of B.

(iii) The construction in (ii) is compatible with localizations, so in general one covers $\mathcal{Y}$ by $\mathcal{X}$-affine schemes $\mathcal{Y}_i$ and glues $\mathcal{Y}^{\mathrm{an}}$ from $\mathcal{Y}_i^{\mathrm{an}}$.

6.3.2. *The analytification map.* There exist natural *analytification maps* $\pi_{\mathcal{Y}}:\mathcal{Y}^{\mathrm{an}}\to\mathcal{Y}$ which can be constructed through the steps (i)–(iii), or directly (ii) and (iii). Let us describe them in the affine case $\mathcal{Y}=\mathrm{Spec}\,B$:

(1) The map is $\mathrm{Spf}\,B\hookrightarrow\mathrm{Spec}\,B$. It is injective and the image is the set of open prime ideals of B.

(2)/(3′) The map $\mathcal{Y}^{\mathrm{an}} \to \mathcal{Y}$ is injective and its image is the set of maximal ideals of B.

(3) The map $\mathcal{Y}^{\mathrm{an}} \to \mathcal{Y}$ is surjective; see [Berkovich 1993, Proposition 2.6.2].

6.3.3. *Sheaves.* The analytification functor also extends to coherent sheaves: for any $\mathcal{X}$-scheme $\mathcal{Y}$ of finite type there exists an analytification functor $\mathrm{Coh}(\mathcal{Y}) \to \mathrm{Coh}(\mathcal{Y}^{\mathrm{an}})$ given by $\mathcal{F}^{\mathrm{an}} = \pi_{\mathcal{Y}}^{*}\mathcal{F}$.

6.3.4. *Properties.* For each $\mathcal{X}$-proper scheme $\mathcal{Y}$ the analytification functor $\mathrm{Coh}(\mathcal{Y}) \xrightarrow{\sim} \mathrm{Coh}(Y)$ is an equivalence of categories. In particular, the analytification functor induces an equivalence between the categories of projective $\mathcal{X}$-schemes and X-spaces. The references are:

(1) Grothendieck's existence theorem [EGA III_1 1961, 5.1.4].

(2) Theorem C.1.1 below.

(3) The analytification was introduced in [Berkovich 1993, Section 2.6], and comparison of coherent sheaves can be found in [Poineau 2010, Theorem A.1].

(3′) Köpf's theorem [1974, Sections 5 and 6]; see also [Conrad 2006, Example 3.2.6].

6.3.5. *Analytification and regularity.* Various properties are respected by analytification, but for our needs we only need to study the situation with regularity.

Proposition 6.3.6. *Assume that X is an affinoid space with $A = \mathcal{O}_X(X)$, $\mathcal{X} = \mathrm{Spec}(A)$, and $\mathcal{Y}$ is an $\mathcal{X}$-scheme of finite type with $Y = \mathcal{Y}^{\mathrm{an}}$, then*:

(i) *If $\mathcal{Y}$ is regular then Y is regular.*

(ii) *Conversely, assume that Y is regular. Then*

(*a*) *in cases* (2), (3) *and* (3′), *$\mathcal{Y}$ is regular*,

(*b*) *in case* (1) *assume also that $\mathcal{Y}$ is $\mathcal{X}$-proper*, *and then $\mathcal{Y}$ is regular.*

Proof. Note that case (3′) follows from (3) since a qcqs rigid space can be enhanced to an analytic space, and the regularity is preserved. We will study cases (1), (2) and (3) separately, but let us first make a general remark. The claims (i) and (iia) are local on $\mathcal{Y}$, so we can assume that $\mathcal{Y} = \mathrm{Spec}\, B$ for a finitely generated A-algebra B in these cases.

Case (1): In this case, A is an I-adic ring and $X = \mathrm{Spf}\, A$. Since A is qe, B is qe and so the I-adic completion homomorphism $B \to \widehat{B}$ is regular. This implies (i) since if B is regular then $\widehat{B}$ is regular, and so $\mathrm{Spf}\, \widehat{B}$ is regular.

Let us prove (ii). Since A is I-adic, I is contained in the Jacobson radical of A, see [Atiyah and Macdonald 1969, Proposition 10.15(iv)], and so any point of $\mathcal{X}$ has a specialization in $\mathcal{X}_s := V(I)$. By the properness of $f : \mathcal{Y} \to \mathcal{X}$, any point of $\mathcal{Y}$ has a specialization in $\mathcal{Y}_s := f^{-1}(\mathcal{X}_s)$; hence it suffices to prove the following claim: if $\mathcal{Y}$ is of finite type over $\mathcal{X}$ and Y is regular, then $\mathcal{Y}$ is regular at any point $y \in \mathcal{Y}_s$.

The latter claim is local around y; hence we can assume, again, that $\mathcal{Y} = \mathrm{Spec}\, B$. Let $m \subset B$ be the ideal corresponding to y; then the m-adic completion $B \to \widehat{B}_m$ factors through the I-adic completion

$B \to \widehat{B}$, and so $\widehat{B}_m$ is the completion of $\widehat{B}$ along $m\widehat{B}$. Since X is qe, $\widehat{B}$ is qe and so $\widehat{B} \to \widehat{B}_m$ is regular. By our assumption $\widehat{B}$ is regular; hence $\widehat{B}_m$ is regular too. The homomorphism $B_m \to \widehat{B}_m$ is faithfully flat; hence B_m is regular and we win.

Case (3): In this case, A is k-affinoid and $X = \mathcal{M}(A)$. Consider a point $y \in Y$ and set $\boldsymbol{y} = \pi_{\mathcal{Y}}(y) \in \mathcal{Y}$. By [Ducros 2018, Lemma 2.4.6(1)], $\mathcal{Y}$ is regular at $\boldsymbol{y}$ if and only if Y is regular at y. Since $\pi_{\mathcal{Y}}$ is surjective this implies that $\mathcal{Y}$ is regular if and only if Y is regular.

Case (2): If $y \in Y$ and $\boldsymbol{y} = \pi_{\mathcal{Y}}(y)$ then it follows easily from Lemma B.6.1 that the homomorphism $f_y : \mathcal{O}_{\mathcal{Y},\boldsymbol{y}} \to \mathcal{O}_{Y,y}$ induces an isomorphism of the completions. A local ring is regular if and only if its completion is regular; hence $\mathcal{O}_{\mathcal{Y},\boldsymbol{y}}$ is regular if and only if $\mathcal{O}_{Y,y}$ is regular. Since the image of $\pi_{\mathcal{Y}}$ contains all closed points, we obtain that Y is regular if and only if $\mathcal{Y}$ is regular. ☐

6.4. *The factorization theorem.*

6.4.1. *Blowings up.* Each of the categories $\mathfrak{Sp}$ has a natural notion of blowings up $f : X' \to X$ along ideals; e.g., see [Temkin 2008, Sections 2.4.4 and 5.1.2]. In fact, $Bl_I(X)$ can be described as follows: if $Y \subset X$ is an affinoid domain, $\mathcal{Y} = \operatorname{Spec}(\mathcal{O}_X(Y))$ and $\mathcal{I} \subset \mathcal{O}_{\mathcal{Y}}$ is induced by I, then the restriction of f onto Y is the analytification of the blowing up $Bl_{\mathcal{I}}(\mathcal{Y}) \to \mathcal{Y}$. We will only consider blowings up with nowhere-dense centers.

6.4.2. *Weak factorization.* By a weak factorization of $X_1 \to X_2$ we mean a diagram

$$X_1 = V_0 \overset{\phi_1}{\dashrightarrow} V_1 \overset{\phi_2}{\dashrightarrow} \cdots \overset{\phi_{l-1}}{\dashrightarrow} V_{l-1} \overset{\phi_l}{\dashrightarrow} V_l = X_2$$

along with subspaces Z_i and ideal sheaves J_i satisfying conditions (1)–(5) of Section 1.2, where in (2) and (4) the word "scheme" is replaced with "space". For brevity of notation, such a datum will be denoted by $(V_\bullet, \phi_\bullet, Z_\bullet)$.

We define the regular surjective category of blowings up $\mathrm{Bl}_{\mathrm{rs}}^{\mathfrak{Sp}}$ in $\mathfrak{Sp}$ and the regular surjective category of weak factorizations $\mathrm{Fact}_{\mathrm{rs}}^{\mathfrak{Sp}}$ on $\mathfrak{Sp}$ as in Definitions 1.3.1 and 1.3.2. By definition, these categories are fibered over the category of regular spaces with regular morphisms, and the fibers over a regular space X will be denoted by $\mathrm{Bl}(X)$ and $\mathrm{Fact}(X)$. Thus, $\mathrm{Bl}(X)$ is the set of blowings up $X' \to X$ with regular X and $\mathrm{Fact}(X)$ is the set of all regular factorizations of blowings up of X.

Lemma 6.4.3. *Let X be an affinoid space, $A = \mathcal{O}_X(X)$ and $\mathcal{X} = \operatorname{Spec} A$. Then the analytification functor $\mathcal{Y} \mapsto \mathcal{Y}^{\mathrm{an}}$ induces bijections* $\mathrm{Bl}(X) \xrightarrow{\sim} \mathrm{Bl}(\mathcal{X})$ *and* $\mathrm{Fact}(X) \xrightarrow{\sim} \mathrm{Fact}(\mathcal{X})$.

Proof. By the relative GAGA, see Section 6.3.4, analytification induces a bijection between the blowings up $X' \to X$ and $\mathcal{X}' \to \mathcal{X}$. By Proposition 6.3.6, X' is regular if and only if $\mathcal{X}'$ is regular; hence $\mathrm{Bl}(X) \xrightarrow{\sim} \mathrm{Bl}(\mathcal{X})$. The second bijection is proved similarly, but this time one also relates regularity of the centers in the factorizations. ☐

6.4.4. *The main theorem.* We are now in position to prove the following analogue of Theorem 1.3.3.

Theorem 6.4.5. *There is a functor* $\mathrm{Bl}_{\mathrm{rs}}^{\mathfrak{Sp}}(\mathrm{char}=0) \to \mathrm{Fact}_{\mathrm{rs}}^{\mathfrak{Sp}}(\mathrm{char}=0)$ *from the regular surjective category of blowings up* $f : X' \to X$ *in characteristic* 0 *to the regular surjective category of factorizations*

$$X' = V_0 \dashrightarrow V_1 \dashrightarrow \cdots \dashrightarrow V_{l-1} \dashrightarrow V_l = X$$

in characteristic 0 *such that the composite*

$$\mathrm{Bl}_{\mathrm{rs}}^{\mathfrak{Sp}}(\mathrm{char}=0) \to \mathrm{Fact}_{\mathrm{rs}}^{\mathfrak{Sp}}(\mathrm{char}=0) \to \mathrm{Bl}_{\mathrm{rs}}^{\mathfrak{Sp}}(\mathrm{char}=0)$$

is the identity. The same holds in positive and mixed characteristics if Hypothetical Statements 2.2.13 and 2.3.6 hold true.

Proof. First, let us construct a factorization of $f : X' \to X$. Fix an admissible affinoid covering $X = \bigcup_{i=1}^n X_i$ and set $X_i' = X_i \times_X X'$. The rings $A_i = \mathcal{O}_X(X_i)$ are qe, see [Temkin 2008, Section 5.2.3], so the scheme $\mathcal{X} = \coprod_{i=1}^n \mathcal{X}_i$ with $\mathcal{X}_i = \mathrm{Spec}(A_i)$ is noetherian and qe. Let I be the ideal defining f and let $I_i \subset A_i$ be its restrictions. Consider the blowings up $F_i : \mathcal{X}_i' \to \mathcal{X}_i$ defined by I_i. The analytification of F_i is the restriction f_i of f over X_i by the relative GAGA; hence $\mathcal{X}_i'$ is regular by Proposition 6.3.6(ii).

Set $\mathcal{X}' = \coprod_{i=1}^n \mathcal{X}_i'$ and consider the factorization $(\mathcal{V}_\bullet, \Phi_\bullet, \mathcal{Z}_\bullet)$ of the blowing up $F : \mathcal{X}' \to \mathcal{X}$. For each i, it induces a factorization $(\mathcal{V}_{i,\bullet}, \Phi_{i,\bullet}, \mathcal{Z}_{i,\bullet})$ of $F_i : \mathcal{X}_i' \to \mathcal{X}_i$ and the analytification of the latter is a factorization of $f_i : X_i' \to X_i$ that will be denoted by $(V_{i,\bullet}, \phi_{i,\bullet}, Z_{i,\bullet})$.

We claim that the latter factorizations glue to a factorization of f. It suffices to prove that for any i, j and an affinoid domain $Y \subset X_i \cap X_j$, the restrictions of $(V_{i,\bullet}, \phi_{i,\bullet}, Z_{i,\bullet})$ and $(V_{j,\bullet}, \phi_{j,\bullet}, Z_{j,\bullet})$ onto Y coincide. Set $B = \mathcal{O}_X(Y)$ and $\mathcal{Y} = \mathrm{Spec}(B)$, and let $G : \mathcal{Y}' \to \mathcal{Y}$ be the blowing up along the ideal induced by I. In particular, the analytification $g : Y' \to Y$ of G is the restriction of f. The regular homomorphisms $A_i \to B$ and $A_j \to B$ induce regular morphisms $h_i, h_j : \mathcal{Y} \to \mathcal{X}$ such that G is the pullback of F with respect to either of these morphisms. The factorizations of G induced from $(\mathcal{V}_\bullet, \Phi_\bullet, \mathcal{Z}_\bullet)$ via h_i and h_j coincide by Lemma 6.4.6 below. It remains to note that the factorizations of g induced from the factorizations of f_i and f_j are the analytifications of these factorizations of G.

We have constructed a factorization of f. The same argument as was used to glue local factorizations to a global one shows that the construction is independent of the affinoid covering. Finally, compatibility of factorization with a regular morphism $h : Y \to X$ is deduced in the same way from Lemma 6.2.8 and compatibility with regular morphisms of factorization for schemes. $\square$

The following result is an analogue of [Temkin 2008, Lemma 2.3.1].

Lemma 6.4.6. *Assume that* $\mathcal{F} : \mathrm{Bl} \to \mathrm{Fact}$ *is a factorization functor*, $f : X' \to X$ *and* $g : Y' \to Y$ *are two blowings up with regular source and target and* $h_i : Y \to X$ *with* $i = 1, 2$ *are two regular morphisms such that* $h_i^*(f) = g$. *Then the pullbacks of* $\mathcal{F}(f)$ *to a factorization of* g *via* h_1 *and* h_2 *coincide.*

Proof. Extend h_i to morphisms $\phi_i : Y \coprod X \to X$ so that the map on X is the identity. Each ϕ_i is a surjective regular morphism; hence the pullback of $\mathcal{F}(f)$ to $Y \coprod X$ via ϕ_i coincides with the factorization of the blowing up $Y' \coprod X' \to Y \coprod X$. Restricting the latter onto Y coincides with $h_i^*(\mathcal{F}(f))$. $\square$

Remark 6.4.7. (i) An analogue of Lemma 6.4.6 holds true in any category $\mathfrak{Sp}$ and the above proof applies verbatim.

(ii) Although $h_i^*(\mathcal{F}(f))$ coincide, they can differ from $\mathcal{F}(g)$ when h_i are not surjective. See also [Temkin 2008, Remark 2.3.2(ii)].

Appendix A: Construction of a birational cobordism via deformation to the normal cone

Proof of Proposition 4.4.1. We follow the construction of [Abramovich et al. 2002, Theorem 2.3.1] word for word, except we make it even more explicit and check functoriality.

Step 1: cobordism $B_{\mathcal{O}}$ for trivial blowing up. We start with

$$B_{\mathcal{O}} = \mathbb{P}^1_{X_2} = \mathbb{P}(\mathcal{O}_{X_2} \cdot T_0 \oplus \mathcal{O}_{X_2} \cdot T_1) =: \mathbb{P}_{X_2}(E_{\mathcal{O}}),$$

with its projection $\pi_0 : B_{\mathcal{O}} \to X_2$. Providing the generators T_0 and T_1 with $\mathbb{G}_m$-weights 0 and 1, the scheme $B_{\mathcal{O}}$ is a birational cobordism for the identity morphism with the trivial ideal (1), with the standard action of $\mathbb{G}_m$ linearized, except that it does not satisfy Assumption 4.2.2. But that may be achieved after the fact by taking the symmetric square. The construction is clearly functorial.

Step 2a: construction of a singular cobordism B_I. Assume X_1 is given as the blowing up of the ideal I on X_2. We blow up the $\mathbb{G}_m$-equivariant ideal $I^B := I \otimes \mathcal{O}_{B_{\mathcal{O}}} + I_{\{0\}}$ on $B_{\mathcal{O}}$, where $I_{\{0\}}$ is the defining ideal of $\{0\} \times X_2$. The ideal is clearly the unit ideal on $\mathbb{P}^1_U$. This blowing up gives rise to a $\mathbb{G}_m$-scheme B_I and projective morphism $\pi_I : B_I \to B_{\mathcal{O}}$; this is evidently functorial in ϕ. The arguments of Section 2.1.10 show that $\pi^{B_I/X_2} := \pi_0 \circ \pi_I : B_I \to X_2$ is projective, again in a functorial manner. In particular $B_I \subset \mathbb{P}(E_I)$ for some functorial $\mathbb{G}_m$-sheaf E_I.

Step 2b: coordinates of B_I. Let us make the construction of the previous step explicit: write $F_I = \pi_{0*} I^B(1) = I \cdot U_0 \oplus \mathcal{O}_{X_2} \cdot U_1$ with U_0, U_1 having corresponding $\mathbb{G}_m$-weights 0 and 1. Let

$$E_I = F_I \otimes E_{\mathcal{O}} = I \cdot U_0 T_0 \oplus (\mathcal{O}_{X_2} \cdot U_1 T_0 \oplus I \cdot U_0 T_1) \oplus \mathcal{O}_{X_2} \cdot U_1 T_1,$$

with corresponding $\mathbb{G}_m$-weights 0, 1 and 2. Again it does not satisfy Assumption 4.2.2, but again that may be achieved after the fact by taking the symmetric square.

We have a surjection $\pi_0^* F_I \to I^B(1)$ where the first coordinate sends $f \cdot U_0 \mapsto f T_0$ and the second sends $U_1 \mapsto T_1$. We thus have $\mathbb{G}_m$-equivariant closed embeddings

$$\begin{aligned} B_I = \mathrm{Bl}_{I^B}(B_{\mathcal{O}}) = \mathrm{Bl}_{I^B(1)}(B_{\mathcal{O}}) \subset \mathbb{P}_{B_{\mathcal{O}}}(\pi_0^* F_I) = \mathbb{P}_{X_2}(F_I) \times_{X_2} B_{\mathcal{O}} = \mathbb{P}_{X_2}(F_I) \times_{X_2} \mathbb{P}_{X_2}(E_{\mathcal{O}}) \\ \subset \mathbb{P}_{X_2}(F_I \otimes E_{\mathcal{O}}) = \mathbb{P}_{X_2}(E_I), \end{aligned}$$

where $\mathrm{Bl}_{I^B(1)}(B_{\mathcal{O}})$ denotes the blowing up of the fractional ideal $I^B(1)$ and the last inclusion is the Segre embedding.

We describe $B_I = \mathrm{Proj}_{X_2} A$ as follows. The algebra

$$A := \bigoplus_d (I^d \cdot T_0^{2d} \oplus I^{d-1} \cdot T_0^{2d-1} T_1 \oplus \cdots \oplus \mathcal{O}_{X_2} \cdot T_0 T_1^{2d-1} \oplus \mathcal{O}_{X_2} \cdot T_1^{2d}),$$

with terms $I^{d-k} \cdot T_0^j T_1^k$ when $j > k$ and $\mathcal{O}_{X_2} \cdot T_0^j T_1^k$ when $j \leq k$, is a graded $\mathbb{G}_m$-weighted quotient $\operatorname{Sym}^\bullet E_I \twoheadrightarrow A$, where we set $U_j = T_j$ and map $I^{\otimes d} \twoheadrightarrow I^d$.

We note that B_I admits an equivariant projection morphism $B_I \to B_{\mathcal{O}} = \mathbb{P}_{X_2}(E_{\mathcal{O}})$ which is an isomorphism away from the divisor (T_1^2), and an equivariant projection morphism $B_I \to \mathbb{P}_{X_2}(F_I)$, whose image is the closed subscheme we denote by

$$\mathbb{P}_{X_2}(F_I)' := \operatorname{Proj}_{X_2} \bigoplus_{n \geq 0} \left(\bigoplus_{j=0}^{n} I^j \right).$$

The morphism $B_I \to \mathbb{P}_{X_2}(F_I)'$ is an isomorphism away from the zero section $\operatorname{Proj}_{X_2} \bigoplus_{n\geq 0} \mathcal{O}_{X_2} \subset \mathbb{P}_{X_2}(F_I)'$, whose complement is the total space $\operatorname{Spec} \operatorname{Sym}((I\,\mathcal{O}_{X_1})^{-1})$ of the invertible sheaf $I\mathcal{O}_{X_1}$ on X_1.

Step 2c: stable and unstable loci for weight 1. The homogeneous Cartier divisor (T_0T_1) is the union of two regular subschemes: $X_1 = \operatorname{Proj}_{X_2} \bigoplus_{n\geq 0}(I^n \cdot T_0^{2n})$, which is the zero locus of (T_0T_1, T_1^2), and $X_2 = \operatorname{Proj}_{X_2} \bigoplus_{n\geq 0}(\mathcal{O}_{X_2} \cdot T_1^{2n})$, which is the zero locus of $(T_0T_1, I \cdot T_0^2)$. Since the zero locus of the "irrelevant ideal" $(I \cdot T_0^2, T_0T_1, T_1^2)$ is empty, these two subschemes are disjoint. In particular each is a regular Cartier divisor. It follows that both X_1 and X_2 lie in the regular locus B_I^{reg}, which is open since B_I is of finite type over the qe scheme X_2.

We have $X_1 = B_I \cap \mathbb{P}_{X_2}((E_I)_0)$ and $X_2 = B_I \cap \mathbb{P}_{X_2}((E_I)_2)$, where the indices 0 and 2 denote the components with given $\mathbb{G}_m$-weight (the variable a in Section 4.2). Their union (T_0T_1) is the unstable locus $(B_I)^{\mathrm{un}}_a 1$. The complement is affine, explicitly

$$\begin{aligned}(B_I)_1^{\mathrm{sst}} &= \operatorname{Spec}_{X_2} A[(T_0T_1)^{-1}]_{\mathrm{degree}=0} \\ &= \operatorname{Spec}_{X_2}\left(\cdots \oplus I^2 \left(\frac{T_0}{T_1}\right)^2 \oplus I\left(\frac{T_0}{T_1}\right) \oplus \mathcal{O}_{X_2} \oplus \mathcal{O}_{X_2}\left(\frac{T_1}{T_0}\right) \oplus \mathcal{O}_{X_2}\left(\frac{T_1}{T_0}\right)^2 \oplus \cdots \right).\end{aligned}$$

This scheme is in general singular, but the quotient is simpler:

$$(B_I)_1^{\mathrm{sst}} /\!\!/ \mathbb{G}_m = \operatorname{Spec}_{X_2} \mathcal{O}_{X_2} = X_2.$$

Step 2d: stable and unstable loci for weight 2. The projective Cartier divisor (T_1^2) can be identified as

$$(B_I)^{\mathrm{un}}_a 2 = \mathbb{P}_{X_2}(I \cdot T_0^2) \cup \mathbb{P}_{Z(I)}(I/I^2 \cdot T_0^2 \oplus \mathcal{O} \cdot T_0T_1) = X_1 \cup C(Z(I)),$$

where $C(Z(I))$ is the normal cone. The complement is again affine, of the form

$$\begin{aligned}(B_I)_2^{\mathrm{sst}} &= \operatorname{Spec}_{X_2} A[T_1^{-1}]_{\mathrm{degree}=0} \\ &= \operatorname{Spec}_{X_2}\left(\cdots \oplus \mathcal{O}_{X_2}\left(\frac{T_0}{T_1}\right)^2 \oplus \mathcal{O}_{X_2}\left(\frac{T_0}{T_1}\right) \oplus \mathcal{O}_{X_2} \right) = \mathbb{A}^1_{X_2}.\end{aligned}$$

Thus,

$$(B_I)_2^{\mathrm{sst}} /\!\!/ \mathbb{G}_m = \operatorname{Spec}_{X_2} \mathcal{O}_{X_2} = X_2$$

and the morphism $(B_I)_2^{\mathrm{sst}} \to X_2$ is smooth. Another way to see this is to notice that the map $B_I \to B_{\mathcal{O}}$ restricts to an open embedding on $(B_I)_2^{\mathrm{sst}}$, and the image is the complement of $\{0\} \times X_2$.

Step 2e: stable and unstable loci for weight 0. The projective zero locus of $(I \cdot T_0)^2$ can be identified as

$$(B_I)^{\mathrm{un}}_a 0 = \mathbb{P}_{X_2}(\mathcal{O}_{X_2} \cdot T_1^2) \cup \mathbb{P}_{Z(I)}(\mathcal{O}_{X_2} \cdot T_0 T_1 \oplus \mathcal{O}_{X_2} \cdot T_1^2) = X_2 \cup \mathbb{P}^1_{Z(I)}.$$

The complement is not necessarily affine, as I is not necessarily principal. However, recalling the sheaf F_I from Step 2b, the morphism $(B_I)_0^{\mathrm{sst}} \to \mathbb{P}_{X_2}(F_I)$ is an open embedding, whose image is the complement of the zero section. So $(B_I)_0^{\mathrm{sst}}$ is the total space of the invertible sheaf $I\mathcal{O}_{X_1}$ on X_1. Thus, $(B_I)_0^{\mathrm{sst}} /\!\!/ \mathbb{G}_m = X_1$ and the morphism $(B_I)_0^{\mathrm{sst}} \to X_1$ is smooth.

Step 3a: resolving (B_I, D_{B_I}). Let $D_{B_I} \subset B_I$ be the preimage of D_2. Applying resolution of pairs to (B_I, D_{B_I}) we obtain a functorial projective $\mathbb{G}_m$-equivariant morphism $B \to B_I$ such that B is regular and the preimage $D_B \subset B$ of D_2 is a simple normal crossings divisor. Here we use Theorem 2.2.10 if the characteristic is 0. In positive and mixed characteristic we may use parts (1) and (2) of Hypothetical Statement 2.2.13 since $\dim B = \dim X_2 + 1$. In addition, $B \to B_I$ is projectively the identity outside of the union of D_{B_I} and the singular locus of B_I, which is included in the preimage of

$$\mathbb{P}_{X_2}((E_I)_1) = \mathbb{P}_{X_2}(\mathcal{O}_{X_2} \cdot U_1 T_0 \oplus I \cdot U_0 T_1).$$

It follows that (B, D_B) is a regular birational cobordism for ϕ.

Step 3b: embedding. By the arguments of Section 2.1.10, the composition $B \to B_I \to B_{\mathcal{O}}$ is functorially a single blowing up of an ideal J. Write $\tilde{J} = J\mathcal{O}_{B_I}$ so that $B = \mathrm{Bl}_{\tilde{J}} B_I$. There is a functorially defined integer d such that $\tilde{J}(d)$ is globally generated on B_I relative to X_2. Using [Hartshorne 1977, II.7.10(b)] we have an equivariant embedding of B inside

$$\mathbb{P}_{X_2}(\widetilde{E}) := \mathbb{P}_{X_2}(\pi_*^{B_I/X_2} \tilde{J}(d)).$$

We claim that $a_{\min}(B) = 0$ and $a_{\max}(B) = 2d$. First, since E_I has weights $a_{\min}(E_I) = 0$ and $a_{\max}(E_I) = 2$, we have $a_{\min}(\mathrm{Sym}^d(E_I)) = 0$ and $a_{\max}(\mathrm{Sym}^d(E_I)) = 2d$. Second, the weights 0 and $2d$ survive in the homogeneous coordinate ring of B_I with respect to $O(d)$ as described in the steps above. Third, the weights in $\pi_*^{B_I/X_2} \tilde{J}(d)$ necessarily lie among those of $\mathrm{Sym}^d(E_I)$, so $a_{\min}(B) \geq 0$ and $a_{\max}(B) \leq 2d$. To show that the weights 0 and $2d$ survive in B it suffices to show this over a dense open set in X_2. Since $B \to B_I$ is projectively the identity over U, the weight-0 and weight-$2d$ components of $\pi_*^{B_I/X_2} \tilde{J}(d)$ are everywhere nonzero, as needed.

Inspecting the description of unstable loci in Section 4.1, equation (1) we note that $B_0^{\mathrm{sst}} = B \times_{B_I} (B_I)_0^{\mathrm{sst}}$ and $B_{2d}^{\mathrm{sst}} = B \times_{B_I} (B_I)_2^{\mathrm{sst}}$.

Step 3c: B is a cobordism for ϕ that respects U. We have shown in Steps 2d and 2e that the morphisms $q_2 : (B_I)_2^{\mathrm{sst}} \to X_2$ and $q_1 : (B_I)_0^{\mathrm{sst}} \to X_1$ are smooth. Functoriality of resolution of pairs with respect to q_i implies that, once restricted to $(B_I)_2^{\mathrm{sst}}$, the morphism $B \to B_I$ is the pullback of the resolution $X_2' \to X_2$ of (X_2, D_2), and once restricted to $(B_I)_0^{\mathrm{sst}}$, the morphism $B \to B_I$ is the pullback of the resolution $X_1' \to X_1$ of (X_1, D_1). It follows that $B \times_{B_I} (B_I)_2^{\mathrm{sst}} /\!\!/ \mathbb{G}_m = X_2'$ and $B \times_{B_I} (B_I)_0^{\mathrm{sst}} /\!\!/ \mathbb{G}_m = X_1'$ and hence B is a cobordism for ϕ. Also, we note that $B \cap \mathbb{P}(\widetilde{E}_0) = X_1'$ and $B \cap \mathbb{P}(\widetilde{E}_{2d}) = X_2'$, so Assumption 4.2.3 applies.

To show that B is compatible with U it suffices to show that both $B \to B_I$ and $B_I \to B_{\mathcal{O}}$ are projectively the identity over U. This is so for the blowing up $B_I \to \mathbb{P}^1_{X_2}$ because $I + I_{\{0\}}$ is the unit ideal on $\mathbb{P}^1_U$, and this is so for the resolution $B \to B_I$ because $\mathbb{P}^1_U$ is regular and disjoint from the preimage of D_2. □

Appendix B: Germs of complex analytic spaces

In this section we use germs to extend the category of complex analytic spaces to include certain Stein compacts. This will be used later to establish a tight connection between scheme theory and complex analytic geometry. In particular, this is needed to develop a relative GAGA theory.

B.1 ***Semianalytic sets.*** We follow the setup of [Frisch 1967]. A subset X of an analytic space $\mathcal{X}$ is called *semianalytic* if its local germs belong to the minimal class of germs, stable under finite unions and complements, generated by inequalities of the form $f(x) < 0$ for real analytic f; see [loc. cit., p. 120]. It is called *a Stein* if X has a fundamental system of neighborhoods of Stein subspaces of $\mathcal{X}$; see [loc. cit., p. 123].

B.2 ***The category of germs.*** A *germ of a complex analytic space* (or, simply, a germ) is a pair $(\mathcal{X}, X)$ consisting of an analytic space $\mathcal{X}$ and a semianalytic subset $X \subset \mathcal{X}$. We call X the *support* of $(\mathcal{X}, X)$ and we call $\mathcal{X}$ a *representative* of $(\mathcal{X}, X)$. Sometimes, we will use the shorter notation $X = (\mathcal{X}, X)$.

A morphism $\phi : (\mathcal{X}, X) \to (\mathcal{Y}, Y)$ consists of a neighborhood $\mathcal{X}'$ of X and an analytic map $f : \mathcal{X}' \to \mathcal{Y}$ taking X to Y. We say that f is a *representative* of ϕ. Note that a morphism $(\mathcal{X}, X) \to (\mathcal{Y}, Y)$ is an isomorphism if it induces a bijection of X and Y and an isomorphism of their neighborhoods.

We identify an analytic space X with the germ (X, X). In particular, the category of analytic spaces becomes a full subcategory of the category of germs.

B.3 ***The structure sheaf.*** Given a germ $(\mathcal{X}, X)$ we provide its support with the *structure sheaf* $\mathcal{O}_X := \mathcal{O}_{\mathcal{X}}|_X = i^*\mathcal{O}_{\mathcal{X}}$, where $i : X \hookrightarrow \mathcal{X}$ is the embedding. In particular, we obtain a functor $\mathcal{F} : (\mathcal{X}, X) \mapsto (X, \mathcal{O}_X)$ from the category of germs to the category of locally ringed spaces.

Remark B.3.1. We do not aim to develop a complete theory of semianalytic germs, so we do not study the natural question of whether $\mathcal{F}$ is fully faithful.

B.4 ***Closed polydiscs and convergent power series.*** Consider an analytic affine space $\mathcal{X} = \mathbb{A}^n_{\mathbb{C}}$ with coordinates $t_1, \dots, t_n$. For any tuple r of numbers $r_1, \dots, r_n \in [0, \infty)$, by the closed polydisc $D = D_r$ of radius r we mean the subset of $\mathcal{X}$ given by the inequalities $|t_i| \le r_i$. Note that r_i can be zero. By $\mathbb{C}\{t_1, \dots, t_n\}^\dagger_r$ we denote the ring of overconvergent series in $t_1, \dots, t_n$ of radius r. It is a noetherian regular excellent ring of dimension n; see [Matsumura 1980, Theorem 102].

Lemma B.4.1. *Let $D = D_r \subset \mathcal{X} = \mathbb{A}^n_{\mathbb{C}}$ be a polydisc and $A = \mathcal{O}_{\mathcal{X}}(D) = \Gamma(\mathcal{O}_D)$. Then:*

(i) $\mathbb{C}\{t_1, \dots, t_n\}^\dagger_r \xrightarrow{\sim} A$.

(ii) $\Gamma(D, \cdot)$ *induces an equivalence between the categories of coherent $\mathcal{O}_D$-modules and finitely generated A-modules, and higher cohomology of coherent $\mathcal{O}_D$-modules vanish.*

(iii) *For any $a \in D$ the ideal $m_a = (t_1 - a_1, \dots, t_n - a_n) \subset A$ is maximal, and any maximal ideal of A is of this form.*

(iv) *The completion of A along m_a is $\mathbb{C}[[t_1 - a_1, \dots, t_n - a_n]]$.*

Proof. The first claim is a classical result of analysis of several complex variables. Assertion (ii) follows from the fact that D is the intersection of open polydiscs containing it, and the latter are Stein spaces. Assertion (iv) follows easily from (iii), so we will only prove (iii).

For any $f \in A$ the quotient

$$g_1 = (f(t_1, \dots, t_n) - f(a_1, t_2, \dots, t_n))/(t_1 - a_1)$$

lies in A, so $f = (t_1 - a_1)g_1 + f_1(t_2, \dots, t_n)$ with $f_1 = f(a_1, t_2, \dots, t_n)$. Applying the same argument to t_2 and f_1, etc., we will obtain in the end a representation $f = f(a_1, \dots, a_n) + \sum_{i=1}^n (t_i - a_i)g_i$. In particular, $A/m_a = \mathbb{C}$ and hence m_a is maximal.

Conversely, assume that $m \subset A$ is maximal. The norm $\|f\| = \max_{x \in D} |f(x)|$ on A induces a norm on the field $\kappa = A/m$; hence the completion $K = \hat{\kappa}$ is a Banach $\mathbb{C}$-field. Thus, $K = \mathbb{C}$ by the Gel'fand–Mazur theorem, and we obtain that $t_i - a_i \in m$ for some $a_i \in \mathbb{C}$. Finally, $|a_i| \le r_i$ as otherwise $t_i - a_i \in A^\times$. □

B.5 ***Classes of morphisms.*** Let $\phi : (\mathcal{Y}, Y) \to (\mathcal{X}, X)$ be a morphism of germs. We say that ϕ is *without boundary* if there exists a representative $f : \mathcal{Y}' \to \mathcal{X}$ such that $Y = f^{-1}(X)$. Let P be one of the following properties: smooth, open immersion, closed immersion. We say that ϕ is P if it is without boundary and has a representative which is P. We say that ϕ is an *embedding of a subdomain* if it possesses a representative which is an open immersion and we say that ϕ is *quasismooth* if it possesses a representative which is smooth.

Remark B.5.1. The above terminology is chosen to match its nonarchimedean analogue as much as possible.

B.6 ***Affinoid germs.*** A germ X is called *affinoid* if it admits a closed immersion into a germ of the form $(\mathbb{C}^n, D)$ where D is a closed polydisc. Such a germ is controlled by the ring $\mathcal{O}_X(X)$ very tightly.

Lemma B.6.1. *Assume that X is an affinoid germ and let $A = \mathcal{O}_X(X)$ and $f : (X, \mathcal{O}_X) \to Y = \mathrm{Spec}(A)$ be the corresponding map of locally ringed spaces. Then:*

(i) *A is a quotient of a ring $\mathbb{C}\{t_1, \dots, t_n\}_r^\dagger$; in particular it is an excellent noetherian ring.*

(ii) *$\Gamma(X, \cdot)$ induces an equivalence between the categories of coherent $\mathcal{O}_X$-modules and finitely generated A-modules, and higher cohomology of coherent $\mathcal{O}_X$-modules vanish.*

(iii) *f establishes a bijection between X and the closed points of Y.*

(iv) *For any point $x \in X$ with $y = f(x)$ the homomorphism $\mathcal{O}_{Y,y} \to \mathcal{O}_{X,x}$ is regular and its completion $\widehat{\mathcal{O}}_{Y,y} \to \widehat{\mathcal{O}}_{X,x}$ is an isomorphism.*

Proof. In the case of a closed polydisc the assertion was proved in Lemma B.4.1. In general, we fix a closed embedding $i : X \hookrightarrow D$ into a closed polydisc. So, $\mathcal{O}_X$ becomes a coherent $\mathcal{O}_D$-algebra such that the homomorphism $\phi : \mathcal{O}_D \to \mathcal{O}_X$ is surjective, and then all assertions except the first half of (iv) follow easily from the case of a polydisc. For example, $\Gamma(X, \mathcal{O}_X)$ is a quotient of $\Gamma(D, \mathcal{O}_D)$ since $H^1(D, \operatorname{Ker}\phi) = 0$, thereby proving (i).

The only new assertion is that $\phi : \mathcal{O}_{Y,y} \to \mathcal{O}_{X,x}$ is regular. This follows from the facts that $\hat{\phi}$ is an isomorphism and the local ring $\mathcal{O}_{Y,y}$ is excellent (since it is a localization of the excellent ring A). □

Appendix C: The complex relative GAGA theorem

C.1 ***Statement of the theorem.*** Let $(\mathcal{X}, X)$ be an affinoid germ as in Appendix B with ring of global analytic functions A, and $r \geq 0$ an integer. Set $\mathbb{P}^r_X = \mathbb{CP}^r \times X$ and endow it with a locally ringed space structure using the sheaf $\mathcal{O}_{\mathbb{P}^r_X} = \mathcal{O}_{\mathbb{P}^r_{\mathcal{X}}}|_{\mathbb{P}^r_X}$. We have a germ $(\mathbb{P}^r_{\mathcal{X}}, \mathbb{P}^r_X)$ and a morphism of locally ringed spaces $h : \mathbb{P}^r_X \to \mathbb{P}^r_A$. The aim of this appendix is to prove the following extension of Lemma B.6.1:

Theorem C.1.1 (Serre's Théorème 3). *Let $(\mathcal{X}, X)$ be an affinoid germ with ring of global analytic functions A, and $r \geq 0$ an integer. Then the pullback functor $h^* : \mathrm{Coh}(\mathbb{P}^r_A) \to \mathrm{Coh}(\mathbb{P}^r_X)$ is an equivalence which induces isomorphisms on cohomology groups.*

Since $(\mathcal{X}, X)$ is closed in $(\mathbb{C}^n, D)$ it suffices to consider the case $(\mathcal{X}, X) = (\mathbb{C}^n, D)$. So from now on we make this assumption, and write A for the ring of holomorphic functions on $X = D$.

We follow the steps of Serre's original proof [1956, §3] in some detail, to alleviate our skepticism that this generalization might actually work. See also [Kedlaya 2009], which sketches Serre's proof. One difficulty is that we do not know if $D \times \mathbb{C}^r$ is Stein in the sense of [Frisch 1967] or [Grauert and Remmert 1979]. The problem is that if $\{D_i\}$ are the open polydiscs containing D then $\{D_i \times \mathbb{C}^r\}$ do not form a *fundamental* family of neighborhoods of $D \times \mathbb{C}^r$, while functions on $D \times \mathbb{C}^r$ are only guaranteed to extend to some member of a fundamental family of neighborhoods. This is circumvented in Lemma C.2.2, which is the only point where we differ from the original arguments.

C.2 ***Cohomology.***

Proposition C.2.1 (Serre's Théorème 1). *Let $\mathcal{F}$ be a coherent sheaf on $\mathbb{P}^r_A$. The homomorphism h^* : $H^i(\mathbb{P}^r_A, \mathcal{F}) \to H^i(\mathbb{P}^r_D, h^*\mathcal{F})$ is an isomorphism.*

Lemma C.2.2. (1) *We have $H^i(\mathbb{P}^r_A, \mathcal{F}) = H^i(\mathbb{P}^r_D, h^*\mathcal{F}) = 0$ for $i > r$ and all $\mathcal{F}$.*

(2) *Proposition C.2.1 holds for $\mathcal{F} = \mathcal{O}_{\mathbb{P}^r_A}$ for all $r \geq 0$.*

Proof. (1) For $H^i(\mathbb{P}^r_A, \mathcal{F}) = 0$ use the standard Čech covering of $\mathbb{P}^r_A$, which has only $r+1$ elements. We need to show $H^i(\mathbb{P}^r_D, h^*\mathcal{F}) = 0$.

On the analytic side we mimic the standard argument for vanishing using Čech cocycles of a covering by closed polydiscs instead of affine spaces. Let $h^*\mathcal{F} \to S^\bullet$ be the standard flabby resolution of $h^*\mathcal{F}$ by discontinuous sections, so $H^i(Y, h^*\mathcal{F}|_Y) = H^i(\Gamma(Y, S^\bullet))$ for any subset $Y \subset \mathbb{P}^r_D$. Let $\mathbb{C}^r \simeq U_i \subset \mathbb{CP}^r$

be the standard open sets and let $D_i \subset U_i$ be the standard closed polydisc of fixed radius > 1. Set $X_i = D \times D_i \subset \mathbb{P}^r_D$ and for each subset $I \subset \{0, \dots, n\}$ let $X_I = \bigcap_{i \in I} X_i$. Then the X_I are complex affinoids for $I \neq \varnothing$; hence $H^i(X_I, h^*\mathcal{F}|_{X_I}) = 0 = H^i(\Gamma(X_I, S^\bullet))$ for $i > 0$ and $I \neq \varnothing$.

On the other hand

$$\mathcal{C}^\bullet(\{X_i\}, S^j) = \Big[\bigoplus_{|I|=1} S^j_{X_I} \to \bigoplus_{|I|=2} S^j_{X_I} \to \cdots \Big]$$

is a flabby resolution of S^j so $H^0(\Gamma(\mathbb{P}^r_D, \mathcal{C}^\bullet(\{X_i\}, S^j))) = \Gamma(\mathbb{P}^r_D, S^j)$ and for $i > 0$ we have

$$H^i(\Gamma(\mathbb{P}^r_D, \mathcal{C}^\bullet(\{X_i\}, S^j))) = 0.$$

Consider the double complex $C^{p,q} = \bigoplus_{|I|=p} \Gamma(X_I, S^q)$ and its two edges $\Gamma(\mathbb{P}^r, S^\bullet)$ and $\check{C}^p = \oplus_{|I|=p} \Gamma(X_I, h^*\mathcal{F})$. We obtain

$$H^i(\mathbb{P}^r_D, h^*\mathcal{F}) = H^i(\Gamma(\mathbb{P}^r, S^\bullet)) = \mathbb{H}^i(C^{\bullet,\bullet}) = H^i(\check{C}^\bullet).$$

The latter is trivial in degrees $> r$.

(2) We have $\Gamma(\mathcal{O}_{\mathbb{P}^r_A}) = A$ and $H^i(\mathcal{O}_{\mathbb{P}^r_A}) = 0$ for $i > 0$ by [Hartshorne 1977, Theorem III.5.1]. It suffices to show that $\pi_*\mathcal{O}_{\mathbb{P}^r_D} = \mathcal{O}_D$ and $R^i\pi_*\mathcal{O}_{\mathbb{P}^r_D} = 0$ for $i > 0$, where $\pi : \mathbb{P}^r_D \to D$ is the projection, since D is Stein. For this note that $\mathcal{O}_{\mathbb{P}^r_D} = j_r^{-1}\mathcal{O}_{\mathbb{P}^r_{\mathbb{CP}^n}}$, where $j_r : \mathbb{P}^r_D \to \mathbb{P}^r_{\mathbb{CP}^n}$ is the inclusion

$$\begin{tikzcd} \mathbb{P}^r_D \arrow[r, "j_r"] \arrow[d, "\pi"'] & \mathbb{P}^r_{\mathbb{CP}^n} \arrow[d, "\varpi"] \\ D \arrow[r, "j_0"'] & \mathbb{CP}^n. \end{tikzcd}$$

By the topological proper push-forward theorem [Iversen 1986, Corollary VII.1.5] we have

$$R^i\pi_*\mathcal{O}_{\mathbb{P}^r_D} = j_0^{-1}R^i\varpi_*\mathcal{O}_{\mathbb{P}^r_{\mathbb{CP}^n}},$$

and the result follows from Serre's original GAGA theorems. □

Lemma C.2.3. *The proposition holds for $\mathcal{F} = \mathcal{O}_{\mathbb{P}^r_A}(n)$ for all $r \geq 0$ and all integers n.*

Proof. Induction identical to [Serre 1956, Section 13, Lemme 5]: the result holds for $r = 0$ since D is Stein. Supposing it holds for $r - 1$ and all n, we have the exact sequence $0 \to \mathcal{O}_{\mathbb{P}^r_D}(n-1) \to \mathcal{O}_{\mathbb{P}^r_D}(n) \to \mathcal{O}_{\mathbb{P}^{r-1}_D}(n) \to 0$ and the corresponding sequence for $\mathbb{P}^r_A$. We obtain a canonical homomorphism of long exact sequences:

$$\begin{tikzcd} H^{i-1}(\mathbb{P}^{r-1}_A, \mathcal{O}(n)) \arrow[r] \arrow[d] & H^i(\mathbb{P}^r_A, \mathcal{O}(n-1)) \arrow[r] \arrow[d] & H^i(\mathbb{P}^r_A, \mathcal{O}(n)) \arrow[r] \arrow[d] & H^i(\mathbb{P}^{r-1}_A, \mathcal{O}(n)) \arrow[d] \\ H^{i-1}(\mathbb{P}^{r-1}_D, \mathcal{O}(n)) \arrow[r] & H^i(\mathbb{P}^r_D, \mathcal{O}(n-1)) \arrow[r] & H^i(\mathbb{P}^r_D, \mathcal{O}(n)) \arrow[r] & H^i(\mathbb{P}^{r-1}_D, \mathcal{O}(n)). \end{tikzcd}$$

The vertical arrows on the right and left are isomorphisms by the inductive assumption. It follows that the result holds for r and $\mathcal{O}(n-1)$ if and only if it holds for $\mathcal{O}(n)$. Since we have proven that it holds for $\mathcal{O}$, it holds for all n. □

Proof of the proposition. The proof is identical to Serre's Théorème 1. We apply descending induction on i for all coherent $\mathbb{P}^r_A$ modules $\mathcal{F}$. The case of $i > r$ is proved by the lemma. Since $\mathcal{F}$ is coherent, there is an epimorphism $\mathcal{E} \to \mathcal{F}$ with $\mathcal{E} = \bigoplus_{i=1}^m \mathcal{O}_{\mathbb{P}^r_A}(-k_i)$. Denoting by $\mathcal{G}$ the kernel, $\mathcal{G}$ is coherent and we have a short exact sequence

$$0 \to \mathcal{G} \to \mathcal{E} \to \mathcal{F} \to 0.$$

Since the map h is flat we have an exact sequence

$$0 \to h^*\mathcal{G} \to h^*\mathcal{E} \to h^*\mathcal{F} \to 0.$$

In the commutative diagram of cohomologies with exact rows

$$\begin{array}{ccccccc} H^i(\mathbb{P}^r_A, \mathcal{E}) & \longrightarrow & H^i(\mathbb{P}^r_A, \mathcal{F}) & \longrightarrow & H^{i+1}(\mathbb{P}^r_A, \mathcal{G}) & \longrightarrow & H^{i+1}(\mathbb{P}^r_A, \mathcal{E}) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ H^i(\mathbb{P}^r_D, h^*\mathcal{E}) & \longrightarrow & H^i(\mathbb{P}^r_D, h^*\mathcal{F}) & \longrightarrow & H^{i+1}(\mathbb{P}^r_D, h^*\mathcal{G}) & \longrightarrow & H^{i+1}(\mathbb{P}^r_D, h^*\mathcal{E}) \end{array}$$

the vertical arrows on the left and right are isomorphisms by Lemma C.2.3. By the induction hypothesis $H^{i+1}(\mathbb{P}^r_A, \mathcal{G}) \to H^{i+1}(\mathbb{P}^r_D, h^*\mathcal{G})$ is an isomorphism as well. By the five lemma the result holds for $H^i(\mathbb{P}^r_A, \mathcal{F}) \to H^i(\mathbb{P}^r_D, h^*\mathcal{F})$ as required. □

C.3 *Homomorphisms.*

Proposition C.3.1 (Serre's Théorème 2). *For any coherent $\mathbb{P}^r_A$-modules $\mathcal{F}, \mathcal{G}$ the natural homomorphism*

$$\mathrm{Hom}_{\mathbb{P}^r_A}(\mathcal{F}, \mathcal{G}) \to \mathrm{Hom}_{\mathbb{P}^r_D}(h^*\mathcal{F}, h^*\mathcal{G})$$

is an isomorphism. In particular the functor h^* is fully faithful.

Lemma C.3.2. *The sheaf homomorphism*

$$h^* \mathcal{H}om_{\mathbb{P}^r_A}(\mathcal{F}, \mathcal{G}) \to \mathcal{H}om_{\mathbb{P}^r_D}(h^*\mathcal{F}, h^*\mathcal{G})$$

is an isomorphism.

Proof. This follows since $\mathcal{O}_{\mathbb{P}^r_D}$ is a flat $\mathcal{O}_{\mathbb{P}^r_A}$-module. Indeed, for a closed point $x \in \mathbb{P}^r_D$ corresponding to a point $x' = h(x) \in \mathbb{P}^r_A$ we have

$$\begin{aligned}(h^* \mathcal{H}om_{\mathbb{P}^r_A}(\mathcal{F}, \mathcal{G}))_x &= \mathcal{H}om_{\mathcal{O}_{x'}}(\mathcal{F}_{x'}, \mathcal{G}_{x'}) \otimes_{\mathcal{O}_{x'}} \mathcal{O}_x \\ &= \mathcal{H}om_{\mathcal{O}_x}(\mathcal{F}_{x'} \otimes_{\mathcal{O}_{x'}} \mathcal{O}_x, \mathcal{G}_{x'} \otimes_{\mathcal{O}_{x'}} \mathcal{O}_x) = \mathcal{H}om_{\mathbb{P}^r_D}(h^*\mathcal{F}, h^*\mathcal{G})_x.\end{aligned}$$ □

Proof of the proposition. By Serre's Théorème 1, h^* preserves cohomology of coherent sheaves. Taking H^0 in the lemma the result follows. □

C.4 *The equivalence.* It remains to show:

Proposition C.4.1. *The functor h^* is essentially surjective.*

Proof. This is an inductive argument on r identical to Serre's Théorème 3 which we repeat below. The case $r = 0$ follows from Lemma B.6.1. Assume the result is known for $r - 1$ and let $\mathcal{F}$ be a coherent sheaf on $\mathbb{P}^r_D$. By Lemma C.4.2 below there is an epimorphism $\phi : \mathcal{O}(-n_0)^{k_0} \to \mathcal{F}$, and applying this again to $\mathrm{Ker}(\phi)$ we get a resolution $\mathcal{O}(-n_1)^{k_1} \xrightarrow{\psi} \mathcal{O}(-n_0)^{k_0} \to \mathcal{F} \to 0$. By Serre's Théorème 2 the homomorphism ψ is the analytification of an algebraic sheaf homomorphism ψ', so the cokernel $\mathcal{F}$ of ψ is also the analytification of the cokernel of ψ'. □

Lemma C.4.2. *Assume the proposition holds for $r-1$. Then for any coherent sheaf $\mathcal{F}$ on $\mathbb{P}^r_D$ there is n_0 so that $\mathcal{F}(n)$ is globally generated whenever $n > n_0$.*

Proof. By compactness it suffices to show that global sections of $\mathcal{F}(n)$ generate $\mathcal{F}(n)_x$ for fixed x. By Nakayama it suffices to show that global sections of $\mathcal{F}(n)$ generate the fiber $\mathcal{F}(n)_x \otimes_{\mathcal{O}_{D,x}} \mathbb{C}_x$.

Picking a hyperplane $\mathbb{P}^{r-1}_D \simeq H \ni x$ we obtain an exact sequence $0 \to \mathcal{O}(-1) \to \mathcal{O} \to \mathcal{O}_H \to 0$, giving an exact sequence $\mathcal{F}(-1) \xrightarrow{\varphi_1} \mathcal{F} \xrightarrow{\varphi_0} \mathcal{F}_H \to 0$. Writing $\mathcal{P}$ for $\mathrm{Ker}(\varphi_0) = \mathrm{Im}(\varphi_1)$ we have two exact sequences

$$0 \to \mathcal{G} \to \mathcal{F}(-1) \to \mathcal{P} \to 0 \quad \text{and} \quad 0 \to \mathcal{P} \to \mathcal{F} \to \mathcal{F}_H \to 0,$$

noting that $\mathcal{G}$ and $\mathcal{F}_H$ are coherent sheaves on H. Twisting by $\mathcal{O}(n)$ gives

$$0 \to \mathcal{G}(n) \to \mathcal{F}(n-1) \to \mathcal{P}(n) \to 0 \quad \text{and} \quad 0 \to \mathcal{P}(n) \to \mathcal{F}(n) \to \mathcal{F}_H(n) \to 0.$$

The long exact cohomology sequence gives

$$H^1(\mathbb{P}^r_D, \mathcal{F}(n-1)) \to H^1(\mathbb{P}^r_D, \mathcal{P}(n)) \to H^2(H, \mathcal{G}(n))$$

and

$$H^1(\mathbb{P}^r_D, \mathcal{P}(n)) \to H^1(\mathbb{P}^r_D, \mathcal{F}(n)) \to H^1(H, \mathcal{F}_H(n)).$$

By the assumption $\mathcal{F}_H$ and $\mathcal{G}$ are analytifications of algebraic sheaves, so for large n the terms on the right vanish by Serre's Théorème 1. It follows that $\dim H^1(\mathbb{P}^r_D, \mathcal{F}(n))$ stabilizes for large n, and when it does the exact sequences above imply that $H^1(\mathbb{P}^r_D, \mathcal{P}(n)) \to H^1(\mathbb{P}^r_D, \mathcal{F}(n))$ is bijective so $H^0(\mathbb{P}^r_D, \mathcal{F}(n)) \to H^0(H, \mathcal{F}_H(n))$ is surjective. Since the result holds for analytifications of algebraic sheaves, $\mathcal{F}_H(n)$ is globally generated for large n, implying that $\mathcal{F}(n)_x \otimes_{\mathcal{O}_{D,x}} \mathbb{C}_x$ is generated by global sections, as needed. □

References

[Abramovich and Temkin 2017] D. Abramovich and M. Temkin, "Torification of diagonalizable group actions on toroidal schemes", *J. Algebra* **472** (2017), 279–338. MR Zbl

[Abramovich and Temkin 2018] D. Abramovich and M. Temkin, "Luna's fundamental lemma for diagonalizable groups", *Algebr. Geom.* **5**:1 (2018), 77–113. MR Zbl

[Abramovich et al. 1999] D. Abramovich, K. Matsuki, and S. Rashid, "A note on the factorization theorem of toric birational maps after Morelli and its toroidal extension", *Tohoku Math. J.* (2) **51**:4 (1999), 489–537. MR Zbl

[Abramovich et al. 2002] D. Abramovich, K. Karu, K. Matsuki, and J. Włodarczyk, "Torification and factorization of birational maps", *J. Amer. Math. Soc.* **15**:3 (2002), 531–572. MR Zbl

[Abramovich et al. 2015] D. Abramovich, L. Caporaso, and S. Payne, "The tropicalization of the moduli space of curves", *Ann. Sci. Éc. Norm. Supér.* (4) **48**:4 (2015), 765–809. MR Zbl

[Atiyah and Macdonald 1969] M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley, Reading, MA, 1969. MR Zbl

[Berkovich 1993] V. G. Berkovich, "Étale cohomology for non-Archimedean analytic spaces", *Inst. Hautes Études Sci. Publ. Math.* **78** (1993), 5–161. MR Zbl

[Bierstone and Milman 1997] E. Bierstone and P. D. Milman, "Canonical desingularization in characteristic zero by blowing up the maximum strata of a local invariant", *Invent. Math.* **128**:2 (1997), 207–302. MR Zbl

[Borisov and Libgober 2005] L. Borisov and A. Libgober, "McKay correspondence for elliptic genera", *Ann. of Math.* (2) **161**:3 (2005), 1521–1569. MR Zbl

[Brion and Procesi 1990] M. Brion and C. Procesi, "Action d'un tore dans une variété projective", pp. 509–539 in *Operator algebras, unitary representations, enveloping algebras, and invariant theory* (Paris, 1989), edited by A. Connes et al., Progr. Math. **92**, Birkhäuser, Boston, 1990. MR Zbl

[Buonerba 2015] F. Buonerba, "Functorial resolution of tame quotient singularities in positive characteristic", preprint, 2015. arXiv

[Conrad 2006] B. Conrad, "Relative ampleness in rigid geometry", *Ann. Inst. Fourier* (*Grenoble*) **56**:4 (2006), 1049–1126. MR Zbl

[Cossart and Piltant 2014] V. Cossart and O. Piltant, "Resolution of singularities of arithmetical threefolds, II", preprint, 2014. arXiv

[Cossart et al. 2009] V. Cossart, U. Jannsen, and S. Saito, "Canonical embedded and non-embedded resolution of singularities for excellent two-dimensional schemes", preprint, 2009. arXiv

[Dolgachev and Hu 1998] I. V. Dolgachev and Y. Hu, "Variation of geometric invariant theory quotients", *Inst. Hautes Études Sci. Publ. Math.* **87** (1998), 5–56. MR Zbl

[Ducros 2009] A. Ducros, "Les espaces de Berkovich sont excellents", *Ann. Inst. Fourier* (*Grenoble*) **59**:4 (2009), 1443–1552. MR Zbl

[Ducros 2018] A. Ducros, *Families of Berkovich spaces*, Astérisque **400**, Soc. Math. France, Paris, 2018. MR Zbl

[EGA II 1961] A. Grothendieck, "Eléments de géométrie algébrique, II: Étude globale élémentaire de quelques classes de morphismes", *Inst. Hautes Études Sci. Publ. Math.* **8** (1961), 5–222. MR Zbl

[EGA III_1 1961] A. Grothendieck, "Eléments de géométrie algébrique, III: Étude cohomologique des faisceaux cohérents, I", *Inst. Hautes Études Sci. Publ. Math.* **11** (1961), 5–167. MR Zbl

[EGA IV_2 1965] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR Zbl

[EGA IV_3 1966] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, III", *Inst. Hautes Études Sci. Publ. Math.* **28** (1966), 5–255. MR Zbl

[Frisch 1967] J. Frisch, "Points de platitude d'un morphisme d'espaces analytiques complexes", *Invent. Math.* **4** (1967), 118–138. MR Zbl

[Gillet and Soulé 2000] H. Gillet and C. Soulé, "Direct images in non-Archimedean Arakelov theory", *Ann. Inst. Fourier* (*Grenoble*) **50**:2 (2000), 363–399. MR Zbl

[Grauert and Remmert 1979] H. Grauert and R. Remmert, *Theory of Stein spaces*, Grundlehren der Math. Wissenschaften **236**, Springer, 1979. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Math. **52**, Springer, 1977. MR Zbl

[Hironaka 1975] H. Hironaka, "Flattening theorem in complex-analytic geometry", *Amer. J. Math.* **97** (1975), 503–547. MR Zbl

[Illusie and Temkin 2014] L. Illusie and M. Temkin, "Gabber's modification theorem (absolute case)", exposé VIII, pp. 103–160 in *Travaux de Gabber sur l'uniformisation locale et la cohomologie étale des schémas quasi-excellents* (Palaiseau, 2006–2008), Astérisque **363-364**, Soc. Math. France, Paris, 2014. MR Zbl

[Iversen 1986] B. Iversen, *Cohomology of sheaves*, Springer, 1986. MR Zbl

[Kato 1994] K. Kato, "Toric singularities", *Amer. J. Math.* **116**:5 (1994), 1073–1099. MR Zbl

[Kedlaya 2009] K. Kedlaya, "Algebraic geometry: GAGA", course notes, Massachusetts Inst. Tech., 2009, available at https://tinyurl.com/kedgaga.

[Kempf et al. 1973] G. Kempf, F. F. Knudsen, D. Mumford, and B. Saint-Donat, *Toroidal embeddings, I*, Lecture Notes in Math. **339**, Springer, 1973. MR Zbl

[Kollár 2007] J. Kollár, *Lectures on resolution of singularities*, Ann. of Math. Studies **166**, Princeton Univ. Press, 2007. MR

[Kontsevich and Soibelman 2006] M. Kontsevich and Y. Soibelman, "Affine structures and non-Archimedean analytic spaces", pp. 321–385 in *The unity of mathematics* (Cambridge, 2003), edited by P. Etingof et al., Progr. Math. **244**, Birkhäuser, Boston, 2006. MR Zbl

[Köpf 1974] U. Köpf, *Über eigentliche Familien algebraischer Varietäten über affinoiden Räumen*, Schr. Math. Inst. Univ. Münster (2) **7**, 1974. MR Zbl

[Matsumura 1980] H. Matsumura, *Commutative algebra*, 2nd ed., Math. Lect. Note Ser. **56**, Benjamin/Cummings, Reading, MA, 1980. MR Zbl

[Morelli 1996] R. Morelli, "The birational geometry of toric varieties", *J. Algebraic Geom.* **5**:4 (1996), 751–782. MR Zbl

[Mumford et al. 1994] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, 3rd ed., Ergebnisse der Mathematik (2) **34**, Springer, 1994. MR Zbl

[Noohi 2005] B. Noohi, "Foundations of topological stacks, I", preprint, 2005. arXiv

[Poineau 2010] J. Poineau, "Raccord sur les espaces de Berkovich", *Algebra Number Theory* **4**:3 (2010), 297–334. MR Zbl

[Popescu 1986] D. Popescu, "General Néron desingularization and approximation", *Nagoya Math. J.* **104** (1986), 85–115. MR Zbl

[Porta and Yu 2016] M. Porta and T. Y. Yu, "Higher analytic stacks and GAGA theorems", *Adv. Math.* **302** (2016), 351–409. Zbl

[Serre 1956] J.-P. Serre, "Géométrie algébrique et géométrie analytique", *Ann. Inst. Fourier* (*Grenoble*) **6** (1956), 1–42. MR Zbl

[Simons 2015] M. Baker and S. Payne (editors), "Simons Symposium 2015: open problem sessions", compilation of submitted problems, 2015, available at https://tinyurl.com/simsymp.

[Simpson 1996] C. Simpson, "Algebraic (geometric) n-stacks", preprint, 1996. arXiv

[Spivakovsky 1999] M. Spivakovsky, "A new proof of D. Popescu's theorem on smoothing of ring homomorphisms", *J. Amer. Math. Soc.* **12**:2 (1999), 381–444. MR Zbl

[Temkin 2008] M. Temkin, "Desingularization of quasi-excellent schemes in characteristic zero", *Adv. Math.* **219**:2 (2008), 488–522. MR Zbl

[Temkin 2012] M. Temkin, "Functorial desingularization of quasi-excellent schemes in characteristic zero: the nonembedded case", *Duke Math. J.* **161**:11 (2012), 2207–2254. MR Zbl

[Temkin 2018] M. Temkin, "Functorial desingularization over $\mathbb{Q}$: boundaries and the embedded case", *Israel J. Math.* **224**:1 (2018), 455–504. MR Zbl

[Thaddeus 1996] M. Thaddeus, "Geometric invariant theory and flips", *J. Amer. Math. Soc.* **9**:3 (1996), 691–723. MR Zbl

[Ulirsch 2015] M. Ulirsch, *Tropical geometry of logarithmic schemes*, Ph.D. thesis, Brown University, 2015, available at https://tinyurl.com/ulirthes.

[Włodarczyk 2000] J. Włodarczyk, "Birational cobordisms and factorization of birational maps", *J. Algebraic Geom.* **9**:3 (2000), 425–449. MR Zbl arXiv

[Włodarczyk 2003] J. Włodarczyk, "Toroidal varieties and the weak factorization theorem", *Invent. Math.* **154**:2 (2003), 223–331. MR Zbl

[Włodarczyk 2006] J. Włodarczyk, "Algebraic Morse theory and the weak factorization theorem", pp. 653–682 in *International Congress of Mathematicians, II* (Madrid, 2006), edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Włodarczyk 2009] J. Włodarczyk, "Simple constructive weak factorization", pp. 957–1004 in *Algebraic geometry, II* (Seattle, 2005), edited by D. Abramovich et al., Proc. Sympos. Pure Math. **80, Part 2**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

[Yu 2018] T. Y. Yu, "Gromov compactness in non-Archimedean analytic geometry", *J. Reine Angew. Math.* **741** (2018), 179–210. MR Zbl

Communicated by Gavril Farkas
Received 2017-12-20 Revised 2018-11-23 Accepted 2019-01-04

abrmovic@math.brown.edu *Department of Mathematics, Brown University, Providence, RI, United States*

temkin@math.huji.ac.il *Einstein Institute of Mathematics, The Hebrew University of Jerusalem, Jerusalem, Israel*

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.425

Effective generation and twisted weak positivity of direct images

Yajnaseni Dutta and Takumi Murayama

We study pushforwards of log pluricanonical bundles on projective log canonical pairs (Y, Δ) over the complex numbers, partially answering a Fujita-type conjecture due to Popa and Schnell in the log canonical setting. We show two effective global generation results. First, when Y surjects onto a projective variety, we show a quadratic bound for generic generation for twists by big and nef line bundles. Second, when Y is fibered over a smooth projective variety, we show a linear bound for twists by ample line bundles. These results additionally give effective nonvanishing statements. We also prove an effective weak positivity statement for log pluricanonical bundles in this setting, which may be of independent interest. In each context we indicate over which loci positivity holds. Finally, using the description of such loci, we show an effective vanishing theorem for pushforwards of certain log-sheaves under smooth morphisms.

1. Introduction

Throughout this paper, all varieties will be over the complex numbers.

Popa and Schnell proposed the following relative version of Fujita's conjecture:

Conjecture 1.1 [Popa and Schnell 2014, Conjecture 1.3]. *Let $f : Y \to X$ be a morphism of smooth projective varieties*, *with* $\dim X = n$, *and let $\mathcal{L}$ be an ample line bundle on X. For each $k \geq 1$, the sheaf*

$$f_*\omega_Y^{\otimes k} \otimes \mathcal{L}^{\otimes \ell}$$

is globally generated for all $\ell \geq k(n+1)$.

MSC2010: primary 14C20; secondary 14D06, 14F05, 14E30, 14Q20, 14J17.
Keywords: pluricanonical bundles, Fujita's conjecture, weak positivity, effective results, Seshadri constants.

Additionally assuming that $\mathcal{L}$ is globally generated, Popa and Schnell proved Conjecture 1.1 more generally for log canonical pairs (Y, Δ). Previously, Deng [2017, Theorem C] and the first author [Dutta 2017, Proposition 1.2] studied this conjecture for klt $\mathbb{Q}$-pairs, and were able to remove the global generation assumption on $\mathcal{L}$ to obtain generic effective generation statements. In this paper, we obtain similar generic generation results, more generally for log canonical pairs (Y, Δ).

First, when X is arbitrarily singular and $\mathcal{L}$ is only big and nef, we obtain the following quadratic bound on ℓ. The case when (Y, Δ) is klt and $k = 1$ is due to de Cataldo [1998, Theorem 2.2].

Theorem A. *Let $f : Y \to X$ be a surjective morphism of projective varieties, where X is of dimension n. Let (Y, Δ) be a log canonical $\mathbb{R}$-pair and let $\mathcal{L}$ be a big and nef line bundle on X. Consider a Cartier divisor P on Y such that $P \sim_{\mathbb{R}} k(K_Y + \Delta)$ for some integer $k \geq 1$. Then, the sheaf*

$$f_*\mathcal{O}_Y(P) \otimes_{\mathcal{O}_X} \mathcal{L}^{\otimes \ell}$$

is generated by global sections on an open set U for every integer $\ell \geq k(n^2 + 1)$.

On the other hand, we have the following linear bound when X is smooth and $\mathcal{L}$ is ample. The statement in (i) extends [Deng 2017, Theorem C] to log canonical pairs. As we were writing this, we learned that a statement similar to (ii) was also obtained by Iwai [2017, Theorem 1.5].

Theorem B. *Let $f : Y \to X$ be a fibration of projective varieties, where X is smooth of dimension n. Let (Y, Δ) be a log canonical $\mathbb{R}$-pair and let $\mathcal{L}$ be an ample line bundle on X. Consider a Cartier divisor P on Y such that $P \sim_{\mathbb{R}} k(K_Y + \Delta)$ for some integer $k \geq 1$. Then, the sheaf*

$$f_*\mathcal{O}_Y(P) \otimes_{\mathcal{O}_X} \mathcal{L}^{\otimes \ell}$$

is globally generated on an open set U for

(i) *every integer $\ell \geq k(n+1) + n^2 - n$; and*

(ii) *every integer $\ell > k(n+1) + \frac{1}{2}(n^2 - n)$ when (Y, Δ) is a klt $\mathbb{Q}$-pair.*

Here, a *fibration* is a morphism whose generic fiber is irreducible.

In both Theorems A and B, when Y is smooth and Δ has simple normal crossing support, we have explicit descriptions of the open set U. See Remark 5.1. Thus, we have descriptions of the loci where global generation holds up to a log resolution.

When X is smooth of dimension ≤ 3 and $\mathcal{L}$ is ample, the bound on ℓ can be improved. This gives the predicted bound in Conjecture 1.1 for surfaces; see Remark 5.2.

Remark 1.2 (effective nonvanishing). Theorems A and B can be interpreted as effective nonvanishing statements. With notation as in the theorems, it follows that $f_*\mathcal{O}_Y(P) \otimes \mathcal{L}^{\otimes \ell}$ admits global sections for all $\ell \geq k(n^2 + 1)$ when $\mathcal{L}$ is big and nef, and for all $\ell \geq k(n+1) + n^2 - n$ when $\mathcal{L}$ is ample and X is smooth. Moreover, just as in Theorem B(ii), the effective bound of the second nonvanishing statement can be improved in the case when (Y, Δ) is a klt $\mathbb{Q}$-pair.

We now state the technical results used in proving Theorems A and B.

An extension theorem. Recall that if $\mu : X' \to X$ is the blow-up of a projective variety X at x with exceptional divisor E, then the *Seshadri constant* of a nef Cartier divisor L at x is

$$\varepsilon(L;x) := \sup\{t \in \mathbb{R}_{\geq 0} \mid \mu^*L - tE \text{ is nef}\}.$$

The following replaces the role of Deng's extension theorem [2017, Theorem 2.11] in our proofs.

Theorem C. *Let $f : Y \to X$ be a surjective morphism of projective varieties, where X is of dimension n and Y is smooth. Let Δ be an $\mathbb{R}$-divisor on Y with simple normal crossing support and coefficients in* $(0,1]$, *and let L be a big and nef $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor on X. Suppose there exists a closed point $x \in U(f,\Delta)$ and a real number $\ell > n/\varepsilon(L;x)$ such that*

$$P_\ell \sim_{\mathbb{R}} K_Y + \Delta + \ell f^*L$$

for some Cartier divisor P_ℓ on Y. Then, the restriction map

$$H^0(Y, \mathcal{O}_Y(P_\ell)) \longrightarrow H^0(Y_x, \mathcal{O}_{Y_x}(P_\ell)) \tag{1}$$

is surjective, and the sheaf $f_*\mathcal{O}_Y(P_\ell)$ is globally generated at x.

See Notation 2.1(a) for the definition of the open set $U(f,\Delta)$.

Remark 1.3 (comments on the proofs). The proofs of Theorems A and B(i) are in a way an algebraization of Deng's techniques, exploiting a generic lower bound for Seshadri constants due to Ein, Küchle, and Lazarsfeld (Theorem 2.20). In the algebraic setting, this lower bound was first used by de Cataldo to prove a version of Theorem A for klt pairs when $k = 1$. One of our main challenges was to extend de Cataldo's theorem to the log canonical case (see Theorem C above).

To obtain the better bound in Theorem B(ii) for klt $\mathbb{Q}$-pairs, we use [Dutta 2017, Proposition 1.2] instead of Seshadri constants.

In Theorems A and C, in order to work with line bundles $\mathcal{L}$ that are big and nef instead of ample, we needed to study the augmented base locus $\boldsymbol{B}_+(\mathcal{L})$ of $\mathcal{L}$ (see Definition 2.22). We used Birkar's generalization of Nakamaye's theorem [Birkar 2017, Theorem 1.4] and [Küronya 2013, Proposition 2.7], which capture how $\mathcal{L}$ fails to be ample.

The proof of Theorem C relies on a cohomological injectivity theorem due to Fujino [2017a, Theorem 5.4.1]. If (Y,Δ) is replaced by an arbitrary log canonical $\mathbb{R}$-pair, then the global generation statement in Theorem C still holds over some open set (Corollary 3.2).

Remark 1.4 (effective vanishing). With the new input of weak positivity, which is discussed next, we give some effective vanishing statements for certain cases of such pushforwards under smooth morphisms (see Theorem 5.3). This improves similar statements in [Dutta 2017] and is in the spirit of [Popa and Schnell 2014, Proposition 3.1], where they showed a similar statement with the assumption that $\mathcal{L}$ is ample and globally generated.

Effective twisted weak positivity. In order to prove Theorem B, we also use the following weak positivity result for log canonical pairs. This may be of independent interest.

In this setting, weak positivity was partially known due to Campana [2004, Theorem 4.13], and later more generally due to Fujino [2017b, Theorem 1.1], but using a slightly weaker notion of weak positivity (see [loc. cit., Definition 7.3] and the comments thereafter). Our result extends these results.

Theorem D (twisted weak positivity). *Let $f : Y \to X$ be a fibration of normal projective varieties such that X is Gorenstein of dimension n. Let Δ be an $\mathbb{R}$-Cartier $\mathbb{R}$-divisor on Y such that (Y, Δ) is log canonical and $k(K_Y + \Delta)$ is $\mathbb{R}$-linearly equivalent to a Cartier divisor for some integer $k \geq 1$. Then, the sheaf*

$$f_*\mathcal{O}_Y(k(K_{Y/X} + \Delta))$$

is weakly positive.

Recall that a torsion-free coherent sheaf $\mathscr{F}$ is *weakly positive* if there exists a nonempty open set U such that for every integer a, there is an integer $b \geq 1$ such that

$$\mathrm{Sym}^{[ab]} \mathscr{F} \otimes H^{\otimes b}$$

is generated by global sections on U for all ample line bundles H. Here, $\cdot^{[s]}$ is the reflexive hull of $\cdot^{s}$ (see Notation 2.6).

Popa and Schnell [2014, Theorem 4.2] showed that if $\Delta = 0$, the morphism f has generically reduced fibers in codimension 1, and $H = \omega_X \otimes \mathcal{L}^{\otimes n+1}$ with $\mathcal{L}$ ample and globally generated, then weak positivity in Theorem D holds over $U(f, 0)$ for all $b \geq k$. In a similar spirit, we prove the following "effective" version of twisted weak positivity when Y is smooth and Δ has simple normal crossing support. Moreover, Theorem D is deduced from this result and therefore we also obtain an explicit description, up to a log resolution, of the locus over which weak positivity holds. This extends [Popa and Schnell 2014, Theorem 4.2] to arbitrary fibrations.

Theorem E (effective weak positivity). *Let $f : Y \to X$ be a fibration of projective varieties, where Y is smooth and X is normal and Gorenstein of dimension n. Let Δ be an $\mathbb{R}$-divisor on Y with simple normal crossing support and with coefficients of Δ^h in $(0, 1]$. Consider a Cartier divisor P on Y such that $P \sim_{\mathbb{R}} k(K_Y + \Delta)$ for some integer $k \geq 1$. Let U be the intersection of $U(f, \Delta)$ with the largest open set over which $f_*\mathcal{O}_Y(P)$ is locally free, and let $H = \omega_X \otimes \mathcal{L}^{\otimes n+1}$ for $\mathcal{L}$ an ample and globally generated line bundle on X. Then, the sheaf*

$$(f_*\mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{[s]} \otimes H^{\otimes \ell}$$

is generated by global sections on U for all integers $\ell \geq k$ and $s \geq 1$.

Here, Δ^h is the *horizontal part* of Δ; see Notation 2.1(b).

When $\lfloor \Delta \rfloor = 0$, one can, in a way, get rid of the assumption that $f_*\mathcal{O}_Y(P)$ is locally free on U using invariance of log plurigenera [Hacon et al. 2018, Theorem 4.2]; see Remark 4.2.

The proof of Theorem E relies on Viehweg's fiber product trick; see [Viehweg 1983, §3], [Popa and Schnell 2014, Theorem 4.2], or [Höring 2010, §3] for an exposition.

2. Definitions and preliminary results

Throughout this paper, a *variety* is an integral separated scheme of finite type over the complex numbers. We will also fix the following notation:

Notation 2.1. Let $f : Y \to X$ be a morphism of projective varieties, where Y is smooth, and let Δ be an $\mathbb{R}$-divisor with simple normal crossing support on Y.

(a) We denote by $U(f, \Delta)$ the largest open subset of X such that

- $U(f, \Delta)$ is contained in the smooth locus X_{reg} of X;
- $f : f^{-1}(U(f, \Delta)) \to U(f, \Delta)$ is smooth; and
- the fibers $Y_x := f^{-1}(x)$ intersect each component of Δ transversely for all closed points $x \in U(f, \Delta)$.

This open set $U(f, \Delta)$ is nonempty by generic smoothness; see [Hartshorne 1977, Corollary III.10.7] and [Lazarsfeld 2004a, Lemma 4.1.11].

(b) We write

$$\Delta = \Delta^v + \Delta^h,$$

where Δ^v and Δ^h do not share any components, such that

- every component of Δ^h is *horizontal* over X, i.e., surjects onto X; and
- Δ^v is *vertical* over X, i.e., $f(\mathrm{Supp}(\Delta^v)) \subsetneq X$.

Note that $U(f, \Delta)$ satisfies $U(f, \Delta) \cap f(\Delta^v) = \varnothing$.

Reflexive sheaves and weak positivity. In this section, fix an integral noetherian scheme X. To prove Theorem E, we need some basic results on reflexive sheaves, which we collect here.

Definition 2.2. A coherent sheaf $\mathscr{F}$ on X is *reflexive* if the natural morphism $\mathscr{F} \to \mathscr{F}^{\vee\vee}$ is an isomorphism, where $\mathscr{G}^{\vee} := \mathcal{H}om_{\mathcal{O}_X}(\mathscr{G}, \mathcal{O}_X)$. In particular, locally free sheaves are reflexive.

A coherent sheaf $\mathscr{F}$ on X is *normal* if the restriction map

$$\Gamma(U, \mathscr{F}) \longrightarrow \Gamma(U \smallsetminus Z, \mathscr{F})$$

is bijective for every open set $U \subseteq X$ and every closed subset Z of U of codimension at least 2.

Proposition 2.3 (see [Hartshorne 1994, Proposition 1.11]). *If X is normal, then every reflexive coherent sheaf $\mathscr{F}$ is normal.*

Lemma 2.4 [Stacks 2018, Tag 0AY4]. *Let $\mathscr{F}$ and $\mathscr{G}$ be coherent sheaves on X, and assume that $\mathscr{F}$ is reflexive. Then, $\mathcal{H}om_{\mathcal{O}_X}(\mathscr{G}, \mathscr{F})$ is also reflexive.*

We will often use these facts to extend morphisms from the complement of codimension at least 2, as recorded in the following:

Corollary 2.5. *Suppose X is normal, and let $\mathscr{F}$ and $\mathscr{G}$ be coherent sheaves on X such that $\mathscr{F}$ is reflexive. If $U \subseteq X$ is an open subset such that* $\operatorname{codim}(X \smallsetminus U) \geq 2$, *then every morphism $\varphi : \mathscr{G}|_U \to \mathscr{F}|_U$ extends uniquely to a morphism $\tilde{\varphi} : \mathscr{G} \to \mathscr{F}$.*

Proof. The morphism φ corresponds to a section of the sheaf $\mathcal{H}om_{\mathcal{O}_X}(\mathscr{G}, \mathscr{F})$ over U. The sheaf $\mathcal{H}om_{\mathcal{O}_X}(\mathscr{G}, \mathscr{F})$ is reflexive by Lemma 2.4; hence the section φ extends uniquely to a section $\tilde{\varphi}$ of $\mathcal{H}om_{\mathcal{O}_X}(\mathscr{G}, \mathscr{F})$ over X by Proposition 2.3. $\square$

We will use the following notation throughout this paper:

Notation 2.6 [Höring 2010, Notation 3.3]. Let $\mathscr{F}$ be a torsion-free coherent sheaf on a normal variety X. Let $i : X^* \hookrightarrow X$ be the largest open set such that $\mathscr{F}|_{X^*}$ is locally free. We define

$$\operatorname{Sym}^{[b]} \mathscr{F} := i_* \operatorname{Sym}^b(\mathscr{F}|_{X^*}) \quad \text{and} \quad \mathscr{F}^{[b]} := i_*((\mathscr{F}|_{X^*})^{\otimes b}).$$

We can also describe these sheaves as follows:

$$\operatorname{Sym}^{[b]} \mathscr{F} \simeq (\operatorname{Sym}^b(\mathscr{F}))^{\vee\vee} \quad \text{and} \quad \mathscr{F}^{[b]} \simeq (\mathscr{F}^{\otimes b})^{\vee\vee}.$$

Indeed, these pairs of reflexive sheaves coincide in codimension 1 and hence are isomorphic (see [Hartshorne 1994, Theorem 1.12]).

We can now define the positivity notion appearing in Theorem D.

Definition 2.7 (weak positivity [Viehweg 1983, Definition 1.2]). Let X be a normal variety, and let $U \subseteq X$ be an open set. A torsion-free coherent sheaf $\mathscr{F}$ on X is said to be *weakly positive on U* if for every positive integer a and every ample line bundle $\mathcal{L}$ on X, there exists an integer $b \geq 1$ such that $\operatorname{Sym}^{[ab]} \mathscr{F} \otimes \mathcal{L}^{\otimes b}$ is globally generated on U. We say $\mathscr{F}$ is *weakly positive* if $\mathscr{F}$ is weakly positive on some open set U.

Dualizing complexes and canonical sheaves. The main reference for this section is [Hartshorne 1966]. We define the following:

Definition 2.8. Let $h : X \to \operatorname{Spec} k$ be an equidimensional scheme of finite type over a field k. Then the *normalized dualizing complex* for X is $\omega_X^\bullet := h^! k$, where $h^!$ is the exceptional pullback of Grothendieck duality [loc. cit., Corollary VII.3.4]. One defines the *canonical sheaf* on X to be the coherent sheaf

$$\omega_X := \boldsymbol{H}^{-\dim X} \omega_X^\bullet.$$

When X is smooth and equidimensional over a field, the canonical sheaf ω_X is isomorphic to the invertible sheaf of volume forms $\Omega_X^{\dim X}$ [loc. cit., III.2].

We will need the explicit description of the exceptional pullback functor for finite morphisms. Let $\nu : Y \to X$ be a finite morphism of equidimensional schemes of finite type over a field. Consider the functor

$$\bar{\nu}^* : \mathsf{Mod}(\nu_* \mathcal{O}_Y) \longrightarrow \mathsf{Mod}(\mathcal{O}_Y)$$

obtained from the morphism $\bar{\nu} : (Y, \mathcal{O}_Y) \to (X, \nu_* \mathcal{O}_Y)$ of ringed spaces. This functor $\bar{\nu}^*$ satisfies the following properties (see [loc. cit., III.6]):

(a) The functor $\bar{\nu}^*$ is exact since the morphism $\bar{\nu}$ of ringed spaces is flat. We define the functor

$$\nu^! : \mathsf{D}^+(\mathsf{Mod}(\mathcal{O}_X)) \longrightarrow \mathsf{D}^+(\mathsf{Mod}(\mathcal{O}_Y)),$$
$$\mathscr{F} \longmapsto \bar{\nu}^* \, \boldsymbol{R}\mathcal{H}om_{\mathcal{O}_X}(\nu_*\mathcal{O}_Y, \mathscr{F}).$$

(b) For every $\mathcal{O}_X$-module $\mathscr{G}$, we have $\nu^*\mathscr{G} \simeq \bar{\nu}^*(\mathscr{G} \otimes_{\mathcal{O}_X} \nu_*\mathcal{O}_Y)$.

(c) If $\omega_X^\bullet$ is the normalized dualizing complex for X, then $\nu^! \omega_Y^\bullet$ is the normalized dualizing complex for Y.

Using the above description, we construct the following *pluri-trace map* for integral schemes over fields, which we will use in the proof of Theorem E. We presume that this construction is already known to the experts, but we could not find a reference.

Lemma 2.9. *Let $d : Y' \to Y$ be a dominant proper birational morphism of integral schemes of finite type over a field, where Y' is normal and Y is Gorenstein. Then, there is a map of pluricanonical sheaves*

$$d_* \omega_{Y'}^{\otimes k} \longrightarrow \omega_Y^{\otimes k}$$

which is an isomorphism where d is an isomorphism.

Proof. By the universal property of normalization [Stacks 2018, Tag 035Q], we can factor d as

$$\begin{tikzcd} Y' \arrow[r, "d'"] \arrow[rr, bend right, "d"'] & \overline{Y} \arrow[r, "\nu"] & Y \end{tikzcd}$$

where ν is the normalization. Note that d' is proper and birational since d is.

We first construct a similar morphism for ν. Let $n = \dim Y$. Since Y is Gorenstein, the canonical sheaf ω_Y is invertible and the normalized dualizing complex is $\omega_Y[n]$ [Hartshorne 1966, Proposition V.9.3]. Using property (c) above we have

$$\begin{aligned} \omega_{\overline{Y}} = \boldsymbol{H}^{-n}(\nu^! \omega_Y^\bullet) &\simeq \bar{\nu}^*(\boldsymbol{R}^{-n}\, \mathcal{H}om_{\mathcal{O}_Y}(\nu_*\mathcal{O}_{\overline{Y}}, \mathcal{O}_Y[n]) \otimes_{\mathcal{O}_Y} \omega_Y) \\ &\simeq \bar{\nu}^*(\mathcal{H}om_{\mathcal{O}_Y}(\nu_*\mathcal{O}_{\overline{Y}}, \mathcal{O}_Y) \otimes_{\mathcal{O}_Y} \omega_Y), \end{aligned}$$

where we get the first isomorphism since $\bar{\nu}^*$ is exact by (a) and since ω_Y is invertible.

Now $\mathcal{H}om_{\mathcal{O}_Y}(\nu_*\mathcal{O}_{\overline{Y}}, \mathcal{O}_Y)$ admits a morphism to $\nu_*\mathcal{O}_{\overline{Y}}$, which makes it the largest ideal in $\nu_*\mathcal{O}_{\overline{Y}}$ that is also an ideal in $\mathcal{O}_Y$. It is the so-called *conductor ideal* of the normalization map [Kollár 2013, (5.2)]. Thus, we get a morphism

$$\omega_{\overline{Y}} \hookrightarrow \bar{\nu}^*(\nu_*\mathcal{O}_{\overline{Y}} \otimes \omega_Y) \simeq \nu^*\omega_Y.$$

The last isomorphism follows from (b) above. By taking the $(k-1)$-fold tensor product of the above morphism we have

$$\omega_{\overline{Y}}^{\otimes(k-1)} \hookrightarrow \nu^* \omega_Y^{\otimes(k-1)}. \tag{2}$$

Finally, we use (2) to construct a map

$$d_* \omega_{Y'}^{\otimes k} \longrightarrow \nu^* \omega_Y^{\otimes (k-1)} \otimes_{\mathcal{O}_{\overline{Y}}} \omega_{\overline{Y}}.$$

First, we construct the above morphism over U, where d' is an isomorphism. Define $V := d'^{-1}(U)$. The identity map

$$\mathrm{id} : d'_* \omega_V^{\otimes k} \longrightarrow \omega_U^{\otimes k}$$

composed with map obtained from (2) gives the map

$$\tau : d'_* \omega_V^{\otimes k} \longrightarrow \nu^* \omega_Y^{\otimes (k-1)}|_U \otimes_{\mathcal{O}_U} \omega_U.$$

Since $\nu^* \omega_Y^{\otimes (k-1)}$ is invertible and $\omega_{\overline{Y}}$ is reflexive, the sheaf $\nu^* \omega_Y^{\otimes (k-1)} \otimes \omega_{\overline{Y}}$ is also reflexive. Now $\operatorname{codim}(Y \smallsetminus U) \geq 2$ by Zariski's main theorem; see [Hartshorne 1977, Theorem V.5.2]. Therefore by Corollary 2.5 we obtain

$$\tilde{\tau} : d'_* \omega_{Y'}^{\otimes k} \longrightarrow \nu^* \omega_Y^{\otimes (k-1)} \otimes_{\mathcal{O}_{\overline{Y}}} \omega_{\overline{Y}}.$$

Composing $\nu_* \tilde{\tau}$ with one copy of the trace morphism $\nu_* \omega_{\overline{Y}} \to \omega_Y$ [Hartshorne 1966, Proposition III.6.5], we get

$$d_* \omega_{Y'}^{\otimes k} \xrightarrow{\nu_* \tilde{\tau}} \nu_* (\nu^* \omega_Y^{\otimes (k-1)} \otimes_{\mathcal{O}_{\overline{Y}}} \omega_{\overline{Y}}) \simeq \omega_Y^{\otimes (k-1)} \otimes_{\mathcal{O}_Y} \nu_* \omega_{\overline{Y}} \xrightarrow{\mathrm{id} \otimes \mathrm{Tr}} \omega_Y^{\otimes k}. \tag{3}$$

The statement about the isomorphism locus of the above morphism holds by construction of the maps above. Indeed, in (3) the trace morphism is compatible with flat base change [Hartshorne 1966, Proposition III.6.6(2)], and hence compatible with restriction to the open set where d is an isomorphism. □

Singularities of pairs. We follow the conventions of [Fujino 2017a, §2.3]; see also [Kollár 2013, §1.1,2.1]. Recall that X_{reg} denotes the regular locus of a scheme X; see Notation 2.1(a).

Definition 2.10 (canonical divisor). Let X be a normal variety of dimension n. A *canonical divisor* K_X on X is a Weil divisor such that

$$\mathcal{O}_{X_{\mathrm{reg}}}(K_X) \simeq \Omega^n_{X_{\mathrm{reg}}}.$$

The choice of a canonical divisor K_X is unique up to linear equivalence. Then one defines $\mathcal{O}_X(K_X)$ to be the reflexive sheaf of rank 1 associated to K_X.

The following lemma allows us to freely pass between divisor and sheaf notation on normal varieties:

Lemma 2.11. *Let X be a normal variety of dimension n. Then, $\mathcal{O}_X(K_X)$ is isomorphic to ω_X.*

Proof. The sheaf $\mathcal{O}_X(K_X)$ is reflexive by definition and the canonical sheaf ω_X is S_2, by [Stacks 2018, Tag 0AWE], and hence reflexive, by [Hartshorne 1994, Theorem 1.9]. Since they are both isomorphic to $\Omega^n_{X_{\mathrm{reg}}}$ on X_{reg} and $\operatorname{codim}(X \smallsetminus X_{\mathrm{reg}}) \geq 2$, we have $\mathcal{O}_X(K_X) \simeq \omega_X$ by [loc. cit., Theorem 1.12]. □

Definition 2.12 (discrepancy). Let (X, Δ) be a pair consisting of a normal variety X and an $\mathbb{R}$-divisor Δ on X such that $K_X + \Delta$ is $\mathbb{R}$-Cartier. Suppose $f : Y \to X$ is a proper birational morphism from a normal variety Y, and choose canonical divisors K_Y and K_X such that $f_* K_Y = K_X$. In this case, we may write

$$K_Y = f^*(K_X + \Delta) + \sum_i a(E_i, X, \Delta) E_i,$$

where the E_i are irreducible Weil divisors. The real number $a(E_i, X, \Delta)$ is called the *discrepancy of E_i* with respect to (X, Δ), and the *discrepancy* of (X, Δ) is

$$\operatorname{discrep}(X, \Delta) = \inf_E \{a(E, X, \Delta) \mid E \text{ is an exceptional divisor over } X\},$$

where the infimum runs over all irreducible exceptional divisors of all proper birational morphisms $f : Y \to X$.

Definition 2.13 (singularities of pairs). Let (X, Δ) be a pair consisting of a normal variety X and an effective $\mathbb{R}$-divisor Δ on X such that $K_X + \Delta$ is $\mathbb{R}$-Cartier. We say that (X, Δ) is *klt* if $\operatorname{discrep}(X, \Delta) > -1$ and $\lfloor \Delta \rfloor = 0$. We say that (X, Δ) is *log canonical* if $\operatorname{discrep}(X, \Delta) \geq -1$.

We will repeatedly use the following results about log resolutions of log canonical $\mathbb{R}$-pairs.

Lemma 2.14. *Let (Y, Δ) be a log canonical (resp. klt) $\mathbb{R}$-pair, and consider a Cartier divisor P on Y such that $P \sim_{\mathbb{R}} k(K_Y + \Delta + H)$ for some integer $k \geq 1$ and some $\mathbb{R}$-Cartier $\mathbb{R}$-divisor H. Then, for every proper birational morphism $\mu : \widetilde{Y} \to Y$ such that $\widetilde{Y}$ is smooth and $\mu^{-1}(\Delta) + \operatorname{exc}(\mu)$ has simple normal crossing support, there exists a divisor $\widetilde{P}$ on $\widetilde{Y}$ and an $\mathbb{R}$-divisor $\tilde{\Delta}$ such that*

(i) *$\tilde{\Delta}$ has coefficients in $(0, 1]$ (resp. $(0, 1)$) and simple normal crossing support;*

(ii) *the divisor $\widetilde{P} - \mu^* P$ is an effective divisor with support in $\operatorname{Supp}(\operatorname{exc}(\mu))$;*

(iii) *the divisor $\widetilde{P}$ satisfies $\widetilde{P} \sim_{\mathbb{R}} k(K_{\widetilde{Y}} + \tilde{\Delta} + \mu^* H)$; and*

(iv) *there is an isomorphism $\mu_* \mathcal{O}_{\widetilde{Y}}(\widetilde{P}) \simeq \mathcal{O}_Y(P)$.*

Proof. On $\widetilde{Y}$, we can write

$$K_{\widetilde{Y}} - \mu^*(K_Y + \Delta) = Q - N,$$

where Q and N are effective $\mathbb{R}$-divisors without common components such that $Q - N$ has simple normal crossing support and Q is μ-exceptional. Note that since (Y, Δ) is log canonical (resp. klt), all coefficients in N are less than or equal to 1 (resp. less than 1). Let

$$\tilde{\Delta} := N + \lceil Q \rceil - Q$$

so that by definition $\tilde{\Delta}$ has simple normal crossing support and coefficients in $(0, 1]$ (resp. $(0, 1)$). Now setting $\widetilde{P} := \mu^* P + k \lceil Q \rceil$, we have

$$\begin{aligned} \widetilde{P} &\sim_{\mathbb{R}} k\mu^*(K_Y + \Delta + H) + k\lceil Q \rceil \\ &\sim_{\mathbb{R}} kK_{\widetilde{Y}} + k(N + \lceil Q \rceil - Q) + \mu^* H = k(K_{\widetilde{Y}} + \tilde{\Delta} + \mu^* H). \end{aligned}$$

Since $\lceil Q \rceil$ is μ-exceptional, we get $\mu_* \mathcal{O}_{\widetilde{Y}}(\widetilde{P}) \simeq \mathcal{O}_Y(P)$ by using the projection formula. $\square$

We also use the following stronger notion of log resolution due to Szabó:

Theorem 2.15 [Kollár 2013, Theorem 10.45.2]. *Let X be a variety, and let D be a Weil divisor on X. Then, there is a log resolution $\mu : \widetilde{X} \to X$ of (X, D) such that μ is an isomorphism over the locus where X is smooth and D has simple normal crossing support.*

A few tools from Popa–Schnell. The following result is a slight generalization of [Popa and Schnell 2014, Variant 1.6]. This will be instrumental in proving Theorems D and E.

Theorem 2.16. *Let $f : Y \to X$ be a morphism of projective varieties, where Y is normal and X is of dimension n. Let Δ be an $\mathbb{R}$-divisor on Y and H a semiample $\mathbb{Q}$-divisor on X such that for some integer $k \geq 1$, there is a Cartier divisor P on Y satisfying*

$$P \sim_{\mathbb{R}} k(K_Y + \Delta + f^*H).$$

Suppose, moreover, that Δ can be written as $\Delta = \Delta' + \Delta^v$, where (Y, Δ') is log canonical and Δ^v is an $\mathbb{R}$-Cartier $\mathbb{R}$-divisor that is vertical over X. Let $\mathcal{L}$ be an ample and globally generated line bundle on X. Then, the sheaf

$$f_*\mathcal{O}_Y(P) \otimes \mathcal{L}^{\otimes \ell}$$

is generated by global sections on some open set U for all $\ell \geq k(n+1)$. Moreover, when Δ' has simple normal crossing support, we have $U = X \smallsetminus f(\mathrm{Supp}(\Delta^v))$.

Proof. Possibly after a log resolution of (Y, Δ), we may assume that $\Delta = \Delta^h + \Delta^v$ in the sense of Notation 2.1(b) such that (Y, Δ^h) is log canonical and Δ has simple normal crossing support. Indeed, let $\mu : \widetilde{Y} \to Y$ be a log resolution of (Y, Δ). Then, by Lemma 2.14 applied to the pair (Y, Δ') and $H = \Delta^v$, we obtain a log canonical $\mathbb{R}$-divisor $\tilde{\Delta}$ with simple normal crossing support on $\widetilde{Y}$ satisfying

$$K_{\widetilde{Y}} + \tilde{\Delta} + \mu^*\Delta^v \sim_{\mathbb{R}} \mu^*(K_Y + \Delta) + N,$$

where N is an effective μ-exceptional divisor. We rename $\widetilde{Y}$ and $\tilde{\Delta} + \mu^*\Delta^v$ as Y and Δ respectively.

Now Δ has simple normal crossing support and Δ^h is log canonical. Moreover, since f^*H is semiample, by Bertini's theorem we can pick a $\mathbb{Q}$-divisor $D \sim_{\mathbb{Q}} f^*H$ with smooth support and satisfying the conditions that $D + \Delta$ has simple normal crossing support and D does not share any components with Δ. Letting $\Delta'' := \Delta^v - \lfloor \Delta^v \rfloor$, we have

$$\Delta = \Delta^h + \Delta'' + \lfloor \Delta^v \rfloor$$

and $(Y, \Delta^h + \Delta'' + D)$ is log canonical. Since $\mathcal{L}$ is ample and globally generated, we therefore obtain that

$$f_*\mathcal{O}_Y(k(K_Y + \Delta^h + \Delta'' + f^*H)) \otimes \mathcal{L}^{\otimes \ell}$$

is generated by global sections for all $\ell \geq k(n+1)$ by [Popa and Schnell 2014, Variant 1.6]. But

$$f_*\mathcal{O}_Y(k(K_Y + \Delta^h + \Delta'' + f^*H)) \otimes \mathcal{L}^{\otimes \ell} \longrightarrow f_*\mathcal{O}_Y(P) \otimes \mathcal{L}^{\otimes \ell},$$

and they have the same stalks at every point $x \in U$. Thus, the sheaf on the right-hand side is generated by global sections at x for all $x \in U$ and for all $\ell \geq k(n+1)$. □

We will also need the following result, which is used in the proof of [loc. cit., Variant 1.6]:

Lemma 2.17 (cf. [Popa and Schnell 2014, p. 2280]). *Let $f : Y \to X$ be a morphism of projective varieties, and let $\mathscr{F}$ be a coherent sheaf on Y such that the image of the counit map*

$$f^* f_* \mathscr{F} \longrightarrow \mathscr{F}$$

of the adjunction $f^* \dashv f_*$ is of the form $\mathscr{F}(-E)$ for some effective Cartier divisor E on Y. Then, for every effective Cartier divisor $E' \preceq E$, we have $f_*(\mathscr{F}(-E')) \simeq f_*\mathscr{F}$.

Proof. We have the factorization

$$f^* f_* \mathscr{F} \longrightarrow \mathscr{F}(-E') \hookrightarrow \mathscr{F}$$

and by applying the adjunction $f^* \dashv f_*$, we have a factorization

$$\begin{tikzcd} f_*\mathscr{F} \arrow[r] \arrow[rr, bend right, "\mathrm{id}"'] & f_*(\mathscr{F}(-E')) \arrow[r, hook] & f_*\mathscr{F} \end{tikzcd}$$

of the identity. □

Finally, we record the following numerical argument that will appear in the proofs of Theorems A and B.

Lemma 2.18 (cf. [Popa and Schnell 2014, Theorem 1.7, Step 2]). *Let X be a smooth projective variety. Let Δ be an effective $\mathbb{R}$-Cartier divisor and E an effective $\mathbb{Z}$-divisor with simple normal crossing support such that $\Delta + E$ also has simple normal crossing support and Δ has coefficients in $(0, 1]$. Let $0 \leq c < 1$ be a real number. Then, there exists an effective Cartier divisor $E' \preceq E$ such that $\Delta + cE - E'$ has simple normal crossing support and coefficients in $(0, 1]$.*

Seshadri constants. The effectivity of our results in Theorems A and B relies on Seshadri constants. These were originally introduced by Demailly to measure local positivity of line bundles and thereby study Fujita-type conjectures. See [Lazarsfeld 2004a, Chapter 5] for more on these invariants.

Definition 2.19. Let X be a projective variety, and let $x \in X$ be a closed point. Let L be a nef $\mathbb{R}$-Cartier $\mathbb{R}$-divisor on X. Denote by $\mu : X' \to X$ the blow-up of X at x with exceptional divisor E. The *Seshadri constant* of L at x is

$$\varepsilon(L; x) := \sup\{t \in \mathbb{R}_{\geq 0} \mid \mu^* L - tE \text{ is nef}\}.$$

If $\mathcal{L}$ is a nef line bundle, then we denote by $\varepsilon(\mathcal{L}; x)$ the Seshadri constant of the associated Cartier divisor L at x.

The following result is crucial in making our results effective.

Theorem 2.20 [Ein et al. 1995, Theorem 1]. *Let X be a projective variety of dimension n. Let L be a big and nef Cartier divisor on X. Then, for every $\delta > 0$, the locus*

$$\left\{ x \in X \;\middle|\; \varepsilon(L; x) > \frac{1}{n+\delta} \right\}$$

contains an open dense set.

Remark 2.21. If in the notation of Theorem 2.20 we also assume that X is smooth and L is ample, then better lower bounds are known if $n = 2, 3$. Under these additional assumptions, the locus

$$\left\{ x \in X \;\middle|\; \varepsilon(L; x) > \frac{1}{(n-1)+\delta} \right\}$$

contains an open dense set if $n = 2$ [Ein and Lazarsfeld 1993, Theorem] or $n = 3$ [Cascini and Nakamaye 2014, Theorem 1.2]. Here, we use [Ein et al. 1995, Lemma 1.4] to obtain results for general points from the cited results, which are stated for very general points. In general, it is conjectured that in the situation of Theorem 2.20, the locus

$$\left\{ x \in X \;\middle|\; \varepsilon(L; x) > \frac{1}{1+\delta} \right\}$$

contains an open dense set [Lazarsfeld 2004a, Conjecture 5.2.5].

The stable and augmented base locus. In order to deal with big and nef line bundles in Theorems A and C, we will need some facts about base loci, following [Ein et al. 2009].

Definition 2.22. Let X be a projective variety. If L is a $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor on X, then the *stable base locus* of L is the closed set

$$\boldsymbol{B}(L) := \bigcap_m \operatorname{Bs}|mL|_{\mathrm{red}},$$

where m runs over all integers such that mL is Cartier. If L is an $\mathbb{R}$-Cartier $\mathbb{R}$-divisor on X, the *augmented base locus* of L is the closed set

$$\boldsymbol{B}_+(L) := \bigcap_A \boldsymbol{B}(L - A),$$

where A runs over all ample $\mathbb{R}$-Cartier $\mathbb{R}$-divisors A such that $L - A$ is $\mathbb{Q}$-Cartier. By definition, if L is a $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor, then

$$\boldsymbol{B}(L) \subseteq \boldsymbol{B}_+(L).$$

Note that $\boldsymbol{B}_+(L) \neq X$ if and only if L is big by Kodaira's lemma [Lazarsfeld 2004a, Proposition 2.2.22].

We will also need the following result, which shows how augmented base loci and Seshadri constants are related. The result follows from [Ein et al. 2009, §6] if the scheme X is a smooth variety, but we will need it more generally for singular varieties.

Corollary 2.23. *Let X be a projective variety, and let $x \in X$ be a closed point. Suppose L is a big and nef $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor. If $\varepsilon(L; x) > 0$, then $x \notin \boldsymbol{B}_+(L)$.*

Proof. If $x \in \boldsymbol{B}_+(L)$, then by [Birkar 2017, Theorem 1.4] there exists a closed subvariety $V \subseteq X$ containing x such that $L^{\dim V} \cdot V = 0$, in which case $\varepsilon(L; x) = 0$ by [Lazarsfeld 2004a, Proposition 5.1.9]. □

3. An extension theorem

We now turn to the proof of Theorem C. The proof relies on the following application of cohomology and base change.

Lemma 3.1. *Let $f : Y \to X$ be a proper morphism of separated noetherian schemes, and let $\mathscr{F}$ be a coherent sheaf on Y. Let $x \in X$ be a point that has an open neighborhood $U \subseteq X$, where $\mathscr{F}|_{f^{-1}(U)}$ is flat over U. Consider the following cartesian square:*

$$\begin{tikzcd} Y_x \arrow[r] \arrow[d] & Y \arrow[d, "f"] \\ \operatorname{Spec}(\kappa(x)) \arrow[r] & X \end{tikzcd}$$

If the restriction map

$$H^0(Y, \mathscr{F}) \longrightarrow H^0(Y_x, \mathscr{F}|_{Y_x})$$

is surjective, then the restriction map

$$H^0(X, f_*\mathscr{F}) \longrightarrow f_*\mathscr{F} \otimes_{\mathcal{O}_X} \kappa(x)$$

is also surjective.

Proof. Let $f_U := f|_{f^{-1}(U)}$ and $\mathscr{F}_U := \mathscr{F}|_{f^{-1}(U)}$. We have the commutative diagram

$$\begin{tikzcd} H^0(X, f_*\mathscr{F}) \arrow[r] \arrow[dd, equal] & f_*\mathscr{F} \otimes_{\mathcal{O}_X} \kappa(x) \arrow[d, "\beta", "\wr"'] \\ & f_{U*}\mathscr{F}_U \otimes_{\mathcal{O}_U} \kappa(x) \arrow[d, "\alpha^0(x)"] \\ H^0(Y, \mathscr{F}) \arrow[r, two heads] & H^0(Y_x, \mathscr{F}|_{Y_x}) \end{tikzcd}$$

where the bottom arrow is surjective by assumption, β is an isomorphism by computing affine-locally, and $\alpha^0(x)$ is the natural base change map [Illusie 2005, (8.3.2.3)]. By the commutativity of the diagram, this map $\alpha^0(x)$ is surjective, and hence is an isomorphism by cohomology and base change [loc. cit., Corollary 8.3.11]. Thus, the top horizontal arrow is also surjective. □

Before proving Theorem C, we first explain how to deduce a generic global generation statement for arbitrary log canonical $\mathbb{R}$-pairs (Y, Δ) from Theorem C by passing to a log resolution.

Corollary 3.2. *Let $f : Y \to X$ be a surjective morphism of projective varieties, where X is of dimension n. Let (Y, Δ) be a log canonical $\mathbb{R}$-pair, and let L be an big and nef $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor on X. Let ℓ be a* *real number for which there exists a Cartier divisor P_ℓ on Y such that*

$$P_\ell \sim_{\mathbb{R}} K_Y + \Delta + \ell f^* L.$$

If $\ell > n/\varepsilon(L;x)$ for general $x \in X$, then the sheaf $f_* \mathcal{O}_Y(P_\ell)$ is generically globally generated.

Proof. Applying Lemma 2.14 for $H = \ell f^* L$ to a log resolution $\mu : \widetilde{Y} \to Y$ of (Y, Δ), we have the following commutative diagram:

$$\begin{array}{ccc} H^0(X, (f \circ \mu)_* \mathcal{O}_{\widetilde{Y}}(\widetilde{P}_\ell)) & \longrightarrow & (f \circ \mu)_* \mathcal{O}_{\widetilde{Y}}(\widetilde{P}_\ell) \otimes \kappa(x) \\ \uparrow \wr & & \uparrow \wr \\ H^0(X, f_* \mathcal{O}_Y(P_\ell)) & \longrightarrow & f_* \mathcal{O}_Y(P_\ell) \otimes \kappa(x) \end{array}$$

where $\widetilde{P}_\ell$ is the divisor on $\widetilde{Y}$ satisfying the properties in Lemma 2.14. Then, Theorem C for $(\widetilde{Y}, \tilde{\Delta})$ implies that for some open subset $U \subseteq X$, the top horizontal arrow is surjective for all closed points $x \in U$ such that $\ell > n/\varepsilon(L;x)$; hence the bottom horizontal arrow is also surjective at these closed points x. We therefore conclude that $f_* \mathcal{O}_Y(P_\ell)$ is generically globally generated. $\square$

To prove Theorem C, we need the following result on augmented base loci.

Lemma 3.3. *Let X be a projective variety of dimension n, and let L be a big and nef $\mathbb{R}$-Cartier $\mathbb{R}$-divisor on X. Let $x \in X$ be a closed point, and suppose $\varepsilon(L;x) > 0$. Let $\mu : X' \to X$ be the blow-up of X at x with exceptional divisor E. For every positive real number $\delta < \varepsilon(L;x)$, we have*

$$\boldsymbol{B}_+(\mu^* L - \delta E) \cap E = \varnothing.$$

In particular, if $\mu^* L - \delta E$ is a $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor, then

$$\mathrm{Bs}|m(\mu^* L - \delta E)| \cap E = \varnothing$$

for all sufficiently large and divisible integers m.

Proof. First, the $\mathbb{R}$-Cartier $\mathbb{R}$-divisor $\mu^* L - \delta E$ is big and nef since

$$\mu^* L - \delta E \sim_{\mathbb{R}} \frac{\delta}{\varepsilon(L;x)}(\mu^* L - \varepsilon(L;x) E) + \left(1 - \frac{\delta}{\varepsilon(L;x)}\right)\mu^* L \tag{4}$$

is the sum of a nef $\mathbb{R}$-Cartier $\mathbb{R}$-divisor and a big and nef $\mathbb{R}$-Cartier $\mathbb{R}$-divisor. Thus, by [Birkar 2017, Theorem 1.4], we know that $\boldsymbol{B}_+(\mu^* L - \delta E)$ is the union of positive-dimensional closed subvarieties V of X' such that $(\mu^* L - \delta E)^{\dim V} \cdot V = 0$.

It suffices to show such a V cannot contain any point $y \in E$. First, if $V \subseteq E$, then

$$(\mu^* L - \delta E)^{\dim V} \cdot V = (-\delta E)^{\dim V} \cdot V = \delta^{\dim V} (-E|_E)^{\dim V} \cdot V > 0,$$

since $\mathcal{O}_E(-E) \simeq \mathcal{O}_E(1)$ is very ample. On the other hand, if $V \not\subseteq E$, then V is the strict transform of some closed subvariety $V_0 \subseteq X$ containing x, and by (4), we have

$$\begin{aligned}(\mu^*L - \delta E)^{\dim V} \cdot V &= \left(\frac{\delta}{\varepsilon(L;x)}(\mu^*L - \varepsilon(L;x)E) + \left(1 - \frac{\delta}{\varepsilon(L;x)}\right)\mu^*L\right)^{\dim V} \cdot V \\ &\geq \left(1 - \frac{\delta}{\varepsilon(L;x)}\right)^{\dim V} (\mu^*L)^{\dim V} \cdot V \\ &= \left(1 - \frac{\delta}{\varepsilon(L;x)}\right)^{\dim V} L^{\dim V} \cdot V_0 > 0,\end{aligned}$$

where the first inequality is by nefness of $\mu^*L - \varepsilon(L;x)E$, and the last inequality is by [Lazarsfeld 2004a, Proposition 5.1.9] and the condition $\varepsilon(L;x) > 0$.

The last statement about base loci follows from the fact that

$$\boldsymbol{B}_+(\mu^*L - \delta E) \supseteq \boldsymbol{B}(\mu^*L - \delta E) = \mathrm{Bs}|m(\mu^*L - \delta E)|_{\mathrm{red}}$$

for all sufficiently large and divisible integers m, where the last equality holds by [loc. cit., Proposition 2.1.21] since $\mu^*L - \delta E$ is a $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor. □

Finally, we need the following cohomological injectivity theorem due to Fujino.

Theorem 3.4 [Fujino 2017a, Theorem 5.4.1]. *Let Y be a smooth complete variety and let Δ be an $\mathbb{R}$-divisor on Y with coefficients in* $(0,1]$ *and simple normal crossing support. Let L be a Cartier divisor on Y and let D be an effective Weil divisor on Y whose support is contained in* $\operatorname{Supp}\Delta$. *Assume that $L \sim_{\mathbb{R}} K_Y + \Delta$. Then, the natural homomorphism*

$$H^i(Y, \mathcal{O}_Y(L)) \longrightarrow H^i(Y, \mathcal{O}_Y(L+D))$$

induced by the inclusion $\mathcal{O}_Y \to \mathcal{O}_Y(D)$ is injective for every i.

We can now prove Theorem C.

Proof of Theorem C. Fix $x \in U$, and consider the cartesian square

$$\begin{array}{ccc} Y' & \xrightarrow{B} & Y \\ {\scriptstyle f'}\downarrow & & \downarrow{\scriptstyle f} \\ X' & \xrightarrow{b} & X \end{array}$$

where b is the blow-up of X at x. Since f is flat in a neighborhood of x, the morphism B can be identified with the blow-up of Y along Y_x, which is a smooth subvariety of codimension n [Stacks 2018, Tag 0805]. Moreover, if E is the exceptional divisor of b and D is the exceptional divisor of B, then $f'^*E = D$. By Lemma 3.1, the surjectivity of (1) in the statement of Theorem C implies the generic global generation statement, so it suffices to show that the map in (1) is surjective.

First, we note that $(Y', B^*\Delta)$ is log canonical: since Y_x intersects every component of Δ transversely, the pullback $B^*\Delta$ of Δ is equal to the strict transform Δ' of Δ [Fulton 1984, Corollary 6.7.2], and so in particular, (Y', Δ') is log canonical.

Since $\varepsilon(L; x) > n/\ell$, we can choose a sufficiently small $\delta > 0$ such that $(n+\delta)/\ell \in \mathbb{Q}$ and $\varepsilon(L; x) > (n+\delta)/\ell$. Thus, using the fact that L is a $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor, for real numbers m of the form m_0/ℓ for sufficiently large and divisible integers m_0, we have that $m(\ell b^*L - (n+\delta)E)$ is Cartier. Lemma 3.3 then implies

$$S := \operatorname{Bs}|m(\ell b^*L - (n+\delta)E)|_{\mathrm{red}}$$

does not intersect E, i.e., $m(\ell b^*L - (n+\delta)E)$ is globally generated away from S, and in particular, is globally generated on an open set containing E. Thus, the pullback $m(\ell B^* f^*L - (n+\delta)D)$ of this divisor is globally generated away from $S' := f'^{-1}(S)$, and in particular is globally generated on an open set containing D. Choose

$$\mathfrak{D}_x \in |m(\ell B^* f^*L - (n+\delta)D)|$$

which is smooth and irreducible away from $f'^{-1}(S)$, and is such that the component of $\mathfrak{D}_x$ not contained in S' intersects each component of the support of Δ' transversely away from S'. Note that such a choice is possible by applying Bertini's theorem [Hartshorne 1977, Corollary III.10.9 and Remark III.10.9.3]. Since $\mathfrak{D}_x$ may have singularities along S', however, we will need to pass to a log resolution before applying Theorem 3.4.

By Theorem 2.15, there exists a common log resolution $\mu : \widetilde{Y} \to Y'$ for $\mathfrak{D}_x$ and (Y', Δ') that is an isomorphism away from $f'^{-1}(S) \subsetneq Y'$. We then write

$$\mu^*\mathfrak{D}_x = D' + F, \quad \mu^*\Delta' = \mu_*^{-1}\Delta' + F_1,$$

where D' is a smooth divisor intersecting Y_x transversely and F, F_1 are supported on $\mu^{-1}(S')$. Define

$$F' := \left\lfloor \frac{1}{m}F + F_1 \right\rfloor, \quad \tilde{\Delta} := \mu^*\Delta' + \frac{1}{m}\mu^*\mathfrak{D}_x - F' + \delta\mu^*D, \quad \widetilde{P}_\ell := \mu^*B^*P_\ell + K_{\widetilde{Y}/Y'}.$$

Note that $\tilde{\Delta}$ has simple normal crossing support containing μ^*D, and has coefficients in $(0, 1]$ by assumption on the log resolution and by definition of F'. Note also that

$$\begin{aligned}
\widetilde{P}_\ell - F' &\sim_{\mathbb{R}} \mu^*B^*(K_Y + \Delta + \ell f^*L) + K_{\widetilde{Y}/Y'} - F' \\
&\sim_{\mathbb{R}} K_{\widetilde{Y}} + \mu^*\Delta' - F' + \mu^*(\ell B^* f^*L - (n-1)D) \\
&\sim_{\mathbb{R}} K_{\widetilde{Y}} + \mu^*\Delta' + \frac{1}{m}\mu^*\mathfrak{D}_x - F' + (1+\delta)\mu^*D \\
&\sim_{\mathbb{R}} K_{\widetilde{Y}} + \tilde{\Delta} + \mu^*D,
\end{aligned}$$

where the second equivalence follows from the fact that B is the blow-up of the smooth subvariety Y_x, which is of codimension n; hence

$$K_{\widetilde{Y}} = \mu^*K_{Y'} + K_{\widetilde{Y}/Y'} = \mu^*B^*K_Y + (n-1)\mu^*D + K_{\widetilde{Y}/Y'}.$$

We can now apply the injectivity result Theorem 3.4 to $\widetilde{P}_\ell - F' - \mu^* D \sim_{\mathbb{R}} K_{\widetilde{Y}} + \tilde{\Delta}$ to see that

$$H^1(\widetilde{Y}, \mathcal{O}_{\widetilde{Y}}(\widetilde{P}_\ell - F' - \mu^* D)) \longrightarrow H^1(\widetilde{Y}, \mathcal{O}_{\widetilde{Y}}(\widetilde{P}_\ell - F')) \tag{5}$$

is injective. Next, consider the following commutative diagram:

$$\begin{tikzcd}
H^0(\widetilde{Y}, \mathcal{O}_{\widetilde{Y}}(\widetilde{P}_\ell - F')) \arrow[r, two heads] \arrow[d, hook] & H^0(\mu^*(D), \mathcal{O}_{\mu^*(D)}(\widetilde{P}_\ell - F')) \arrow[d, hook, "\wr"] \\
H^0(\widetilde{Y}, \mathcal{O}_{\widetilde{Y}}(\widetilde{P}_\ell)) \arrow[r] & H^0(\mu^*(D), \mathcal{O}_{\mu^*(D)}(\widetilde{P}_\ell)) \\
H^0(Y', \mathcal{O}_{Y'}(B^* P_\ell)) \arrow[r] \arrow[u, "\wr"] & H^0(D, \mathcal{O}_D(B^* P_\ell)) \arrow[u, "\wr"] \\
H^0(Y, \mathcal{O}_Y(P_\ell)) \arrow[r] \arrow[u, "\wr"] & H^0(Y_x, \mathcal{O}_{Y_x}(P_\ell)) \arrow[u, "\wr"]
\end{tikzcd}$$

The top right vertical arrow is an isomorphism since F' is disjoint from $\mu^*(D)$. The bottom right vertical arrow is an isomorphism since $B|_D$ realizes D as a projective bundle over Y_x; hence $(B|_D)_* \mathcal{O}_D \simeq \mathcal{O}_{Y_x}$. The other vertical isomorphisms follow from the projection formula and the fact that μ and B are birational. Finally, the top horizontal arrow is surjective by the long exact sequence on cohomology and the injectivity of (5). The commutativity of the diagram implies the bottom row is surjective, which is exactly the map in (1). □

4. Effective twisted weak positivity

We now prove Theorem E using Viehweg's fiber product trick. This trick enables us to reduce the global generation of the reflexivized s-fold tensor product $f_* \mathcal{O}_Y(k(K_Y + \Delta))^{[s]}$ to $s = 1$ with Y replaced by a suitable $\widetilde{Y}^s$. The main obstacle is picking a suitable boundary divisor on $\widetilde{Y}^s$. We tackle this using Theorem 2.16. Readers are encouraged to consult [Popa and Schnell 2014, §4], [Viehweg 1983, §3], or [Höring 2010, §3].

Throughout the proof we use $\mathcal{O}_X(K_X)$ and ω_X interchangeably whenever X is a normal variety. We can do so by Lemma 2.11.

Proof of Theorem E. For every positive integer s, let Y^s denote the reduction of the unique irreducible component of

$$\underbrace{Y \times_X Y \times_X \cdots \times_X Y}_{s \text{ times}}$$

that surjects onto X; note that it is unique since f has irreducible generic fiber. Setting $V := f^{-1}(U)$, we define V^s similarly.

Let $d : Y^{(s)} \to Y^s$ be a desingularization of Y^s, and note that d is an isomorphism over V^s. We will also denote by V^s the image of V^s under any birational modification of Y^s which is an isomorphism along V^s. Define $d_i = \pi_i \circ d$ for $i \in \{1, 2, \dots, s\}$, where $\pi_i : Y^s \to Y$ is the i-th projection. Since d_i is a surjective morphism between integral varieties, the pullback $d_i^* \Delta_j$ of the Cartier divisor Δ_j is well-defined for every component Δ_j of Δ; see [Stacks 2018, Tag 02OO(1)].

Let $\mu : \widetilde{Y}^s \to Y^{(s)}$ be a log resolution as in Theorem 2.15 of the pair $\bigl(Y^{(s)}, \sum_i d_i^* \Delta\bigr)$ so that μ is an isomorphism over V^s. Define

$$\tilde{\Delta} = \mu^* \sum_i d_i^* \Delta.$$

Claim 4.1. *There exists a map*

$$\tilde{f}^s_* \mathcal{O}_{\widetilde{Y}^s}(k(K_{\widetilde{Y}^s/X} + \tilde{\Delta})) \longrightarrow (f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{[s]} \tag{6}$$

which is an isomorphism over U.

Let X_0 be the open set in X such that

- the map f is flat over X_0;
- the regular locus of X contains X_0; and
- the sheaf $f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta))$ is locally free over X_0.

Then, $\operatorname{codim}(X \smallsetminus X_0) \geq 2$. Indeed, X is normal and both $f_* \mathcal{O}_Y$ and $f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta))$ are torsion-free. Now by construction, we have $U \subseteq X_0$. Since $(f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{[s]}$ is reflexive and is isomorphic to $(f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{\otimes s}$ on X_0, a map

$$\tilde{f}^s_* \mathcal{O}_{\widetilde{Y}^s}(k(K_{\widetilde{Y}^s/X} + \tilde{\Delta})) \longrightarrow (f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{\otimes s}$$

over X_0 will extend to a map of the form in (6) on X by Corollary 2.5. This, together with flat base change [Hartshorne 1977, Proposition III.9.3], implies that it suffices to construct a map

$$\tilde{f}^s_* \mathcal{O}_{\widetilde{Y}^s_0}(k(K_{\widetilde{Y}^s_0/X_0} + \tilde{\Delta}|_{\widetilde{Y}^s_0})) \longrightarrow (f_* \mathcal{O}_{Y_0}(k(K_{Y_0/X_0} + \Delta|_{Y_0})))^{\otimes s}$$

which is an isomorphism over U.

Define $Y_0 := f^{-1}(X_0)$. In this case, by [Höring 2010, Corollary 5.24] we know that

$$Y_0^s := \underbrace{Y_0 \times_X Y_0 \times_X \cdots \times_X Y_0}_{s \text{ times}} \simeq \underbrace{Y_0 \times_{X_0} Y_0 \times_{X_0} \cdots \times_{X_0} Y_0}_{s \text{ times}}$$

and that Y_0^s is Gorenstein. We can therefore apply Lemma 2.9 to $d \circ \mu$, to obtain a morphism

$$(d \circ \mu)_* \omega^{\otimes k}_{\widetilde{Y}^s_0/X_0} \longrightarrow \omega^{\otimes k}_{Y_0^s/X_0}$$

which is an isomorphism over V^s. Here $\omega_{Y_0^s/X_0} := \omega_{Y_0} \otimes f^{s*} \omega_{X_0}^{-1}$ and we define $\omega_{\widetilde{Y}^s_0/X_0}$ similarly. This induces a map

$$\tilde{f}^s_* \mathcal{O}_{\widetilde{Y}^s_0}(k(K_{\widetilde{Y}^s_0/X_0} + \tilde{\Delta}|_{\widetilde{Y}^s_0})) \longrightarrow f^s_* \Bigl(\omega^{\otimes k}_{Y_0^s/X_0} \otimes \bigotimes_i \pi_i^* \mathscr{M}|_{Y_0^s}\Bigr) \tag{7}$$

which is an isomorphism over U, where $\mathscr{M} := \mathcal{O}_Y(P - kK_Y)$ is the line bundle associated to the Cartier divisor $P - kK_Y \sim_{\mathbb{R}} k\Delta$.

We will now show that the sheaf on the right-hand side of (7) admits an isomorphism to

$$(f_*\mathcal{O}_{Y_0}(k(K_{Y_0/X_0} + \Delta|_{Y_0})))^{\otimes s}.$$

Note that this would show Claim 4.1, since (7) is an isomorphism over U. We proceed by induction, adapting the argument in [Höring 2010, Lemma 3.15] to our twisted setting. Note that the case $s = 1$ is clear, since in this case $Y^s = Y$ and the sheaves in question are equal.

By [loc. cit., Corollary 5.24] we have

$$\omega^{\otimes k}_{Y_0^s/X_0} \otimes \bigotimes_i \pi_i^*(\mathscr{M}|_{Y_0}) \simeq \pi_s^*(\omega^{\otimes k}_{Y_0/X_0} \otimes \mathscr{M}|_{Y_0}) \otimes \pi'^*(\omega^{\otimes k}_{Y_0^{s-1}/X_0} \otimes \mathscr{M}^{s-1}|_{Y_0^{s-1}}),$$

where $\pi' : Y^s \to Y^{s-1}$ and $\mathscr{M}^{s-1} := \bigotimes_{i=1}^{s-1} \pi_i^*\mathscr{M}$. Since $\omega^{\otimes k}_{Y_0^{s-1}/X_0} \otimes \mathscr{M}^{s-1}|_{Y_0^{s-1}}$ is locally free, by the projection formula we obtain

$$f_*^s\Big(\omega^{\otimes k}_{Y_0^s/X_0} \otimes \bigotimes_{i=1}^{s} \pi_i^*\mathscr{M}|_{Y_0}\Big) \simeq f_*\big((\omega^{\otimes k}_{Y_0/X_0} \otimes \mathscr{M}|_{Y_0}) \otimes \pi_{s*}\pi'^*(\omega^{\otimes k}_{Y_0^{s-1}/X_0} \otimes \mathscr{M}^{s-1}|_{Y_0^{s-1}})\big).$$

Now by flat base change [Hartshorne 1977, Proposition III.9.3],

$$\pi_{s*}\pi'^*(\omega^{\otimes k}_{Y_0^{s-1}/X_0} \otimes \mathscr{M}^{s-1}|_{Y_0^{s-1}}) \simeq f^* f_*^{s-1}(\omega^{\otimes k}_{Y_0^{s-1}/X_0} \otimes \mathscr{M}^{s-1}|_{Y_0^{s-1}}).$$

By induction the latter is isomorphic to

$$f^*(f_*\mathcal{O}_{Y_0}(k(K_{Y_0/X_0} + \Delta|_{Y_0}))^{\otimes s-1}).$$

Therefore

$$f_*^s\Big(\omega^{\otimes k}_{Y_0^s/X_0} \otimes \bigotimes_i \pi_i^*\mathscr{M}|_{Y_0}\Big) \simeq f_*\big(\omega^{\otimes k}_{Y_0/X_0} \otimes \mathscr{M}|_{Y_0} \otimes f^*\big(f_*\mathcal{O}_{Y_0}(k(K_{Y_0/X_0} + \Delta|_{Y_0}))^{\otimes s-1}\big)\big).$$

Since $f_*\mathcal{O}_Y(k(K_{Y/X} + \Delta))$ is locally free over X_0, we can apply the projection formula to obtain

$$f_*^s\Big(\omega^{\otimes k}_{Y_0^s/X_0} \otimes \bigotimes_i \pi_i^*\mathscr{M}|_{Y_0}\Big) \simeq (f_*\mathcal{O}_{Y_0}(k(K_{Y_0/X_0} + \Delta|_{Y_0})))^{\otimes s}.$$

This concludes the proof of Claim 4.1.

We now use Theorem 2.16 to finish the proof of Theorem E.

We first claim $\tilde{\Delta}$ satisfies the hypothesis of Theorem 2.16. To do so, first note that on π_i is flat over Y_0, and therefore by flat pullback of cycles we have

$$\pi_i^*(\Delta_j)|_{Y_0^s} = \pi_i^{-1}(\Delta_j|_{Y_0}) = Y_0 \times_{X_0} \cdots \times_{X_0} \underbrace{\Delta_j}_{i\text{-th position}} \times_{X_0} \cdots \times_{X_0} Y_0.$$

Since $Y_0 \supseteq V$ and both d and μ are isomorphisms over V^s, the pullbacks $\mu^*(\pi_i \circ d)^*\Delta_j^h|_{V^s}$ of the horizontal components of Δ are smooth above U for all $i \in \{1, 2, \dots, s\}$. In other words, the components of $\tilde{\Delta}$ either do not intersect V^s, or intersect the fiber over x transversely for all $x \in U$. Thus,

$$\tilde{\Delta}|_{V^s} = \mu^{-1} d^{-1} \sum_i \pi_i^{-1}(\Delta^h|_V).$$

In particular, using Notation 2.1(b), we have that the horizontal part $\tilde{\Delta}^h$ equals the closure $\overline{\tilde{\Delta}|_{V^s}}$ of $\tilde{\Delta}|_{V^s}$ in $\widetilde{Y}^s$. We can therefore write

$$\tilde{\Delta} = \tilde{\Delta}^h + \tilde{\Delta}^v,$$

where by construction, the coefficients of $\tilde{\Delta}^h$ are in $(0, 1]$ and $\tilde{f}^s(\tilde{\Delta}^v) \cap U = \varnothing$.

Finally, we note from Mori's cone theorem [Kollár and Mori 1998, Theorem 1.24] that $H = \omega_X \otimes \mathcal{L}^{\otimes n+1}$ is nef and hence semiample by the base point free theorem [loc. cit., Theorem 3.3]. Therefore $f^* H^{\otimes(\ell-k)}$ is also semiample for all $\ell \geq k$. Using H again to denote a divisor class of H, we argue that since

$$\tilde{f}^s_* \mathcal{O}_{\widetilde{Y}^s}(k(K_{\widetilde{Y}^s/X} + \tilde{\Delta})) \otimes H^{\otimes \ell} \simeq \tilde{f}^s_* \mathcal{O}_{\widetilde{Y}^s}\bigl(k(K_{\widetilde{Y}^s} + \tilde{\Delta} + (\ell - k)\tilde{f}^{s*} H)\bigr) \otimes \mathcal{L}^{\otimes k(n+1)}, \tag{8}$$

with $\mathcal{L}$ ample and globally generated, we can apply Theorem 2.16 to conclude that the sheaf above in (8) is generated by global sections over U for all $\ell \geq k$. Now fix a closed point $x \in U$. We have the commutative diagram

$$\begin{array}{ccc} H^0\bigl(X, \tilde{f}^s_* \mathcal{O}_{\widetilde{Y}^s}(k(K_{\widetilde{Y}^s/X} + \tilde{\Delta})) \otimes H^{\otimes \ell}\bigr) & \twoheadrightarrow & \bigl(\tilde{f}^s_* \mathcal{O}_{\widetilde{Y}^s}(k(K_{\widetilde{Y}^s/X} + \tilde{\Delta})) \otimes H^{\otimes \ell}\bigr) \otimes \kappa(x) \\ \downarrow & & \downarrow \wr \\ H^0\bigl(X, (f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{[s]} \otimes H^{\otimes \ell}\bigr) & \longrightarrow & \bigl((f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{[s]} \otimes H^{\otimes \ell}\bigr) \otimes \kappa(x) \end{array}$$

where the vertical arrows are induced by the map (6) from Claim 4.1, and the top horizontal arrow is surjective by the global generation of the sheaves in (8) over U. Since (6) is an isomorphism over U, the right vertical arrow is an isomorphism; hence by the commutativity of the diagram, the bottom horizontal arrow is surjective. We therefore conclude that

$$(f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{[s]} \otimes H^{\otimes \ell}$$

is generated by global sections over U for all $\ell \geq k$. □

Remark 4.2. When $\lfloor \Delta \rfloor = 0$, if we moreover take $U(f, \Delta)$ to be an open set over which every stratum of (Y, Δ) is smooth, then applying invariance of log plurigenera [Hacon et al. 2018, Theorem 4.2], we can assert that $f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta))|_{U(f,\Delta)}$ is locally free. In this case we can take X_0 to be simply the locus inside X_{reg} over which f is flat. Moreover, the isomorphism

$$(f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{\otimes s} \simeq (f_* \mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{[s]}$$

automatically holds over $U(f, \Delta)$. Thus, Theorem E holds more generally over $U(f, \Delta)$.

We now deduce Theorem D from Theorem E.

Proof of Theorem D. Using Lemma 2.14, we assume that Y is smooth and Δ has simple normal crossing support. Then, Theorem E implies

$$(f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta)))^{[s]} \otimes H^{\otimes \ell}$$

is generated by global sections for all $\ell \geq k$ on an open set $U \subseteq X$. Since $f_*\mathcal{O}_Y(k(K_{Y/X}+))$ is locally free over U, the map

$$(f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta)))^{[s]} \longrightarrow \mathrm{Sym}^{[s]}(f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta)))$$

is surjective over U; hence

$$\mathrm{Sym}^{[s]}(f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta))) \otimes H^{\otimes \ell}$$

is also generated by global sections for all $\ell \geq k$ on U.

Note that for any ample line bundle $\mathcal{L}$, there is an integer $b \geq 1$ such that $H^{\otimes -k} \otimes \mathcal{L}^{\otimes b}$ is globally generated. For such a b, the sheaf

$$\mathrm{Sym}^{[s]}(f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta))) \otimes \mathcal{L}^{\otimes b}$$

is also generated by global sections on U. Since b depends only on k and H and is independent of s, we can set $s = ab$. This implies weak positivity of $f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta))$ over U. $\square$

Remark 4.3. The proof of Theorem D shows that when Y is smooth and Δ has simple normal crossing support, the sheaf $f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta))$ is weakly positive over the open set in the statement of Theorem E.

5. Generic generation for pluricanonical sheaves

Proof of Theorem A. We now prove Theorem A, following the strategy in [Popa and Schnell 2014, Theorem 1.7] and [Dutta 2017, Theorem A]. The idea is to reduce to the case where Y is smooth and Δ has simple normal crossing support, and then maneuver into a situation to which Theorem C applies.

Proof of Theorem A. We start with some preliminary reductions.

Step 0: We may assume that the image of the counit morphism

$$f^* f_*\mathcal{O}_Y(P) \longrightarrow \mathcal{O}_Y(P) \tag{9}$$

for the adjunction $f^* \dashv f_*$ is nonzero.

Suppose the image of (9) is the zero sheaf. Then, the natural isomorphism

$$\mathrm{Hom}_{\mathcal{O}_Y}(f^* f_*\mathcal{O}_Y(P), \mathcal{O}_Y(P)) \simeq \mathrm{Hom}_{\mathcal{O}_X}(f_*\mathcal{O}_Y(P), f_*\mathcal{O}_Y(P))$$

from the adjunction $f^* \dashv f_*$ implies that the identity morphism $\mathrm{id}: f_*\mathcal{O}_Y(P) \to f_*\mathcal{O}_Y(P)$ is the zero morphism. This implies $f_*\mathcal{O}_Y(P) = 0$; hence the conclusion of Theorem A trivially holds.

Step 1 (cf. [Popa and Schnell 2014, Theorem 1.7, Step 1]): We can reduce to the case where

(a) Y is smooth;

(b) Δ has simple normal crossing support and coefficients in $(0, 1]$; and

(c) the image of (9) is of the form $\mathcal{O}_Y(P-E)$ for a divisor E such that $\Delta+E$ has simple normal crossing support.

A priori, the image of the counit (9) is of the form $\mathfrak{b}\cdot\mathcal{O}_Y(P)$, where $\mathfrak{b}\subseteq\mathcal{O}_Y$ is the *relative base ideal* of $\mathcal{O}_Y(P)$. By Step 0, this ideal is nonzero, and so consider a simultaneous log resolution $\mu:\widetilde{Y}\to Y$ of $\mathfrak{b}$ and (Y,Δ). The image of the counit morphism

$$\mu^* f^* f_*\mathcal{O}_Y(P)\longrightarrow\mu^*\mathcal{O}_Y(P)=\mathcal{O}_{Y'}(\mu^*P) \tag{10}$$

is the sheaf $\mathcal{O}_{Y'}(\mu^*P-E')$ [Lazarsfeld 2004b, Generalization 9.1.17].

We then apply Lemma 2.14 to μ. With the notation of the lemma we note that on $\widetilde{Y}$ the counit morphism (10) becomes the surjective morphism

$$(f\circ\mu)^*(f\circ\mu)_*\mathcal{O}_{\widetilde{Y}}(\widetilde{P})\twoheadrightarrow\mathcal{O}_{\widetilde{Y}}(\mu^*P-E')=\mathcal{O}_{\widetilde{Y}}(\widetilde{P}-(\widetilde{P}-\mu^*P)-E').$$

Setting $E:=(\widetilde{P}-\mu^*P)+E'$, we see that (c) holds for $\widetilde{P}$.

Finally, Theorem A for $(\widetilde{Y},\tilde{\Delta})$ and $\widetilde{P}$ implies that

$$(f\circ\mu)_*\mathcal{O}_{\widetilde{Y}}(\widetilde{P})\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\simeq f_*\mathcal{O}_Y(P)\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}$$

is generated by global sections on some open set U for $\ell\geq k(n^2+1)$. This concludes Step 1.

Henceforth, we work in the situation of Step 1. Before moving on to Step 2, we fix some notation. Let L denote the divisor class of $\mathcal{L}$. Let U be the subset of $U(f,\Delta+E)$ where

$$\varepsilon(\mathcal{L};x)>\frac{1}{n+\frac{1}{kn}}$$

for every $x\in U$, which is nonempty by Notation 2.1(a) and Theorem 2.20.

We set m to be the smallest positive integer such that $f_*\mathcal{O}_Y(P)\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes m}$ is globally generated on U. This integer m exists by [Küronya 2013, Proposition 2.7] since $U\cap\boldsymbol{B}_+(L)=\varnothing$ by Corollary 2.23.

Finally, we set $B:=\operatorname{Bs}|P-E+mf^*L|_{\mathrm{red}}\subsetneq Y$ and note that $B\cap f^{-1}(U)=\varnothing$.

Step 2: Reducing the problem to $k=1$ and a suitable pair.

From now on, fix a closed point $x\in U$.

The surjection

$$f^*f_*\mathcal{O}_Y(P)\otimes_{\mathcal{O}_Y}f^*\mathcal{L}^{\otimes m}\twoheadrightarrow\mathcal{O}_Y(P-E)\otimes_{\mathcal{O}_Y}f^*\mathcal{L}^{\otimes m}$$

implies that $\mathcal{O}_Y(P-E)\otimes_{\mathcal{O}_Y}f^*\mathcal{L}^{\otimes m}$ is globally generated on $f^{-1}(U)$. Choose a general member

$$\mathfrak{D}_x\in|P-E+mf^*L|.$$

By Bertini's theorem [Hartshorne 1977, Corollary III.10.9 and Remark III.10.9.3], we may assume that $\mathfrak{D}_x$ is smooth away from the base locus B of the linear system $|P-E+mf^*L|$. We may also assume that

$\mathfrak{D}_x$ intersects the fiber Y_x transversely, and the support of Δ and E transversely away from B [Lazarsfeld 2004a, Lemma 4.1.11]. We then have

$$k(K_Y+\Delta)\sim_{\mathbb{R}} K_Y+\Delta+\frac{k-1}{k}\mathfrak{D}_x+\frac{k-1}{k}E-\frac{k-1}{k}mf^*L;$$

hence for every integer ℓ,

$$k(K_Y+\Delta)+\ell f^*L\sim_{\mathbb{R}} K_Y+\Delta+\frac{k-1}{k}\mathfrak{D}_x+\frac{k-1}{k}E+\left(\ell-\frac{k-1}{k}m\right)f^*L.$$

We now adjust the coefficients of Δ and E so they do not share any components. Applying Lemma 2.18 to $c=(k-1)/k$, we see that there exists an effective divisor $E'\preceq E$ such that

$$\Delta' := \Delta+\frac{k-1}{k}E-E'$$

is effective with simple normal crossing support, with components intersecting Y_x transversely, and with coefficients in $(0,1]$. We can then write

$$P-E'+\ell f^*L\sim_{\mathbb{R}} K_Y+\Delta'+\frac{k-1}{k}\mathfrak{D}+\left(\ell-\frac{k-1}{k}m\right)f^*L. \tag{11}$$

Step 3: Applying Theorem C to obtain global generation.

By Lemma 2.17, we have $f_*\mathcal{O}_X(P-E')\simeq f_*\mathcal{O}_X(P)$. It therefore suffices to show that

$$f_*\mathcal{O}_Y(P-E')\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell} \tag{12}$$

is globally generated at x. We first modify $\mathfrak{D}_x$ to allow us to apply Theorem C. By Theorem 2.15, there exists a common log resolution $\mu:\widetilde{Y}\to Y$ for $\mathfrak{D}_x$ and (Y,Δ) that is an isomorphism away from $B\subsetneq Y$. We then write

$$\mu^*\mathfrak{D}_x=D+F,\quad \mu^*\Delta'=\mu_*^{-1}\Delta'+F_1,$$

where D is a smooth prime divisor intersecting the fiber over x transversely and F, F_1 are supported on $\mu^{-1}(B)$. Define

$$F' := \left\lfloor\frac{k-1}{k}F+F_1\right\rfloor,\quad \tilde{\Delta} := \mu^*\Delta'+\frac{k-1}{k}\mu^*\mathfrak{D}_x-F',\quad \widetilde{P} := \mu^*P+K_{\widetilde{Y}/Y}.$$

Note that $\tilde{\Delta}$ has simple normal crossing support and coefficients in $(0,1]$ by assumption on the log resolution and by definition of F'. Moreover, the support of $\tilde{\Delta}$ intersects the fiber over x transversely. Pulling back the decomposition in (11) and adding $K_{\widetilde{Y}/Y}-F'$ yields

$$\begin{aligned}\widetilde{P}-\mu^*E'-F'+\ell(f\circ\mu)^*L &\sim_{\mathbb{R}} K_{\widetilde{Y}}+\mu^*\Delta'+\frac{k-1}{k}\mu^*\mathfrak{D}_x-F'+\left(\ell-\frac{k-1}{k}m\right)(f\circ\mu)^*L\\ &\sim_{\mathbb{R}} K_{\widetilde{Y}}+\tilde{\Delta}+\left(\ell-\frac{k-1}{k}m\right)(f\circ\mu)^*L.\end{aligned} \tag{13}$$

We now claim that it suffices to show

$$(f \circ \mu)_* \mathcal{O}_{\widetilde{Y}}(\widetilde{P} - \mu^* E' - F') \otimes_{\mathcal{O}_X} \mathcal{L}^{\otimes \ell} \tag{14}$$

is globally generated at x. Consider the commutative diagram

$$\begin{tikzcd}
H^0\big(X, (f\circ\mu)_*\mathcal{O}_{\widetilde{Y}}(\widetilde{P}-\mu^*E'-F')\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\big) \arrow[r] \arrow[d, hook] & \big((f\circ\mu)_*\mathcal{O}_{\widetilde{Y}}(\widetilde{P}-\mu^*E'-F')\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\big)\otimes\kappa(x) \arrow[d, hook, "\wr"] \\
H^0\big(X, (f\circ\mu)_*\mathcal{O}_{\widetilde{Y}}(\widetilde{P}-\mu^*E')\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\big) \arrow[r] \arrow[d, "\wr"] & \big((f\circ\mu)_*\mathcal{O}_{\widetilde{Y}}(\widetilde{P}-\mu^*E')\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\big)\otimes\kappa(x) \arrow[d, "\wr"] \\
H^0\big(X, f_*(\mu_*\mathcal{O}_{\widetilde{Y}}(\widetilde{P})\otimes\mathcal{O}_Y(-E'))\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\big) \arrow[r] \arrow[d, "\wr"] & \big(f_*(\mu_*\mathcal{O}_{\widetilde{Y}}(\widetilde{P})\otimes\mathcal{O}_Y(-E'))\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\big)\otimes\kappa(x) \arrow[d, "\wr"] \\
H^0\big(X, f_*\mathcal{O}_Y(P-E')\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\big) \arrow[r] & f_*\mathcal{O}_Y(P-E')\otimes_{\mathcal{O}_X}\mathcal{L}^{\otimes\ell}\otimes\kappa(x)
\end{tikzcd}$$

where the top right isomorphism holds since F' is supported away from $(f \circ \mu)^{-1}(U)$; hence the stalks of the two sheaves are isomorphic, and the other isomorphisms follow from the projection formula and the fact that $K_{\widetilde{Y}/Y}$ is μ-exceptional. If the top horizontal arrow is surjective, then the commutativity of the diagram implies that the bottom horizontal arrow is also surjective, i.e., the sheaf in (12) is globally generated at x.

We now apply Theorem C to the decomposition (13) to see that the sheaf in (14) is globally generated at x for all

$$\ell - \frac{k-1}{k} m > \frac{n}{\varepsilon(\mathcal{L}; x)}.$$

By choice of U, we know that

$$\varepsilon(\mathcal{L}; x) > \frac{1}{n + \frac{1}{kn}}$$

at all $x \in U$, and so by applying the same argument used so far to all $x \in U$, we see $f_* \mathcal{O}_Y(P) \otimes_{\mathcal{O}_X} \mathcal{L}^{\otimes \ell}$ is globally generated on U for all

$$\ell > n\left(n + \frac{1}{kn}\right) + \frac{k-1}{k} m = n^2 + \frac{1}{k} + \frac{k-1}{k} m.$$

By the minimality of m, we know

$$m \le \left\lfloor n^2 + \frac{1}{k} + \frac{k-1}{k} m \right\rfloor + 1 \le n^2 + \frac{k-1}{k} m + 1.$$

The inequality between the leftmost and rightmost quantities is equivalent to $m \le k(n^2 + 1)$; that is, $f_* \mathcal{O}_Y(P) \otimes_{\mathcal{O}_X} \mathcal{L}^{\otimes \ell}$ is globally generated on U for $\ell \ge k(n^2 + 1)$. $\square$

Proof of Theorem B. Restricting to X smooth and $\mathcal{L}$ ample, we now show a slightly better bound. The strategy of Theorem B is the same as that for Theorem A: We first reduce to the case when Y is smooth and Δ has simple normal crossing support. Then, using twisted weak positivity this time, we maneuver to a situation in which we can apply Theorem C or [Dutta 2017, Proposition 1.2].

Proof of Theorem B. We begin with Steps 0 and 1 of the proof of Theorem A to reduce to a situation where Y is smooth and Δ has simple normal crossing support. Following Step 1, we also assume that there exists an effective divisor E with simple normal crossing support such that

$$f^*f_*\mathcal{O}_Y(P) \longrightarrow \mathcal{O}_Y(P-E) \tag{15}$$

is surjective.

Step 2: Reducing the problem to $k=1$ and a suitable pair.

Unless otherwise mentioned, throughout this proof we fix U to denote the intersection of $U(f,\Delta+E)$ with the open set over which $f_*\mathcal{O}_Y(P)$ is locally free.

In the diagram

$$\begin{array}{ccc} f^*\big((f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta)))^{\otimes b}\big) & \twoheadrightarrow & \mathcal{O}_Y(bk(K_{Y/X}+\Delta)-bE) \\ \downarrow & & \| \\ f^*\big((f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta)))^{[b]}\big) & \dashrightarrow & \mathcal{O}_Y(bk(K_{Y/X}+\Delta)-bE) \end{array}$$

the dashed map exists making the diagram commute. Indeed, the map exists over the locus X_1 where $f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta))$ is locally free. Since X_1 has a complement of codimension ≥ 2, and the bottom right sheaf is locally free, we can extend the dashed map to all of X (Corollary 2.5).

Now the top arrow is the surjective map obtained by taking the b-th tensor power of (15). Then the commutativity of the diagram implies that the bottom arrow is also surjective. By Theorem E we know that over U,

$$f_*\mathcal{O}_Y(k(K_{Y/X}+\Delta))^{[b]}\otimes\mathcal{L}^{\otimes b}$$

is generated by global sections for $b \gg 1$. Therefore so is $\mathcal{O}_Y(bk(K_{Y/X}+\Delta)-bE)\otimes f^*\mathcal{L}^{\otimes b}$ over $f^{-1}(U)$.

We now fix a point $x \in U$.

Letting L denote a Cartier divisor class of $\mathcal{L}$, we can apply Bertini's theorem to choose a divisor

$$D \in |bk(K_{Y/X}+\Delta)-bE+bf^*L|$$

such that on $f^{-1}(U)$, D is smooth, $D+\Delta+E$ has simple normal crossing support, D is not contained in the support of $\Delta+E$, and D intersects the fiber over x transversely. Then write

$$\frac{1}{b}D \sim_{\mathbb{R}} k(K_{Y/X}+\Delta)-E+f^*L.$$

Multiplying both sides by $\frac{k-1}{k}$, and then adding $K_{Y/X} + \Delta + \frac{k-1}{k}E$, we have

$$K_{Y/X} + \Delta + \frac{k-1}{kb}D + \frac{k-1}{k}E \sim_{\mathbb{R}} k(K_{Y/X} + \Delta) + \frac{k-1}{k}f^*L. \tag{16}$$

Now applying Lemma 2.18 for $c = \frac{k-1}{k}$, there exists an effective divisor $E' \preceq E$ such that

$$\Delta' := \Delta + \frac{k-1}{k}E - E'$$

has coefficients in $(0, 1]$. Subtracting $E' + \frac{k-1}{k}f^*L$ from both sides in (16), we can therefore write

$$K_{Y/X} + \frac{k-1}{kb}D + \Delta' - \frac{k-1}{k}f^*L \sim_{\mathbb{R}} k(K_{Y/X} + \Delta) - E'.$$

Let us now denote by H the line bundle $\omega_X \otimes \mathcal{L}^{\otimes n+1}$ and a divisor class in it at the same time. For a positive integer ℓ, we add $f^*K_X + (k-1)f^*H + (\ell - (k-1)(n+1))f^*L$ to both sides to obtain

$$K_Y + \frac{k-1}{kb}D + \Delta' + (k-1)f^*H + \left(\ell - \frac{k-1}{k} - (k-1)(n+1)\right)f^*L \sim_{\mathbb{R}} P - E' + \ell f^*L. \tag{17}$$

As noted earlier $E' \preceq E$ is an effective Cartier divisor and therefore

$$f_*\mathcal{O}_Y(P - E') \simeq f_*\mathcal{O}_Y(P)$$

by Lemma 2.17. Moreover since the right-hand side of (17) is a Cartier divisor, it is enough to tackle the generation of the left side.

<u>Step 3</u>: Applying Theorem C to obtain global generation.

First, we need to modify D to be able to apply Theorem C.

Let $\mu : Y' \to Y$ be a log resolution of $\frac{k-1}{kb}D + \Delta'$ as in Theorem 2.15. Such a modification is an isomorphism over $f^{-1}(U)$ by choice of D. Write

$$\mu^*D = \widetilde{D} + F, \quad \mu^*\Delta' = \tilde{\Delta}' + F_1,$$

where $\widetilde{D}$ is the strict transform of the components of D that lie above U and $\tilde{\Delta}'$ is the strict transform of Δ'. Note that both F and F_1 has support outside of $f^{-1}(U)$.

Define

$$F' := \left\lfloor \frac{k-1}{kb}F + F_1 \right\rfloor, \quad \tilde{\Delta} := \mu^*D + \mu^*\Delta' - F', \quad \widetilde{P} := \mu^*P + K_{Y'/Y}.$$

By definition $\tilde{\Delta}$ has coefficients in $(0, 1]$. Now pulling back (17) and adding $K_{Y'/Y} - F'$ we and rewrite (17) as

$$K_{Y'} + \tilde{\Delta} + (k-1)\mu^*f^*H + \left(\ell - \frac{k-1}{k} - (k-1)(n+1)\right)\mu^*f^*L \sim_{\mathbb{R}} \widetilde{P} - \mu^*E' + \ell\mu^*f^*L - F'.$$

This can be compared to (13). By the arguments following (13) we can say that it is enough to show global generation for the pushforward of the left side under $f \circ \mu$ to deduce desired global generation for $f_*\mathcal{O}_Y(P) \otimes \mathcal{L}^{\otimes \ell}$ for suitable ℓ.

To do so, we note once again that from Mori theory it follows that $H = \omega_X \otimes \mathcal{L}^{\otimes n+1}$ is semiample. Therefore $(k-1)\mu^* f^* H$ is also semiample. Applying Bertini's theorem one more time we can pick an effective fractional $\mathbb{Q}$-divisor $D' \sim_{\mathbb{Q}} (k-1)\mu^* f^* H$ with smooth support and its support intersects components of $\tilde{\Delta} + D'$ and the fiber over x transversely. We can now rewrite the linear equivalence as

$$K_{Y'} + \tilde{\Delta} + D' + \left(\ell - \frac{k-1}{k} - (k-1)(n+1)\right)\mu^* f^* L \sim_{\mathbb{R}} \widetilde{P} - \mu^* E' + \ell\mu^* f^* L - F'. \tag{18}$$

Note that $\tilde{\Delta} + D'$ on the left-hand side of (18) has simple normal crossing support with coefficients in $(0, 1]$ and $\operatorname{Supp}(\tilde{\Delta} + D')$ intersects the fiber over x transversely. Thus, we can apply Theorem C on the left-hand side to conclude that

$$f_*\mathcal{O}_Y(P) \otimes \mathcal{L}^{\otimes \ell}$$

is generated by global sections over U for all

$$\ell > \frac{n}{\varepsilon(L; x)} + k(n+1) - n - \frac{1}{k}.$$

After possibly shrinking U we assume that

$$\varepsilon(\mathcal{L}; x) > \frac{1}{n + \frac{1}{n(k+1)}}$$

for all points $x \in U$, and hence

$$\ell > n\left(n + \frac{1}{n(k+1)}\right) + k(n+1) - n - \frac{1}{k} = k(n+1) + n^2 - n - \frac{1}{k(k+1)}.$$

Therefore, $\ell \geq k(n+1) + n^2 - n$. This proves (i).

Step 4: The case of klt $\mathbb{Q}$-pairs.

When Δ is a klt $\mathbb{Q}$-pair, we apply [Dutta 2017, Proposition 1.2] on the left-hand side of (18). To do so, we first trace the construction of $\tilde{\Delta} + D'$ to note that its coefficients lie in $(0, 1)$. We then apply the proposition with

$$H = \frac{1}{k}\mu^* f^* L, \quad A = (\ell - k(n+1) + n)\mu^* f^* L$$

to obtain global generation on U for all $\ell > k(n+1) + \frac{1}{2}(n^2 - n)$. This proves (ii). □

We summarize below the locus of global generation for Theorems A and B:

Remark 5.1. When Y is smooth and the relative base locus of P is an effective divisor E such that $\Delta + E$ has simple normal crossing support, the open set U for Theorem A contains the largest open subset of $U(f, \Delta + E)$ such that $\varepsilon(\mathcal{L}; x) > \left(n + \frac{1}{kn}\right)^{-1}$ and for Theorem B(i), U contains the intersection of $U(f, \Delta + E)$, the locus where $f_*\mathcal{O}_Y(P)$ is locally free, and the open set where

$$\varepsilon(\mathcal{L}; x) > \left(n + \frac{1}{n(k+1)}\right)^{-1}.$$

Finally, for Theorem B(ii), U contains the intersection of $U(f, \Delta + E)$ and the locus where $f_*\mathcal{O}_Y(P)$ is locally free.

Remark 5.2. Using the better bounds in Remark 2.21 for low dimensions ($n = 2, 3$), one can show that the lower bounds $\ell \geq k(n^2 - n + 1)$ in Theorem A and $\ell \geq k(n+1) + n^2 - 2n$ in Theorem B suffice when X is smooth and $\mathcal{L}$ is ample. In particular, Conjecture 1.1 for generic global generation holds for $n = 2$. In the klt case, the conjectured lower bound in fact holds when $n \leq 4$ as was observed in [Dutta 2017].

If the conjectured lower bound for Seshadri constants in Remark 2.21 holds, then Theorem A would hold for the lower bound $\ell \geq k(n+1)$, thereby proving this generic version of Conjecture 1.1 in higher dimensions for big and nef line bundles.

An effective vanishing theorem. With the help of our effective twisted weak positivity, we improve the effective vanishing statement in [Dutta 2017, Theorem 3.1]:

Theorem 5.3. *Let $f : Y \to X$ be a smooth fibration of smooth projective varieties with* $\dim X = n$. *Let Δ be a $\mathbb{Q}$-divisor with simple normal crossing support with coefficients in* $[0, 1)$ *such that every stratum of (Y, Δ) is smooth and dominant over X, and let $\mathcal{L}$ be an ample line bundle on X. Assume also that for some fixed integer $k \geq 1$, $k(K_Y + \Delta)$ is Cartier and $\mathcal{O}_Y(k(K_Y + \Delta))$ is relatively base point free. Then, for every $i > 0$ and all $\ell \geq k(n+1) - n$, we have*

$$H^i\big(X, f_*\mathcal{O}_Y(k(K_Y + \Delta)) \otimes \mathcal{L}^{\otimes \ell}\big) = 0.$$

Moreover, if K_X is semiample, for every $i > 0$ and every ample line bundle $\mathcal{L}$ we have

$$H^i\big(X, f_*\mathcal{O}_Y(k(K_Y + \Delta)) \otimes \mathcal{L}\big) = 0.$$

Proof. The hypothesis on f and Δ ensures invariance of log plurigenera, as noted in Remark 4.2; hence $f_*\mathcal{O}_Y(k(K_{Y/X} + \Delta))$ is locally free. This means $U(f, \Delta) = X$. Furthermore, by the description of the open set in the proof of Theorem D, we have that there exists a positive integer b such that

$$(f_*\mathcal{O}_Y(k(K_{Y/X} + \Delta)))^{[b]} \otimes \mathcal{L}^{\otimes b}$$

is globally generated everywhere on X. Now since $\mathcal{O}_Y(k(K_Y + \Delta))$ is relatively base point free, we can choose a divisor $\frac{1}{b}D \sim_{\mathbb{R}} k(K_{Y/X} + \Delta) + f^*L$ satisfying the Bertini-type properties as in Step 2 of Theorem B. Define $H := K_X + (n+1)L$, which is semiample by Mori's cone theorem and the base point free theorem. As before, we then write

$$K_Y + \Delta + \frac{k-1}{kb}D + (k-1)f^*H + \left(\ell - \frac{k-1}{k} - (k-1)(n+1)\right)f^*L \sim_{\mathbb{R}} k(K_Y + \Delta) + \ell f^*L.$$

Since the divisor $\Delta + \frac{k-1}{kb} D$ is klt and $(k-1)H + \big(\ell - \frac{k-1}{k} - (k-1)(n+1)\big)L$ is ample for all $\ell \geq k(n+1) - n$, by Kollár's vanishing theorem [1995, Theorem 10.19] we obtain

$$H^i\big(X, f_*\mathcal{O}_Y(k(K_Y+\Delta)) \otimes \mathcal{L}^{\otimes \ell}\big) = 0$$

for all $\ell \geq k(n+1) - n$ and for all $i > 0$.

Moreover, when K_X is already semiample, we take $H = K_X$. In this case, the linear equivalence above looks as follows:

$$K_Y + \Delta + \frac{k-1}{kb} D + (k-1) f^*H + \left(\ell - \frac{k-1}{k}\right) f^*L \sim_{\mathbb{R}} k(K_Y+\Delta) + \ell f^*L.$$

Then, we obtain the desired vanishing for all $\ell \geq 1$ and $i > 0$. $\square$

Acknowledgments

We would like to thank our advisors Mihnea Popa and Mircea Mustaţă for their unconditional support, many enlightening conversations, and useful comments. We are also grateful to Karl Schwede for helpful discussions regarding Grothendieck duality theory. Finally, we would like to thank Lawrence Ein, Mihai Fulger, Emanuel Reinecke and the anonymous referee for some insightful comments.

References

[Birkar 2017] C. Birkar, "The augmented base locus of real divisors over arbitrary fields", *Math. Ann.* **368**:3-4 (2017), 905–921. MR Zbl

[Campana 2004] F. Campana, "Orbifolds, special varieties and classification theory", *Ann. Inst. Fourier* (*Grenoble*) **54**:3 (2004), 499–630. MR Zbl

[Cascini and Nakamaye 2014] P. Cascini and M. Nakamaye, "Seshadri constants on smooth threefolds", *Adv. Geom.* **14**:1 (2014), 59–79. MR Zbl

[de Cataldo 1998] M. A. A. de Cataldo, "Effective nonvanishing, effective global generation", *Ann. Inst. Fourier* (*Grenoble*) **48**:5 (1998), 1359–1378. MR Zbl

[Deng 2017] Y. Deng, "Applications of the Ohsawa–Takegoshi extension theorem to direct image problems", preprint, 2017. arXiv

[Dutta 2017] Y. Dutta, "On the effective freeness of the direct images of pluricanonical bundles", preprint, 2017. arXiv

[Ein and Lazarsfeld 1993] L. Ein and R. Lazarsfeld, "Seshadri constants on smooth surfaces", pp. 177–186 in *Journées de Géométrie Algébrique d'Orsay* (Orsay, 1992), Astérisque **218**, Soc. Math. France, Paris, 1993. MR Zbl

[Ein et al. 1995] L. Ein, O. Küchle, and R. Lazarsfeld, "Local positivity of ample line bundles", *J. Differential Geom.* **42**:2 (1995), 193–219. MR Zbl

[Ein et al. 2009] L. Ein, R. Lazarsfeld, M. Mustaţă, M. Nakamaye, and M. Popa, "Restricted volumes and base loci of linear series", *Amer. J. Math.* **131**:3 (2009), 607–651. MR Zbl

[Fujino 2017a] O. Fujino, *Foundations of the minimal model program*, MSJ Memoirs **35**, Math. Soc. Japan, Tokyo, 2017. MR Zbl

[Fujino 2017b] O. Fujino, "Notes on the weak positivity theorems", pp. 73–118 in *Algebraic varieties and automorphism groups* (Kyoto, 2014), edited by K. Masuda et al., Adv. Stud. Pure Math. **75**, Math. Soc. Japan, Tokyo, 2017. MR Zbl

[Fulton 1984] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik (3) **2**, Springer, 1984. MR Zbl

[Hacon et al. 2018] C. D. Hacon, J. M$^{\text{c}}$Kernan, and C. Xu, "Boundedness of moduli of varieties of general type", *J. Eur. Math. Soc.* **20**:4 (2018), 865–901. MR Zbl

[Hartshorne 1966] R. Hartshorne, *Residues and duality*, Lecture Notes in Math. **20**, Springer, 1966. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Math. **52**, Springer, 1977. MR Zbl

[Hartshorne 1994] R. Hartshorne, "Generalized divisors on Gorenstein schemes", *K-Theory* **8**:3 (1994), 287–339. MR Zbl

[Höring 2010] A. Höring, "Positivity of direct image sheaves: a geometric point of view", *Enseign. Math.* (2) **56**:1-2 (2010), 87–142. MR Zbl

[Illusie 2005] L. Illusie, "Grothendieck's existence theorem in formal geometry", Chapter 8, pp. 179–233 in *Fundamental algebraic geometry*, Math. Surveys Monogr. **123**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Iwai 2017] M. Iwai, "On the global generation of direct images of pluri-adjoint line bundles", preprint, 2017. To appear in *Math. Z.* arXiv

[Kollár 1995] J. Kollár, *Shafarevich maps and automorphic forms*, Princeton Univ. Press, 1995. MR Zbl

[Kollár 2013] J. Kollár, *Singularities of the minimal model program*, Cambridge Tracts in Math. **200**, Cambridge Univ. Press, 2013. MR Zbl

[Kollár and Mori 1998] J. Kollár and S. Mori, *Birational geometry of algebraic varieties*, Cambridge Tracts in Math. **134**, Cambridge Univ. Press, 1998. MR Zbl

[Küronya 2013] A. Küronya, "Positivity on subvarieties and vanishing of higher cohomology", *Ann. Inst. Fourier* (*Grenoble*) **63**:5 (2013), 1717–1737. MR Zbl

[Lazarsfeld 2004a] R. Lazarsfeld, *Positivity in algebraic geometry, I: Classical setting: line bundles and linear series*, Ergebnisse der Mathematik (3) **48**, Springer, 2004. MR Zbl

[Lazarsfeld 2004b] R. Lazarsfeld, *Positivity in algebraic geometry, II: Positivity for vector bundles, and multiplier ideals*, Ergebnisse der Mathematik (3) **49**, Springer, 2004. MR Zbl

[Popa and Schnell 2014] M. Popa and C. Schnell, "On direct images of pluricanonical bundles", *Algebra Number Theory* **8**:9 (2014), 2273–2295. MR Zbl

[Stacks 2018] The Stacks Project contributors, "The Stacks Project", online reference, 2018, https://stacks.math.columbia.edu/.

[Viehweg 1983] E. Viehweg, "Weak positivity and the additivity of the Kodaira dimension for certain fibre spaces", pp. 329–353 in *Algebraic varieties and analytic varieties* (Tokyo, 1981), edited by S. Iitaka, Adv. Stud. Pure Math. **1**, North-Holland, Amsterdam, 1983. MR Zbl

Communicated by János Kollár
Received 2018-02-06 Revised 2018-10-23 Accepted 2018-11-24

ydutta@math.northwestern.edu *Department of Mathematics, Northwestern University, Evanston, IL, United States*

takumim@umich.edu *Department of Mathematics, University of Michigan, Ann Arbor, MI, United States*

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.455

Lovász–Saks–Schrijver ideals and coordinate sections of determinantal varieties

Aldo Conca and Volkmar Welker

Motivated by questions in algebra and combinatorics we study two ideals associated to a simple graph G:

- the Lovász-Saks-Schrijver ideal defining the d-dimensional orthogonal representations of the graph complementary to G, and
- the determinantal ideal of the $(d+1)$-minors of a generic symmetric matrix with 0 in positions prescribed by the graph G.

In characteristic 0 these two ideals turn out to be closely related and algebraic properties such as being radical, prime or a complete intersection transfer from the Lovász–Saks–Schrijver ideal to the determinantal ideal. For Lovász–Saks–Schrijver ideals we link these properties to combinatorial properties of G and show that they always hold for d large enough. For specific classes of graphs, such a forests, we can give a complete picture and classify the radical, prime and complete intersection Lovász–Saks–Schrijver ideals.

1. Introduction

Let $\Bbbk$ be a field, $n \geq 1$ be an integer and set $[n] = \{1, \ldots, n\}$. For a simple graph $G = ([n], E)$ with vertex set $[n]$ and edge set E we study the following two classes of ideals associated to G.

- *Lovász–Saks–Schrijver ideals*: For an integer $d \geq 1$ we consider the polynomial ring

  $$S = \Bbbk\big[y_{i\ell} : i \in [n],\ \ell \in [d]\big].$$

  For every edge $e = \{i, j\} \in \binom{[n]}{2}$ we set

  $$f_e^{(d)} = \sum_{\ell=1}^{d} y_{i\ell} y_{j\ell}.$$

  The ideal

  $$L_G^{\Bbbk}(d) = (\, f_e^{(d)} : e \in E\,) \subseteq S$$

  is called the Lovász–Saks–Schrijver ideal, LSS-ideal for short, of G with respect to $\Bbbk$. The ideal $L_G^{\Bbbk}(d)$ defines the variety of orthogonal representations of the graph complementary to G. We refer the reader to [Lovász et al. 1989; Lovász 2009] for background on orthogonal representations and

The work on this paper was partly supported by a DAAD Vigoni project and by INdAM.
MSC2010: primary 05E40; secondary 05C62, 13P10.
Keywords: determinantal rings, complete intersections, ideals associated to graphs, Gröbner bases.

results on the geometry of the variety of orthogonal representations which provided intuition for some of our results.

- *Coordinate sections of generic (symmetric) determinantal ideals*: Consider the polynomial ring $S = \Bbbk[x_{ij} : 1 \le i \le j \le n]$ and let X be the generic $n \times n$ symmetric matrix, that is, the (i, j)-th entry of X is x_{ij} if $i \le j$ and x_{ji} if $i > j$. Let X_G^{sym} be the matrix obtained from X by replacing the entries in positions (i, j) and (j, i) for $\{i, j\} \in E$ with 0. For an integer d let $I_d^{\Bbbk}(X_G^{\mathrm{sym}}) \subseteq S$ be the ideal of d-minors of X_G^{sym}. The ideal $I_d^{\Bbbk}(X_G^{\mathrm{sym}})$ defines a coordinate hyperplane section of the generic symmetric determinantal variety. Similarly, we consider ideals defining coordinate hyperplane sections of the generic determinantal varieties and the generic skew-symmetric Pfaffian varieties.

We observe in Section 7 that the ideal $I_{d+1}^{\Bbbk}(X_G^{\mathrm{sym}})$ and the ideal $L_G^{\Bbbk}(d)$ are closely related. Indeed, if $\Bbbk$ has characteristic 0, classical results from invariant theory can be employed to show that $I_{d+1}^{\Bbbk}(X_G^{\mathrm{sym}})$ is radical (resp. is prime, has the expected height) provided $L_G^{\Bbbk}(d)$ is radical (resp. is prime, is a complete intersection). We also exhibit similar relations between variants of $L_G^{\Bbbk}(d)$ and ideals defining coordinate sections of determinantal and Pfaffian ideals.

These facts turn the focus on algebraic properties of the LSS-ideals $L_G^{\Bbbk}(d)$. In particular, we analyze the questions: When is $L_G^{\Bbbk}(d)$ a radical ideal? When is it a complete intersection? When is it a prime ideal? Other properties of ideals such as defining a normal ring or a UFD are interesting as well but will not be treated here. In Section 4 we prove the following:

Theorem 1.1. *Let $G = ([n], E)$ be a graph. Then*:

(1) *If $L_G^{\Bbbk}(d)$ is prime then $L_G^{\Bbbk}(d)$ is a complete intersection.*

(2) *If $L_G^{\Bbbk}(d)$ is a complete intersection then $L_G^{\Bbbk}(d+1)$ is prime.*

As an immediate consequence we have:

Corollary 1.2. *Let $G = ([n], E)$ be a graph. Then*:

(1) *If $L_G^{\Bbbk}(d)$ is prime (resp. complete intersection) then $L_G^{\Bbbk}(d+1)$ is prime (resp. complete intersection).*

(2) *If $L_G^{\Bbbk}(d)$ is prime (resp. complete intersection) then $L_{G'}^{\Bbbk}(d)$ is prime (resp. complete intersection) for every subgraph G' of G.*

In Section 5 we use these results to show that for d large enough $L_G^{\Bbbk}(d)$ is prime and complete intersection. To this end, for a graph $G = ([n], E)$ we define a graph theoretic invariant $\mathrm{pmd}(G) \in \mathbb{N}$, called the positive matching decomposition number of G. We prove in Lemma 5.4 that $\mathrm{pmd}(G) \le \min\{2n - 3, |E|\}$ and that $\mathrm{pmd}(G) \le \min\{n - 1, |E|\}$ if G is bipartite. We show the following:

Theorem 1.3. *Let $G = ([n], E)$ be a graph. Then for $d \ge \mathrm{pmd}(G)$ the ideal $L_G^{\Bbbk}(d)$ is a radical complete intersection. In particular, $L_G^{\Bbbk}(d)$ is prime if $d \ge \mathrm{pmd}(G) + 1$.*

The fact that $L_G^{\Bbbk}(d)$ is a complete intersection for large d also follows from [Sam and Weyman 2015, Theorems 3.5 and 3.8] or using the theory of strength from results in [Ananyan and Hochster 2016]. To have an explicit bound in Theorem 1.3 is crucial in order to use this result and the connection between $I_{d+1}^{\Bbbk}(X_G^{\mathrm{sym}})$ and $L_G^{\Bbbk}(d)$. Indeed, for deducing meaningful results, we need to single out cases where we can say something about $L_G^{\Bbbk}(d)$ for $d \le n-1$. The results described in the following paragraph can be seen as steps in this direction.

Already in Section 4 we give necessary conditions for $L_G^{\Bbbk}(d)$ to be prime in terms of subgraphs of G, see Proposition 4.4. In particular, we prove that if $L_G^{\Bbbk}(d)$ is prime then G does not contain a complete bipartite subgraph $K_{a,b}$ with $a+b=d+1$ (i.e., $\bar{G}$ is $(n-d)$-connected). Similar results are obtained for complete intersections. In general these conditions are only necessary but in Section 6 we show that for small values of d they can be used to characterize the properties. For $d=1$ the characterization is obvious and in [Herzog et al. 2015] it is proved that $L_G^{\Bbbk}(2)$ is prime if and only if G is a matching. We obtain the following:

Theorem 1.4. *Let G be a graph. Then*:

(1) $L_G^{\Bbbk}(3)$ *is prime if and only if G does not contain* $K_{1,3}$ *and does not contain* $K_{2,2}$.

(2) $L_G^{\Bbbk}(2)$ *is a complete intersection if and only if G does not contain* $K_{1,3}$ *and does not contain* C_{2m} *for some* $m \ge 2$.

Here C_n denotes the cycle with n vertices. Finally for forests (i.e., graphs without cycles) we can give a complete picture.

Theorem 1.5. *Let G be a forest and denote by* $\Delta(G)$ *the maximal degree of a vertex in G. Then*:

(1) $L_G^{\Bbbk}(d)$ *is radical for all d.*

(2) $L_G^{\Bbbk}(d)$ *is a complete intersection if and only if* $d \ge \Delta(G)$.

(3) $L_G^{\Bbbk}(d)$ *is prime if and only if* $d \ge \Delta(G)+1$.

In Section 7 we demonstrate in characteristic 0 the above mentioned connection between $L_G^{\Bbbk}(d)$ and $I_{d+1}^{\Bbbk}(X_G^{\mathrm{sym}})$. Using the results from the preceding sections we deduce sufficient conditions for $I_{d+1}^{\Bbbk}(X_G^{\mathrm{sym}})$ to be radical, prime or of expected height. Similar results are obtained for coordinate hyperplane sections of the generic determinantal varieties and the generic skew-symmetric Pfaffian varieties. To our knowledge coordinate sections of determinantal varieties have been systematically studied only in the case of maximal minors, see for example the results in [Boocher 2012; Eisenbud 1988; Giusti and Merle 1982].

In Section 8 we use the results from Section 4 and Section 7 to formulate obstructions that prevent $L_G^{\Bbbk}(d)$ to be prime or a complete intersection. We also study the exact asymptotics in terms of the number of vertices of the least d such that $L_G^{\Bbbk}(d)$ is prime for G a complete and a complete bipartite graph. Finally, in Section 9 we pose open problems, formulate conjectures and exhibit a relation between hypergraph LSS-ideals and coordinate sections of bounded rank tensor varieties.

To complete the outline of the paper we mention that Section 2 sets up the graph theory and Gröbner theory. Section 3 recalls results from [Herzog et al. 2015] for the case $d = 2$ which in particular show that $L_G^{\Bbbk}(2)$ is always radical if char $\Bbbk \neq 2$. We then exhibit and discuss counterexamples which demonstrate that this is not the case for $d = 3$.

2. Notations and generalities

2A. ***Graph and hypergraph theory.*** In the following we introduce graph theory notation. We mostly follow the conventions from [Diestel 1997]. For us a graph $G = (V, E)$ is a simple graph on a finite vertex set V. In particular, E is a subset of the set of 2-element subsets $\binom{V}{2}$ of V. In most of the cases we assume that $V = [n] = \{1, \ldots, n\}$. A subgraph of a graph $G = (V, E)$ is a graph $G' = (V', E')$ such that $V' \subseteq V$ and $E' \subseteq E$. Given two graphs G and G' we say that G contains G' if G has a subgraph isomorphic to G'.

More generally, a hypergraph $H = (V, E)$ is a pair consisting of a finite set of vertices V and a set E of subsets of V. We are only interested in the situation when the sets in E are inclusionwise incomparable. Such a set of subsets is called a clutter.

For $m, n > 0$ we will use the following notation:

- K_n denotes the complete graph on n vertices, i.e., $K_n = ([n], \{\{i, j\} : 1 \leq i < j \leq n\})$,
- $K_{m,n}$ denotes the complete bipartite graph $([m] \cup [\tilde{n}], \{\{i, \tilde{j}\} : i \in [m], \tilde{j} \in [\tilde{n}]\})$ with bipartition $[m]$ and $[\tilde{n}] = \{\tilde{1}, \ldots, \tilde{n}\}$.
- B_n denotes the subgraph of $K_{n,n}$ obtained by removing the edges $\{i, \tilde{i}\}$ for $i = 1, \ldots, n$.
- For $n > 2$ we denote by C_n the cycle with n vertices, i.e., the subgraph of K_n with edges $\{1, 2\}$, $\{2, 3\}, \ldots, \{n-1, n\}, \{n, 1\}$.
- For $n > 1$ we denote by P_n the path with n vertices, i.e., the subgraph of K_n with edges $\{1, 2\}$, $\{2, 3\}, \ldots, \{n-1, n\}$.

We denote by $\bar{G} = (V, \bar{E})$ with $\bar{E} = \binom{V}{2} \setminus E$ the graph complementary to $G = (V, E)$. Let $W \subseteq V$. We write $G_W = (W, \{e \in E : e \subseteq W\})$ for the graph induced by G on vertex set W and $G - W$ for the subgraph induced by G on $V \setminus W$. In case $W = \{v\}$ for some $v \in V$ we simply write $G - v$ for $G - \{v\}$.

A graph $G = ([n], E)$ with $n \geq k + 1$ is called k-(vertex)connected if for every $W \subset V$ with $|W| = k - 1$ the graph $G - W$ is connected. The degree $\deg(v)$ of a vertex v of G is $|\{e \in E : v \in e\}|$ and $\Delta(G) = \max_{v \in V} \deg(v)$. Clearly, if $G = ([n], E)$ is k-connected then every vertex has degree at least k and $\Delta(\bar{G}) \leq n - k - 1$. We denote by $\omega(G)$ the clique number of G, i.e., the largest a such that G contains K_a. The following well known fact follows directly from the definitions.

Lemma 2.1. *Given a graph $G = ([n], E)$ and an integer $1 \leq d \leq n$ the following conditions are equivalent*:

(1) *$\bar{G}$ is $(n-d)$-connected.*

(2) *G does not contain $K_{a,b}$ with $a + b = d + 1$.*

2B. ***Basics on LSS-ideals and their generalization to hypergraphs.*** Let $H = ([n], E)$ be a hypergraph. For an integer $d \geq 1$ we consider the polynomial ring $S = \Bbbk[y_{i\ell} : i \in [n], \ell \in [d]]$. We define for $e \in E$

$$f_e^{(d)} = \sum_{\ell=1}^{d} \prod_{i \in e} y_{i\ell}.$$

If E is a clutter we call the ideal

$$L_H^{\Bbbk}(d) = (\, f_e^{(d)} : e \in E\,) \subseteq S$$

the LSS-ideal of the hypergraph H.

It will sometimes be useful to consider $L_H^{\Bbbk}(d)$ as a multigraded ideal. For that we equip S with the multigrading induced by $\deg(y_{i\ell}) = \mathfrak{e}_i$ for the i-th unit vector $\mathfrak{e}_i$ in $\mathbb{Z}^n$ and $(i, \ell) \in [n] \times [d]$. Clearly, for $e \in E$ the polynomial $f_e^{(d)}$ is multigraded of degree $\sum_{i \in e} \mathfrak{e}_i$. In particular, $L_H^{\Bbbk}(d)$ is $\mathbb{Z}^n$-multigraded. The following remark is an immediate consequence of the fact that if E is a clutter the two polynomials $f_e^{(d)}$ and $f_{e'}^{(d)}$ corresponding to distinct edges $e, e' \in E$ have incomparable multidegrees.

Remark 2.2. Let $H = ([n], E)$ be a hypergraph such that E is clutter. The generators $f_e^{(d)}$, $e \in E$, of $L_H^{\Bbbk}(d)$ form a minimal system of generators. In particular, $L_H^{\Bbbk}(d)$ is a complete intersection if and only if the polynomials $f_e^{(d)}$, $e \in E$, form a regular sequence.

The following alternative description of $L_G^{\Bbbk}(d)$ for a graph G turns out to be helpful in some places.

Remark 2.3. Let $G = ([n], E)$ be a graph. Consider the $n \times d$ matrix $Y = (y_{i\ell})$. Then $L_G^{\Bbbk}(d)$ is the ideal generated by the entries of the matrix YY^T in positions (i, j) with $\{i, j\} \in E$. Here Y^T denotes the transpose of Y.

Similarly, for a bipartite graph G, say a subgraph of $K_{m,n}$, one considers two sets of variables y_{ij} with $(i, j) \in [m] \times [d]$, z_{ij} with $(i, j) \in [d] \times [n]$ and the matrices $Y = (y_{ij})$ and $Z = (z_{ij})$. Then $L_G^{\Bbbk}(d)$ coincides (after renaming the variables in the obvious way) with the ideal generated by the entries of the product matrix YZ in positions (i, j) for $\{i, \tilde{j}\} \in E$.

2C. ***Gröbner bases.*** We use the following notations and facts from Gröbner bases theory, see for example [Bruns and Conca 2003]. Consider the polynomial ring $S = \Bbbk[x_1, \ldots, x_m]$. For a vector

$$\mathfrak{w} = (w_i : i \in [m]) \in \mathbb{R}^m$$

and a nonzero polynomial

$$f = \sum_{\alpha \in \mathbb{N}^{[m]}} a_\alpha x^\alpha$$

we set $m_{\mathfrak{w}}(f) = \max_{a_\alpha \neq 0}\{\alpha \cdot \mathfrak{w}\}$ and

$$\mathrm{in}_{\mathfrak{w}}(f) = \sum_{\alpha \cdot \mathfrak{w} = m_{\mathfrak{w}}(f)} a_\alpha x^\alpha.$$

The latter is called the initial form of f with respect to $\mathfrak{w}$. For an ideal I we denote by $\mathrm{in}_{\mathfrak{w}}(I)$ the ideal generated by $\mathrm{in}_{\mathfrak{w}}(f)$ with $f \in I \setminus \{0\}$. For a term order $\prec$ we denote similarly by $\mathrm{in}_{\prec}(f)$ the largest term of f and by $\mathrm{in}_\prec(I)$ the ideal generated by $\mathrm{in}_\prec(f)$ with $f \in I \setminus \{0\}$. The following will allow us to deduce properties of ideals from properties of their initial ideals.

Proposition 2.4. *Let I be a homogeneous ideal in the polynomial ring S and let τ be either a term order $\prec$ or a vector $\mathfrak{w} \in \mathbb{R}^m$. If $\mathrm{in}_\tau(I)$ is radical or a complete intersection or prime then so is I. Moreover, if $I = (f_1, \dots, f_r)$ and the elements $\mathrm{in}_\tau(f_1), \dots, \mathrm{in}_\tau(f_r)$ form a regular sequence then $f_1, \dots, f_r$ form a regular sequence and $\mathrm{in}_\tau(I) = (\mathrm{in}_\tau(f_1), \dots, \mathrm{in}_\tau(f_r))$.*

3. Known results and counterexamples for Lovász–Saks–Schrijver ideals

We recall results from [Herzog et al. 2015] and present examples showing that $L_G^{\Bbbk}(3)$ is not radical in general. First observe that, for obvious reasons, $L_G^{\Bbbk}(1)$ is radical, it is a complete intersection if and only if G is a matching and it is prime if and only if G has no edges. For $d = 2$ the following result from [Herzog et al. 2015] gives a complete answer for two of the three properties under discussion.

Theorem 3.1 [Herzog et al. 2015, Theorems 1.1, 1.2 and Corollary 5.3]. *Let $G = ([n], E)$ be a graph. If $\mathrm{char}\,\Bbbk \neq 2$ then the ideal $L_G^{\Bbbk}(2)$ is radical. If $\mathrm{char}\,\Bbbk = 2$ then $L_G^{\Bbbk}(2)$ is radical if and only if G is bipartite. Furthermore, $L_G^{\Bbbk}(2)$ is prime if and only if G is a matching.*

Indeed, in [Herzog et al. 2015] the characterization of the graphs G for which $L_G^{\Bbbk}(2)$ is prime is given under the assumption that $\mathrm{char}\,\Bbbk \neq 1, 2 \bmod (4)$ but it turns out that the statement holds as well in arbitrary characteristic (see Proposition 4.4 for the missing details).

The next examples show that $L_G^{\Bbbk}(3)$ need not be radical. In the examples we assume that $\Bbbk$ has characteristic 0 but we consider it very likely that the ideals are not radical over any field.

A quick criterion implying that an ideal J in a ring S is not radical is to identify an element $g \in S$ such that $J : g \neq J : g^2$. We call such a g a witness (of the fact that J is not radical). Of course the potential witnesses must be sought among the elements that are "closely related" to J. Alternatively, one can try to compute the radical of J or even its primary decomposition directly and read off whether J is radical. But these direct computations are extremely time consuming for LSS-ideals and did not terminate on our computers in the examples below. Nevertheless, in all examples we have quickly identified witnesses.

Example 3.2. We present three examples of graphs G such that $L_G^{\Bbbk}(3)$ is not radical over any field $\Bbbk$ of characteristic 0. The first example has 6 vertices and 9 edges and it is the smallest example we have found (both in terms of edges and vertices). The second example has 7 vertices and 10 edges and it is a complete intersection. This shows that $L_G^{\Bbbk}(3)$ can be a complete intersection without being radical. The third example is bipartite, a subgraph of $K_{5,4}$, with 12 edges, and is the smallest bipartite example we have found. In all cases, since the LSS-ideal $L_G^{\Bbbk}(3)$ has integral coefficients, we may assume that $\Bbbk = \mathbb{Q}$ and exhibit a witness g, i.e., a polynomial g such that $L_G^{\Bbbk}(3) : g \neq L_G^{\Bbbk}(3) : g^2$. The latter inequality can be checked with the help of CoCoA [Abbott et al. 2018] or Macaulay 2 [Grayson and Stillman 1993].

(1) Let G be the graph with 6 vertices and 9 edges depicted in Figure 1, left, i.e., with edges

$$E = \big\{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{2, 3\}, \{2, 4\}, \{2, 6\}, \{3, 5\}, \{4, 6\}\big\}.$$

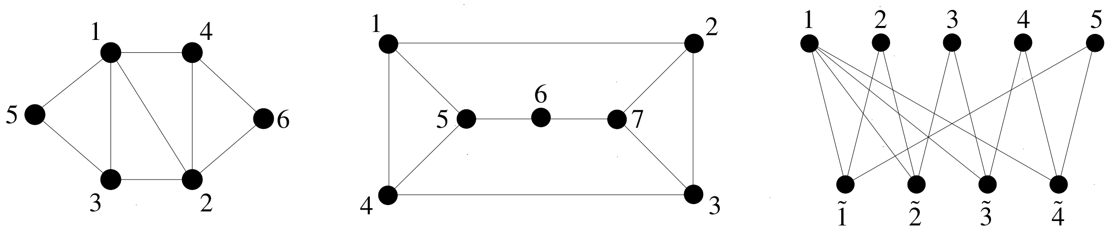

Figure 1. Graphs G with nonradical $L_G^{\Bbbk}(3)$.

Here the witness can be chosen as follows. Denote by $Y=(y_{ij})$ a generic 6×3 matrix. As discussed in Remark 2.3 the ideal $L_G^{\mathbb{Q}}(3)$ is generated by the entries of YY^T corresponding to the positions in E. Now g can be taken as the 3-minor of Y with row indices $1,5,6$.

(2) Let G be the graph with 7 vertices and 10 edges depicted in Figure 1, middle, i.e., with edges

$$E=\big\{\{1,2\},\{1,4\},\{1,5\},\{2,3\},\{2,7\},\{3,4\},\{3,7\},\{4,5\},\{5,6\},\{6,7\}\big\}.$$

Here the witness can be chosen as follows. Denote by $Y=(y_{ij})$ a generic 7×3 matrix. Again as discussed in Remark 2.3 the ideal $L_G^{\mathbb{Q}}(3)$ is generated by the entries of YY^T corresponding to the positions in E. Now g can be taken as the 3-minor of Y with row indices $1,2,4$. The fact that $L_G^{\mathbb{Q}}(3)$ is a complete intersection can be checked quickly with CoCoA [Abbott et al. 2018] or Macaulay 2 [Grayson and Stillman 1993].

(3) Let G be the subgraph of the complete bipartite graph $K_{5,4}$ depicted in Figure 1, right, i.e., with edges

$$E=\big\{\{1,\tilde{1}\},\{1,\tilde{2}\},\{1,\tilde{3}\},\{1,\tilde{4}\},\{2,\tilde{1}\},\{2,\tilde{2}\},\{3,\tilde{2}\},\{3,\tilde{3}\},\{4,\tilde{3}\},\{4,\tilde{4}\},\{5,\tilde{1}\},\{5,\tilde{4}\}\big\}.$$

Denote by $X=(x_{ij})$ a generic 5×3 matrix and by $Y=(y_{ij})$ a generic 3×4 matrix. As explained in Remark 2.3 the ideal $L_G^{\mathbb{Q}}(3)$ is generated by the entries of XY corresponding to the positions in E. Now the witness g can be taken to be the 3-minor of X corresponding to the column indices $1,2,4$.

4. Stabilization of algebraic properties of $L_G^{\Bbbk}(d)$

In this section we prove Theorem 1.1 and state some of its consequences. We recall first some facts on the symmetric algebra of a module stating the results in the way that suit our needs best.

Recall that, given a ring R and an R-module M presented as the cokernel of an R-linear map

$$f:R^m\to R^n$$

the symmetric algebra $\mathrm{Sym}_R(M)$ of M is (isomorphic to) the quotient of $\mathrm{Sym}_R(R^n)=R[x_1,\dots,x_n]$ by the ideal J generated by the entries of $A\,(x_1,\dots,x_n)^T$, where A is the $m\times n$ matrix representing f. Vice versa every quotient of $R[x_1,\dots,x_n]$ by an ideal J generated by homogeneous elements of degree 1 in the x_i's is the symmetric algebra of an R-module.

Part (1) of the following is a special case of [Avramov 1981, Proposition 3] and part (2) a special case of [Huneke 1981, Theorem 1.1]. Here and in the rest of the paper for a matrix A with entries in a ring R and a number t we denote by $I_t(A)$ the ideal of R generated by the t-minors of A.

Theorem 4.1. *Let R be a complete intersection. Then*:

(1) $\mathrm{Sym}_R(M)$ *is a complete intersection if and only if* height $I_t(A) \geq m-t+1$ *for all* $t = 1, \dots, m$.

(2) $\mathrm{Sym}_R(M)$ *is a domain and* $I_m(A) \neq 0$ *if and only if R is a domain, and* height $I_t(A) \geq m-t+2$ *for all* $t = 1, \dots, m$.

The equivalent conditions of (2) *imply those of* (1).

Remark 4.2. Let $G = ([n], E)$ be a graph. The ideal $L_G^{\Bbbk}(d) \subseteq S = \Bbbk[y_{i,j} : i \in [n],\ j \in [d]]$ is generated by elements that have degree at most one in each block of variables. Hence $L_G^{\Bbbk}(d)$ can be seen as an ideal defining a symmetric algebra in various ways.

For example, set $G_1 = G - n$, $U = \{i \in [n-1] : \{i, n\} \in E\}$, $u = |U|$, $S' = \Bbbk[y_{i,j} : i \in [n-1],\ j \in [d]]$ and $R = S'/L_{G_1}^{\Bbbk}(d)$. Then $S/L_G^{\Bbbk}(d)$ is the symmetric algebra of the cokernel of the R-linear map

$$R^u \to R^d$$

associated to the $u \times d$ matrix $A = (y_{ij})$ with $i \in U$ and $j = 1, \dots, d$.

Remark 4.3. In order to apply Theorem 4.1 to the case described in Remark 4.2 it is important to observe that for every G no minors of the matrix $(y_{ij})_{(i,j)\in[n]\times[d]}$ vanish modulo $L_G^{\Bbbk}(d)$. This is because $L_G^{\Bbbk}(d)$ is contained in the ideal J generated by the monomials $y_{ik}y_{jk}$ and the terms in the minors of (y_{ij}) do not belong to J for obvious reasons.

Proposition 4.4. *Let $G = ([n], E)$ be a graph. If $L_G^{\Bbbk}(d)$ is prime then G does not contain $K_{a,b}$ with $a+b > d$.*

Proof. Suppose by contradiction that $L_G^{\Bbbk}(d)$ is prime and G contains $K_{a,b}$ for some $a+b > d$. We may decrease either a or b or both and assume right away that $a+b = d+1$ with $a, b \geq 1$. In particular $a, b \leq d$ and $a+b \leq n$. We may assume that $K_{a,b}$ is a subgraph of G with edges $\{i, a+j\}$ for $i \in [a]$ and $j \in [b]$. Set $R = S/L_G^{\Bbbk}(d)$ and $Y = (y_{i\,\ell}) \in R^{a\times d}$ and $Z = (z_{\ell,i}) \in R^{d\times b}$ with $z_{\ell,i} = y_{i+a,\ell}$. Since $K_{a,b}$ is a subgraph of G we have $YZ = 0$ in R. By assumption R is a domain and $YZ = 0$ can be seen as a matrix identity over the field of fractions of R. Hence

$$\mathrm{rank}(Y) + \mathrm{rank}(Z) \leq d.$$

From $a+b = d+1$ it follows that $\mathrm{rank}(Y) < a$ or $\mathrm{rank}(Z) < b$. This implies that $I_a(Y) = 0$ or $I_b(Z) = 0$ as ideals of R. But by Remark 4.3 none of the minors of Y and Z are in $L_G^{\Bbbk}(d)$. This is a contradiction and hence $L_G^{\Bbbk}(d)$ is not prime. □

Lemma 4.5. *Let A be an $m \times n$ matrix with entries in a Noetherian ring R. Assume $m \leq n$. Let $S = R[x] = R[x_1, \dots, x_m]$ be a polynomial ring over R and let B be the $m \times (n+1)$ matrix with entries* *in S obtained by adding the column $(x_1, \dots, x_m)^T$ to A. Then we have* height $I_1(B) =$ height $I_1(A) + m$ *and*

$$\text{height } I_t(B) \geq \min\bigl\{\text{height } I_{t-1}(A),\ \text{height } I_t(A) + m - t + 1\bigr\}$$

for all $1 < t \leq m$.

Proof. Set $u = \min\bigl\{\text{height } I_{t-1}(A),\ \text{height } I_t(A) + m - t + 1\bigr\}$. Let P be a prime ideal of S containing $I_t(B)$. We have to prove that height $P \geq u$. If $P \supseteq I_{t-1}(A)$ then height $P \geq$ height $I_{t-1}(A) \geq u$. If $P \not\supseteq I_{t-1}(A)$ then we may assume that the $(t-1)$-minor F corresponding to the first $(t-1)$ rows and columns of A is not in P. Hence, height $P =$ height $P R_F[x]$ and $P R_F[x]$ contains $I_t(A) R_F[x]$ and $\bigl(x_j - F^{-1} G_j : j = t, \dots, m\bigr)$ with $G_j \in R[x_1, \dots, x_{t-1}]$. Since the elements $x_j - F^{-1} G_j$ are algebraically independent over R_F we have

$$\text{height } P R_F[x] \geq \text{height } I_t(A) R_F + (m - t + 1) \geq \text{height } I_t(A) + (m - t + 1). \qquad \square$$

Proof of Theorem 1.1. To prove (1) we argue by induction on n. The induction base $n \leq 2$ is obvious. Assume $n > 2$. We use the notation from Remark 4.2 and set

$$S = \Bbbk\bigl[y_{ij} : i \in [n],\ j \in [d]\bigr], \quad S' = \Bbbk\bigl[y_{i,j} : i \in [n-1],\ j \in [d]\bigr],$$
$$G_1 = G - n, \quad U = \{i \in [n-1] : \{i, n\} \in E\}, \quad u = |U|.$$

Note, that $S'/L^{\Bbbk}_{G_1}(d)$ is an algebra retract of $S/L^{\Bbbk}_G(d)$. Therefore $L^{\Bbbk}_{G_1}(d) = L^{\Bbbk}_G(d) \cap S'$ and so $L^{\Bbbk}_{G_1}(d)$ is prime. By induction, it follows that $L^{\Bbbk}_{G_1}(d)$ is a complete intersection. Since u is the degree of the vertex n in G we have that $K_{1,u} \subset G$. Since $L^{\Bbbk}_G(d)$ is prime Proposition 4.4 implies $1 + u < d + 1$, i.e., $u < d$. By virtue of Remark 4.3 we have that the minors of the matrix A are nonzero in $S'/L^{\Bbbk}_{G_1}(d)$. In particular, $I_u(A) \neq 0$ in $S'/L^{\Bbbk}_{G_1}(d)$ and hence (2) in Theorem 4.1 holds. Then (1) in Theorem 4.1 holds as well, i.e., $L^{\Bbbk}_G(d)$ is a complete intersection.

To prove (2) we again argue by induction on n. For $n \leq 2$ the assertion is obvious. Assume $n > 2$. We again use the notation $G_1 = G - n$, $U = \{i \in [n] : \{i, n\} \in E\}$, $u = |U|$. In addition we set $Y = (y_{ij})_{(i,j) \in U \times [d+1]}$, $S = \Bbbk[y_{ij} : i \in [n],\ j \in [d+1]]$, $S' = \Bbbk[y_{ij} : i \in [n-1],\ j \in [d+1]]$ and $R = S'/L^{\Bbbk}_{G_1}(d+1)$. By construction, $S/L^{\Bbbk}_G(d+1)$ is the symmetric algebra of the R-module presented as the cokernel of the map $R^u \to R^{d+1}$ associated to Y.

By assumption, $L^{\Bbbk}_G(d)$ is a complete intersection and hence $L^{\Bbbk}_{G_1}(d)$ is a complete intersection as well. It then follows by induction that $L^{\Bbbk}_{G_1}(d+1)$ is prime and hence R is a domain. Since the polynomials $f^{(d)}_{\{i,n\}}$ with $i \in U$ are a regular sequence contained in the ideal $(y_{nj} : 1 \leq j \leq d)$ we have $u \leq d$ and by Remark 4.3 $I_u(Y) \neq 0$ in R. Therefore, by Theorem 4.1(2) we have

$$L^{\Bbbk}_G(d+1) \text{ is prime} \quad \Longleftrightarrow \quad \text{height } I_t(Y) \geq u - t + 2 \text{ in } R \quad \text{for every } t = 1, \dots, u.$$

Equivalently, we have to prove that

$$\text{height}\bigl(I_t(Y) + L^{\Bbbk}_{G_1}(d+1)\bigr) \geq u - t + 2 + g \text{ in } S' \quad \text{for every } t = 1, \dots, u,$$

where $g = \text{height}\, L^{\Bbbk}_{G_1}(d+1) = |E| - u$.

Consider the weight vector $\mathfrak{w} \in \mathbb{R}^{n\times(d+1)}$ defined by $\mathfrak{w}_{ij} = 1$ and $\mathfrak{w}_{i\,d+1} = 0$ for all $j \in [d]$ and $i \in [n]$. By construction the initial forms of the standard generators of $\text{in}_{\mathfrak{w}}(L^{\Bbbk}_{G_1}(d+1))$ are the standard generators of $L^{\Bbbk}_{G_1}(d)$. Since the standard generators of $I_t(Y)$ coincide with their initial forms with respect to $\mathfrak{w}$ it follows that $\text{in}_{\mathfrak{w}}(I_t(Y)) \supseteq I_t(Y)$ (indeed equality holds but we do not need this fact).

Therefore, $\text{in}_{\mathfrak{w}}(I_t(Y) + L^{\Bbbk}_{G_1}(d+1)) \supseteq I_t(Y) + L^{\Bbbk}_{G_1}(d)$ and it is enough to prove that

$$\text{height}\big(I_t(Y) + L^{\Bbbk}_{G_1}(d)\big) \geq u - t + 2 + g \text{ in } S' \quad \text{for every } t = 1, \dots, u$$

or, equivalently,

$$\text{height}\, I_t(Y) \geq u - t + 2 \text{ in } R' \quad \text{for every } t = 1, \dots, u,$$

where $R' = S'/L^{\Bbbk}_{G_1}(d)$.

The variables $y_{1\,d+1}, \dots, y_{n-1\,d+1}$ do not appear in the generators of $L^{\Bbbk}_{G_1}(d)$. Hence

$$R' = R''[y_{1\,d+1}, \dots, y_{n-1\,d+1}] \quad \text{with } R'' = \Bbbk\big[y_{ij} : (i,j) \in [n-1] \times [d]\big]/L^{\Bbbk}_{G_1}(d).$$

Let Y' be the matrix Y with the $(d+1)$-st column removed. Then $S/L^{\Bbbk}_{G}(d)$ can be regarded as the symmetric algebra of the R''-module presented as the cokernel of the map

$$(R'')^u \xrightarrow{Y'} (R'')^d. \tag{1}$$

By assumption $S/L^{\Bbbk}_{G}(d)$ is a complete intersection. Hence by Theorem 4.1(1) we know

$$\text{height}\, I_t(Y') \geq u - t + 1 \text{ in } R'' \quad \text{for every } t = 1, \dots, u$$

Since Y is obtained from Y' by adding a column of variables over R'' by Lemma 4.5 we have

$$\text{height}\, I_t(Y) \geq \min\big\{\text{height}\, I_{t-1}(Y'),\ \text{height}\, I_t(Y') + u - t + 1\big\} \geq u - t + 2$$

in R' and for all $t = 1, \dots, u$. □

Proof of Corollary 1.2. Assertion (1) in Corollary 1.2 is a formal consequence of Theorem 1.1. Assertion (2) is obvious for complete intersections. Finally assume that $L^{\Bbbk}_{G}(d)$ is prime. Then by Theorem 1.1 $L^{\Bbbk}_{G}(d)$ is a complete intersection. The statement now follows from a general fact: if a regular sequence generates a prime ideal in a standard graded algebra or in a local ring then so does every subset of the sequence. □

5. Positive matching decompositions

In this section we introduce positive matching decompositions and prove Theorem 1.3.

Definition 5.1. Given a hypergraph $H = (V, E)$ a positive matching of H is a subset $M \subset E$ of pairwise disjoint sets (i.e., a matching) such that there exists a weight function $w : V \to \mathbb{R}$ satisfying:

$$\sum_{i\in A} w(i) > 0 \quad \text{if } A \in M, \qquad \sum_{i\in A} w(i) < 0 \quad \text{if } A \in E \setminus M. \tag{2}$$

The next lemma summarizes some elementary properties of positive matchings.

Lemma 5.2. *Let $H = (V, E)$ be a hypergraph such that E is a clutter, $M \subseteq E$ and $V_M = \bigcup_{A \in M} A$.*

(1) *M is a positive matching for H if and only if M is a positive matching for the induced hypergraph $(V_M, \{A \in E : A \subseteq V_M\})$.*

(2) *Assume M is a positive matching on H and $A \in E$ is such that $M_1 = M \cup \{A\}$ is a matching. Assume also there is a vertex $a \in A$ such that*

$$\big\{B \in E : B \subset V_{M_1} \text{ and } a \in B\big\} = \{A\}.$$

Then $M \cup \{A\}$ is a positive matching of H.

(3) *If H is a bipartite graph with bipartition $V = V_1 \cup V_2$ then M is a positive matching if and only if M is a matching and directing the edges $e \in E$ from V_1 to V_2 if $e \in M$ and from V_2 to V_1 if $e \in E \setminus M$ yields an acyclic orientation.*

Proof. (1) Set $H_1 = (V_M, \{A \in E : A \subseteq V_M\})$. Clearly a weight function on V for which M is a positive matching restricts to V_M making M a positive matching of H_1. Conversely, assume we are given a weight function w on V_M that makes M a positive matching. Then we extend w to V by assigning to the vertices in $V \setminus V_M$ a weight sufficiently negative to induce a negative weight on the elements of E which contain at least one element from $V \setminus V_M$. For example, one can set $w(i) = -|V| \max\{w(j) : j \in V_M\}$ for every $i \in V \setminus V_M$. Such an extension makes M a positive matching for H.

(2) Let w be a weight that makes M a positive matching of H. In view of (1), it is enough to prove that there is a weight v defined on V_{M_1} making M_1 a positive matching for the restriction of H to V_{M_1}. We set $v(i) = w(i)$ if $i \in V_{M_1}$ and $i \neq a$ and we give $v(a)$ a high enough value to have $v(A) > 0$, i.e., $v(a) > -\sum_{i \in A\ i \neq a} w(i)$. Since there are no elements in E other than A that are contained in V_{M_1} and contain a the resulting weight v has the desired properties.

(3) We change the coordinates $w(i)$ to $-w(i)$ for $i \in V_2$ in the inequalities defining a positive matchings. As a simple reformulation of (2) we get that in these coordinates a matching M is positive if and only if there is a weight function such that for $\{i, j\} \in E,\ i \in V_1,\ j \in V_2$ we have

$$w(i) > w(j) \quad \text{if } \{i, j\} \in M, \qquad w(i) < w(j) \quad \text{if } \{i.j\} \in E \setminus M. \tag{3}$$

This is equivalent to the existence of a region in the arrangement of hyperplanes $w(i) = w(j)$ for $\{i, j\} \in E$ in $\mathbb{R}^V$ satisfying (3). But it is well known that the regions in this arrangement are in one to one correspondence with the acyclic orientations of G (see [Greene and Zaslavsky 1983, Lemma 7.1]). □

Now we are in position to introduce the key concept of this section.

Definition 5.3. Let $H = (V, E)$ be a hypergraph for which E is a clutter. A positive matching decomposition (or pm-decomposition) of G is a partition $E = \bigcup_{i=1}^{p} E_i$ into pairwise disjoint subsets such that E_i is a positive matching on $\big(V, E \setminus \bigcup_{j=1}^{i-1} E_j\big)$ for $i = 1, \dots, p$. The E_i are called the parts of the pm-decomposition. The smallest p for which G admits a pm-decomposition with p parts will be denoted by $\operatorname{pmd}(H)$.

Note that one has $\operatorname{pmd}(H) \le |E|$ because of the obvious pm-decomposition $\bigcup_{A\in E}\{A\}$. On the other hand $\operatorname{pmd}(G)$ is smaller than $|E|$ for most clutters. For graphs we have:

Lemma 5.4. *Let $G = ([n], E)$ be a graph. Then*:

(1) $\operatorname{pmd}(G) \le \min(2n-3, |E|)$.

(2) *If G is bipartite then* $\operatorname{pmd}(G) \le \min(n-1, |E|)$.

(3) $\operatorname{pmd}(G) \ge \Delta(G)$ *with equality if G is a forest.*

Proof. (1) Since we have already argued that $\operatorname{pmd}(G) \le |E|$ to prove the first statement we have to show that $\operatorname{pmd}(G) \le 2n-3$. To this end we may assume that G is the complete graph K_n because any pm-decomposition of K_n induces a pm-decomposition on its subgraphs. For $\ell = 1, \dots, 2n-3$ we set $E_\ell = \{\{i, j\} : i + j = \ell + 2\}$. Clearly one has $E = \bigcup_{\ell=1}^{2n-3} E_\ell$. So to prove that this is a pm-decomposition of K_n we have to prove that E_t is a positive matching on $G_t = \big([n], \bigcup_{\ell=t}^{2n-3} E_\ell\big)$. To this end we build E_t by inserting the edges one by one starting from those that involve vertices with smaller indices and repeatedly use Lemma 5.2(2) to prove that we actually get a positive matching. For example for $n = 8$, to prove that E_7 is a positive matching on G_7 we order the elements in E_7 as follows $\{4, 5\}, \{3, 6\}, \{2, 7\}, \{1, 9\}$. We assume we know already that $\{\{4, 5\}, \{3, 6\}\}$ is a positive matching and use Lemma 5.2(2) with $A = \{2, 7\}$ and $a = 2$ to prove that $\{\{4, 5\}, \{3, 6\}, \{2, 7\}\}$ is a positive matching as well.

(2) In this case it is enough to prove that $\operatorname{pmd}(K_{m,n}) \le n + m - 1$. For $\ell = 1, \dots, m + n - 1$ we set $E_\ell = \{\{i, \tilde{j}\} : i + j = \ell + 1\}$. Clearly one has $E = \bigcup_{\ell=1}^{m+n-1} E_\ell$. So to prove that this is a positive matching decomposition of $K_{m,n}$ we have to prove that E_ℓ is a positive matching on $E \setminus \bigcup_{k=1}^{\ell-1} E_k$ for $\ell = 1, \dots, m + n - 1$.

For $\ell = 1$ the assertion is obvious since E_1 contains a single edge. Now assume $\ell \ge 2$. By Lemma 5.2(3) it suffices to show that directing the edges in E_ℓ from $[m]$ to $[\tilde{n}]$ and the edges in $E \setminus \bigcup_{k=1}^{\ell} E_k$ in the other direction yields an acyclic orientation. Assume the resulting directed graph has a directed cycle. Let $\{i, \tilde{j}\} \in E_\ell$ be the edge from E_ℓ in this directed cycle for which j is minimal. The directed edge following the edge $i \to \tilde{j}$ in the directed cycle is of the form $\tilde{j} \to i'$ for some i' with $i' + j > \ell + 1$. This implies $i' > i$. Now let $i' \to \tilde{j}'$ be the edge following $\tilde{j} \to i'$ in the directed cycle. Then $\{i', \tilde{j}'\} \in E_\ell$ and $i' + j' = \ell + 1$. But this yields $j' < j$ which contradicts the minimality of j. Hence there is no directed cycle and E_ℓ is a positive matching on $E \setminus \bigcup_{k=1}^{\ell-1} E_k$.

(3) The inequality $\Delta(G) \le \operatorname{pmd}(G)$ is obvious. To prove that equality holds if G is a forest we argue by induction on the number of vertices. We may assume $\{n-1, n\} \in E$ and that n is a leaf of G. Hence $G_1 = G - n$ is a forest on $n-1$ vertices and by induction there exists a positive matching decomposition $E_1, \dots, E_p$ of G_1 with $p = \Delta(G_1)$. If $\Delta(G_1) < \Delta(G)$ we may simply set $E_{p+1} = \{\{n-1, n\}\}$ and note that, by virtue of Lemma 5.2(1), $E_1, \dots, E_{p+1}$ is a positive matching decomposition of G. If instead $\Delta(G_1) = \Delta(G)$ then there exists i such that $n - 1 \notin V_{E_i}$ and hence $E_i' = E_i \cup \{\{n-1, n\}\}$ is a matching. Using (1) and (2) of Lemma 5.2 one easily checks that the resulting decomposition $E_1, \dots, E_{i-1}, E_i', E_{i+1} \dots, E_p$ is a positive matching decomposition of G. $\square$

Next we connect positive matching decompositions to algebraic properties of LSS-ideals.

Lemma 5.5. *Let $H=(V,E)$ be a hypergraph such that E is a clutter, $d \geq p = \mathrm{pmd}(H)$ and $E = \bigcup_{\ell=1}^{p} E_\ell$ a positive matching decomposition. Then there exists a term order $<$ on S such that for every ℓ and every $A \in E_\ell$ we have*

$$\mathrm{in}_<(f_A^{(d)}) = \prod_{i \in A} y_{i\ell}. \tag{4}$$

Proof. To define $<$ we first define weight vectors $\mathfrak{w}_1, \dots, \mathfrak{w}_p \in \mathbb{R}^{V \times [d]}$. For that purpose we use the weight functions $w_\ell : V \to \mathbb{R}$, associated to each matching E_ℓ, $\ell = 1, \dots, p$. The weight vector $\mathfrak{w}_\ell$ is defined as follows:

- $\mathfrak{w}_\ell(y_{ik}) = 0$ if $k \neq \ell$, and
- $\mathfrak{w}_\ell(y_{i\ell}) = w_\ell(i)$.

By construction it follows that:

$$\mathrm{in}_{\mathfrak{w}_1}(f_A^{(d)}) = \begin{cases} \prod_{i \in A} y_{i1} & \text{if } A \in E_1, \\ \sum_{k=2}^{d} \prod_{i \in A} y_{ik} & \text{if } A \in E \setminus \{E_1\}. \end{cases} \tag{5}$$

We define the term order $<$ as follows: $y^\alpha < y^\beta$ if

(1) $|\alpha| < |\beta|$, or

(2) $|\alpha| = |\beta|$ and $\mathfrak{w}_\ell(y^\alpha) < \mathfrak{w}_\ell(y^\beta)$ for the smallest ℓ such that $\mathfrak{w}_\ell(y^\alpha) \neq \mathfrak{w}_\ell(y^\beta)$, or

(3) $|\alpha| = |\beta|$ and $\mathfrak{w}_\ell(y^\alpha) = \mathfrak{w}_\ell(y^\beta)$ for all ℓ and $y^\alpha <_0 y^\beta$ for an arbitrary but fixed term order $<_0$.

Now a simple induction shows that for all ℓ and for all $A \in E_\ell$ we have $\mathrm{in}_<(f_A^{(d)}) = \prod_{i \in A} y_{i\ell}$. ☐

Proof of Theorem 1.3. Let $d \geq p = \mathrm{pmd}(G)$ and $E = \bigcup_{\ell=1}^{p} E_\ell$ a pm-decomposition of G. By Lemma 5.5 there is a term order $<$ satisfying (4). Since each E_ℓ, $\ell = 1, \dots, p$, is a matching (4) implies that the initial monomials of the generators $f_A^{(d)}$ of $L_H^{\Bbbk}(d)$ are pairwise coprime and square free. Then the assertion follows from Proposition 2.4. The rest follows from Theorem 1.1. ☐

The following is an immediate consequence of Theorem 1.3 and Lemma 5.4:

Corollary 5.6. *Let $G = ([n], E)$ be a graph. Then $L_G^{\Bbbk}(d)$ is a radical complete intersection for $d \geq \min\{2n-3, |E|\}$ and prime for $d \geq \min\{2n-3, |E|\} + 1$. If G is bipartite then $L_G^{\Bbbk}(d)$ is a radical complete intersection for $d \geq \min\{n-1, |E|\}$ and prime for $d \geq \min\{n-1, |E|\} + 1$.*

6. Proofs of Theorem 1.4 and Theorem 1.5

Proof of Theorem 1.4. (1) By Proposition 4.4 if $L_G^{\Bbbk}(3)$ is prime then G does not contain $K_{1,3}$ and $K_{2,2}$. Now assume G does not contain $K_{1,3}$ and $K_{2,2}$. In addition, we may assume that $\Bbbk$ is algebraically closed. Since the tensor product over $\Bbbk$ of $\Bbbk$-algebras that are domains is a domain (see the Corollary to Proposition 1 in Bourbaki's Algebra [Bourbaki 1990, Chapter V, 17]) we may also assume that the graph is connected. A connected graph not containing $K_{1,3}$ and $K_{2,2}$ is either an isolated vertex or a path

P_n on $n > 1$ vertices or a cycle C_n with n vertices for $n = 3$ or $n \geq 5$. For an isolated vertex we have $L_G^{\Bbbk}(3) = (0)$. Hence we have to prove that $L_G^{\Bbbk}(3)$ is prime when $G = P_n$ for $n \geq 2$ or $G = C_n$ for $n = 3$ or $n \geq 5$. If $G = P_n$ then by Lemma 5.4 $\operatorname{pmd}(P_n) = \Delta(P_n) \leq 2$. Hence by Theorem 1.3 it follows that $L_{P_n}^{\Bbbk}(3)$ is prime.

Now let $G = C_n$ for $n = 3$ or $n \geq 5$ and set $m = n-1$. To prove that $L_{C_n}^{\Bbbk}(3)$ is prime we use the symmetric algebra perspective. Observe that $C_n - n$ is $P_m = P_{n-1}$. Set $J = L_{P_m}^{\Bbbk}(3)$, $S = \Bbbk[y_{ij} : i \in [m],\ j \in [3]]$ and $R = S/J$. We have already proved that J is a prime complete intersection of height $m-1$. We have to prove that the symmetric algebra of the cokernel of the R-linear map

$$R^2 \xrightarrow{Y} R^3 \quad \text{with } Y = \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{m1} & y_{m2} & y_{m3} \end{pmatrix}$$

is a domain. Since by Remark 4.3 $I_2(Y) \neq 0$ in R, taking into consideration Remark 4.2 we may apply Theorem 4.1. Therefore, it is enough to prove that

$$\text{height } I_1(Y) \geq 3 \quad \text{and} \quad \text{height } I_2(Y) \geq 2 \qquad \text{in } R.$$

Equivalently, it is enough to prove that in S

$$\text{height } I_1(Y) + J \geq m+2, \tag{6}$$

$$\text{height } I_2(Y) + J \geq m+1. \tag{7}$$

First we prove (6). Since height $I_1(Y) = 6$ in S then (6) is obvious for $m \leq 4$. For $m > 4$ observe that $I_1(Y) + J$ can be written as $I_1(Y) + H$, where H is the LSS-ideal of the path with vertices $2, 3, \ldots, m-1$. Because $I_1(Y)$ and H use disjoint set of variables, we have

$$\text{height } I_1(Y) + H = 6 + m - 3 = m + 3$$

and this proves (6). Now we note that the condition height $I_2(Y) \geq 1$ holds in R because R is a domain and $I_2(Y) \neq 0$. Hence we deduce from Theorem 4.1(1) that $L_{C_n}^{\Bbbk}(3)$ is a complete intersection for all $n \geq 3$.

It remains to prove (7). Since $I_2(Y)$ is a prime ideal of S of height 2 and $J \not\subset I_2(Y)$ the ideal $I_2(Y) + J$ has height at least 3. Hence the assertion (7) is obvious for $m = 2$, i.e., $n = 3$. Therefore, we may assume $m \geq 4$ (here we use $n \neq 4$). Let P be a prime ideal of S containing $I_2(Y) + J$. We have to prove that height $P \geq m+1$. If P contains $I_1(Y)$ then height $P \geq m+2$ by (6). So we may assume that P does not contain $I_1(Y)$, say $y_{11} \notin P$, and prove that height $PS_x \geq m+1$, where $x = y_{11}$. Since $I_2(Y)S_x = (y_{m2} - x^{-1}y_{m1}y_{12},\ y_{m3} - x^{-1}y_{m1}y_{13})$ we have

$$\begin{aligned} f_{m-1,m}^{(3)} &= y_{m-1,1}y_{m1} + y_{m-1,2}y_{m2} + y_{m-1,3}y_{m3} \\ &= y_{m-1,1}y_{m1} + y_{m-1,2}x^{-1}y_{m1}y_{12} + y_{m-1,3}x^{-1}y_{m1}y_{13} \\ &= x^{-1}y_{m1}f_{1,m-1}^{(3)} \mod I_2(Y)S_x. \end{aligned}$$

From $f^{(3)}_{m-1,m} \in J$ it follows that $y_{m1} f^{(3)}_{1,m-1} \in PS_x$. This implies that either $y_{m1} \in PS_x$ or $f^{(3)}_{1,m-1} \in PS_x$. In the first case PS_x contains y_{m1}, y_{m2}, y_{m3} and the LSS-ideal associated to the path with vertices $1, \dots, m-1$. Hence height $PS_x \geq 3 + m - 2 = m + 1$ as desired. Finally, if $f^{(3)}_{1,m-1} \in PS_x$ we have that PS_x contains the ideal $L^{\Bbbk}_{C_{m-1}}(3)$ associated to the cycle with vertices $1, \dots, m-1$ and we have already observed that this ideal is a complete intersection. Since $y_{m2} - x^{-1} y_{m1} y_{12}$, $y_{m3} - x^{-1} y_{m1} y_{13}$ are in PS_x as well it follows that height $PS_x \geq 2 + m - 1 = m + 1$.

(2) For the "only if" part we note that if $L^{\Bbbk}_G(2)$ is a complete intersection then $L^{\Bbbk}_G(3)$ is prime by Theorem 1.1 and hence G does cannot contain $K_{1,3}$ by Proposition 4.4. Suppose, by contradiction, that G contains C_{2m} for some $m \geq 2$. Hence $L^{\Bbbk}_{C_{2m}}(2)$ is a complete intersection of height $2m$. But the generators of $L^{\Bbbk}_{C_{2m}}(2)$ are (up to sign) among the 2-minors of the matrix

$$\begin{pmatrix} y_{11} & -y_{22} & y_{31} & \cdots & y_{2m-1,1} & -y_{2m,2} \\ y_{12} & y_{21} & y_{32} & \cdots & y_{2m-1,2} & y_{2m,1} \end{pmatrix}$$

and the ideal of 2-minors of such a matrix has height $2m - 1$, a contradiction.

For the converse implication, we may assume that $\Bbbk$ is algebraically closed. Since the tensor product over a perfect field $\Bbbk$ of reduced $\Bbbk$-algebras is reduced [Bourbaki 1990, Theorem 3, Chapter V, 15], we may also assume that G is connected. A connected graph satisfying the assumptions is either an isolated vertex, or a path or a cycle with an odd number of vertices. We have already observed that $\operatorname{pmd}(P_n) = \Delta(P_n) \leq 2$. By Theorem 1.3 it follows that $L^{\Bbbk}_{P_n}(2)$ is a complete intersection. It remains to prove that $L^{\Bbbk}_{C_{2m+1}}(2)$ is a complete intersection (of height $2m+1$). Note that $L^{\Bbbk}_{P_{2m+1}}(2) \subset L^{\Bbbk}_{C_{2m+1}}(2)$ and we know already that $L^{\Bbbk}_{P_{2m+1}}(2)$ is a complete intersection of height $2m$. Hence it remains to prove that $f^{(2)}_{1,2m+1}$ does not belong to any minimal prime of $L^{\Bbbk}_{P_{2n+1}}(2)$. The generators of $L^{\Bbbk}_{P_{2n+1}}(2)$ are (up to sign) the adjacent 2-minors of the matrix

$$\begin{pmatrix} y_{11} & -y_{22} & y_{31} & \cdots & y_{2m-1,1} & -y_{2m,2} & y_{2m+1,1} \\ y_{12} & y_{21} & y_{32} & \cdots & y_{2m-1,2} & y_{2m,1} & y_{2m+1,2} \end{pmatrix}.$$

The minimal primes of $L^{\Bbbk}_{P_{2n+1}}(2)$ are described in the proof of [Diaconis et al. 1998, Theorem 4.3], see also [Hoşten and Sullivant 2004; Herzog et al. 2010]. By the description given in [Diaconis et al. 1998] it is easy to see that all minimal primes of $L^{\Bbbk}_{P_{2n+1}}(2)$ with the exception of $I_2(Y)$ are contained in the ideal $Q = (y_{ij} : 2 < i < 2m+1,\ 1 \leq j \leq 2)$. Clearly, $f^{(2)}_{1,2m+1} \notin Q$. Finally, one has $f^{(2)}_{1,2m+1} \notin I_2(Y)$ since the monomial $y_{11} y_{2m+1,1}$ is divisible by no monomials in the support of the generators of $I_2(Y)$. □

We proceed with the proof of Theorem 1.5. We first formulate a more general statement. For this we need to introduce the concept of Cartwright–Sturmfels ideals. This concept was coined in [Conca et al. 2016] inspired by earlier work in [Conca et al. 2015; Cartwright and Sturmfels 2010]. It was further developed and applied to various classes of ideals in [Conca et al. 2017; 2018].

Consider for $d_1, \dots, d_n \geq 1$ the polynomial ring $S = \Bbbk[y_{ij} : i \in [n],\ j \in [d_i]]$ with multigrading $\deg y_{ij} = \mathfrak{e}_i \in \mathbb{Z}^n$. The group $G = \mathrm{GL}_{d_1}(\Bbbk) \times \cdots \times \mathrm{GL}_{d_n}(\Bbbk)$ acts naturally on S as the group of $\mathbb{Z}^n$-graded K-algebra automorphism. The Borel subgroup of G is $B = U_{d_1}(\Bbbk) \times \cdots \times U_{d_n}(\Bbbk)$, where $U_d(\Bbbk)$ denotes the subgroup of upper triangular matrices in $\mathrm{GL}_d(\Bbbk)$. A $\mathbb{Z}^n$-graded ideal J is Borel fixed if $g(J) = J$ for every $g \in B$. A $\mathbb{Z}^n$-graded ideal I of S is called a Cartwright–Sturmfels ideal if there exists a radical Borel fixed ideal J with the same multigraded Hilbert-series.

Theorem 6.1. *For $d_1, \dots, d_n \geq 1$ let $S = \Bbbk[y_{ij} : i \in [n],\ j \in [d_i]]$ be the polynomial ring with $\mathbb{Z}^n$ multigrading induced by* $\deg y_{ij} = \mathfrak{e}_i \in \mathbb{Z}^n$ *and $G = (V, E)$ be a forest. For each $e = \{i, j\} \in E$ let $f_e \in S$ be a $\mathbb{Z}^n$-graded polynomial of degree $\mathfrak{e}_i + \mathfrak{e}_j$. Then $I = (f_e : e \in E)$ is a Cartwright–Sturmfels ideal. In particular, I and all its initial ideals are radical.*

Proof. First, we observe that we may assume that the generators f_e of I form a regular sequence. To this end we introduce new variables and for each $e = \{i, j\} \in E$ we add to f_e a monomial m_e in the new variables of degree $e_i + e_j$ so that m_e and $m_{e'}$ are coprime if $e \neq e'$. The new polynomials $f_e + m_e$ with $e \in E$ form a regular sequence by Proposition 2.4 since their initial terms with respect to an appropriate term order are the pairwise coprime monomials m_e. The ideal I arises as a multigraded linear section of the ideal $(f_e + m_e : e \in E)$ by setting all new variables to 0. By [Conca et al. 2015, Theorem 1.16(5)] the family of Cartwright–Sturmfels ideals is closed under any multigraded linear section. Hence it is enough to prove the statement for the ideal $(f_e + m_e : e \in E)$. Equivalently we may assume right away that the generators f_e of I form a regular sequences.

The multigraded Hilbert series of a multigraded S-module M can by written as

$$\frac{K_M(z_1, \dots, z_n)}{\prod_{i=1}^n (1 - z_i)^{d_i}}.$$

The numerator $K_M(z_1, \dots, z_n)$ is a Laurent polynomial with integral coefficients called the K-polynomial of M. Since the f_e's form a regular sequence the K-polynomial of S/I is the polynomial

$$F(z) = F(z_1, \dots, z_n) = \prod_{\{i,j\} \in E} (1 - z_i z_j) \in \mathbb{Q}[z_1, \dots, z_n].$$

To prove that I is Cartwright–Sturmfels we have to prove that there is a Borel-fixed radical ideal J such that the K-polynomial of S/J is $F(z)$. Taking into consideration the duality between Cartwright–Sturmfels ideals and Cartwright–Sturmfels* ideals discussed in [Conca et al. 2016], it is enough to exhibit a monomial ideal J whose generators are in the polynomial ring $S' = \Bbbk[y_1, y_2, \dots, y_n]$ equipped with the (fine) $\mathbb{Z}^n$-grading $\deg y_i = \mathfrak{e}_i \in \mathbb{Z}^n$ such that the K-polynomial of J regarded as an S'-module is $F(1 - z_1, \dots, 1 - z_n)$, that is,

$$\prod_{\{i,j\} \in E} (z_i + z_j - z_i z_j).$$

We claim that, under the assumption that $([n], E)$ is a forest, the ideal

$$J = \prod_{\{i,j\} \in E} (y_i, y_j)$$

has the desired property. In other words, we have to prove that the tensor product

$$T_E = \bigotimes_{\{i,j\}\in E} T_{\{i,j\}}$$

of the truncated Koszul complexes

$$T_{\{i,j\}} : 0 \to S'(-\mathfrak{e}_i - \mathfrak{e}_j) \to S'(-\mathfrak{e}_i) \oplus S'(-\mathfrak{e}_j) \to 0$$

associated to y_i, y_j resolves the ideal J. Consider a leaf $\{a, b\}$ of E. Set $E' = E \setminus \{\{a, b\}\}$,

$$J' = \prod_{\{i,j\}\in E'} (y_i, y_j)$$

and $J'' = (y_a, y_b)$. Then by induction on the number of edges we have that $T_{E'}$ resolves the ideal J'. Then the homology of T_E is $\operatorname{Tor}_*^{S'}(J', J'')$. Since $\{a, b\}$ is a leaf, one of the two variables y_a, y_b does not appear at all in the generators of J'. Hence y_a, y_b forms a regular J'-sequence. Then $\operatorname{Tor}_{\geq 1}^{S'}(J', J'') = 0$ and hence T_E resolves $J' \otimes J''$. Finally, $J' \otimes J'' = J'J''$ since $\operatorname{Tor}_1^{S'}(J', S/J'') = 0$. This concludes the proof that the ideal I is a Cartwright–Sturmfels ideal. Every initial ideal of a Cartwright–Sturmfels ideal is a Cartwright–Sturmfels ideal as well because this property just depends on the Hilbert series. In particular, every initial ideal of a Cartwright–Sturmfels ideal is radical. □

Proof of Theorem 1.5. Setting $d_1 = \cdots = d_n = d$ and $f_e = f_e^{(d)}$ in Theorem 6.1 we have that $L_G^{\Bbbk}(d)$ is a Cartwright–Sturmfels ideal and hence radical. Assertions (2) and (3) follow from Lemma 5.4, Theorem 1.3, Proposition 4.4 and Theorem 1.1. □

7. Invariant theory, determinantal ideals of matrices with 0's and their relation to LSS-ideals

The first goal of this section is to recall some classical results from invariant theory, see for example the paper by De Concini and Procesi [1976]. In particular, we recall how determinantal/Pfaffian rings arise as invariant rings of group actions. We assume throughout this section that the base field $\Bbbk$ is of characteristic 0. After the recap of invariant theory we will establish the connection to LSS-ideals.

7A. ***Generic determinantal rings as rings of invariants* (gen).** We take an $m \times n$ matrix of variables $X_{m,n}^{\text{gen}} = (x_{ij})$ and consider the ideal $I_{d+1}^{\Bbbk}(X_{m,n}^{\text{gen}})$ of $S^{\text{gen}} = \Bbbk[x_{ij} : (i, j) \in [m] \times [n]]$ generated by the $(d+1)$-minors of $X_{m,n}^{\text{gen}}$. Consider two matrices of variables Y and Z of size $m \times d$ and $d \times n$ and the following action of $\mathfrak{G} = \mathrm{GL}_d(\Bbbk)$ on the polynomial ring $\Bbbk[Y, Z]$: The matrix $A \in \mathfrak{G}$ acts by the $\Bbbk$-algebra automorphism of $\Bbbk[Y, Z]$ that sends $Y \to YA$ and $Z \to A^{-1}Z$. The entries of the product matrix YZ are clearly invariant under this action. Hence the ring of invariants $\Bbbk[Y, Z]^{\mathfrak{G}}$ contains the subalgebra $\Bbbk[YZ]$ generated by the entries of the product YZ. The *first main theorem of invariant theory* for this action says that $\Bbbk[Y, Z]^{\mathfrak{G}} = \Bbbk[YZ]$. We have a surjective $\Bbbk$-algebra map

$$\phi : S^{\text{gen}} \to \Bbbk[Y, Z]^{\mathfrak{G}} = \Bbbk[YZ]$$

sending X to YZ. Clearly the product matrix YZ has rank d and hence we have $I_{d+1}^{\Bbbk}(X_{m,n}^{\text{gen}}) \subseteq \operatorname{Ker}\phi$. The *second main theorem of invariant theory* says that $I_{d+1}^{\Bbbk}(X_{m,n}^{\text{gen}}) = \operatorname{Ker}\phi$. Hence

$$S/I_{d+1}^{\Bbbk}(X_{m,n}^{\text{gen}}) \simeq \Bbbk[YZ]. \tag{8}$$

7B. ***Generic symmetric determinantal rings as rings of invariants (*sym*).*** We take an $n \times n$ symmetric matrix of variables $X_n^{\text{sym}} = (x_{ij})$ and consider the ideal $I_{d+1}^{\Bbbk}(X_n^{\text{sym}})$ in $S^{\text{sym}} = \Bbbk[x_{ij} : 1 \le i \le j \le n]$ generated by the $(d+1)$-minors of X_n^{sym}. Consider a matrix of variables Y of size $n \times d$ and the following action of the orthogonal group $\mathfrak{G} = \mathcal{O}_d(\Bbbk) = \{A \in \mathrm{GL}_d(\Bbbk) : A^{-1} = A^T\}$ on the polynomial ring $\Bbbk[Y]$: any $A \in \mathfrak{G}$ acts by the $\Bbbk$-algebra automorphism of $\Bbbk[Y]$ that sends Y to YA. The entries of the product matrix YY^T are invariant under this action and hence the ring of invariants contains the subalgebra $\Bbbk[YY^T]$ generated by the entries of YY^T. The *first main theorem of invariant theory* for this action asserts that $\Bbbk[Y]^G = \Bbbk[YY^T]$. Then we have a surjective presentation

$$\phi : S^{\text{sym}} \to \Bbbk[YY^T]$$

sending X to YY^T. Since the product matrix YY^T has rank d we have $I_{d+1}(X) \subseteq \operatorname{Ker}\phi$. The *second main theorem of invariant theory* then says that $I_{d+1}(X) = \operatorname{Ker}\phi$. Hence

$$S^{\text{sym}}/I_{d+1}^{\Bbbk}(X_n^{\text{sym}}) \simeq \Bbbk[YY^T]. \tag{9}$$

7C. ***Generic Pfaffian rings as rings of invariants (*skew*).*** We take an $n \times n$ skew-symmetric matrix of variables $X_n^{\text{skew}} = (x_{ij})$ and consider the ideal $\mathrm{Pf}_{2d+2}^{\Bbbk}(X)$ generated by the Pfaffians of size $(2d+2)$ of X_n^{skew} in $S^{\text{skew}} = \Bbbk[x_{ij} : 1 \le i < j \le n]$. Consider a matrix of variables Y of size $n \times 2d$ and let J be the $2d \times 2d$ block matrix with d blocks

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

on the diagonal and 0 in the other positions. The symplectic group $\mathfrak{G} = \mathrm{Sp}_{2d}(\Bbbk) = \{A \in \mathrm{GL}_{2t}(\Bbbk) : AJA^T = J\}$ acts on the polynomial ring $\Bbbk[Y]$ as follows: an $A \in \mathfrak{G}$ acts on $\Bbbk[Y]$ by the automorphism that sends $Y \to YA$. The entries of the product matrix YJY^T are invariant under this action and hence the ring of invariants contains the subalgebra $\Bbbk[YJY^T]$ generated by the entries of YJY^T. The *first main theorem of invariant theory* for the current action says that $\Bbbk[Y]^G = \Bbbk[YJY^T]$. Then we have a surjective presentation: $\phi : S^{\text{skew}} \to \Bbbk[YY^T]$ sending X to YJY^T. The product matrix YJY^T has rank $2d$ and hence we have $\mathrm{Pf}_{2d+2}^{\Bbbk}(X) \subseteq \operatorname{Ker}\phi$. The *second main theorem of invariant theory* for this action says that $\mathrm{Pf}_{2d+2}^{\Bbbk}(X) = \operatorname{Ker}\phi$. Hence

$$S^{\text{skew}}/\,\mathrm{Pf}_{2d+2}^{\Bbbk}(X_n^{\text{skew}}) \simeq \Bbbk[YJY^T]. \tag{10}$$

7D. ***Determinantal ideals of matrices with* 0*'s and their relation to LSS-ideals.*** The classical invariant theory point of view shows that the generic determinantal and Pfaffian ideals are prime as they are kernels of ring maps whose codomains are integral domains. Their height is also well known (see for example [Bruns and Vetter 1988]):

(gen) The height of the ideal $I_d^{\Bbbk}(X_{m,n}^{\mathrm{gen}})$ of d-minors of a $m\times n$ matrix of variables is $(n+1-d)(m+1-d)$.

(sym) The height of the ideal $I_d^{\Bbbk}(X_n^{\mathrm{sym}})$ of d-minors of a symmetric $n\times n$ matrix of variables is $\binom{n-d+2}{2}$.

(skew) The height of the ideal of Pfaffians $\mathrm{Pf}_{2d}^{\Bbbk}(X_n^{\mathrm{skew}})$ of size $2d$ (and degree d) of an $n\times n$ skew-symmetric matrix of variables is $\binom{n-2d+2}{2}$.

If one replaces the entries of the matrices with general linear forms in, say, u variables, then Bertini's theorem in combination with the fact that the generic determinantal/Pfaffian rings are Cohen–Macaulay implies that the determinantal/Pfaffian ideals remain prime as long as $u\ge 2+\text{height}$ and radical if $u\ge 1+\text{height}$.

But what about the case of special linear sections of determinantal ideals of matrices? And what about the case of coordinate sections? Are the corresponding ideals prime or radical? To describe coordinate sections we employ the following notation.

(gen) In the generic case we take a bipartite graph $G=([m]\cup[\tilde{n}],E)$ and denote by X_G^{gen} the matrix obtained from the $m\times n$ matrix of variables by replacing the entries in position (i,j) with 0 for all $\{i,\tilde{j}\}\in E$.

(sym) In the generic symmetric case we take a subgraph $G=([n],E)$ of K_n and denote by X_G^{sym} the matrix obtained from the $n\times n$ symmetric matrix of variables by replacing with 0 the entries in position (i,j) and (j,i) for all $\{i,j\}\in E$.

(skew) In the generic skew-symmetric case we take a subgraph $G=([n],E)$ of K_n and denote by X_G^{skew} the matrix obtained from the skew-symmetric matrix of variables by replacing with 0 the entries in position (i,j) and (j,i) for all $\{i,j\}\in E$.

In this terminology $I_d^{\Bbbk}(X_G^{\mathrm{gen}})$ is the ideal of d-minors of X_G^{gen} in S^{gen} and similarly in the symmetric case. We write $\mathrm{Pf}_{2d}^{\Bbbk}(X_G^{\mathrm{skew}})$ for the ideal of Pfaffians of size $2d$ of X_G^{skew} in S^{skew}. We ask for conditions on G that imply that $I_d^{\Bbbk}(X_G^{\mathrm{gen}})$, $I_d^{\Bbbk}(X_G^{\mathrm{sym}})$ or $\mathrm{Pf}_{2d}^{\Bbbk}(X_G^{\mathrm{skew}})$ is radical or prime or has the expected height.

Clearly, special linear sections of generic determinantal ideals can give nonprime and nonradical ideals. On the positive side, for maximal minors, we have the following results:

Remark 7.1. (1) Eisenbud [1988] proved that the ideal of maximal minors of a 1-generic $m\times n$ matrix of linear forms is prime and remains prime even after modding out any set of $\le m-2$ linear forms. In particular, the ideal of maximal minors of an $m\times n$ matrix of linear forms is prime provided the ideal generated by the entries of the matrix has at least $m(n-1)+2$ generators.

(2) Giusti and Merle [1982] studied the ideal of maximal minors of coordinate sections in the generic case. One of their main results, [Giusti and Merle 1982, Theorem 1.6.1] characterizes, in combinatorial terms, the subgraphs G of $K_{m,n}$, $m\le n$, such that the variety associated to $I_m^{\Bbbk}(X_G^{\mathrm{gen}})$ is irreducible, i.e., the radical of $I_m^{\Bbbk}(X_G^{\mathrm{gen}})$ is prime.

(3) Boocher [2012] proved that for any subgraph G of $K_{m,n}$, $m \le n$, the ideal $I_m^{\Bbbk}(X_G^{\text{gen}})$ is radical. Combining his result with the result of Giusti and Merle, one obtains a characterization of the graphs G such that $I_m^{\Bbbk}(X_G^{\text{gen}})$ is prime.

(4) Generalizing the result of Boocher, in [Conca et al. 2015; 2016] it is proved that ideals of maximal minors of a matrix of linear forms that is either row or column multigraded is radical.

In the generic case every nonzero minor of a matrix of type X_G^{gen} has no multiple factors because its multidegree is square-free. This explains, at least partially, why the determinantal ideals of X_G^{gen} have the tendency to be radical. However, the following example shows that they are not radical in general.

Example 7.2. Let X_G^{gen} be the 6×6 matrix associated to the graph from Example 3.2(3). That is, in the 6×6 generic matrix we set to 0 the entries in positions

$$(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (3,2), (3,3), (4,3), (4,4), (5,1), (5,4).$$

Then $I_4^{\Bbbk}(X_G^{\text{gen}})$ is not radical over a field of characteristic 0 and very likely over any field. Here the "witness" is $g = x_{1,5}$, i.e., $I_4^{\Bbbk}(X_G^{\text{gen}}) : g \neq I_4^{\Bbbk}(X_G^{\text{gen}}) : g^2$. Since G is contained in $K_{5,4}$ one can consider as well $I_4^{\Bbbk}(X_G^{\text{gen}})$ in the 5×5 matrix but that ideal turns out to be radical.

Similarly for symmetric matrices we have:

Example 7.3. Let X_G^{sym} be the 7×7 generic symmetric matrix associated to the graph from Example 3.2(1). That is, in the 7×7 generic symmetric matrix we set to 0 the entries in positions

$$\{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{2,3\}, \{2,4\}, \{2,6\}, \{3,5\}, \{4,6\}$$

as well as in the symmetric positions. Then $I_4^{\Bbbk}(X_G^{\text{sym}})$ is not radical over a field of characteristic 0. The witness here is $g = x_{1,6}$. Since G is contained in K_6 one can consider as well $I_4^{\Bbbk}(X_G^{\text{sym}})$ in the 6×6 matrix but that ideal turns out to be radical.

It turns out that Examples 3.2, 7.2 and 7.3 are indeed closely related as we now explain.

Let $G = ([m] \cup [\tilde{n}], E)$ be a subgraph of the complete bipartite graph $K_{m,n}$. In view of the isomorphism (8) we have that

$$S^{\text{gen}} / \big(I_{d+1}^{\Bbbk}(X_{m,n}^{\text{gen}}) + (x_{ij} : \{i, \tilde{j}\} \in E) \big) \simeq \Bbbk[YZ] / J_G(d),$$

where $Y = (y_{ij})$, $Z = (z_{ij})$ are respectively $m \times d$ and $d \times n$ matrices of variables and $J_G(d)$ is the ideal of $\Bbbk[YZ]$ generated by $(YZ)_{i,j}$ with $\{i, \tilde{j}\} \in E$. Furthermore

$$I_{d+1}^{\Bbbk}(X_{m,n}^{\text{gen}}) + (x_{ij} : \{i, \tilde{j}\} \in E) = I_{d+1}^{\Bbbk}(X_G^{\text{gen}}) + (x_{ij} : \{i, \tilde{j}\} \in E).$$

The LSS-ideal $L_G^{\Bbbk}(d) \subset \Bbbk[Y, Z]$ is indeed equal to $J_G(d)\Bbbk[Y, Z]$. Now it is a classical result in invariant theory (derived from the fact that linear groups are reductive in characteristic 0), that $\Bbbk[YZ]$ is a direct summand of $\Bbbk[Y, Z]$ in characteristic 0. This implies that

$$J_G(d) = L_G^{\Bbbk}(d) \cap \Bbbk[YZ].$$

The next proposition is an immediate consequence.

Proposition 7.4. *Let $\Bbbk$ be a field of characteristic* 0, $d \geq 1$ *and* $G = ([m] \cup [\tilde{n}], E)$ *be a subgraph of* $K_{m,n}$. *If* $L_G^{\Bbbk}(d)$ *is radical* (*resp. is a complete intersection*, *is prime*) *then* $I_{d+1}^{\Bbbk}(X_G^{\text{gen}})$ *is radical* (*resp. has maximal height*, *is prime*).

Now we start from a subgraph G of K_n. For $d+1 \leq n$ we may consider the coordinate section $I_{d+1}^{\Bbbk}(X_G^{\text{sym}})$ of $I_{d+1}^{\Bbbk}(X_n^{\text{sym}})$. Using the isomorphism (9) we obtain:

Proposition 7.5. *Let $\Bbbk$ be a field of characteristic* 0 *and* $G = ([n], E)$ *a graph. If* $L_G^{\Bbbk}(d)$ *is radical* (*resp. is a complete intersection*, *is prime*) *then* $I_{d+1}^{\Bbbk}(X_G^{\text{sym}})$ *is radical* (*resp. has maximal height*, *is prime*).

For $2d+2 \leq n$ we may consider the coordinate section $\mathrm{Pf}_{2d+2}^{\Bbbk}(X_G^{\text{skew}})$ of $\mathrm{Pf}_{2d+2}^{\Bbbk}(X_n^{\text{skew}})$. We may as well consider the associated twisted LSS-ideal $\hat{L}_G^{\Bbbk}(d)$ defined as follows. For every $i \in [n]$ we consider $2d$ indeterminates $y_{i\,1}, \ldots, y_{i\,2d}$. For $e = \{i, j\}$, $1 \leq i < j \leq n$ we set $\hat{f}_e^{(d)}$ to be the entry of the matrix YJY^T in row i and column j, i.e.,

$$\hat{f}_e^{(d)} = \sum_{k=1}^{d} \big(y_{i\,2k-1} y_{j\,2k} - y_{i\,2k} y_{j\,2k-1} \big).$$

Then we define the twisted LSS-ideal associated to G as follows:

$$\hat{L}_G^{\Bbbk}(d) = (\hat{f}_e^{(d)} : e \in E).$$

For $d = 1$ the twisted LSS-ideal coincides with the so-called binomial edge ideal defined and studied in [Herzog et al. 2010; Kiani and Saeedi Madani 2016; Matsuda and Murai 2013; Ohtani 2011].

Remark 7.6. Assume G is bipartite with bipartition $[n] = V_1 \cup V_2$ then the coordinate transformation (see [Bolognini et al. 2018, Corollary 6.2])

$$\begin{aligned} y_{i\,2k-1} &\mapsto y_{i\,2k-1} \quad \text{and} \quad & y_{i\,2k} &\mapsto y_{i\,2k} & &\text{for } i \in V_1, \\ y_{j\,2k} &\mapsto y_{j\,2k-1} \quad \text{and} \quad & y_{j\,2k-1} &\mapsto -y_{j\,2k} & &\text{for } j \in V_2, \end{aligned}$$

sends $\hat{L}_G^{\Bbbk}(d)$ to $L_G^{\Bbbk}(2d)$. In particular, for a bipartite graph G we have that $\hat{L}_G^{\Bbbk}(d)$ is radical (resp. prime) if and only if $L_G^{\Bbbk}(2d)$ is radical (resp. prime).

Using the isomorphism (10) we obtain:

Proposition 7.7. *Let $\Bbbk$ be a field of characteristic* 0 *and* $G = ([n], E)$ *a graph. If* $\hat{L}_G^{\Bbbk}(d)$ *is radical* (*resp. is a complete intersection*, *is prime*) *then* $\mathrm{Pf}_{2d+2}^{\Bbbk}(X_G^{\text{skew}})$ *is radical* (*resp. has maximal height*, *is prime*).

Now, in characteristic 0, the results that we have established for LSS-ideals can be turned into statements concerning coordinate sections of determinantal ideals.

Theorem 7.8. *Let $\Bbbk$ be a field of characteristic* 0.

(1) *For every subgraph G of $K_{m,n}$ the ideals* $I_2^{\Bbbk}(X_G^{\text{gen}})$ *and* $I_3^{\Bbbk}(X_G^{\text{gen}})$ *are radical.*

(2) *For every subgraph G of K_n the ideals* $I_2^{\Bbbk}(X_G^{\text{sym}})$ *and* $I_3^{\Bbbk}(X_G^{\text{sym}})$ *are radical.*

(3) *For every subgraph G of K_n the ideal* $\mathrm{Pf}_4^{\Bbbk}(X_G^{\mathrm{skew}})$ *is radical.*

Furthermore if G is a forest then:

(4) $I_d^{\Bbbk}(X_G^{\mathrm{gen}})$, $I_d^{\Bbbk}(X_G^{\mathrm{sym}})$ *and* $\mathrm{Pf}_{2d}^{\Bbbk}(X_G^{\mathrm{skew}})$ *are radical for all d.*

(5) $I_d^{\Bbbk}(X_G^{\mathrm{gen}})$ *and* $I_d^{\Bbbk}(X_G^{\mathrm{sym}})$ *have maximal height if* $d \geq \Delta(G)+1$.

(6) $I_d^{\Bbbk}(X_G^{\mathrm{gen}})$ *and* $I_d^{\Bbbk}(X_G^{\mathrm{sym}})$ *are prime if* $d \geq \Delta(G)+2$.

Proof. The statements for ideals of 2-minors follow from Propositions 7.4 and 7.5 using the fact that the edge ideal of a graph is radical. Indeed these results hold over a field of arbitrary characteristic as the corresponding ideals are "toric."

By [Herzog et al. 2015, Theorem 1.1] the ideal $L_G^{\Bbbk}(2)$ is radical for all graphs G. Using Propositions 7.4 and 7.5 this implies that $I_3^{\Bbbk}(X_G^{\mathrm{gen}})$ is radical for bipartite graphs G and $I_3^{\Bbbk}(X_G^{\mathrm{sym}})$ is radical for arbitrary graphs.

By [Herzog et al. 2010, Corollary 2.2] the ideal $\hat{L}_G^{\Bbbk}(1)$ is radical for all graphs G. Using Proposition 7.7 this implies that $\mathrm{Pf}_4^{\Bbbk}(X_G^{\mathrm{skew}})$ is radical for arbitrary graphs.

Finally, for a forest G the results in the case of minors are derived from Propositions 7.4, 7.5 and Theorem 1.5. In the Pfaffian case they follow using Theorem 6.1 and Proposition 7.7. □

The following corollary is an immediate consequence of assertion (3) in Theorem 7.8.

Corollary 7.9. *Let $G(2,n)$ be the coordinate ring of the Grassmannian of 2-dimensional subspaces in $\Bbbk^n$ in its standard Plücker coordinates. Then any subset of the Plücker coordinates generates a radical ideal in $G(2,n)$.*

A statement analogous to Corollary 7.9 for higher order Grassmannians is not true. Indeed, the point is that a set of m-minors of a generic matrix $m \times n$ does not generate a radical ideal in general (as it does for $m = 2$). For example, in the Grassmannian $G(3,6)$ modulo [123], [124], [135], [236] the class of [125][346] is a nonzero nilpotent.

Next we look into necessary conditions for $I_d^{\Bbbk}(X_G^{\mathrm{gen}})$ and $I_d^{\Bbbk}(X_G^{\mathrm{sym}})$ to be prime.

Lemma 7.10. *Let $G = ([n], G)$ be a graph.*

(1) *If $I_{d+1}^{\Bbbk}(X_G^{\mathrm{sym}})$ is prime then G does not contain $K_{a,b}$ for $a+b > d$ (i.e., $\bar{G}$ is $(n-d)$-connected).*

(2) *If $G = B_d$ with $d \geq 4$ and X is the generic $(d+2)\times(d+2)$ matrix then $I_{d+1}^{\Bbbk}(X_G^{\mathrm{gen}})$ is not prime.*

Proof. (1) Assume by contradiction that G contains $K_{a,b}$ for $a+b = d+1$. We may assume that the corresponding set of vertices are $[a]$ and $\{a+j : j \in [b]\}$. But then the submatrix of X_G^{sym} of the first $d+1$ rows and columns is block-diagonal with (at least) two blocks. Hence its determinant is nonzero, is reducible and has degree $d+1$. Since all the generators of $I_{d+1}^{\Bbbk}(X_G^{\mathrm{sym}})$ have degree $d+1$ it follows that $I_{d+1}^{\Bbbk}(X_G^{\mathrm{sym}})$ cannot be prime.

(2) Set $Y_d = X^{\text{gen}}_{B_d}$, i.e.,

$$Y_d = \begin{pmatrix} x_{11} & 0 & \cdots & 0 & x_{1,d+1} & x_{1,d+2} \\ 0 & x_{22} & \cdots & 0 & x_{2,d+1} & x_{2,d+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & x_{dd} & \vdots & \vdots \\ x_{d+1,1} & x_{d+1,2} & \cdots & \cdots & x_{d+1,d+1} & x_{d+1,d+2} \\ x_{d+2,1} & x_{d+2,2} & \cdots & \cdots & x_{d+2,d+1} & x_{d+2,d+2} \end{pmatrix}.$$

and $J = I_{d+1}(Y_d)$ and let S be the polynomial ring whose indeterminates are the nonzero entries of Y_d. First, we prove that for every $d \geq 1$ the ideal J has the expected height, i.e., height $J = 4$. For $d = 1, 2, 3$ the ideal J is indeed prime of height 4: for $d = 1$ this is obvious because Y_1 is the generic 3×3 matrix; for $d = 2$ and $d = 3$ it follows from the fact that the corresponding LSS-ideal is prime by virtue of Proposition 7.4. For $d > 3$ let P be a prime containing J. If P contains $(x_{11}, x_{22}, x_{33}, x_{44})$ then height $P \geq 4$. If P does not contain $(x_{11}, x_{22}, x_{33}, x_{44})$ we may assume $x_{11} \notin P$. Inverting x_{11} and using the standard localization trick for determinantal ideals one sees that $PS_{x_{11}}$ contains, up to a change of variables, $I_d(Y_{d-1})$. Hence height $P =$ height $PS_x \geq 4$. Now that we know that J has height 4 to prove that J is not prime for $d \geq 4$ it is enough to observe that $J \subset (x_{11}, x_{22}, x_{33}, x_{44})$. The latter is straightforward since mod $(x_{11}, x_{22}, x_{33}, x_{44})$ the submatrix of Y consisting of the first 4-rows has rank 2. □

8. Obstructions to algebraic properties and asymptotic behavior

In this section we return to the study of LSS-ideals $L^{\Bbbk}_G(d)$. Using results from Section 4 and results about $I_{d+1}(X^{\text{gen}}_{B_d})$ from Section 7 we derive necessary conditions for $L^{\Bbbk}_G(d)$ to be a complete intersection or prime. In addition, we discuss the exact asymptotic behavior of these properties for complete and complete bipartite graphs. To this end it is convenient to introduce the following notation. Given an algebraic property $\mathcal{P}$ of ideals and a graph G we set

$$\text{asym}_{\Bbbk}(\mathcal{P}, G) = \inf\big\{d : L^{\Bbbk}_G(d') \text{ has property } \mathcal{P} \text{ for all } d' \geq d\big\}.$$

Here we are interested in the properties $\mathcal{P} \in \{\text{radical, c.i., prime}\}$. By Theorem 1.1, Corollary 1.2 and Theorem 1.3 we know that for every graph G we have

$$\begin{aligned} \text{asym}_{\Bbbk}(\text{c.i.}, G) &= \min\{d : L^{\Bbbk}_G(d) \text{ is c.i.}\} &&\leq \text{pmd}(G), \\ \text{asym}_{\Bbbk}(\text{prime}, G) &= \min\{d : L^{\Bbbk}_G(d) \text{ is prime}\} &&\leq \text{pmd}(G) + 1, \\ \text{asym}_{\Bbbk}(\text{c.i.}, G) &\leq \text{asym}_{\Bbbk}(\text{prime}, G) &&\leq \text{asym}_{\Bbbk}(\text{c.i.}, G) + 1. \end{aligned}$$

Furthermore there are graphs such that $\text{asym}_{\Bbbk}(\text{prime}, G) = \text{asym}_{\Bbbk}(\text{c.i.}, G) + 1$ (e.g., odd cycles or forests) and others such that $\text{asym}_{\Bbbk}(\text{prime}, G) = \text{asym}_{\Bbbk}(\text{c.i.}, G)$ (e.g., even cycles). We have the following obstructions:

Proposition 8.1. *Let $G = ([n], E)$. Then:*

(1) *If $L^{\Bbbk}_G(d)$ is prime then G does not contain $K_{a,b}$ with $a+b=d+1$. Furthermore, if $d>3$ and* $\operatorname{char}\Bbbk=0$ *then G does not contain B_d.*

(2) *If $L^{\Bbbk}_G(d)$ is a complete intersection then G does not contain $K_{a,b}$ with $a+b=d+2$. Furthermore, if $d>2$ and* $\operatorname{char}\Bbbk=0$ *then G does not contain B_{d+1}.*

Proof. (1) The first assertion has been already proved in Proposition 4.4. For the second let $\operatorname{char}\Bbbk=0$ and $d>3$. By contradiction, assume G contains B_d. Then by Corollary 1.2 we know that $L^{\Bbbk}_{B_d}(d)$ is prime because $L^{\Bbbk}_G(d)$ is prime. Then Proposition 7.4 implies that $I_{d+1}(X^{\text{gen}}_{B_d})$ is prime for a generic matrix X of arbitrary size and this contradicts Lemma 7.10(2).

(2) Assertion (2) follows from (1) by using Theorem 1.1. □

Another obstruction is described in the following proposition.

Proposition 8.2. *Let $\Bbbk$ be a field of characteristic 0 and $n\in\mathbb{N}$. Let w_n be the largest positive integer such that $\binom{w_n}{2}\le n$. Then:*

(1) *$L^{\Bbbk}_{K_n}(d)$ is not prime for $d=n+\binom{w_n-2}{2}-1$.*

(2) *$L^{\Bbbk}_{K_n}(d)$ is not a complete intersection for $d=n+\binom{w_{n+1}-2}{2}-2$.*

Proof. (1) We set $h_n=\binom{w_n}{2}$ and $m_n=w_n+d-1$. The numbers are chosen so that, using the formulas for the height of determinantal ideals mentioned in Section 7, the ideal $I_{d+1}(X)$ of $(d+1)$-minors of a generic symmetric $m_n\times m_n$ matrix X has height h_n. Consider K_n as the graph $\big([m_n],\binom{[n]}{2}\big)$ on m_n vertices where the vertices $n+1,\dots,m_n$ do not appear in edges. Assume, by contradiction, that the ideal $L^{\Bbbk}_{K_n}(d)$ is prime. Then by Proposition 7.5 the ideal $I^{\Bbbk}_{d+1}(X^{\text{sym}}_{K_n})$ is prime and of height h_n. But one has

$$I^{\Bbbk}_{d+1}(X^{\text{sym}}_{K_n})\subset(x_{11},x_{22},\dots,x_{h_nh_n}) \tag{11}$$

which is a contradiction. To check (11) it is enough to prove that the rank of the matrix

$$X^{\text{sym}}_{K_n} \mod (x_{11},x_{22},\dots,x_{h_nh_n})$$

is at most d. That is, we have to check that the rank of an $(m_n\times m_n)$-matrix with block decomposition

$$\begin{pmatrix}0 & A\\ B & C\end{pmatrix},$$

where 0 is the zero matrix of size $(h_n\times n)$, is at most d. Since $d=m_n-n+m_n-h_n$ the latter is obvious.

(2) We set $h_n=\binom{w_{n+1}}{2}$ and $m_n=w_{n+1}+d-1$. As above, the numbers are chosen so that the ideal $I_{d+1}(X)$ of $(d+1)$-minors of a generic symmetric $m_n\times m_n$ matrix X has height h_n.

Assume, by contradiction, that $L^{\Bbbk}_{K_n}(d)$ is a complete intersection. From Proposition 7.5 it follows that $I^{\Bbbk}_{d+1}(X^{\text{sym}}_{K_n})$ has height h_n. But

$$I^{\Bbbk}_{d+1}(X^{\text{sym}}_{K_n})\subset(x_{11},x_{22},\dots,x_{h_n-1,h_n-1}) \tag{12}$$

which is a contradiction. As above (12) boils down to an obvious statement about the rank of a matrix with a zero submatrix of a certain size. □

Using this result we can now analyze the asymptotic behavior of $\operatorname{asym}_{\Bbbk}(\text{c.i.}, K_n)$ and $\operatorname{asym}_{\Bbbk}(\text{prime}, K_n)$.

Corollary 8.3. *Let $\Bbbk$ be a field of characteristic 0. Then*

$$\lim_{n\to\infty} \frac{\operatorname{asym}_{\Bbbk}(\text{c.i.}, K_n)}{n} = \lim_{n\to\infty} \frac{\operatorname{asym}_{\Bbbk}(\text{prime}, K_n)}{n} = 2. \tag{13}$$

Proof. By Corollary 5.6 we have $\operatorname{asym}_{\Bbbk}(\text{prime}, K_n) \le 2n-2$. By Proposition 8.2 we have

$$n + \binom{w_{n+1}-2}{2} - 1 \le \operatorname{asym}_{\Bbbk}(\text{c.i.}, K_n) \le \operatorname{asym}_{\Bbbk}(\text{prime}, K_n). \tag{14}$$

Hence the equalities in (13) follow from the fact that

$$\lim_{n\to\infty} \frac{\binom{w_{n+1}-2}{2}}{n} = 1. \qquad \square$$

Using Proposition 8.2 and Theorem 1.1 we obtain further obstructions.

Corollary 8.4. *Let G be a graph on n vertices and $\Bbbk$ a field of characteristic 0 and denote by $\alpha = \omega(G)$ the clique number of G. Then $L_G^{\Bbbk}(d)$ is not a complete intersection for $d \le \alpha + \binom{w_{\alpha+1}-2}{2} - 2$ and $L_G^{\Bbbk}(d)$ is not prime for $d \le \alpha + \binom{w_\alpha - 2}{2} - 1$, where w_α is defined as in Proposition 8.2.*

To give an actual feeling for the obstruction, we present one example:

Example 8.5. For $n = 15$ one has $w_n = 6$ and $L_{K_n}^{\Bbbk}(d)$ is not prime for $d = 15 + \binom{6-2}{2} - 1 = 20$. Therefore $L_G^{\Bbbk}(20)$ is not prime if G contains K_{15}, i.e., $\omega(G) \ge 15$.

For the case of complete bipartite graphs $K_{m,n}$ results of De Concini and Strickland [1981] or Musili and Seshadri [1983] on the varieties of complexes imply the following:

Theorem 8.6. *Let $G = K_{m,n}$. Then:*

(1) *$L_G^{\Bbbk}(d)$ is radical for every d.*

(2) *$L_G^{\Bbbk}(d)$ is a complete intersection if and only if $d \ge m+n-1$.*

(3) *$L_G^{\Bbbk}(d)$ is prime if and only if $d \ge m+n$.*

(4) $\operatorname{pmd}(G) = m+n-1$.

Proof. Taking into account Remark 2.3, the assertions (1), (2), and (3) follow form general results on the variety of complexes proved in [De Concini and Strickland 1981] and, with different techniques, in [Musili and Seshadri 1983]. It has been observed by Tchernev [2001] that the assertions in [De Concini and Strickland 1981] that refer to the so-called Hodge algebra structure of the variety of complexes are not correct. However, those assertions can be replaced with statements concerning Gröbner bases as it is done, for example, in a similar case in [Tchernev 2001]. Hence, (1), (2) and (3) can still be deduced from the arguments in [De Concini and Strickland 1981].

Alternative proofs of (2) and (3) are obtained by combining Proposition 8.1 and Corollary 5.6. Finally (4) is a consequence of Lemma 5.4 and Proposition 8.1. $\square$

9. Questions and open problems

We have seen that for the properties "complete intersection" and "prime" of $L_G^{\Bbbk}(d)$ there is persistence along the parameter d but Example 3.2 shows persistence does not occur in general for the property of being radical.

Question 9.1. What patterns can occur in the set $\{d : L_G^{\Bbbk}(d) \text{is radical}\}$ for a graph G?

Since the complete intersection property and prime property of $L_G^{\Bbbk}(d)$ for a given d are inherited by subgraphs, the properties can be characterized by means of forbidden subgraphs. We have explicitly identified the forbidden subgraphs in Theorem 1.4 for $d = 2$ and complete intersection and for $d = 3$ and prime. For $d = 3$ and complete intersection we do not even have a conjecture for the set of forbidden graphs. For $d = 4$ results from Lovász's book [2009, Chapter 9.4] suggest the following:

Question 9.2. Is it true that $L_G^{\Bbbk}(4)$ is prime if and only if G does not contain $K_{a,b}$ for $a+b=5$ and B_4?

Via the fact that primeness of $L_G^{\Bbbk}(d)$ implies primeness of $I_{d+1}^{\Bbbk}(X_G)$ a result by Giusti and Merle [1982, Theorem 1.6.1] guides the intuition behind the following question.

Question 9.3. Let G be a subgraph of $K_{m,n}$ graph and assume $m \le n$. Is it true that $L_G^{\Bbbk}(m-1)$ is prime if and only if G does not contain $K_{a,b}$ for $a+b \ge m$?

By Propositions 7.4 and 7.5 we know that if $L_G^{\Bbbk}(d)$ is radical or prime then so are $I_{d+1}^{\Bbbk}(X_G^{\text{gen}})$ and $I_{d+1}^{\Bbbk}(X_G^{\text{sym}})$ respectively. But our general bounds for $\text{asym}_{\Bbbk}(\text{radical}, G)$ and $\text{asym}_{\Bbbk}(\text{prime}, G)$ from Corollary 5.6 are not good enough to make use of this implication. Indeed, Corollary 8.3 shows that for the properties complete intersection and prime and n large enough there are graphs G for which Proposition 7.5 does not prove primality of an interesting ideal. On the other hand the use of Theorem 1.5 in Theorem 7.8 shows that one can take advantage of this connection in some cases. It would be interesting to exhibit classes different from forests where this is possible.

Question 9.4. Are there more interesting classes of graphs $G = ([n], E)$ for which $\text{asym}_{\Bbbk}(\text{c.i.}, G) \le n-1$ or $\text{asym}_{\Bbbk}(\text{prime}, G) \le n$?

Despite the fact that Proposition 8.2 destroys the hope for using Theorem 7.8 for general graphs, it would be interesting to replace the asymptotic result by an actual value. By Corollary 8.3 for n large we have $\text{asym}_{\Bbbk}(\text{prime}, K_n) = 2n - c_n$ for some numbers $c_n \in o(n)$ which using the notation of Proposition 8.2 satisfy $n - \binom{w_n-2}{2} + 1 \ge c_n \ge 2$. But we have no conjecture for an actual formula for c_n.

Question 9.5. What is the exact value of $\text{asym}_{\Bbbk}(\text{prime}, K_n)$?

For radicality we have a concrete conjecture in the case $G = K_n$.

Conjecture 9.6.

$$\text{asym}_{\Bbbk}(\text{radical}, K_n) = 1 \quad (\textit{at least if } \text{char}\,\Bbbk = 0).$$

In other words, given a matrix of variables X of size $n \times d$ we conjecture the ideal of the off-diagonal entries of XX^T is radical for all n, d.

It would also be interesting to study the ideal generated by all the entries of XX^T. We note that the symplectic version of this problem has been investigated by De Concini [1979].

Next we turn to open problems about hypergraph LSS-ideals. We know from Theorem 1.3 that for a hypergraph $H=(V,E)$ for which E is a clutter the ideal $L_H^{\Bbbk}(d)$ is a radical complete intersection for $d \geq \operatorname{pmd}(G)$. But we prove in Theorem 1.3 that $L_H^{\Bbbk}(d)$ is prime for $d \geq \operatorname{pmd}(H)+1$ only in the case that H is a graph.

Question 9.7. Is it true that for a hypergraph $H=(V,E)$, where E is a clutter, we have $L_H^{\Bbbk}(d)$ is prime for $d \geq \operatorname{pmd}(H)+1$?

Similarly, the persistence results from Theorem 1.1 ask for generalizations.

Question 9.8. Let $H=(V,E)$ be a hypergraph, where E is a clutter. Is it true that if $L_H^{\Bbbk}(d)$ is a complete intersection (resp. prime) then so is $L_H^{\Bbbk}(d+1)$?

For a number $r \geq 1$ we call a hypergraph $H=(V,E)$ an r-uniform graph if every element of E has cardinality r. In particular, E is a clutter. We say that an r-uniform graph $H=(V,E)$ is r-partite if there is a partition $V=V_1 \cup \cdots \cup V_r$ such that $\#(A \cap V_i)=1$ for all $i \in [r]$ and for all $A \in E$. Now we connect the study of ideal $L_H^{\Bbbk}(d)$ for r-uniform (r-partite) graphs with the study of coordinate sections of the variety of tensors with a given rank. We consider two mappings:

(ϕ) Let $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$ be the standard basis vectors of $\Bbbk^n$. For vectors $v_i=(v_{i,j})_{j\in[d]} \in \Bbbk^d$, $i \in [r]$, consider the map ϕ that sends $(v_1, \ldots, v_r) \in (\Bbbk^d)^n$ to the tensor

$$\sum_{j=1}^{d} \sum_{\sigma \in S_r} v_{\sigma(1),j} \cdots v_{\sigma(r),j}\, \mathfrak{e}_{\sigma(1)} \otimes \cdots \otimes \mathfrak{e}_{\sigma(r)} \in \underbrace{\Bbbk^n \otimes \cdots \otimes \Bbbk^n}_{r}.$$

We take the sums over the different tensors arising from $\mathfrak{e}_{i_1} \otimes \cdots \otimes \mathfrak{e}_{i_r}$, for numbers $1 \leq i_1 \leq \cdots \leq i_r \leq n$, by permuting the positions as standard basis of the space of symmetric tensors.

(ψ) Let $n=n_1+\cdots+n_r$ for natural numbers $n_1, \ldots, n_r \geq 1$. Let $\mathfrak{e}_i^{(j)} \in \Bbbk^{n_j}$ be the i-th standard basis vector of $\Bbbk^{n_j}$, $i \in [n_j]$ and $j \in [r]$. For vectors $v_i^{(j)}=(v_{i,j})_{j\in[d]} \in \Bbbk^d$ for $i \in [n_j]$ and $j \in [r]$ consider the map ψ that sends $(v_i^{(j)})_{(i,j)\in[n_j]\times[r]}$ to

$$\sum_{(i_1,\ldots,i_r)\in[n_1]\times\cdots\times[n_r]} v_{i_1}^{(1)} \cdots v_{i_r}^{(r)}\, e_{i_1}^{(1)} \otimes \cdots \otimes \mathfrak{e}_{i_r}^{(r)} \in \Bbbk^{n_1} \otimes \cdots \otimes \Bbbk^{n_r}.$$

We take the tensors $\mathfrak{e}_{i_1}^{(1)} \otimes \cdots \otimes \mathfrak{e}_{i_r}^{(r)}$ for numbers $i_j \in [n_j]$, $j \in [r]$ as the standard basis of $\Bbbk^{n_1} \otimes \cdots \otimes \Bbbk^{n_r}$.

Recall that a (symmetric) tensor has (symmetric) rank $\leq d$ it can be written as a sum of $\leq d$ decomposable (symmetric) tensors. For more details on tensor rank and the geometry of bounded rank tensors we refer the reader to [Landsberg 2012]. Let $H=(V,E)$ be a hypergraph. We write $\mathcal{V}(L_H^{\Bbbk}(d))$ for the vanishing locus of $L_H^{\Bbbk}(d)$. The definition of the maps ϕ and ψ immediately implies the following proposition.

Proposition 9.9. *Let $H=([n],E)$ be an r-uniform hypergraph and $\Bbbk$ an algebraically closed field.*

(1) *Then the restriction of the map ϕ to $\mathcal{V}(L_H^{\Bbbk}(d))$ is a parametrization of the variety of symmetric tensors in $\Bbbk^n \otimes \cdots \otimes \Bbbk^n$ (with r factors $\Bbbk^n$) of rank $\leq d$ which when expanded in the standard basis has coefficient zero for the basis elements indexed by $1 \leq i_1 < \cdots < i_r \leq n$ and $\{i_1, \ldots, i_r\} \in E$. In particular, the Zariski-closure of the image of the restriction is irreducible if $L_H^{\Bbbk}(d)$ is prime.*

(2) *If H is r-partite with respect to the partition $V = V_1 \cup \cdots \cup V_r$, where $|V_i| = n_i$, $i \in [r]$, then the restriction of the map ψ to $\mathcal{V}(L_H^{\Bbbk}(d))$ is a parametrization of the variety of tensors in $\Bbbk^{n_1} \otimes \cdots \otimes \Bbbk^{n_r}$ of rank $\leq d$ which when expanded in the standard basis have coefficient zero for the basis elements indexed by $i_1, \ldots, i_r$ where $\{i_1, \ldots, i_r\} \in E$. In particular, the Zariski-closure of the image of the restriction is irreducible if $L_H^{\Bbbk}(d)$ is prime.*

Proposition 9.9 gives further motivation to Question 9.7. Indeed, it suggests to strengthen Question 9.4.

Question 9.10. Let $\Bbbk$ be an algebraically closed field. Can one describe classes of r-uniform hypergraphs H for which $L_H^{\Bbbk}(d)$ is prime for some d bounded from above by the maximal symmetric rank of a symmetric tensor in $\Bbbk^n \otimes \cdots \otimes \Bbbk^n$ (with r factors $\Bbbk^n$)?

An analogous question can be asked for r-partite r-uniform hypergraphs and tensors of bounded rank.

Acknowledgments

We thank Alessio D'Ali, Alessio Sammartano and Lorenzo Venturello for useful discussion concerning the material presented. We thank also an anonymous referee for several useful suggestions and comments concerning the earlier version of the paper.

References

[Abbott et al. 2018] J. Abbott, A. M. Bigatti, and L. Robbiano, "CoCoA-5: a system for doing computations in commutative algebra", 2018, Available at http://cocoa.dima.unige.it.

[Ananyan and Hochster 2016] T. Ananyan and M. Hochster, "Small subalgebras of polynomial rings and Stillman's conjecture", preprint, 2016. arXiv

[Avramov 1981] L. L. Avramov, "Complete intersections and symmetric algebras", *J. Algebra* **73**:1 (1981), 248–263. MR Zbl

[Bolognini et al. 2018] D. Bolognini, A. Macchia, and F. Strazzanti, "Binomial edge ideals of bipartite graphs", *European J. Combin.* **70** (2018), 1–25. MR Zbl

[Boocher 2012] A. Boocher, "Free resolutions and sparse determinantal ideals", *Math. Res. Lett.* **19**:4 (2012), 805–821. MR Zbl

[Bourbaki 1990] N. Bourbaki, *Algebra, II: Chapters 4–7*, Springer, 1990. MR Zbl

[Bruns and Conca 2003] W. Bruns and A. Conca, "Gröbner bases and determinantal ideals", pp. 9–66 in *Commutative algebra, singularities and computer algebra* (Sinaia, Romania, 2002), edited by J. Herzog and V. Vuletescu, NATO Sci. Ser. II Math. Phys. Chem. **115**, Kluwer Acad. Publ., Dordrecht, 2003. MR Zbl

[Bruns and Vetter 1988] W. Bruns and U. Vetter, *Determinantal rings*, Lecture Notes in Math. **1327**, Springer, 1988. MR Zbl

[Cartwright and Sturmfels 2010] D. Cartwright and B. Sturmfels, "The Hilbert scheme of the diagonal in a product of projective spaces", *Int. Math. Res. Not.* **2010**:9 (2010), 1741–1771. MR Zbl

[Conca et al. 2015] A. Conca, E. De Negri, and E. Gorla, "Universal Gröbner bases for maximal minors", *Int. Math. Res. Not.* **2015**:11 (2015), 3245–3262. MR Zbl

[Conca et al. 2016] A. Conca, E. De Negri, and E. Gorla, "Universal Gröbner bases and Cartwright–Sturmfels ideals", 2016. To appear in *Int. Math. Res. Not.* arXiv

[Conca et al. 2017] A. Conca, E. De Negri, and E. Gorla, "Multigraded generic initial ideals of determinantal ideals", pp. 81–96 in *Homological and computational methods in commutative algebra* (Cortona, Italy, 2016), edited by A. Conca et al., Springer INdAM Ser. **20**, Springer, 2017. MR

[Conca et al. 2018] A. Conca, E. De Negri, and E. Gorla, "Cartwright–Sturmfels ideals associated to graphs and linear spaces", *J. Comb. Algebra* **2**:3 (2018), 231–257. MR Zbl

[De Concini 1979] C. De Concini, "Symplectic standard tableaux", *Adv. Math.* **34**:1 (1979), 1–27. MR Zbl

[De Concini and Procesi 1976] C. De Concini and C. Procesi, "A characteristic free approach to invariant theory", *Adv. Math.* **21**:3 (1976), 330–354. MR Zbl

[De Concini and Strickland 1981] C. De Concini and E. Strickland, "On the variety of complexes", *Adv. Math.* **41**:1 (1981), 57–77. MR Zbl

[Diaconis et al. 1998] P. Diaconis, D. Eisenbud, and B. Sturmfels, "Lattice walks and primary decomposition", pp. 173–193 in *Mathematical essays in honor of Gian-Carlo Rota* (Cambridge, 1996), edited by B. E. Sagan and R. P. Stanley, Progr. Math. **161**, Birkhäuser, Boston, 1998. MR Zbl

[Diestel 1997] R. Diestel, *Graph theory*, Graduate Texts in Math. **173**, Springer, 1997. MR Zbl

[Eisenbud 1988] D. Eisenbud, "Linear sections of determinantal varieties", *Amer. J. Math.* **110**:3 (1988), 541–575. MR Zbl

[Giusti and Merle 1982] M. Giusti and M. Merle, "Singularités isolées et sections planes de variétés déterminantielles, II: Sections de variétés déterminantielles par les plans de coordonnées", pp. 103–118 in *Algebraic geometry* (La Rábida, Spain, 1981), edited by J. M. Aroca et al., Lecture Notes in Math. **961**, Springer, 1982. MR Zbl

[Grayson and Stillman 1993] D. R. Grayson and M. E. Stillman, "Macaulay2, a software system for research in algebraic geometry", 1993, Available at http://www.math.uiuc.edu/Macaulay2.

[Greene and Zaslavsky 1983] C. Greene and T. Zaslavsky, "On the interpretation of Whitney numbers through arrangements of hyperplanes, zonotopes, non-Radon partitions, and orientations of graphs", *Trans. Amer. Math. Soc.* **280**:1 (1983), 97–126. MR Zbl

[Herzog et al. 2010] J. Herzog, T. Hibi, F. Hreinsdóttir, T. Kahle, and J. Rauh, "Binomial edge ideals and conditional independence statements", *Adv. Appl. Math.* **45**:3 (2010), 317–333. MR Zbl

[Herzog et al. 2015] J. Herzog, A. Macchia, S. Saeedi Madani, and V. Welker, "On the ideal of orthogonal representations of a graph in $\mathbb{R}^2$", *Adv. Appl. Math.* **71** (2015), 146–173. MR Zbl

[Hoşten and Sullivant 2004] S. Hoşten and S. Sullivant, "Ideals of adjacent minors", *J. Algebra* **277**:2 (2004), 615–642. MR Zbl

[Huneke 1981] C. Huneke, "On the symmetric algebra of a module", *J. Algebra* **69**:1 (1981), 113–119. MR Zbl

[Kiani and Saeedi Madani 2016] D. Kiani and S. Saeedi Madani, "The Castelnuovo–Mumford regularity of binomial edge ideals", *J. Combin. Theory Ser. A* **139** (2016), 80–86. MR Zbl

[Landsberg 2012] J. M. Landsberg, *Tensors: geometry and applications*, Graduate Studies in Math. **128**, Amer. Math. Soc., Providence, RI, 2012. MR Zbl

[Lovász 2009] L. Lovász, "Geometric representations of graphs", preprint, 2009, Available at https://tinyurl.com/lovgerep.

[Lovász et al. 1989] L. Lovász, M. Saks, and A. Schrijver, "Orthogonal representations and connectivity of graphs", *Linear Algebra Appl.* **114/115** (1989), 439–454. MR Zbl

[Matsuda and Murai 2013] K. Matsuda and S. Murai, "Regularity bounds for binomial edge ideals", *J. Commut. Algebra* **5**:1 (2013), 141–149. MR Zbl

[Musili and Seshadri 1983] C. Musili and C. S. Seshadri, "Schubert varieties and the variety of complexes", pp. 329–359 in *Arithmetic and geometry, II*, edited by M. Artin and J. Tate, Progr. Math. **36**, Birkhäuser, Boston, 1983. MR Zbl

[Ohtani 2011] M. Ohtani, "Graphs and ideals generated by some 2-minors", *Comm. Algebra* **39**:3 (2011), 905–917. MR Zbl

[Sam and Weyman 2015] S. V. Sam and J. Weyman, "Littlewood complexes and analogues of determinantal varieties", *Int. Math. Res. Not.* **2015**:13 (2015), 4663–4707. MR Zbl

[Tchernev 2001] A. B. Tchernev, "Universal complexes and the generic structure of free resolutions", *Michigan Math. J.* **49**:1 (2001), 65–96. MR Zbl

Communicated by Joseph Gubeladze
Received 2018-02-08 Revised 2018-11-03 Accepted 2018-12-24

conca@dima.unige.it *Dipartimento di Matematica, Università di Genova, 16146 Genova, Italy*

welker@mathematik.uni-marburg.de *Fachbereich Mathematik und Informatik, Philipps-Universität Marburg, 35032 Marburg, Germany*

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.485

On rational singularities and counting points of schemes over finite rings

Itay Glazer

We study the connection between the singularities of a finite type $\mathbb{Z}$-scheme X and the asymptotic point count of X over various finite rings. In particular, if the generic fiber $X_{\mathbb{Q}} = X \times_{\operatorname{Spec}\mathbb{Z}} \operatorname{Spec}\mathbb{Q}$ is a local complete intersection, we show that the boundedness of $|X(\mathbb{Z}/p^n\mathbb{Z})|/p^{n \dim X_{\mathbb{Q}}}$ in p and n is in fact equivalent to the condition that $X_{\mathbb{Q}}$ is reduced and has rational singularities. This paper completes a recent result of Aizenbud and Avni.

1. Introduction

1A. ***Motivation.*** Given a finite type $\mathbb{Z}$-scheme X, the study of the quantity $|X(\mathbb{Z}/m\mathbb{Z})|$ and its asymptotic behavior is a fundamental question in number theory. The case when $m = p$, or more generally the quantity $|X(\mathbb{F}_q)|$ with $q = p^n$, has been studied by many authors, most famously by Lang and Weil [1954], Dwork [1960], Grothendieck [1965] and Deligne [1974; 1980]. The Lang–Weil estimates (see [Lang and Weil 1954]) give a good asymptotic description of $|X(\mathbb{F}_q)|$:

$$|X(\mathbb{F}_q)| = q^{\dim X_{\mathbb{F}_q}}(C_X + O(q^{-\frac{1}{2}})),$$

where C_X is the number of top dimension irreducible components of $X_{\overline{\mathbb{F}}_q}$ that are defined over $\mathbb{F}_q$. From these estimates and the fact that

$$|X(F)| = |U(F)| + |(X \setminus U)(F)|, \tag{1-1}$$

for any open subscheme $U \subseteq X$ and any finite field F, one can see that the asymptotics of $|X(\mathbb{F}_{p^n})|$, in p or in n, does not depend on the singularity properties of X. For finite rings, however, (1-1) is no longer true (e.g., $|\mathbb{A}^1(A)| = |A|$ and $|(\mathbb{A}^1 - \{0\})(A)| = |A^{\times}|$) and indeed, the number $|X(\mathbb{Z}/m\mathbb{Z})|$ and its asymptotics have much to do with the singularities of X. The case when $m = p^n$ is a prime power was studied by Borevich and Shafarevich, among others (see the works of Denef [1991], Igusa [2000], du Sautoy and Grunewald [2000], and a recent overview by Mustaţă [2011]).

For a finite ring A, set $h_X(A) := |X(A)|/|A|^{\dim X_{\mathbb{Q}}}$. If $X_{\mathbb{Q}}$ is smooth, one can show that for almost every prime p, we have $h_X(\mathbb{Z}/p^n\mathbb{Z}) = h_X(\mathbb{Z}/p\mathbb{Z})$ for all n, which by the Lang–Weil estimates is uniformly bounded. On the other hand, if $X_{\mathbb{Q}}$ is singular, then $h_X(\mathbb{Z}/p^n\mathbb{Z})$ need not be bounded in n or in p. The goal of this paper is to investigate this phenomena and to complete the main result presented in [Aizenbud and Avni 2018], which we describe next.

MSC2010: primary 14B05; secondary 14G05.
Keywords: rational singularities, complete intersection, analysis on p-adic varieties, asymptotic point count.

1B. ***Related work.*** Aizenbud and Avni [2018] proved the following:

Theorem 1.1 [Aizenbud and Avni 2018, Theorem 3.0.3]. *Let X be a finite type $\mathbb{Z}$-scheme such that $X_{\mathbb{Q}}$ is equidimensional and a local complete intersection. Then the following are equivalent*:

(i) *For any n, $\lim_{p\to\infty} h_X(\mathbb{Z}/p^n\mathbb{Z}) = 1$.*

(ii) *There exists a finite set of prime numbers S and a constant C, such that $|h_X(\mathbb{Z}/p^n\mathbb{Z}) - 1| < Cp^{-\frac{1}{2}}$ for any prime $p \notin S$ and any $n \in \mathbb{N}$.*

(iii) *$X_{\overline{\mathbb{Q}}}$ is reduced, irreducible and has rational singularities.*

Definition 1.2 [Aizenbud and Avni 2016, 1.2, Definition II]. Let X and Y be smooth varieties over a field k of characteristic 0. We say that a morphism $\varphi : X \to Y$ is (FRS) if it is flat and any geometric fiber is reduced and has rational singularities. We say that φ is (FRS) at $x \in X(k)$ if there exists a Zariski open neighborhood U of x such that $U \times_Y \{\varphi(x)\}$ is reduced and has rational singularities.

Aizenbud and Avni introduced an analytic criterion for a morphism φ to be (FRS), which played a key role in the proof of Theorem 1.1:

Theorem 1.3 [Aizenbud and Avni 2016, Theorem 3.4]. *Let $\varphi : X \to Y$ be a map between smooth algebraic varieties defined over a finitely generated field k of characteristic* 0, *and let $x \in X(k)$. Then the following conditions are equivalent*:

(a) *φ is* (FRS) *at x.*

(b) *There exists a Zariski open neighborhood $x \in U \subseteq X$, such that for any non-Archimedean local field $F \supseteq k$ and any Schwartz measure m on $U(F)$, the measure $(\varphi|_{U(F)})_*(m)$ has continuous density* (*see Definition 2.5 for the notion of Schwartz measure and continuous density of a measure*).

(c) *For any finite extension k'/k, there exists a non-Archimedean local field $F \supseteq k'$ and a nonnegative Schwartz measure m on $X(F)$ that does not vanish at x such that $\varphi_*(m)$ has continuous density.*

1C. ***Main results.*** In this paper, we strengthen Theorem 1.1 as follows:

Theorem 1.4. *Let X be a finite type $\mathbb{Z}$-scheme such that $X_{\mathbb{Q}}$ is equidimensional and a local complete intersection. Then* (i), (ii) *and* (iii) *in Theorem 1.1 are also equivalent to*:

(iv) *$X_{\overline{\mathbb{Q}}}$ is irreducible and there exists $C > 0$ such that $h_X(\mathbb{Z}/p^n\mathbb{Z}) < C$ for any prime p and any $n \in \mathbb{N}$.*

(v) *$X_{\overline{\mathbb{Q}}}$ is irreducible and there exists a finite set of primes S, such that for any $p \notin S$, the sequence $n \mapsto h_X(\mathbb{Z}/p^n\mathbb{Z})$ is bounded.*

Remark. In fact, one can drop the demand that $X_{\overline{\mathbb{Q}}}$ is irreducible in conditions (iii), (iv) and (v), such that they will stay equivalent. For a slightly stronger statement, see Theorem 4.1.

There are two main difficulties in the proof of Theorem 1.4. The first one is portrayed in the fact that condition (v) seems a-priori too weak, as it requires the bound on $h_X(\mathbb{Z}/p^n\mathbb{Z})$ to be uniform only in n, while in condition (ii), the demand is that the bound is uniform both in p and in n.

In order to show that condition (v) implies the other conditions, we first reduce to the case when $X_{\mathbb{Q}}$ is a complete intersection in an affine space, and thus can be written as the fiber at 0 of a morphism $\varphi : \mathbb{A}^M_{\mathbb{Q}} \to \mathbb{A}^N_{\mathbb{Q}}$, which is flat above 0. We can then translate condition (iii), i.e., the condition that $X_{\overline{\mathbb{Q}}}$ is reduced and has rational singularities, to the condition that $\varphi : \mathbb{A}^M_{\mathbb{Q}} \to \mathbb{A}^N_{\mathbb{Q}}$ is (FRS) above 0, i.e., at any point $x \in (\varphi^{-1}(0))(\overline{\mathbb{Q}})$. After some technical argument, one can show that condition (v) implies the following:

Condition 1.5. For any finite extensions $k/\mathbb{Q}$ and k'/k, and any $x \in (\varphi^{-1}(0))(k)$, there exists a prime p with $k' \hookrightarrow \mathbb{Q}_p$, $x \in (\varphi^{-1}(0))(\mathbb{Z}_p)$, such that the sequence $n \mapsto \varphi_*(\mu)(p^n\mathbb{Z}_p^N)/p^{-nN}$ is bounded, where μ is the normalized Haar measure on $\mathbb{Z}_p^M$.

Hence, we would like to strengthen Theorem 1.3, such that Condition 1.5 will imply the (FRS) property of φ above 0.

The measure $\varphi_*(m)$ as in Condition 1.5 is said to be bounded with respect to the local basis $\{p^n\mathbb{Z}_p^N\}_n$ for the topology of $\mathbb{Q}_p^N$ at 0 (see Definition 3.1).We introduce the notion of bounded eccentricity of a local basis to the topology of an F-analytic manifold (Section 3A), and prove the following stronger version of Theorem 1.3:

Theorem 1.6. *Let $\varphi : X \to Y$ be a map between smooth algebraic varieties defined over a finitely generated field k of characteristic* 0, *and let $x \in X(k)$. Then* (a), (b), (c) *in Theorem 1.3 are also equivalent to*:

(c′) *For any finite extension k'/k, there exists a non-Archimedean local field $F \supseteq k'$ and a nonnegative Schwartz measure m on $X(F)$ that does not vanish at x, such that $\varphi_*(m)$ is bounded with respect to some local basis $\mathcal{N}$ of bounded eccentricity at $\varphi(x)$.*

We then use Theorem 1.6 and the fact that the local basis $\{p^n\mathbb{Z}_p^N\}_n$ is of bounded eccentricity to show that (v) implies condition (iii).

The second difficulty is to show that if $h_X(\mathbb{Z}/p^n\mathbb{Z})$ is bounded for almost any prime p, then it is in fact bounded for any p. We first prove this for the case that X is a complete intersection in an affine space, denoted (CIA) (Proposition 4.5). We then deal with the case when $X_{\mathbb{Q}}$ is a (CIA), by constructing a finite type $\mathbb{Z}$-scheme $\widehat{X}$, which is a (CIA) and a morphism $\psi : X \longrightarrow \widehat{X}$, such that $\psi_{\mathbb{Q}} : X_{\mathbb{Q}} \longrightarrow \widehat{X}_{\mathbb{Q}}$ is an isomorphism (Lemma 4.6). We prove this case by showing the existence of $c, N \in \mathbb{N}$ such that

$$|X(\mathbb{Z}/p^n\mathbb{Z})| \leq p^{Nc} \cdot |\widehat{X}(\mathbb{Z}/p^n\mathbb{Z})|,$$

(Lemma 4.7). For the general case, we first cover $X_{\mathbb{Q}}$ by affine $\mathbb{Q}$-schemes $\{U_i\}$ such that U_i is a (CIA), and then consider a collection of $\mathbb{Z}$-schemes $\{\widetilde{U}_i\}$, such that $\widetilde{U}_i \simeq U_i$ over $\mathbb{Q}$. Finally, using the explicit construction of $\widetilde{U}_i$ we show that

$$h_X(\mathbb{Z}/p^n\mathbb{Z}) \leq \sum_i h_{\widetilde{U}_i}(\mathbb{Z}/p^n\mathbb{Z}),$$

and since $(\widetilde{U}_i)_{\mathbb{Q}} \simeq U_i$ is a (CIA), we are done by the last case.

2. Preliminaries

In this section, we recall some definitions and facts in algebraic geometry and the theory of F-analytic manifolds, for a non-Archimedean local field F.

2A. ***Preliminaries in algebraic geometry.*** Let A be a commutative ring. A sequence $x_1, \dots, x_r \in A$ is called a *regular sequence* if x_i is not a zero-divisor in $A/(x_1, \dots, x_{i-1})$ for each i, and we have a proper inclusion $(x_1, \dots, x_r) \subsetneq A$. If $(A, \mathfrak{m})$ is a Noetherian local ring then the *depth* of A, denoted $\operatorname{depth}(A)$, is defined to be the length of the longest regular sequence with elements in $\mathfrak{m}$. It follows from Krull's principal ideal theorem that $\operatorname{depth}(A)$ is smaller or equal to $\dim(A)$, the Krull dimension of A. A Noetherian local ring $(A, \mathfrak{m})$ is *Cohen–Macaulay* if $\operatorname{depth}(A) = \dim(A)$. A locally Noetherian scheme X is said to be *Cohen–Macaulay* if for any $x \in X$, the local ring $\mathcal{O}_{X,x}$ is Cohen–Macaulay.

Let X be an algebraic variety over a field k. We say that X has a *resolution of singularities*, if there exists a proper morphism $p : \widetilde{X} \to X$ such that $\widetilde{X}$ is smooth and p is a birational equivalence. A *strong resolution of singularities* of X is a resolution of singularities $p : \widetilde{X} \to X$ which is an isomorphism over the smooth locus of X, denoted X^{sm}. It is a theorem of Hironaka [1964], that any variety X over a field k of characteristic zero admits a strong resolution of singularities $p : \widetilde{X} \to X$.

For the following definition, see [Kempf et al. 1973, I.3 pages 50–51] or [Aizenbud and Avni 2016, Definition 6.1]; a variety X over a field k of characteristic zero is said to have *rational singularities* if for any (or equivalently, for some) resolution of singularities $p : \widetilde{X} \longrightarrow X$, the natural morphism $\mathcal{O}_X \to \mathrm{R}p_*(\mathcal{O}_{\widetilde{X}})$ is a quasi-isomorphism, where $\mathrm{R}p_*$ is the higher direct image. A point $x \in X(k)$ is a *rational singularity* if there exists a Zariski open neighborhood $U \subseteq X$ of x that has rational singularities.

We denote by Ω_X^r the sheaf of differential r-forms on X and by $\Omega_X^r[X]$ (resp. $\Omega_X^r(X)$) the regular (resp. rational) r-forms. The following lemma gives a local characterization of rational singularities:

Lemma 2.1 [Aizenbud and Avni 2016, Proposition 6.2]. *An affine k-variety X has rational singularities if and only if X is Cohen–Macaulay, normal, and for any, or equivalently, some strong resolution of singularities $p : \widetilde{X} \to X$ and any top differential form $\omega \in \Omega_{X^{\mathrm{sm}}}^{\mathrm{top}}[X^{\mathrm{sm}}]$, there exists a top differential form $\widetilde{\omega} \in \Omega_{\widetilde{X}}^{\mathrm{top}}[\widetilde{X}]$ such that $\omega = \widetilde{\omega}|_{X^{\mathrm{sm}}}$.*

Let X be a finite type scheme over a ring R. Then X is called:

(1) A *complete intersection* (CI) if there exists an affine scheme Y, a smooth morphism $Y \to \operatorname{Spec} R$, a closed embedding $X \hookrightarrow Y$ over $\operatorname{Spec} R$, and a regular sequence $f_1, \dots, f_r \in \mathcal{O}_Y(Y)$, such that the ideal of X in Y is generated by the $\{f_i\}$. In this case, we say that X is a *complete intersection in Y*.

(2) A *local complete intersection* (LCI) if there is an open cover $\{U_i\}$ of X such that each U_i is a (CI).

(3) A *complete intersection in an affine space* (CIA) if X is a complete intersection in Y, with $Y = \mathbb{A}_R^n$ an affine space.

(4) A *local complete intersection in an affine space* (LCIA) if there is an open affine cover $\{U_i\}$ of X such that each U_i is a (CIA).

Remark 2.2. For an affine k-variety, the notion of (CIA) is not equivalent to (CI) (e.g., consider X to be any affine smooth k-variety which is not a (CIA)). On the other hand, the notion of (LCI) is equivalent to (LCIA) for finite type k-schemes. We will therefore use the notation (LCI) for both notions.

The following Proposition 2.3 and Proposition 2.4 are a consequence of the above remark and the miracle flatness theorem (e.g., [Vakil 2017, Theorems 26.2.10 and 26.2.11]).

Proposition 2.3. *Let X be k-variety. If X is an* (LCI) *then there exists an open affine cover $\{U_i\}$ of X and morphisms φ_i, ψ_i, where $\varphi_i : \mathbb{A}_k^{m_i} \longrightarrow \mathbb{A}_k^{n_i}$ is flat above* 0, *and $\psi_i : U_i \hookrightarrow \mathbb{A}_k^{m_i}$ is a closed embedding that induces a k-isomorphism $\psi_i : U_i \simeq \varphi_i^{-1}(0)$.*

Proposition 2.4. *Let X be a finite type $\mathbb{Z}$-scheme. If X is a* (CIA) *then there exist $\mathbb{Z}$-morphisms φ, ψ, where $\varphi : \mathbb{A}^m_{\mathbb{Z}} \longrightarrow \mathbb{A}^n_{\mathbb{Z}}$ is flat above* 0, *and $\psi : X \hookrightarrow \mathbb{A}^m_{\mathbb{Z}}$ is a closed embedding that induces a $\mathbb{Z}$-isomorphism* $\psi : X \simeq \varphi^{-1}(0)$.

A commutative Noetherian local ring A is called *Gorenstein* if it has finite injective dimension as an A-module. A locally Noetherian scheme X is said to be *Gorenstein* if all its local rings are Gorenstein. Any locally Noetherian scheme X which is a local complete intersection is also Gorenstein.

2B. ***Some facts on F-analytic manifolds.*** Let X be a d-dimensional smooth algebraic k-variety and $F \supseteq k$ be a non-Archimedean local field, with ring of integers $\mathcal{O}_F$. Then $X(F)$ has a structure of an F-analytic manifold. Given $\omega \in \Omega^{\mathrm{top}}_X(X)$, we can define a measure $|\omega|_F$ on $X(F)$ as follows. For a compact open set $U \subseteq X(F)$ and an F-analytic diffeomorphism ϕ between an open subset $W \subseteq F^d$ and U, we can write $\phi^*\omega = g \cdot dx_1 \wedge \cdots \wedge dx_n$, for some $g : W \to F$, and define

$$|\omega|_F(U) = \int_W |g|_F \, d\lambda,$$

where $|\cdot|_F$ is the normalized absolute value on F and λ is the normalized Haar measure on F^d. Note that this definition is independent of the diffeomorphism ϕ, and that this uniquely defines a measure on $X(F)$.

Definition 2.5. (1) A measure m on $X(F)$ is called *smooth* if every point $x \in X(F)$ has an analytic neighborhood U and an F-analytic diffeomorphism $f : U \to \mathcal{O}^d_F$ such that f_*m is some Haar measure on $\mathcal{O}^d_F$.

(2) A measure on $X(F)$ is called *Schwartz* if it is smooth and compactly supported.

(3) We say that a measure μ on $X(F)$ has *continuous density*, if there is a smooth measure m and a continuous function $f : X(F) \to \mathbb{C}$ such that $\mu = f \cdot m$.

The following proposition characterizes Schwartz measures and measures with continuous density:

Proposition 2.6 [Aizenbud and Avni 2016, Proposition 3.3]. *Let X be a smooth variety over a non-Archimedean local field F.*

(1) *A measure m on $X(F)$ is Schwartz if and only if it is a linear combination of measures of the form $f|\omega|_F$, where f is a Schwartz function* (*i.e., locally constant and compactly supported*) *on $X(F)$, and $\omega \in \Omega^{\mathrm{top}}_X(X)$ has no zeros or poles in the support of f.*

(2) *A measure μ on $X(F)$ has continuous density if and only if for every point $x \in X(F)$ there is an analytic neighborhood U of x, a continuous function $f : U \to \mathbb{C}$, and $\omega \in \Omega^{\mathrm{top}}_X(X)$ with no poles in U such that $\mu = f|\omega|_F$.*

Proposition 2.7 [Aizenbud and Avni 2016, Proposition 3.5]. *Let $\varphi : X \to Y$ be a smooth map between smooth varieties defined over a non-Archimedean local field F.*

(1) *If m is a Schwartz measure on $X(F)$, then φ_*m is a Schwartz measure on $Y(F)$.*

(2) *Assume that $\omega_X \in \Omega^{\mathrm{top}}_X[X]$ and $\omega_Y \in \Omega^{\mathrm{top}}_Y[Y]$, where ω_Y is nowhere vanishing, and that f is a Schwartz function on $X(F)$. Then the measure $\varphi_*(f|\omega_X|_F)$ is absolutely continuous with respect to $|\omega_Y|_F$, and its density at a point $y \in Y(F)$ is $\int_{\varphi^{-1}(y)(F)} f \cdot \big|(\omega_X/\varphi^*\omega_Y)|_{\varphi^{-1}(y)}\big|_F$.*

3. An analytic criterion for the (FRS) property

Our goal in this section is to prove Theorem 1.6, which is a stronger version of Theorem 1.3, and the main ingredient in the proof of the implication (v) $\Rightarrow$ (iii) of Theorem 1.4. As discussed in the introduction, we want to relax condition (c) of Theorem 1.3, and get a weaker condition (c′) that is similar to Condition 1.5, such that it will imply the (FRS) property (condition (a) of Theorem 1.3).

Definition 3.1. Let F be a non-Archimedean local field, X be an F-analytic manifold and μ be a measure on X. Let $\mathcal{N} = \{N_i\}_{i\in I}$ be a local basis for the topology of X at a point $x \in X$. We say that μ is *bounded with respect to* $\mathcal{N}$, if there exists a smooth measure λ on X and an open analytic neighborhood U of x, such that $|\mu(N_i)/\lambda(N_i)|$ is uniformly bounded on $\mathcal{N}_U := \{N_i \in \mathcal{N} \mid N_i \subseteq U\}$.

Let $\varphi : X \to Y$, m and F be as in Theorem 1.3. A possible relaxation (c′) of (c), is to require $\varphi_*(m)$ to be bounded with respect to any local basis of the topology of $Y(F)$ at $\varphi(x)$. While this condition is equivalent to (a) and (b) it is still not weak enough for our purpose of proving Theorem 1.4. A much weaker condition (c″) is to demand that $\varphi_*(m)$ is bounded with respect to *some* local basis at $\varphi(x)$. Unfortunately, the following example shows that the latter demand is too weak:

Example. Consider the map $\varphi : \mathbb{A}^2_{\mathbb{Q}} \longrightarrow \mathbb{A}_{\mathbb{Q}}$ defined by $(x, y) \longmapsto x^2$. The fiber over 0 is not reduced, and thus φ is not (FRS) over 0. Fix a finite extension $k/\mathbb{Q}$ and embed k in $\mathbb{Q}_p$ for some prime p (see Lemma 4.3). Let λ_1, λ_2 be the normalized Haar measure on $\mathbb{Q}_p, \mathbb{Q}_p^2$ and let $m = 1_{\mathbb{Z}_p^2} \cdot \lambda_2$ be a Schwartz measure. Now consider the following collection $\mathcal{N}$ of sets B_n constructed as follows. Define $B_n^1 := \{x \in \mathbb{Z}_p \mid |x| \le p^{-2n^2}\}$ and $B_n^2 := \{x \in \mathbb{Z}_p \mid |x - a_n| \le p^{-4n}\}$, where $a_n = p^{2n+1}$. Note that any $x \in B_n^2$ has norm p^{-2n-1} and thus is not a square, so $\varphi^{-1}(B_n^2) = \varnothing$. Denote $B_n = B_n^1 \cup B_n^2$ and notice that $\mathcal{N} := \{B_n\}_{n=1}^{\infty}$ is a local basis at 0 and that:

$$\lim_{n\to\infty} \frac{\varphi_* m(B_n)}{\lambda_1(B_n)} = \lim_{n\to\infty} \frac{m(\varphi^{-1}(B_n^1))}{p^{-2n^2} + p^{-4n}} = \lim_{n\to\infty} \frac{p^{-n^2}}{p^{-2n^2} + p^{-4n}} \to 0.$$

This shows that φ satisfies condition (c″) but is not (FRS) at (0, 0).

Luckily, we can relax (c) by demanding that $\varphi_*(m)$ is bounded with respect to *some* local basis at $\varphi(x)$, if this basis is nice enough. In order to define precisely what we mean, we introduce the notion of a local basis of bounded eccentricity.

3A. *Local basis of bounded eccentricity.*

Definition 3.2. Let F be a local field, and λ be a Haar measure on F^n.

(1) A collection of sets $\mathcal{N} = \{N_i\}_{i\in I}$ in F^n is said to have *bounded eccentricity* at $x \in F^n$, if there exists a constant $C > 0$ such that $\sup_i(\lambda(B_{\min_i}(x))/\lambda(B_{\max_i}(x))) \le C$, where $B_{\max_i}(x)$ is the maximal ball around x that is contained in N_i and $B_{\min_i}(x)$ is the minimal ball around x that contains N_i.

(2) We call $\mathcal{N} = \{N_i\}_{i\in I}$ a *local basis of bounded eccentricity* at x, if it is a local basis of the topology of F^n at x, and there exists $\epsilon > 0$, such that $\mathcal{N}_\epsilon := \{N_i \in \mathcal{N} \mid N_i \subseteq B_\epsilon(x)\}$ has bounded eccentricity.

Remark. Note that $\mathcal{N}_\epsilon \neq \varnothing$ for any $\epsilon > 0$ since it is a local basis at x.

Lemma 3.3. *Let $\phi : F^n \to F^n$ be an F-analytic diffeomorphism. Let $\mathcal{N} = \{N_i\}_{i\in\alpha}$ be a local basis of bounded eccentricity at $x \in F^n$. Then $\phi(\mathcal{N})$ is a local basis of bounded eccentricity at $\phi(x)$.*

Proof. Let $d\phi_x = A$ be the differential of ϕ at x. Since ϕ is a diffeomorphism, then for any $C > 1$, there exists $\delta, \delta' > 0$ such that for any $y \in B_\delta(x)$:

$$\frac{1}{C} < \frac{|\phi(y) - \phi(x)|_F}{|A \cdot (y-x)|_F} < C,$$

and for any $z \in B_{\delta'}(\phi(x))$ we have:

$$\frac{1}{C} < \frac{|\phi^{-1}(z) - x|_F}{|A^{-1} \cdot (z - \phi(x))|_F} < C.$$

We can choose small enough δ, δ' such that $\mathcal{N}_\delta$ is a collection of sets which has bounded eccentricity and $\phi(\mathcal{N}_\delta) \supseteq \phi(\mathcal{N})_{\delta'}$. We now claim that $\mathcal{M}_{\delta'} := \phi(\mathcal{N})_{\delta'}$ is a collection of sets which has bounded eccentricity at $\phi(x)$. Let $B_{\min_i}(x)$ be the minimal ball that contains $N_i \in \mathcal{N}_\delta$ and $B_{\max_i}(x)$ be the maximal ball that is contained in N_i. Notice that for any $y \in B_{\min_i}(x) \subseteq B_\delta(x)$ we have

$$|\phi(y) - \phi(x)|_F < C \cdot |A \cdot (y-x)|_F \leq C \cdot \|A\| \cdot |y - x|_F \leq C \cdot \min_i \cdot \|A\|,$$

thus $\phi(N_i) \subseteq \phi(B_{\min_i}(x)) \subseteq B_{C \cdot \min_i \cdot \|A\|}(\phi(x))$. Similarly, for any $z \in B_{\max_i/(C \cdot \|A^{-1}\|)}(\phi(x))$ we have that

$$|\phi^{-1}(z) - x|_F < C \cdot |A^{-1} \cdot (z - \phi(x))|_F \leq C \cdot \|A^{-1}\| \cdot \frac{\max_i}{C \cdot \|A^{-1}\|} = \max_i .$$

Therefore, $\phi^{-1}\big(B_{\max_i/(C \cdot \|A^{-1}\|)}(\phi(x))\big) \subseteq B_{\max_i}(x) \subseteq N_i$ and thus $B_{\max_i/(C \cdot \|A^{-1}\|)}(\phi(x)) \subseteq \phi(N_i)$. Thus we get that

$$B_{\max_i/(C \cdot \|A^{-1}\|)}(\phi(x)) \subseteq \phi(N_i) \subseteq B_{C \cdot \min_i \cdot \|A\|}(\phi(x)),$$

for any i. By assumption, there exists some $D > 0$ such that $\lambda(B_{\min_i}(x))/\lambda(B_{\max_i}(x)) < D$ for any set $N_i \in \mathcal{N}_\delta$. Hence:

$$\frac{\lambda\big(B_{C \cdot \min_i \cdot \|A\|}(\phi(x))\big)}{\lambda\big(B_{\max_i/(C \cdot \|A^{-1}\|)}(\phi(x))\big)} \leq C^{2n} \|A\|^n \cdot \|A^{-1}\|^n \cdot D,$$

and $\mathcal{M}_{\delta'}$ has bounded eccentricity. □

Lemma 3.3 implies that the following notion is well defined:

Definition 3.4. Let X be an F-analytic manifold and λ be a Haar measure on F^n.

(1) A local basis $\mathcal{N}$ at $x \in X$ is said to have *bounded eccentricity* if given an F-analytic diffeomorphism ϕ between an open subset $W \subseteq F^n$ and an open neighborhood U of x, we have that

$$\widetilde{\mathcal{N}} = \{\phi^{-1}(N) \mid N \in \mathcal{N},\ N \subseteq U\}$$

is a local basis of bounded eccentricity.

(2) A measure m on X is said to be $\mathcal{N}$*-bounded*, if there exists $\epsilon > 0$ such that:

$$\sup_{N \in \mathcal{N}_\epsilon} \frac{m(N)}{\lambda(N)} < \infty.$$

3B. ***Proof of Theorem 1.6.*** It is easy to see that (c) $\Rightarrow$ (c′). The proof of the implication (c′) $\Rightarrow$ (a) is a variation of the proof of (c) $\Rightarrow$ (a) of Theorem 1.3 (see [Aizenbud and Avni 2016, Section 3.7]). Let k be a finitely generated field of characteristic 0, $\varphi : X \to Y$ be a morphism of smooth k-varieties X, Y and let $x \in X(k)$. Assume that condition (c′) of Theorem 1.6 holds. Let $Z = \varphi^{-1}(\varphi(x))$ and denote by X^S the smooth locus of φ. The following lemma is a slight variation of [Aizenbud and Avni 2016, Claim 3.19]. Since we use the constructions presented in the proof of [loc. cit.], and for the convenience of the reader, we write the full steps and use similar notation as well.

Lemma 3.5. *There exists a Zariski neighborhood U of x such that $Z \cap X^S \cap U$ is a dense subvariety of $Z \cap U$.*

Proof. Let $Z_1, \dots, Z_n$ be the absolutely irreducible components of Z containing x. After restricting to an open neighborhood of x that does not intersect the other irreducible components, it is enough to show that $Z_i \cap X^S$ is Zariski dense in Z_i for any i. Since X^S is open, it is enough to show that $Z_i \cap X^S$ is nonempty for any i.

Assume that $Z_i \cap X^S = \varnothing$ for some i. Then $\dim \ker d\varphi_z > \dim X - \dim Y$ for any $z \in Z_i(\bar{k})$. By the upper semicontinuity of $\dim \ker d\varphi$, there is a nonempty open set $W_i \subseteq Z_i$ and an integer $r \geq 1$ such that $\dim \ker d\varphi|_z = \dim X - \dim Y + r$ for all $z \in W_i(\bar{k})$ and such that $W_i \cap Z_j = \varnothing$ for any $j \neq i$. Let k'/k be a finite extension such that both Z_i, W_i are defined over k' and $W_i^{\mathrm{sm}}(k') \neq \varnothing$. By [Aizenbud and Avni 2016, Lemma 3.14], we can choose k' such that $x \in \overline{W_i^{\mathrm{sm}}(F)}$ for any non-Archimedean local field $F \supseteq k'$.

By our assumption, there exists a non-Archimedean local field $F \supseteq k'$ and a nonnegative Schwartz measure m on $X(F)$ that does not vanish at x and such that $\varphi_* m$ is bounded with respect to some local basis $\mathcal{N}$ (at $\varphi(x)$) of bounded eccentricity. Since $x \in \overline{W_i^{\mathrm{sm}}(F)}$, there exists a point $p \in W_i^{\mathrm{sm}}(F) \cap \mathrm{supp}(m)$.

By the implicit function theorem, there exist neighborhoods $U_X \subseteq X(F)$ and $U_Y \subseteq Y(F)$ of p and $\varphi(x) = \varphi(p)$ respectively, analytic diffeomorphisms $\alpha_X : U_X \to \mathcal{O}_F^{\dim X}$, $\alpha_Y : U_Y \to \mathcal{O}_F^{\dim Y}$ and $\alpha_{Z_i} : U_X \cap W_i^{\mathrm{sm}}(F) \to \mathcal{O}_F^{\dim Z_i}$ such that $\alpha_X(p) = 0$, $\alpha_Y(\varphi(p)) = 0$, and an analytic map $\psi : \mathcal{O}_F^{\dim X} \to \mathcal{O}_F^{\dim Y}$ such that the following diagram commutes:

$$\begin{tikzcd}
U_X \cap W_i^{\mathrm{sm}}(F) \arrow[r, hook] \arrow[d, "\alpha_{Z_i}"'] & U_X \arrow[r, "\varphi|_{U_X}"] \arrow[d, "\alpha_X"] & U_Y \arrow[d, "\alpha_Y"] \\
\mathcal{O}_F^{\dim Z_i} \arrow[r, hook, "j"] & \mathcal{O}_F^{\dim X} \arrow[r, "\psi"] & \mathcal{O}_F^{\dim Y}
\end{tikzcd}$$

where $j : \mathcal{O}_F^{\dim Z_i} \to \mathcal{O}_F^{\dim X}$ is the inclusion to the first $\dim Z_i$ coordinates. After an analytic change of coordinates we may assume that:

$$\ker d\psi_z = \mathrm{span}\{e_1, \dots, e_{\dim X - \dim Y + r}\},$$

for any $z \in \mathcal{O}_F^{\dim Z_i}$. By Lemma 3.3, we have that $\mathcal{M} := \alpha_Y(\mathcal{N})$ is a local basis of bounded eccentricity at $0 \in \mathcal{O}_F^{\dim Y}$. Note that $\mu := (\alpha_X)_*(1_{U_X} \cdot m)$ is a nonnegative Schwartz measure that does not vanish at 0, and that $\psi_*(\mu)$ is $\mathcal{M}$-bounded. By Proposition 2.6, after restricting to a small enough ball around 0 and applying a homothety, we can assume that μ is the normalized Haar measure.

As part of the data, for any $M_j \in \mathcal{M}$ we are given by $B_{\max_j}(0)$ and $B_{\min_j}(0)$, and there exists $\delta, C > 0$ such that for any $M_j \in \mathcal{M}_\delta := \{M_j \in \mathcal{M} \mid M_j \subseteq B_\delta(0)\}$, we have $B_{\max_j}(0) \subseteq M_j \subseteq B_{\min_j}(0)$ and

$\lambda(B_{\min_j})/\lambda(B_{\max_j}) \le C$. For any $0 < \epsilon < 1$, set

$$A_\epsilon := \left\{(x_1, \dots, x_{\dim X}) \in \mathcal{O}_F^{\dim X} \mid |x_k| < \epsilon^{n_k}\right\},$$

where $n_k = 0$ if $1 \le k \le \dim Z_i$; $n_k = 1$ for $\dim Z_i + 1 \le k \le \dim X - \dim Y + r$; and $n_k = 2$ for $\dim X - \dim Y + r + 1 \le k \le \dim X$.

By choosing δ small enough, we may find a constant $D > 0$ such that $\psi(A_{D\sqrt{\epsilon}}) \subseteq B_\epsilon(0)$ for every $\epsilon < \delta$. In particular, for any $M_j \in \mathcal{M}_\delta$ we get that $\psi(A_{D\cdot\sqrt{\max_j}}) \subseteq B_{\max_j}(0)$, so $\psi^{-1}(B_{\max_j}(0)) \supseteq A_{D\cdot\sqrt{\max_j}}$. Denote $\sqrt{\max_j}$ by ϵ_j and notice that there exists a constant $L > 0$ such that for any j with $M_j \in \mathcal{M}_\delta$, it holds that

$$\begin{aligned}\mu(A_{D\epsilon_j}) &\ge L \cdot (D\epsilon_j)^{\dim X - \dim Y + r - \dim Z_i + 2(\dim Y - r)}\\ &= D' \cdot \epsilon_j^{\dim X + \dim Y - r - \dim Z_i}\\ &\ge D' \cdot \epsilon_j^{2\dim Y - r},\end{aligned}$$

where D' is some positive constant. Altogether, we have:

$$\begin{aligned}\frac{\psi_*(\mu)(M_j)}{\lambda(M_j)} &\ge \frac{\psi_*(\mu)(B_{\max_j}(0))}{\lambda(B_{\min_j}(0))} \ge \frac{1}{C}\frac{\psi_*(\mu)(B_{\max_j}(0))}{\lambda(B_{\max_j}(0))}\\ &\ge \frac{1}{C}\frac{\mu(A_{D\epsilon_j})}{\lambda(B_{\max_j}(0))} \ge \frac{D'}{C}\frac{\epsilon_j^{2\dim Y - r}}{\epsilon_j^{2\dim Y}} \ge \frac{D'}{C}\epsilon_j^{-r}.\end{aligned}$$

Since $\mathcal{M}_\delta$ is a local basis, the above equation is true for arbitrary small ϵ_j, so we have a contradiction to the $\mathcal{M}$-boundedness of $\psi_*(\mu)$. □

Corollary 3.6. *We have that φ is flat at x, and that there is a Zariski neighborhood U_0 of x such that $Z \cap U_0$ is reduced and a local complete intersection* (LCI).

Proof. Let $Z_1, \dots, Z_n$ be the absolutely irreducible components of Z containing x. By the previous lemma, each Z_i contains a smooth point of φ, so $\dim_x Z := \max_i \dim Z_i = \dim X - \dim Y$. Hence, we may find a neighborhood U_0 of x such that $\varphi|_{U_0}$ is flat over $\varphi(x)$ (and in particular flat at x). As a consequence, we get that $Z \cap U_0$ is an (LCI), and in particular Cohen–Macaulay. Since $Z \cap X^S \cap U_0$ is dense in $Z \cap U_0$ and $Z \cap X^S = Z^{\mathrm{sm}}$ (see, e.g., [Hartshorne 1977, III.10.2]) it follows that $Z \cap U_0$ is generically reduced. Since $Z \cap U_0$ is also Cohen–Macaulay, it now follows from (e.g., [Vakil 2017, Exercise 26.3.B]) that it is reduced. □

Without loss of generality, we assume $X = U_0$. The following lemma implies that φ is (FRS) at x, and thus finishes the proof of Theorem 1.6:

Lemma 3.7. *The element x is a rational singularity of Z.*

Proof. After further restricting to Zariski open neighborhoods of x and $\varphi(x)$, we may assume that X and Y are affine, with $\Omega_X^{\mathrm{top}}, \Omega_Y^{\mathrm{top}}$ free. Fix invertible top forms $\omega_X \in \Omega_X^{\mathrm{top}}[X]$, $\omega_Y \in \Omega_Y^{\mathrm{top}}[Y]$. We may find an invertible section $\eta \in \Omega_Z^{\mathrm{top}}[Z]$, such that $\eta|_{Z^{\mathrm{sm}}} = \omega_X|_{X^S}/\varphi^*(\omega_Y)$ (for more details see the last part of the proof of [Aizenbud and Avni 2016, Theorem 3.4]). We denote $\omega_Z := \eta|_{Z^{\mathrm{sm}}}$.

Fix a finite extension k'/k. By assumption, there exists a non-Archimedean local field $F \supseteq k'$ and a nonnegative Schwartz measure m on $X(F)$ that does not vanish at x, such that $\varphi_*(m)$ is bounded with respect to a local basis $\mathcal{N}$ of bounded eccentricity. Write m as $m = f \cdot |\omega_X|_F$. Since Z is an

(LCI), it is also Gorenstein, so by [Aizenbud and Avni 2016, Corollary 3.15], it is enough to prove that $\int_{X^S\cap Z(F)} f|\omega_Z|_F < \infty$ for any such k'/k and F.

Fix some embedding of X into an affine space, and let d be the metric on $X(F)$ induced from the valuation metric. Define a function $h_\epsilon : X(F) \to \mathbb{R}$ by $h_\epsilon(x') = 1$ if $d(x', (X^S(F))^C) \geq \epsilon$ and $h_\epsilon(x') = 0$ otherwise. Notice that h_ϵ is smooth, and $f \cdot h_\epsilon$ is a Schwartz function whose support lies in $X^S(F)$.

Using Proposition 2.7, we have $\varphi_*(f \cdot h_\epsilon|\omega_X|_F) = g_\epsilon|\omega_Y|_F$, where $g_\epsilon(\varphi(x)) = \int_{X^S\cap Z(F)} f \cdot h_\epsilon|\omega_Z|_F$. Note that f is nonnegative and $f \cdot h_\epsilon$ is monotonically increasing when $\epsilon \to 0$, and converges pointwise to f. By Lebesgue's monotone convergence theorem we have:

$$\int_{X^S\cap Z(F)} f|\omega_Z|_F = \lim_{\epsilon\to 0}\int_{X^S\cap Z(F)} f h_\epsilon|\omega_Z|_F = \lim_{\epsilon\to 0} g_\epsilon(\varphi(x)).$$

It is left to show that $g_\epsilon(\varphi(x))$ is bounded in ϵ and we are done. By our assumption, $\varphi_*(f \cdot |\omega_X|_F)$ is $\mathcal{N}$-bounded, so there exists $\delta > 0$ and $M > 0$ such that for all $N_i \in \mathcal{N}_\delta$,

$$\sup_i \frac{\varphi_*(f|\omega_X|_F)(N_i)}{|\omega_Y|_F(N_i)} < M.$$

Note that we used the fact that for small enough δ, $|\omega_Y|_F$ is just the normalized Haar measure up to homothety. Finally, we obtain:

$$\int_{X^S\cap Z(F)} f|\omega_Z|_F = \lim_{\epsilon\to 0} g_\epsilon(\varphi(x))= \lim_{\epsilon\to 0}\left(\lim_{i\to\infty} \frac{\varphi_*(f\cdot h_\epsilon|\omega_X|_F)(N_i)}{|\omega_Y|_F(N_i)}\right) \leq \left(\sup_i \frac{\varphi_*(f|\omega_X|_F)(N_i)}{|\omega_Y|_F(N_i)}\right) < M.$$

□

4. Proof of the main theorem

For any prime power $q = p^r$, we denote the unique unramified extension of $\mathbb{Q}_p$ of degree r by $\mathbb{Q}_q$, its ring of integers by $\mathbb{Z}_q$, and the maximal ideal of $\mathbb{Z}_q$ by $\mathfrak{m}_q$. Recall that for a finite type $\mathbb{Z}$-scheme X and a finite ring A, we have defined $h_X(A) := |X(A)|/|A|^{\dim X_{\mathbb{Q}}}$. In this section we prove the following slightly stronger version of Theorem 1.4:

Theorem 4.1. *Let X be a scheme of finite type over $\mathbb{Z}$ such that $X_{\mathbb{Q}}$ is equidimensional and a local complete intersection. Then the following conditions are equivalent*:

(i) *For any $n \in \mathbb{N}$, $\lim_{p\to\infty} h_X(\mathbb{Z}/p^n\mathbb{Z}) = 1$.*

(ii) *There is a finite set S of prime numbers and a constant C, such that $|h_X(\mathbb{Z}/p^n\mathbb{Z}) - 1| < Cp^{-\frac{1}{2}}$ for any prime $p \notin S$ and any $n \in \mathbb{N}$.*

(iii) *$X_{\overline{\mathbb{Q}}}$ is reduced, irreducible and has rational singularities.*

(iv) *$X_{\overline{\mathbb{Q}}}$ is irreducible and there exists $C > 0$ such that $h_X(\mathbb{Z}/p^n\mathbb{Z}) < C$ for any prime p and $n \in \mathbb{N}$.*

(iv′) *$X_{\overline{\mathbb{Q}}}$ is irreducible and for any prime power q, the sequence $n \mapsto h_X(\mathbb{Z}_q/\mathfrak{m}_q^n)$ is bounded.*

(v) *$X_{\overline{\mathbb{Q}}}$ is irreducible and there exists a finite set S of primes, such that for any $p \notin S$, the sequence $n \mapsto h_X(\mathbb{Z}/p^n\mathbb{Z})$ is bounded.*

Moreover, conditions (iii), (iv), (iv′) *and* (v) *are equivalent without demanding that $X_{\overline{\mathbb{Q}}}$ is irreducible.*

We divide the proof of the theorem into two main parts that correspond to the implications (v) ⇒ (iii) (Section 4A) and (iii) ⇒ (iv′) (Section 4B). Theorem 4.1 can then be deduced as follows; the equivalence of conditions (i), (ii) and (iii) was proved in [Aizenbud and Avni 2018, Theorem 3.0.3] (see Theorem 1.1). The implications (ii) $\Rightarrow$ (v) and (iv$'$) $\Rightarrow$ (v) are trivial, so it follows that conditions (i), (ii), (iii), (iv$'$) and (v) are equivalent. The implication (iv) $\Rightarrow$ (v) is also trivial. Finally, (iv) follows from the rest of the conditions by first setting $q = p$ in (iv$'$) and getting that $\{h_X(\mathbb{Z}/p^n\mathbb{Z})\}_{n\in\mathbb{N}}$ is bounded for any prime p, and then by using (ii) to obtain a bound on $\{h_X(\mathbb{Z}/p^n\mathbb{Z})\}_{n\in\mathbb{N}}$ which is uniform over all primes p.

Lemma 4.2 [Aizenbud and Avni 2018, Lemma 3.1.1]. *Let $X = U_1 \cup U_2$ be an open cover of a scheme. Then for any finite local ring A, we have*:

(1) $|X(A)| = |U_1(A)| + |U_2(A)| - |U_1 \cap U_2(A)|$.

(2) $|X(A)| \geq |U_1(A)|$.

The following lemma is a consequence of Chebotarev's density theorem and Hensel's lemma.

Lemma 4.3 [Glazer and Hendel 2018, Lemma 3.15]. *Let X be a finite type $\mathbb{Z}$-scheme and let $x \in X(\overline{\mathbb{Q}})$. Then*:

(1) *There exists a finite extension k of $\mathbb{Q}$, such that $x \in X(k)$.*

(2) *For any finite extension $k/\mathbb{Q}$ as in* (1), *there exist infinitely many primes p with $i_p : k \hookrightarrow \mathbb{Q}_p$ such that $i_{p*}(x) \in X(\mathbb{Z}_p)$, where $i_{p*} : X(k) \hookrightarrow X(\mathbb{Q}_p)$.*

4A. *Boundedness implies rational singularities.*

Theorem 4.4. *Let X be a finite type $\mathbb{Z}$-scheme such that $X_{\mathbb{Q}}$ is a local complete intersection. Assume that there exists a finite set of primes S, such that for any $p \notin S$, the sequence $n \mapsto h_X(\mathbb{Z}/p^n\mathbb{Z})$ is bounded. Then $X_{\overline{\mathbb{Q}}}$ is reduced and has rational singularities.*

Proof. Step 1: Reduction to the case when $X_{\mathbb{Q}}$ is a complete intersection in an affine space (CIA).

Let $\bigcup_{i=1}^{l} \overline{X}_i$ be an affine cover of $X_{\mathbb{Q}}$, with each $\overline{X}_i$ a (CIA). For any i, there is a finite set S_i of primes, such that $\overline{X}_i$ is defined over $\mathbb{Z}[S_i^{-1}]$ and thus it has a finite type $\mathbb{Z}$-model, denoted X_i. By Lemma 4.2, for each $p \notin S_i$ we have $|X_i(\mathbb{Z}/p^n\mathbb{Z})| \leq |X(\mathbb{Z}/p^n\mathbb{Z})|$ and thus $n \mapsto h_{X_i}(\mathbb{Z}/p^n\mathbb{Z})$ is bounded for each $p \notin S_i \cup S$. By our assumption, this implies that each $(X_i)_{\overline{\mathbb{Q}}}$ is reduced and has rational singularities, and thus also $X_{\overline{\mathbb{Q}}}$.

Step 2: Proof for the case when $X_{\mathbb{Q}}$ is a (CIA).

By Proposition 2.3 we have an inclusion $\overline{\psi} : X_{\mathbb{Q}} \hookrightarrow \mathbb{A}_{\mathbb{Q}}^M$ and a morphism $\overline{\varphi} : \mathbb{A}_{\mathbb{Q}}^M \to \mathbb{A}_{\mathbb{Q}}^N$, flat over 0, such that $\overline{\psi} : X_{\mathbb{Q}} \simeq \overline{\varphi}^{-1}(0)$. As in Step 1, there exists a set S_1 of primes, and morphisms $\varphi : \mathbb{A}^M_{\mathbb{Z}[S_1^{-1}]} \to \mathbb{A}^N_{\mathbb{Z}[S_1^{-1}]}$ and $\psi : X_{\mathbb{Z}[S_1^{-1}]} \hookrightarrow \mathbb{A}^M_{\mathbb{Z}[S_1^{-1}]}$, such that $\varphi_{\mathbb{Q}} = \overline{\varphi}$, $\psi_{\mathbb{Q}} = \overline{\psi}$, φ is flat over 0, and $\psi : X_{\mathbb{Z}[S_1^{-1}]} \simeq \varphi^{-1}(0)$.

It is enough to prove that for any finite extension $k/\mathbb{Q}$ and any $y \in (\varphi^{-1}(0))(k)$, the map $\varphi_k : \mathbb{A}_k^M \to \mathbb{A}_k^N$ is (FRS) at y.

Fix $y \in (\varphi^{-1}(0))(k)$ and let k' be a finite extension of k. By Lemma 4.3, there exists an infinite set of primes T such that for any $p \in T$ we have an inclusion $i_p : k' \hookrightarrow \mathbb{Q}_p$ and $i_{p*}(y) \in \mathbb{Z}_p^M$. Choose $p \in T \setminus (S \cup S_1)$ and consider the local basis of balls $\{p^n\mathbb{Z}_p^N\}_n$ at 0, which clearly has bounded eccentricity. Let μ be the normalized Haar measure on $\mathbb{Z}_p^M$ and notice that μ does not vanish at y. By Theorem 1.6, in order to prove that $\varphi_k : \mathbb{A}_k^M \to \mathbb{A}_k^N$ is (FRS) at y it is enough to show that the sequence

$$n \mapsto \frac{((\varphi_{\mathbb{Z}_p})_*\mu)(p^n\mathbb{Z}_p^N)}{\lambda(p^n\mathbb{Z}_p^N)}$$

is bounded (for any k' and p as above), where λ is the normalized Haar measure on $\mathbb{Q}_p^N$. Consider $\pi_{N,n}:\mathbb{Z}_p^N\to(\mathbb{Z}/p^n\mathbb{Z})^N$ and notice that the following diagram is commutative:

$$\begin{tikzcd} \mathbb{Z}_p^M \arrow[r, "\varphi_{\mathbb{Z}_p}"] \arrow[d, "\pi_{M,n}"'] & \mathbb{Z}_p^N \arrow[d, "\pi_{N,n}"] \\ (\mathbb{Z}/p^n\mathbb{Z})^M \arrow[r, "\varphi_{\mathbb{Z}/p^n}"] & (\mathbb{Z}/p^n\mathbb{Z})^N \end{tikzcd}$$

Therefore we have

$$\mu(\varphi_{\mathbb{Z}_p}^{-1}(p^n\mathbb{Z}_p^N))=\mu(\varphi_{\mathbb{Z}_p}^{-1}\circ\pi_{N,n}^{-1}(0))=\mu(\pi_{M,n}^{-1}\circ\varphi_{\mathbb{Z}/p^n}^{-1}(0))=p^{-Mn}\cdot|\varphi_{\mathbb{Z}/p^n}^{-1}(0)|=p^{-Mn}\cdot|X(\mathbb{Z}/p^n\mathbb{Z})|,$$

and hence

$$\frac{((\varphi_{\mathbb{Z}_p})_*\mu)(p^n\mathbb{Z}_p^N)}{\lambda(p^n\mathbb{Z}_p^N)}=\frac{|X(\mathbb{Z}/p^n\mathbb{Z})|}{p^{(M-N)\cdot n}}=h_X(\mathbb{Z}/p^n\mathbb{Z})$$

is bounded and we are done. □

4B. ***Rational singularities implies boundedness.*** In the last section we proved the implication (v) ⇒ (iii) of Theorem 4.1. In this subsection we prove that (iii) implies (iv′). We divide the proof into three cases:

(1) X is a (CIA).

(2) $X_{\mathbb{Q}}$ is a (CIA).

(3) $X_{\mathbb{Q}}$ is an (LCI).

4B1. *Proof for the case that X is a* (CIA).

Proposition 4.5. *If X is a* (CIA), *then* (iii) ⇒ (iv′).

Proof. By Proposition 2.4, there exists an inclusion $X\hookrightarrow\mathbb{A}_{\mathbb{Z}}^M$ and a morphism $\varphi:\mathbb{A}_{\mathbb{Z}}^M\to\mathbb{A}_{\mathbb{Z}}^N$, flat over 0, such that $X\simeq\varphi^{-1}(0)$. Consider $\varphi_{\mathbb{Q}}:\mathbb{A}_{\mathbb{Q}}^M\to\mathbb{A}_{\mathbb{Q}}^N$ and notice that $\varphi_{\mathbb{Q}}$ is (FRS) at any $x\in\varphi_{\mathbb{Q}}^{-1}(0)(\overline{\mathbb{Q}})$, as $X_{\overline{\mathbb{Q}}}$ has rational singularities.

Let μ be the normalized Haar measure on $\mathbb{Z}_q^M$. As in the proof of Step 2 of Theorem 4.4, we have the following commutative diagram:

$$\begin{tikzcd} \mathbb{Z}_q^M \arrow[r, "\varphi_{\mathbb{Z}_q}"] \arrow[d, "\pi_{n,M}"'] & \mathbb{Z}_q^N \arrow[d, "\pi_{n,N}"] \\ (\mathbb{Z}_q/\mathfrak{m}_q^n)^M \arrow[r, "\varphi_{\mathbb{Z}_q/\mathfrak{m}_q^n}"] & (\mathbb{Z}_q/\mathfrak{m}_q^n)^N \end{tikzcd}$$

In order to show that $h_X(\mathbb{Z}_q/\mathfrak{m}_q^n)$ is bounded, it is enough to show that $(\varphi_{\mathbb{Z}_q})_*\mu$ has bounded density with respect to the local basis $\{p^n\mathbb{Z}_q^N\}_n$.

After base change to $\mathbb{Q}_q$, we have a map $\varphi_{\mathbb{Q}_q}:\mathbb{A}_{\mathbb{Q}_q}^M\to\mathbb{A}_{\mathbb{Q}_q}^N$, which is (FRS) at any point $x\in X(\mathbb{Q}_q)$.

For any $t\in\mathbb{N}$, consider the set $U_t=\varphi_{\mathbb{Z}_q}^{-1}(p^t\mathbb{Z}_q^N)$ and note that it is open, closed and compact. We claim that there exists $R\in\mathbb{N}$, such that for any $t>R$ we have that φ is (FRS) at any point $y\in U_t$. Indeed, otherwise we may construct a sequence $x_t\in U_t$ such that φ is not (FRS) at x_t. By a theorem of Elkik [1978] (see also [Aizenbud and Avni 2016, Theorem 6.3]), the (FRS) locus of φ is an open set. After choosing a convergent subsequence $\{x_{t_j}\}$, we obtain that $\varphi_{\mathbb{Q}_q}$ is not (FRS) at the limit $x_0 \in \mathbb{Z}_q^M$. But $\varphi_{\mathbb{Q}_q}(x_0) \in \bigcap_t \varphi_{\mathbb{Q}_q}(U_t) = \{0\}$ so $x_0 \in X(\mathbb{Q}_q)$ and we get a contradiction.

Finally, by Theorem 1.3, the measure $(\varphi_{\mathbb{Z}_q})_*\mu|_{U_R}$ has continuous density, and in particular bounded with respect to the local basis $\{p^n\mathbb{Z}_q^N\}_n$. Hence, from the definition of U_R, we have for $n > R$:

$$h_X(\mathbb{Z}_q/\mathfrak{m}_q^n) = \frac{(\varphi_{\mathbb{Z}_q})_*\mu(p^n\mathbb{Z}_q^N)}{q^{-nN}} = \frac{(\varphi_{\mathbb{Z}_q})_*\mu|_{U_R}(p^n\mathbb{Z}_q^N)}{q^{-nN}} < C,$$

for some constant $C > 0$ and we are done. □

4B2. *Some constructions.* Let X be an affine $\mathbb{Z}$-scheme with a coordinate ring

$$\mathbb{Z}[X] := \mathbb{Z}[x_1, \dots, x_c]/(f_1, \dots, f_m),$$

and fix $K \in \mathbb{N}$.

(1) For any $g \in \mathbb{Z}[x_1, \dots, x_c]$ denote by $g_K \in \mathbb{Q}[x_1, \dots, x_c]$ the function $g_K(x_1, \dots, x_c) := g\left(\frac{x_1}{K}, \dots, \frac{x_l}{K}\right)$.

(2) For any $\varphi : \mathbb{A}_{\mathbb{Z}}^M \to \mathbb{A}_{\mathbb{Z}}^N$ of the form $\varphi = (\varphi_1, \dots, \varphi_N)$, we denote by $\varphi_K : \mathbb{A}_{\mathbb{Q}}^M \to \mathbb{A}_{\mathbb{Q}}^N$ the morphism $\varphi_K := ((\varphi_1)_K, \dots, (\varphi_N)_K)$.

(3) Let $r(K) \in \mathbb{N}$ be minimal such that $K^{r(K)}(f_i)_K$ has integer coefficients for any i. Denote by $\widetilde{X}_K$ the $\mathbb{Z}$-scheme with the following coordinate ring:

$$\mathbb{Z}[\widetilde{X}_K] := \mathbb{Z}[x_1, \dots, x_c]/\left(K^{r(K)}(f_1)_K, \dots, K^{r(K)}(f_m)_K\right).$$

(4) For any $\mathbb{Q}$-morphism $\psi : X_{\mathbb{Q}} \to \mathbb{A}_{\mathbb{Q}}^M$ of the form $\psi = (\psi_1, \dots, \psi_N)$ let $K\psi$ denote $(K \cdot \psi_1, \dots, K \cdot \psi_N)$.

(5) For any affine $\mathbb{Q}$-scheme Z, with $\mathbb{Q}[Z] = \mathbb{Q}[y_1, \dots, y_d]/(g_1, \dots, g_k)$ and a $\mathbb{Q}$-morphism $\phi : Z \to X_{\mathbb{Q}}$, we may define a morphism $K\phi : Z \to (\widetilde{X}_K)_{\mathbb{Q}}$ by $K\phi(y_1, \dots, y_d) := K \cdot \phi(y_1, \dots, y_d)$.

4B3. *Proof for the case that $X_{\mathbb{Q}}$ is a* (CIA). In this case, we have an inclusion $\psi : X_{\mathbb{Q}} \hookrightarrow \mathbb{A}_{\mathbb{Q}}^M$ and a morphism $\varphi : \mathbb{A}_{\mathbb{Q}}^M \to \mathbb{A}_{\mathbb{Q}}^N$, flat over 0, such that $X_{\mathbb{Q}} \simeq \varphi^{-1}(0)$.

Lemma 4.6. *Let X be a finite type $\mathbb{Z}$-scheme, such that $X_{\mathbb{Q}}$ is a* (CIA), *defined by the morphisms φ, ψ as above. Then there exists a $\mathbb{Z}$-scheme $\widehat{X}_{\varphi,\psi}$, which is a* (CIA), *and a $\mathbb{Z}$-morphism $\phi : X \to \widehat{X}_{\varphi,\psi}$, such that $\phi_{\mathbb{Q}}$ is an isomorphism.*

Proof. Let $\mathbb{Z}[X] := \mathbb{Z}[x_1, \dots, x_c]/(f_1, \dots, f_m)$ be the coordinate ring of X. Denote by $S = \{p_1, \dots, p_s\}$ the set of all prime numbers that appear in the denominators of the polynomial maps ψ and φ, and set $P' := \prod_{p_i \in S} p_i$. Let $t \in \mathbb{N}$ be minimal such that $(P')^t\psi$ has integer coefficients. Denote $P := (P')^t$ and notice that $P\psi$ is a $\mathbb{Z}$-morphism. Let $\varphi_P : \mathbb{A}_{\mathbb{Q}}^M \to \mathbb{A}_{\mathbb{Q}}^N$ be as defined in 4B2. Notice that there exists $m \in \mathbb{N}$ such that $P^m\varphi_P$ has coefficients in $\mathbb{Z}$. We now have the following $\mathbb{Z}$-morphisms:

$$X \xrightarrow{P\psi} \mathbb{A}_{\mathbb{Z}}^M \xrightarrow{P^m\varphi_P} \mathbb{A}_{\mathbb{Z}}^N.$$

Set $\widehat{X}_{\varphi,\psi}$ to be the fiber $(P^m\varphi_P)^{-1}(0)$ and notice that $\phi := P\psi$ is a $\mathbb{Z}$-morphism from X to $\widehat{X}_{\varphi,\psi}$, such that $\phi_{\mathbb{Q}}$ is an isomorphism, and $\widehat{X}_{\varphi,\psi}$ is a (CIA). □

Lemma 4.7. *Let X and Y be affine $\mathbb{Z}$-schemes and $\phi : X \to Y$ be a $\mathbb{Z}$-morphism, such that $\phi_{\mathbb{Q}}$ is an isomorphism. Then there exist $c, N \in \mathbb{N}$, such that for any prime power q and any n*:

$$|X(\mathbb{Z}_q/\mathfrak{m}_q^n)| \leq q^{N \cdot c} \cdot |Y(\mathbb{Z}_q/\mathfrak{m}_q^n)|.$$

Proof. The morphism ϕ induces a map $\phi_n : X(\mathbb{Z}_q/\mathfrak{m}_q^n) \to Y(\mathbb{Z}_q/\mathfrak{m}_q^n)$. It is enough to show that ϕ_n has fibers of size at most $q^{N\cdot c}$. Assume that $\mathbb{Z}[X] = \mathbb{Z}[x_1, \dots, x_c]/(f_1, \dots, f_m)$. As in Section 4B2, we may choose $K, r(K) \in \mathbb{N}$ such that $\widetilde{X}_K$ is a $\mathbb{Z}$-scheme with a coordinate ring

$$\mathbb{Z}[\widetilde{X}_K] := \mathbb{Z}[x_1, \dots, x_c]/\big(K^{r(K)}(f_1)_K, \dots, K^{r(K)}(f_m)_K\big),$$

and $K\phi^{-1} : Y \to \widetilde{X}_K$ is a $\mathbb{Z}$-morphism. The map $(K\phi^{-1} \circ \phi) : X \to \widetilde{X}_K$ is just coordinatewise multiplication by K. Thus $(K\phi^{-1})_n \circ \phi_n : X(\mathbb{Z}_q/\mathfrak{m}_q^n) \to \widetilde{X}_K(\mathbb{Z}_q/\mathfrak{m}_q^n)$ sends $(a_1, \dots, a_c) \in X(\mathbb{Z}_q/\mathfrak{m}_q^n)$ to $(Ka_1, \dots, Ka_c) \in \widetilde{X}_K(\mathbb{Z}_q/\mathfrak{m}_q^n)$.

For any prime p, let $N(p)$ be the maximal integer such that $p^{N(p)}|K$. Note that the map $(a_1, \dots, a_n) \mapsto (Ka_1, \dots, Ka_n)$ from $(\mathbb{Z}_q/\mathfrak{m}_q^n)^c$ to $(\mathbb{Z}_q/\mathfrak{m}_q^n)^c$ has fibers of size $q^{N(p)\cdot c}$ for $n > N(p)$. Indeed, for $(b_1, \dots, b_c) \in (\mathbb{Z}_q/\mathfrak{m}_q^n)^c$, $(Ka_1, \dots, Ka_c) = (b_1, \dots, b_c)$ if and only if $Ka_i = b_i$ for any $1 \le i \le c$. Since $K/p^{N(p)}$ is invertible in $\mathbb{Z}_q/\mathfrak{m}_q^n$, it is equivalent to demand that $p^{N(p)}a_i = c_i$ for some multiple c_i of b_i by an invertible element. Hence, we can reduce to the case of the map $(a_1, \dots, a_c) \mapsto (p^{N(p)}a_1, \dots, p^{N(p)}a_c)$, which clearly has fibers of size $q^{N(p)\cdot c}$ for $n > N(p)$. Note that for any $y \in Y(\mathbb{Z}_q/\mathfrak{m}_q^n)$ we have $|\phi_n^{-1}(y)| \le \big|((K\phi^{-1})_n \circ \phi_n)^{-1}(x)\big|$, where $x = (K\phi^{-1})_n(y)$. Since the fibers of $(K\phi^{-1})_n \circ \phi_n$ are of size bounded by $q^{N(p)c}$, so are the fibers of ϕ_n. We may take $N := K > N(p)$ and we are done. $\square$

Corollary 4.8. *Let X be a finite type $\mathbb{Z}$-scheme such that $X_{\mathbb{Q}}$ is a* (CIA). *Then condition* (iii) *of Theorem 4.1 implies condition* (iv′).

Proof. By Lemma 4.6, we may choose a $\mathbb{Z}$-scheme $\widehat{X}$, which is a (CIA), and a $\mathbb{Z}$-morphism $\phi : X \to \widehat{X}$, such that $\phi_{\mathbb{Q}}$ is an isomorphism. By Proposition 4.5 and Lemma 4.7, there exists $c, N \in \mathbb{N}$, such that for any prime power q, there exists $C > 0$ such that:

$$h_X(\mathbb{Z}_q/\mathfrak{m}_q^n) = \frac{|X(\mathbb{Z}_q/\mathfrak{m}_q^n)|}{q^{n \dim X_{\mathbb{Q}}}} \le q^{c\cdot N} \cdot \frac{|\widehat{X}(\mathbb{Z}_q/\mathfrak{m}_q^n)|}{q^{n \dim X_{\mathbb{Q}}}} \le q^{c\cdot N} \cdot C,$$

and hence condition (iv′) holds. $\square$

4B4. *Proof for the case when $X_{\mathbb{Q}}$ is an* (LCI). Using Lemma 4.2, we may reduce to the case when X is affine, with coordinate ring $\mathbb{Z}[X] := \mathbb{Z}[x_1, \dots, x_c]/(f_1, \dots, f_m)$. Since $X_{\mathbb{Q}}$ is an (LCI), we have an affine open cover $\{\beta_i : U_i \hookrightarrow X_{\mathbb{Q}}\}_i$ of $X_{\mathbb{Q}}$ with inclusions $\psi_i : U_i \hookrightarrow \mathbb{A}_{\mathbb{Q}}^{M_i}$ and maps $\varphi_i : \mathbb{A}_{\mathbb{Q}}^{M_i} \to \mathbb{A}_{\mathbb{Q}}^{N_i}$, flat over 0, such that $\psi_i : U_i \simeq \varphi_i^{-1}(0)$. We may assume that U_i is isomorphic to a basic open set $D(g_i)$ for $g_i \in \mathbb{Q}[X]$ and $\beta_i^* : \mathbb{Q}[X] \to \mathbb{Q}[X, t]/(g_i t - 1)$ is the natural map. Since $\{D(g_i)\}_i$ is a cover of $X_{\mathbb{Q}}$, there exist $c_i' \in \mathbb{Z}[X]$ and $d_i \in \mathbb{Z}$ such that $\sum c_i' \cdot g_i/d_i = 1$. Thus, by multiplying by all the d_i's, we obtain $\sum c_i g_i = D$ for some $c_i \in \mathbb{Z}[X]$ and $D \in \mathbb{Z}$. Choose large enough $P \in \mathbb{N}$ such that the following algebra

$$\mathbb{Z}[x_1, \dots, x_c, t]/(f_1, \dots, f_m, Pg_i t - D \cdot P)$$

is a coordinate ring of a $\mathbb{Z}$-scheme $\widetilde{U}_i$, for any i. Moreover, notice that $\widetilde{U}_i \simeq U_i$ over $\mathbb{Q}$.

Lemma 4.9. *There exists $N \in \mathbb{N}$, such that for any prime power $q = p^r$ and any $n > N$ we have*

$$|X(\mathbb{Z}_q/\mathfrak{m}_q^n)| \le \sum_i |\widetilde{U}_i(\mathbb{Z}_q/\mathfrak{m}_q^n)|.$$

Proof. Let $N(p)$ be the maximal integer such that $p^{N(p)} \mid D \cdot P$. We first claim that for any $n > N(p) + 1$ and $(a_1, \dots, a_c) \in X(\mathbb{Z}_q/\mathfrak{m}_q^n)$, there exists some i such that $Pg_i(a_1, \dots, a_c) \notin \mathfrak{m}_q^{N(p)+1}/\mathfrak{m}_q^n$. Indeed, if

$Pg_i(a_1,\dots,a_c)\in\mathfrak{m}_q^{N(p)+1}/\mathfrak{m}_q^n$ for any i, then $\sum Pg_i(a_1,\dots,a_c)\cdot c_i(a_1,\dots,a_c)=D\cdot P\in\mathfrak{m}_q^{N(p)+1}/\mathfrak{m}_q^n$ and hence $p^{N(p)+1}\mid D\cdot P$ leading to a contradiction. Set $N:=D\cdot P+1$ and notice that $N>N(p)+1$ for any prime p. Fix $n>N$ and let i such that $Pg_i(a_1,\dots,a_c)\notin\mathfrak{m}_q^{N(p)+1}/\mathfrak{m}_q^n$. We now claim that the equation $Pg_i(a_1,\dots,a_c)t-PD=0$ has a solution in $\mathbb{Z}_q/\mathfrak{m}_q^n$. Indeed, if $Pg_i(a_1,\dots,a_c)$ is invertible in $\mathbb{Z}_q/\mathfrak{m}_q^n$, we are done. Otherwise, we have that $Pg_i(a_1,\dots,a_c)=p^l\cdot b\in\mathfrak{m}_q^l/\mathfrak{m}_q^n$ for some $l\le N(p)$, where b is invertible. Write $PD=p^l\cdot a$. We can rewrite the equation as $p^l\cdot(bt-a)=0$, which has a solution $d\in\mathbb{Z}_q/\mathfrak{m}_q^n$ since b is invertible. We see that for any $n>N$ and any $(a_1,\dots,a_c)\in X(\mathbb{Z}_q/\mathfrak{m}_q^n)$ there exists i and $d\in\mathbb{Z}_q/\mathfrak{m}_q^n$ such that $(a_1,\dots,a_c,d)\in\widetilde{U}_i(\mathbb{Z}_q/\mathfrak{m}_q^n)$. This implies the lemma. ☐

Since $(\widetilde{U}_i)_{\mathbb{Q}}\simeq U_i$ is a (CIA) for any i, we obtain

$$h_X(\mathbb{Z}_q/\mathfrak{m}_q^n)=q^{-n\dim X_{\mathbb{Q}}}\cdot|X(\mathbb{Z}_q/\mathfrak{m}_q^n)|\le\sum_i q^{-n\dim X_{\mathbb{Q}}}\cdot|\widetilde{U}_i(\mathbb{Z}_q/\mathfrak{m}_q^n)|<\sum C_i,$$

where $C_i=\sup_n h_{\widetilde{U}_i}(\mathbb{Z}_q/\mathfrak{m}_q^n)$. The implication (iii) $\Rightarrow$ (iv′) of Theorem 4.1 now follows.

Acknowledgements

I would like to thank my advisor Avraham Aizenbud for presenting me with this problem, teaching and helping me in this work. I hold many thanks to Nir Avni for helpful discussions and for hosting me at Northwestern University in July 2016, during which a large part of this work was done. I also thank Yotam Hendel for fruitful talks. This work was partially supported by the ISF grant [687/13], the BSF grant [2012247] and the Minerva Foundation grant.

References

[Aizenbud and Avni 2016] A. Aizenbud and N. Avni, "Representation growth and rational singularities of the moduli space of local systems", *Invent. Math.* **204**:1 (2016), 245–316. MR Zbl

[Aizenbud and Avni 2018] A. Aizenbud and N. Avni, "Counting points of schemes over finite rings and counting representations of arithmetic lattices", *Duke Math. J.* **167**:14 (2018), 2721–2743. MR Zbl

[Deligne 1974] P. Deligne, "La conjecture de Weil, I", *Inst. Hautes Études Sci. Publ. Math.* **43** (1974), 273–307. MR Zbl

[Deligne 1980] P. Deligne, "La conjecture de Weil, II", *Inst. Hautes Études Sci. Publ. Math.* **52** (1980), 137–252. MR Zbl

[Denef 1991] J. Denef, "Report on Igusa's local zeta function", exposé 741, pp. 359–386 in *Séminaire Bourbaki,* 1990/1991, Astérisque **201–203**, 1991. MR Zbl

[Dwork 1960] B. Dwork, "On the rationality of the zeta function of an algebraic variety", *Amer. J. Math.* **82** (1960), 631–648. MR Zbl

[Elkik 1978] R. Elkik, "Singularités rationnelles et déformations", *Invent. Math.* **47**:2 (1978), 139–147. MR Zbl

[Glazer and Hendel 2018] I. Glazer and Y. I. Hendel, "On singularity properties of convolutions of algebraic morphisms", preprint, 2018. to appear in *Selecta Mathematica*. arXiv

[Grothendieck 1965] A. Grothendieck, "Formule de Lefschetz et rationalité des fonctions L", exposé in *Séminaire Bourbaki,* 1964/1965, W. A. Benjamin, Amsterdam, 1965. Reprinted as pp. 41–55 in *Séminaire Bourbaki* **9**, Soc. Math. France, Paris, 1995. MR

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, 1977. MR Zbl

[Hironaka 1964] H. Hironaka, "Resolution of singularities of an algebraic variety over a field of characteristic zero, I, II", *Ann. of Math.* (2) **79** (1964), 109–326. MR Zbl

[Igusa 2000] J.-i. Igusa, *An introduction to the theory of local zeta functions*, AMS/IP Studies in Advanced Mathematics **14**, American Mathematical Society, Providence, RI, 2000. MR Zbl

[Kempf et al. 1973] G. Kempf, F. F. Knudsen, D. Mumford, and B. Saint-Donat, *Toroidal embeddings, I*, Lecture Notes in Mathematics **339**, Springer, 1973. MR Zbl

[Lang and Weil 1954] S. Lang and A. Weil, "Number of points of varieties in finite fields", *Amer. J. Math.* **76** (1954), 819–827. MR Zbl

[Mustaţă 2011] M. Mustaţă, "Zeta functions in algebraic geometry", lecture notes, University of Michigan, 2011, Available at http://www.math.lsa.umich.edu/~mmustata/zeta_book.pdf.

[du Sautoy and Grunewald 2000] M. du Sautoy and F. Grunewald, "Analytic properties of zeta functions and subgroup growth", *Ann. of Math.* (2) **152**:3 (2000), 793–833. MR Zbl

[Vakil 2017] R. Vakil, "The rising sea: foundations of algebraic geometry", preprint, 2017, Available at http://math.stanford.edu/~vakil/216blog/FOAGnov1817public.pdf.

Communicated by Hélène Esnault
Received 2018-03-03 Revised 2018-08-27 Accepted 2018-12-24

itay.glazer@weizmann.ac.il *Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 7610001, Israel*

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.501

The Maillot–Rössler current and the polylogarithm on abelian schemes

Guido Kings and Danny Scarponi

We give a structural proof of the fact that the realization of the degree-zero part of the polylogarithm on abelian schemes in analytic Deligne cohomology can be described in terms of the Bismut–Köhler higher analytic torsion form of the Poincaré bundle. Furthermore, we provide a new axiomatic characterization of the arithmetic Chern character of the Poincaré bundle using only invariance properties under isogenies. For this we obtain a decomposition result for the arithmetic Chow group of independent interest.

Introduction

In an important contribution Maillot and Rössler constructed a Green current $\mathfrak{g}_{\mathscr{A}^\vee}$ for the zero section of an abelian scheme $\mathscr{A}$ which is norm compatible (i.e., $[n]_*\mathfrak{g}_{\mathscr{A}^\vee} = \mathfrak{g}_{\mathscr{A}^\vee}$) and is the push-forward of the arithmetic Chern character of the (canonically metrized) Poincaré bundle. In particular, on the complement of the zero section the Green current $\mathfrak{g}_{\mathscr{A}^\vee}$ is the degree-$(g-1)$ part of the analytic torsion form of the Poincaré bundle. Moreover, certain linear combinations of translates of these currents are even motivic in the sense that their classes in analytic Deligne cohomology are in the image of the regulator from motivic cohomology.

In the special case of a family of elliptic curves, the current $\mathfrak{g}_{\mathscr{A}^\vee}$ is described by a Siegel-function whose usefulness for many arithmetic problems (in particular for special values of L-functions and Iwasawa theory) is well known and one could hope that the Maillot–Rössler current plays a similar role for abelian schemes.

On the other hand the first author has constructed the motivic polylogarithm $\mathrm{pol}^0 \in H^{2g-1}_{\mathscr{M}}(\mathscr{A}\setminus\mathscr{A}[N], g)$ of the abelian scheme $\mathscr{A}$ without its N-torsion points $\mathscr{A}\setminus\mathscr{A}[N]$ [Kings and Rössler 2017]. Here g is the relative dimension of $\mathscr{A}$. The polylogarithm is also norm-compatible $[n]_*\,\mathrm{pol}^0 = \mathrm{pol}^0$ for n coprime to N, and in the elliptic case it is directly related to Siegel functions and modular units.

It is natural to ask how pol^0 is related to $\mathfrak{g}_{\mathscr{A}^\vee}$. This question was answered completely in [Kings and Rössler 2017], and it turns out that the image of $-2\,\mathrm{pol}^0$ in analytic Deligne cohomology is the Maillot–Rössler current $[N]^*\mathfrak{g}_{\mathscr{A}^\vee} - N^{2g}\mathfrak{g}_{\mathscr{A}^\vee}$. Due to the fact that in analytic Deligne cohomology there is no residue sequence, the proof of this fact in [Kings and Rössler 2017] was much more complicated than

MSC2010: primary 11G55; secondary 14G40.
Keywords: abelian polylogarithm, arithmetic Chow groups, Arakelov geometry, Deligne cohomology.

it should be and proceeded by a reduction to the case of a product of elliptic curves via the moduli space of abelian varieties and an explicit computation.

In this paper we give a much simpler and very structural proof of the identity between the polylogarithm and the Maillot–Rössler current (see Theorem 8). We circumvent the difficulties of the approach in [Kings and Rössler 2017] by working in Betti cohomology instead of analytic Deligne cohomology. As a result one only has to compare the residues of the classes. In fact we achieve much more and give an axiomatic characterization of the Maillot–Rössler current which does not involve the Poincaré bundle (see Theorem 11). More precisely, we prove that any class $\hat{\xi} \in \widehat{\mathrm{CH}}^g(\mathscr{A})_{\mathbb{Q}}$ in the arithmetic Chow group which satisfies that its image in the Chow group $\mathrm{CH}^g(\mathscr{A})_{\mathbb{Q}}$ is the zero section and such that

$$([n]^* - n^{2g})(\hat{\xi}) = 0 \quad \text{in } \widehat{\mathrm{CH}}^g(\mathscr{A})_{\mathbb{Q}}$$

holds for some $n \geq 2$ is in fact equal to $(-1)^g p_{1,*}(\widehat{\mathrm{ch}}(\overline{\mathscr{P}}))^{[g]}$. This characterization of the Maillot–Rössler current relies on a decomposition into generalized eigenspaces for the action of $[n]^*$ on the arithmetic groups $\widehat{\mathrm{CH}}^g(\mathscr{A})_{\mathbb{Q}}$, which might be of independent interest (see Corollary 10).

Here is a short synopsis of our paper. In Section 1 we give some background on motivic cohomology and arithmetic Chow groups. In Section 2 we review the polylogarithm and the Maillot–Rössler current. In Section 3 we carry out the comparison between the Maillot–Rössler current and the polylogarithm. In Section 4 we prove a decomposition of the arithmetic Chow group, and in Section 5 we give an axiomatic characterization of the Maillot–Rössler current.

1. Preliminaries on motivic cohomology, Arakelov theory, and Deligne cohomology

Motivic cohomology. Let $\pi : \mathscr{A} \to S$ be an abelian scheme of relative dimension g, let $\varepsilon : S \to \mathscr{A}$ be the zero section, let $N > 1$ be an integer, and let $\mathscr{A}[N]$ be the finite group scheme of N-torsion points. Here S is smooth over a subfield k of the complex numbers. We will write S_0 for the image of ε in $\mathscr{A}$. We denote by $\mathscr{A}^\vee$ the dual abelian scheme of $\mathscr{A}$ and by $\varepsilon^\vee$ its zero section.

C. Soulé [1985] and A. Beilinson [1985] defined motivic cohomology for any variety V over a field

$$H^i_{\mathscr{M}}(V, j) := \mathrm{Gr}^j_\gamma K_{2j-i}(V) \otimes \mathbb{Q}.$$

Remark. In this paper we work with the above rather old-fashioned definition of motivic cohomology for the compatibility with earlier references. This and the requirement that S is smooth over a base field k are not necessary. The latter condition does no harm as we are mainly interested in the arithmetic Chow groups and Deligne cohomology. For a much more general setting we refer to the paper [Huber and Kings 2018] where also the decomposition of motivic cohomology is considered in an up-to-date fashion.

For any integer $a > 1$ and any $W \subseteq \mathscr{A}$ open subscheme such that

$$j : [a]^{-1}(W) \hookrightarrow W$$

is an open immersion (here $[a]:\mathcal{A}\to\mathcal{A}$ is the a-multiplication on $\mathcal{A}$), the trace map with respect to a is defined as

$$\mathrm{tr}_{[a]}: H^{\cdot}_{\mathcal{M}}(W,*) \xrightarrow{j^*} H^{\cdot}_{\mathcal{M}}([a]^{-1}(W),*) \xrightarrow{[a]_*} H^{\cdot}_{\mathcal{M}}(W,*). \tag{1}$$

For any integer r we let

$$H^{\cdot}_{\mathcal{M}}(W,*)^{(r)} := \{\psi \in H^{\cdot}_{\mathcal{M}}(W,*) \mid (\mathrm{tr}_{[a]} - a^r \,\mathrm{Id})^k \psi = 0 \text{ for some } k \geq 1\}$$

be the generalized eigenspace of $\mathrm{tr}_{[a]}$ of degree (or weight) r. One can prove that there is a decomposition into $\mathrm{tr}_{[a]}$-eigenspaces

$$H^{\cdot}_{\mathcal{M}}(\mathcal{A},*) \cong \bigoplus_{r=0}^{2g} H^{\cdot}_{\mathcal{M}}(\mathcal{A},*)^{(r)}$$

which is independent of a and that

$$H^{\cdot}_{\mathcal{M}}(\mathcal{A}\setminus S_0,*)^{(0)} = 0$$

(see Proposition 2.2.1 in [Kings and Rössler 2017]).

Arithmetic varieties. An arithmetic ring is a triple (R, Σ, F_∞) where

- R is an excellent regular Noetherian integral domain,
- Σ is a finite nonempty set of monomorphisms $\sigma : R \to \mathbb{C}$,
- F_∞ is an antilinear involution of the $\mathbb{C}$−algebra $\mathbb{C}^\Sigma := \mathbb{C} \underbrace{\times \cdots \times}_{|\Sigma|} \mathbb{C}$, such that the diagram

$$\begin{array}{ccc} R & \xrightarrow{\delta} & \mathbb{C}^\Sigma \\ {\scriptstyle \mathrm{Id}}\downarrow & & \downarrow{\scriptstyle F_\infty} \\ R & \xrightarrow{\delta} & \mathbb{C}^\Sigma \end{array}$$

  commutes (here by δ we mean the natural map to the product induced by the family of maps Σ).

An arithmetic variety X over R is a scheme of finite type over R, which is flat, quasiprojective, and regular. As usual we write

$$X(\mathbb{C}) := \coprod_{\sigma\in\Sigma} (X \times_{R,\sigma} \mathbb{C})(\mathbb{C}).$$

Note that F_∞ induces an involution $F_\infty : X(\mathbb{C}) \to X(\mathbb{C})$.

Arithmetic Chow groups. Let $p \in \mathbb{N}$. We denote by

- $E^{p,p}(X_\mathbb{R})$ the $\mathbb{R}$-vector space of smooth real forms ω on $X(\mathbb{C})$ of type (p,p) such that $F_\infty^*\zeta = (-1)^p\omega$,
- $\widetilde{E}^{p,p}(X_\mathbb{R})$ the quotient $E^{p,p}(X_\mathbb{R})/(\mathrm{Im}\,\partial + \mathrm{Im}\,\bar{\partial})$,
- $D^{p,p}(X_\mathbb{R})$ the $\mathbb{R}$-vector space of real currents ζ on $X(\mathbb{C})$ of type (p,p) such that $F_\infty^*\zeta = (-1)^p\zeta$,
- $\widetilde{D}^{p,p}(X_\mathbb{R})$ the quotient $D^{p,p}(X_\mathbb{R})/(\mathrm{Im}\,\partial + \mathrm{Im}\,\bar{\partial})$.

If ω or ζ is a form in $E^{p,p}(X_{\mathbb{R}})$ or a current in $D^{p,p}(X_{\mathbb{R}})$, we write $\tilde{\omega}$ or $\tilde{\zeta}$ for its class in $\widetilde{E}^{p,p}(X_{\mathbb{R}})$ or $\widetilde{D}^{p,p}(X_{\mathbb{R}})$, respectively.

We briefly recall the definition of the arithmetic Chow groups of X, as given in [Gillet and Soulé 1990, §3.3]. Let $Z^q(X)$ denote the group of cycles of codimension q in X and $\mathrm{CH}^q(X)$ denote the q-th Chow group of X. We write $\widehat{Z}^q(X)$ for the subgroup of

$$Z^q(X) \oplus \widetilde{D}^{q-1,q-1}(X_{\mathbb{R}})$$

consisting of pairs $(z, \tilde{h})$ where $z \in Z^q(X)$ and $h \in D^{q-1,q-1}(X_{\mathbb{R}})$ satisfy

$$\mathrm{dd}^c\, h + \delta_z \in E^{q,q}(X_{\mathbb{R}}).$$

By definition, the class $\tilde{h}$ is then a Green current for z. Note that if $\tilde{h}$ is a Green current for z, the form $\mathrm{dd}^c\, h + \delta_z$ is closed.

For any codimension-$(q-1)$ integral subscheme $i : W \hookrightarrow X$ and any $f \in k(W)^*$, one can verify, by means of the Poincaré–Lelong lemma, that the pair

$$\widehat{\mathrm{div}}(f) := (\mathrm{div}(f), -i_* \log|f|^2)$$

is an element in $\widehat{Z}^q(X)$. Then the q-th arithmetic Chow group of X is the quotient

$$\widehat{\mathrm{CH}}^q(X) := \widehat{Z}^q(X)/\widehat{R}^q(X)$$

where $\widehat{R}^q(X)$ is the subgroup generated by all pairs $\widehat{\mathrm{div}}(f)$, for any $f \in k(W)^*$ and any $W \subset X$ as above. If $Z^{q,q}(X_{\mathbb{R}}) \subseteq E^{q,q}(X_{\mathbb{R}})$ denotes the subspace of closed forms, we have a well defined map

$$\omega : \widehat{\mathrm{CH}}^q(X) \to Z^{q,q}(X_{\mathbb{R}})$$

sending the class of $(z, \tilde{h})$ to $\mathrm{dd}^c\, h + \delta_z$. Finally we have a map

$$\zeta : \widehat{\mathrm{CH}}^q(X) \to \mathrm{CH}^q(X)$$

sending the class of $(z, \tilde{h})$ to the class of z.

Analytic Deligne cohomology of arithmetic varieties. If X is an arithmetic variety over R we write

$$\mathrm{H}^q_{\mathcal{D}^{\mathrm{an}}}(X_{\mathbb{R}}, \mathbb{R}(p)) := \{\gamma \in \mathrm{H}^q_{\mathcal{D}^{\mathrm{an}}}(X(\mathbb{C}), \mathbb{R}(p)) \mid F_\infty^* \gamma = (-1)^p \gamma\},$$

where $\mathrm{H}^*_{\mathcal{D}^{\mathrm{an}}}(X(\mathbb{C}), \mathbb{R}(p))$ is the analytic Deligne cohomology of the complex manifold $X(\mathbb{C})$, i.e., the hypercohomology of the complex

$$0 \to (2\pi i)^p \mathbb{R} \to \mathcal{O}_{X(\mathbb{C})} \xrightarrow{d} \Omega^1_{X(\mathbb{C})} \to \cdots \to \Omega^{p-1}_{X(\mathbb{C})} \to 0$$

($\Omega^*_{X(\mathbb{C})}$ denotes the de Rham complex of holomorphic forms on $X(\mathbb{C})$). In the following sections we will need the characterization [Burgos 1997, §2]

$$\mathrm{H}^{2p-1}_{\mathcal{D}^{\mathrm{an}}}(X_{\mathbb{R}}, \mathbb{R}(p)) = \{\tilde{x} \in (2\pi i)^{p-1} \widetilde{E}^{p-1,p-1}(X_{\mathbb{R}}) \mid \partial\bar{\partial} x = 0\}. \tag{2}$$

Analytic Deligne cohomology and Betti cohomology. By definition of analytic Deligne cohomology there is a canonical map to Betti cohomology

$$\phi_B : H^{2g-1}_{D^{\mathrm{an}}}((\mathscr{A} \setminus \mathscr{A}[N])_{\mathbb{R}}, \mathbb{R}(g)) \to H^{2g-1}_B((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), \mathbb{R}(g)).$$

Later we will need an explicit description of this map: first we compute the group $H^{2g-1}_B((\mathscr{A}\setminus\mathscr{A}[N])(\mathbb{C}), g)$ with the cohomology of the complex of currents $D^*((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), g) := (2\pi i)^g D^*((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}))$, so

$$H^{2g-1}_B((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), g) = \frac{\{\eta \in D^{2g-1}((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), g) \mid d\eta = 0\}}{\{d\omega \mid \omega \in D^{2g-2}((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), g)\}}.$$

Lemma 1. *Using the description (2), the map ϕ_B sends the class $\tilde{x}$ of $x \in (2\pi i)^{g-1} E^{g-1,g-1}((\mathscr{A}\setminus\mathscr{A}[N])_{\mathbb{R}})$ with $\partial\bar{\partial}x = 0$ to*

$$\phi_B(\tilde{x}) = [4\pi i \mathrm{d}^c x]. \tag{3}$$

Proof. This is [Burgos 1997, Theorem 2.6]. □

We also need an explicit description of the connecting homomorphism, which we call the residue homomorphism

$$\mathrm{res}_B : H^{2g-1}_B((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), \mathbb{R}(g)) \to H^{2g}_{B,\mathscr{A}[N]\setminus S_0}((\mathscr{A} \setminus S_0)(\mathbb{C}), \mathbb{R}(g)).$$

For this we compute $H^{2g}_{B,\mathscr{A}[N]\setminus S_0}((\mathscr{A} \setminus S_0)(\mathbb{C}), \mathbb{R}(g))$ with the cohomology of the simple complex of the restriction morphism $D^*((\mathscr{A} \setminus S_0)(\mathbb{C}), g) \to D^*((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), g)$ and get

$$\begin{aligned} &H^{2g}_{B,\mathscr{A}[N]\setminus S_0}((\mathscr{A} \setminus S_0)(\mathbb{C}), \mathbb{R}(g)) \\ &\quad = \frac{\{(\xi, \tau) \in D^{2g}((\mathscr{A} \setminus S_0)(\mathbb{C}), g) \oplus D^{2g-1}((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), g) \mid d\xi = 0 \text{ and } \xi|_{\mathscr{A}\setminus\mathscr{A}[N]} = d\tau\}}{\{(d\theta, \theta|_{\mathscr{A}\setminus\mathscr{A}[N]} - d\alpha) \mid \theta \in D^{2g-1}((\mathscr{A} \setminus S_0)(\mathbb{C}), g),\ \alpha \in D^{2g-2}((\mathscr{A} \setminus \mathscr{A}[N])(\mathbb{C}), g)\}}. \end{aligned}$$

Note that we are using the simple complex as in [Burgos 1997, §1] (and not the cone in the sense of Verdier) of the restriction morphism to compute cohomology with support. From the definitions one gets immediately:

Lemma 2. *The residue* res_B *sends the class of* η*, which we denote by* $[\eta]$*, to*

$$\mathrm{res}_B([\eta]) = [0, -\eta],$$

where $[0, -\eta]$ *denotes the class of* $(0, -\eta)$.

2. Review of the polylog and the Maillot–Rössler current

***The axiomatic definition of* $\mathbf{pol}^0$.** G. Kings and D. Rössler [2017] have provided a simple axiomatic description of the degree-zero part of the polylogarithm on abelian schemes. We briefly recall it here.

The degree-zero part of the motivic polylogarithm is by definition a class in motivic cohomology

$$\mathrm{pol}^0 \in H^{2g-1}_{\mathscr{M}}(\mathscr{A} \setminus \mathscr{A}[N], g)^{(0)}.$$

To describe it more precisely, consider the residue map along $\mathscr{A}[N]$

$$H^{2g-1}_{\mathscr{M}}(\mathscr{A} \setminus \mathscr{A}[N], g) \to H^{0}_{\mathscr{M}}(\mathscr{A}[N] \setminus S_0, 0).$$

This map induces an isomorphism

$$\mathrm{res} : H^{2g-1}_{\mathscr{M}}(\mathscr{A} \setminus \mathscr{A}[N], g)^{(0)} \cong H^{0}_{\mathscr{M}}(\mathscr{A}[N] \setminus S_0, 0)^{(0)}$$

(see Corollary 2.2.2 in [Kings and Rössler 2017]).

Definition 3. The degree-zero part of the polylog pol^0 is the unique element of $H^{2g-1}_{\mathscr{M}}(\mathscr{A} \setminus \mathscr{A}[N], g)^{(0)}$ mapping under res to the fundamental class 1°_N of $\mathscr{A}[N] \setminus S_0$.

We recall now that we have a map $\mathrm{reg}_{\mathrm{an}}$ defined as the composition

$$H^{2g-1}_{\mathscr{M}}(\mathscr{A} \setminus \mathscr{A}[N], g) \xrightarrow{\mathrm{reg}} \mathrm{H}^{2g-1}_{D}((\mathscr{A} \setminus \mathscr{A}[N])_{\mathbb{R}}, \mathbb{R}(g)) \xrightarrow{\mathrm{forget}} \mathrm{H}^{2g-1}_{D^{\mathrm{an}}}((\mathscr{A} \setminus \mathscr{A}[N])_{\mathbb{R}}, \mathbb{R}(g))$$

where reg is the regulator map into Deligne–Beilinson cohomology and the second map is the forgetful map from Deligne–Beilinson cohomology to analytic Deligne cohomology.

The Maillot–Rössler current $\mathfrak{g}_{\mathscr{A}^\vee}$. V. Maillot and D. Rössler [2015] proved the following theorem.

Theorem 4 [Maillot and Rössler 2015, Theorem 1.1]. *There exists a unique a class of currents $\mathfrak{g}_{\mathscr{A}^\vee} \in \widetilde{D}^{g-1,g-1}(\mathscr{A}_{\mathbb{R}})$ which satisfies the following three properties*:

(i) $\mathfrak{g}_{\mathscr{A}^\vee}$ *is a Green current for* S_0,

(ii) $(S_0, \mathfrak{g}_{\mathscr{A}^\vee}) = (-1)^g p_{1,*}(\widehat{\mathrm{ch}}(\overline{\mathscr{P}}))^{[g]}$ *in the group* $\widehat{\mathrm{CH}}^g(\mathscr{A})_{\mathbb{Q}}$,

(iii) $[n]_*\mathfrak{g}_{\mathscr{A}^\vee} = \mathfrak{g}_{\mathscr{A}^\vee}$ *for all* $n > 0$.

Here we take $\overline{\mathscr{P}}$ to be the Poincaré bundle on $\mathscr{A} \times_S \mathscr{A}^\vee$ equipped with a canonical hermitian metric, $p_1 : \mathscr{A} \times_S \mathscr{A}^\vee \to \mathscr{A}$ is the first projection, and $\widehat{\mathrm{CH}}^g(\mathscr{A})$ denotes the g-th arithmetic Chow group of $\mathscr{A}$. The term $\widehat{\mathrm{ch}}(\overline{\mathscr{P}}) \in \bigoplus_i \widehat{\mathrm{CH}}^i(\mathscr{A} \times_S \mathscr{A}^\vee)$ is the arithmetic Chern character of $\overline{\mathscr{P}}$, and $p_{1,*}(\widehat{\mathrm{ch}}(\overline{\mathscr{P}}))^{[g]}$ denotes the homogeneous component of degree g of $p_{1,*}(\widehat{\mathrm{ch}}(\overline{\mathscr{P}}))$ in the graded ring $\bigoplus_i \widehat{\mathrm{CH}}^i(\mathscr{A})$.

We now consider the arithmetic cycle

$$({}_N S_0, {}_N\, \mathfrak{g}_{\mathscr{A}^\vee}) := ([N]^* - N^{2g})(S_0, \mathfrak{g}_{\mathscr{A}^\vee}).$$

Thanks to the geometry of the Poincaré bundle, one can show that the class of $({}_N S_0, {}_N\, \mathfrak{g}_{\mathscr{A}^\vee})$ in $\widehat{\mathrm{CH}}^g(\mathscr{A})_{\mathbb{Q}}$ is zero [Scarponi 2017, Proposition 5.2]. In particular, $dd^c({}_N\mathfrak{g}_{\mathscr{A}^\vee}|_{\mathscr{A} \setminus \mathscr{A}[N]}) = 0$, and by Theorem 1.2.2(i) in [Gillet and Soulé 1990], there exists a smooth form in the class of currents ${}_N\mathfrak{g}_{\mathscr{A}^\vee}|_{\mathscr{A} \setminus \mathscr{A}[N]}$. Equivalently, ${}_N\mathfrak{g}_{\mathscr{A}^\vee}|_{\mathscr{A} \setminus \mathscr{A}[N]}$ lies in the image of the inclusion

$$\widetilde{E}^{g-1,g-1}((\mathscr{A} \setminus \mathscr{A}[N])_{\mathbb{R}}) \hookrightarrow \widetilde{D}^{g-1,g-1}((\mathscr{A} \setminus \mathscr{A}[N])_{\mathbb{R}}).$$

The group $H^{2g-1}_{D^{\mathrm{an}}}((\mathscr{A} \setminus \mathscr{A}[N])_{\mathbb{R}}, \mathbb{R}(g))$ can be represented by classes in $(2\pi i)^{g-1}\widetilde{E}^{g-1,g-1}((\mathscr{A} \setminus \mathscr{A}[N])_{\mathbb{R}})$ with dd^c equal to zero by (2), so we get:

Lemma 5. *The Maillot–Rössler current defines a class*

$$(2\pi i)^{g-1}({}_N\mathfrak{g}_{\mathscr{A}^\vee})|_{\mathscr{A}\setminus\mathscr{A}[N]} \in H^{2g-1}_{D^{\mathrm{an}}}((\mathscr{A}\setminus\mathscr{A}[N])_{\mathbb{R}}, \mathbb{R}(g)).$$

The exact sequence (see the theorem and remark in [Gillet and Soulé 1990, §3.3.5])

$$H^{2g-1}_{\mathscr{M}}(\mathscr{A}\setminus\mathscr{A}[N], g) \xrightarrow{\mathrm{reg}_{\mathrm{an}}} H^{2g-1}_{D^{\mathrm{an}}}((\mathscr{A}\setminus\mathscr{A}[N])_{\mathbb{R}}, \mathbb{R}(g)) \xrightarrow{r} \widehat{\mathrm{CH}}^g(\mathscr{A}\setminus\mathscr{A}[N])_{\mathbb{Q}}$$

where r sends $\tilde{x}$ to the class of $(0, \tilde{x}/(2\pi i)^{g-1})$, with the vanishing of $({}_N S_0, {}_N\,\mathfrak{g}_{\mathscr{A}^\vee})$ in $\widehat{\mathrm{CH}}^g(\mathscr{A}\setminus\mathscr{A}[N])_{\mathbb{Q}}$, then implies that the Maillot–Rössler current is motivic, i.e.,

$$(2\pi i)^{g-1}({}_N\mathfrak{g}_{\mathscr{A}^\vee})|_{\mathscr{A}\setminus\mathscr{A}[N]} \in \mathrm{reg}_{\mathrm{an}}(H^{2g-1}_{\mathscr{M}}(\mathscr{A}\setminus\mathscr{A}[N], g)).$$

Since the operator $\mathrm{tr}_{[a]}$ defined in (1) obviously operates on analytic Deligne cohomology and the map $\mathrm{reg}_{\mathrm{an}}$ intertwines this operator with $\mathrm{tr}_{[a]}$, we deduce from Theorem 4(iii) the fact:

Lemma 6. *The Maillot–Rössler current is in the image of the regulator from* $H^{2g-1}_{\mathscr{M}}(\mathscr{A}\setminus\mathscr{A}[N], g)^{(0)}$:

$$(2\pi i)^{g-1}({}_N\mathfrak{g}_{\mathscr{A}^\vee})|_{\mathscr{A}\setminus\mathscr{A}[N]} \in \mathrm{reg}_{\mathrm{an}}(H^{2g-1}_{\mathscr{M}}(\mathscr{A}\setminus\mathscr{A}[N], g)^{(0)}).$$

3. The comparison between pol^0 and the class $\mathfrak{g}_{\mathscr{A}^\vee}$

In this section we give an easy conceptual proof of the comparison result between pol^0 and the class $\mathfrak{g}_{\mathscr{A}^\vee}$.

A commutative diagram. The following lemma, proved by Rössler and Kings, is the key for the proof of our comparison result.

Lemma 7 [Kings and Rössler 2017, Lemma 4.2.6]. *The diagram*

$$\begin{tikzcd}
H^{2g-1}_{\mathscr{M}}(\mathscr{A}\setminus\mathscr{A}[N], g)^{(0)} \arrow[rr, "\mathrm{res}", "\simeq"'] \arrow[d, "\mathrm{reg}_{\mathrm{an}}"] & & H^{2g}_{\mathscr{M},\mathscr{A}[N]\setminus S_0}(\mathscr{A}\setminus S_0, g)^{(0)} \arrow[dd, "\mathrm{reg}_{\mathrm{B}}"] \\
H^{2g-1}_{D^{\mathrm{an}}}((\mathscr{A}\setminus\mathscr{A}[N])_{\mathbb{R}}, \mathbb{R}(g)) \arrow[d, "\phi_B"] & & \\
H^{2g-1}_{B}((\mathscr{A}\setminus\mathscr{A}[N])(\mathbb{C}), \mathbb{R}(g)) \arrow[rr, "\mathrm{res}_B"] & & H^{2g}_{B,\mathscr{A}[N]\setminus S_0}((\mathscr{A}\setminus S_0)(\mathbb{C}), \mathbb{R}(g))
\end{tikzcd}$$

is commutative, and the map $\mathrm{reg}_{\mathrm{B}}$ *is injective.*

The comparison result. We are now ready to reprove the comparison result of Kings and Rössler.

Theorem 8 [Kings and Rössler 2017]. *We have the equality*

$$-2\cdot\mathrm{reg}_{\mathrm{an}}(\mathrm{pol}^0) = (2\pi i)^{g-1}({}_N\mathfrak{g}_{\mathscr{A}^\vee})|_{\mathscr{A}\setminus\mathscr{A}[N]}.$$

Proof. Let $\psi \in H^{2g-1}_{\mathscr{M}}(\mathscr{A}\setminus\mathscr{A}[N], g)^{(0)}$ be such that $\mathrm{reg}_{\mathrm{an}}(\psi) = -\frac{1}{2}(2\pi i)^{g-1}({}_N\mathfrak{g}_{\mathscr{A}^\vee})|_{\mathscr{A}\setminus\mathscr{A}[N]}$. By Lemma 7, it is sufficient to show that ψ and pol^0 have the same image under $\mathrm{reg}_{\mathrm{B}}\circ\mathrm{res}$.

Now, by definition of pol^0 we have

$$\mathrm{reg}_{\mathrm{B}}(\mathrm{res}(\mathrm{pol}^0)) = \mathrm{reg}_{\mathrm{B}}(1_N^\circ) = [(2\pi i)^g \delta_{1_N^\circ}, 0],$$

and by the description of res_B in Lemma 2 we have

$$\begin{aligned}\mathrm{reg}_{\mathrm{B}}(\mathrm{res}(\psi)) = \mathrm{res}_B(\phi_B(\mathrm{reg}_{\mathrm{an}}(\psi))) &= \mathrm{res}_B([-(2\pi i)^g \mathrm{d}^c({}_N\mathfrak{g}_{\mathscr{A}^\vee}|_{\mathscr{A}\setminus\mathscr{A}[N]})]) \\ &= [0, (2\pi i)^g \mathrm{d}^c({}_N\mathfrak{g}_{\mathscr{A}^\vee})|_{\mathscr{A}\setminus\mathscr{A}[N]}].\end{aligned}$$

The difference $\mathrm{reg}_{\mathrm{B}}(\mathrm{res}((\mathrm{pol}^0 - \psi)))$ is then represented by the pair

$$((2\pi i)^g \delta_{1_N^\circ}, -(2\pi i)^g \mathrm{d}^c({}_N\mathfrak{g}_{\mathscr{A}^\vee})|_{\mathscr{A}\setminus\mathscr{A}[N]}),$$

which is a coboundary, since (by [Scarponi 2017, Proposition 5.2])

$$(2\pi i)^g \delta_{1_N^\circ} + (2\pi i)^g\, \mathrm{dd}^c({}_N\mathfrak{g}_{\mathscr{A}^\vee}|_{\mathscr{A}\setminus S_0}) = 0. \qquad \square$$

4. A decomposition of the arithmetic Chow group

Recall the exact sequence (see the theorem and remark in [Gillet and Soulé 1990, §3.3.5])

$$H_{\mathscr{M}}^{2p-1}(\mathscr{A}, p) \to \widetilde{E}^{p-1,p-1}(\mathscr{A}_{\mathbb{R}}) \to \widehat{\mathrm{CH}}^p(\mathscr{A})_{\mathbb{Q}} \to \mathrm{CH}^p(\mathscr{A})_{\mathbb{Q}} \to 0. \tag{4}$$

The endomorphism $[n]^*$ acts on this sequence, and we want to study the decomposition into generalized eigenspaces. Denote by $E_{\mathscr{A}}^{p,q}$ the sheaf of p, q-forms on $\mathscr{A}(\mathbb{C})$. For the next result observe that we have an isomorphism of sheaves $\pi^*\varepsilon^* E_{\mathscr{A}}^{p,q} \cong E_{\mathscr{A}}^{p,q}$, which identifies the pull-back of sections of $\varepsilon^* E_{\mathscr{A}}^{p,q}$ on the base with the translation invariant differential forms on $\mathscr{A}$. For a $\mathscr{C}^\infty$ section $a : S(\mathbb{C}) \to \mathscr{A}(\mathbb{C})$, we denote by $\tau_a : \mathscr{A}(\mathbb{C}) \to \mathscr{A}(\mathbb{C})$ the translation by a. A differential form ω is translation invariant, if $\tau_a^*\omega = \omega$ for all sections a.

Theorem 9. *Let $n \geq 2$ and $\omega \in \widetilde{E}^{p,q}(\mathscr{A}_{\mathbb{R}})$. Assume that ω is a generalized eigenvector for $[n]^*$ with eigenvalue λ, i.e., $([n]^* - \lambda)^k\omega = 0$, for some $k \geq 1$. Then the form ω is translation invariant. In particular, there is a section $\eta \in \varepsilon^* E_{\mathscr{A}}^{p,q}(S_{\mathbb{R}})$ with $\omega = \pi^*\eta$. Moreover, one has $[n]^*\omega = n^{p+q}\omega$, i.e., ω is an eigenvector with eigenvalue n^{p+q}.*

Proof. The statement that ω is translation invariant does not depend on the complex structure. We use that locally on the base the family of complex tori $\pi : \mathscr{A}(\mathbb{C}) \to S(\mathbb{C})$ is as a $\mathscr{C}^\infty$-manifold of the form $U \times (S^1)^{2g}$, where $U \subset S(\mathbb{C})$ is open and $S^1 = \mathbb{R}/\mathbb{Z}$ is a real torus. In this situation it suffices to show that ω is translation invariant under a dense subset of points of $(S^1)^{2g}$.

We start to prove the following claim: if the form $\eta := ([n]^* - \lambda)\omega$ is translation invariant, then ω is translation invariant.

First note that $\lambda \neq 0$ because $[n]_*[n]^*\omega = n^{2g}\omega$, which implies that $[n]^*$ is injective. As the set $\{a \in (S^1)^{2g} \mid [n^r](a) = 0 \text{ for some } r \geq 0\}$ is dense in $(S^1)^{2g}$, by induction over r it suffices to show that

$\tau_a^*\omega = \omega$ for a with $[n^r](a) = 0$. The case $r = 0$ is trivial because then $a = 0$. Suppose we know that ω is translation invariant for all b with $[n^{r-1}](b) = 0$, and let a be such that $[n^r](a) = 0$. We compute

$$\lambda\tau_a^*\omega = \tau_a^*([n]^*\omega - \eta) = [n]^*\tau_{[n]a}^*\omega - \tau_a^*\eta = [n]^*\omega - \eta = \lambda\omega.$$

As $\lambda \neq 0$, it follows that $\tau_a^*\omega = \omega$. This completes the induction step.

We now show by induction on k that ω with $([n]^* - \lambda)^k\omega = 0$ is translation invariant. For $k = 1$ this follows from the claim by setting $\eta = 0$. Suppose that all forms η with $([n]^* - \lambda)^{k-1}\eta = 0$ are translation invariant. Then $\eta := ([n]^* - \lambda)\omega$ is translation invariant, and it follows from the claim that also ω is translation invariant.

For the final statement we just observe that $[n]^*$ acts via n^{p+q}-multiplication on the bundle $\varepsilon^* E_{\mathscr{A}}^{p,q}$ whose sections identify with the translation invariant forms on $\mathscr{A}(\mathbb{C})$. □

For the next result we have to consider generalized eigenspaces for $[n]^*$, and to distinguish these from the generalized eigenspaces for $[n]_*$, we write

$$V(a) := \{v \in V \mid ([n]^* - n^a)^k v = 0 \text{ for some } k \geq 1\}.$$

Corollary 10. *For each $a = 0, \ldots, 2g$ there is an exact sequence*

$$H_{\mathscr{M}}^{2p-1}(\mathscr{A}, p)(a) \to \widetilde{E}^{p-1,p-1}(\mathscr{A}_{\mathbb{R}})(a) \to \widehat{\mathrm{CH}}^p(\mathscr{A})_{\mathbb{Q}}(a) \to \mathrm{CH}^p(\mathscr{A})_{\mathbb{Q}}(a)$$

of generalized $[n]^*$-eigenspaces for the eigenvalue n^a. In particular, for $a \neq 2(p-1)$ one has an injection

$$\widehat{\mathrm{CH}}^p(\mathscr{A})_{\mathbb{Q}}(a) \hookrightarrow \mathrm{CH}^p(\mathscr{A})_{\mathbb{Q}}(a).$$

Proof. The sequence (4) is a sequence of modules under the principal ideal domain $\mathbb{C}[X]$, where X acts as $[n]^*$. Note that taking the torsion submodule

$$TM := \ker(M \to M \otimes_{\mathbb{C}[X]} \operatorname{Quot}\mathbb{C}[X])$$

is a left exact functor on short exact sequences

$$0 \to M' \to M \to M'' \to 0.$$

If M' is torsion, the functor T is even exact. As $H_{\mathscr{M}}^{2p-1}(\mathscr{A}, p)$ is torsion, the exact sequence (4) gives rise to an exact sequence

$$0 \to T\,\mathrm{im}(H_{\mathscr{M}}^{2p-1}(\mathscr{A}, p)) \to T\,\widetilde{E}^{p-1,p-1}(\mathscr{A}_{\mathbb{R}}) \to T\,\widehat{\mathrm{CH}}^p(\mathscr{A})_{\mathbb{Q}} \to T\,\mathrm{CH}^p(\mathscr{A})_{\mathbb{Q}}.$$

As a torsion $\mathbb{C}[X]$-module is the direct sum of its generalized eigenspaces, the first claim follows. The second statement follows from the first, as $\widetilde{E}^{p-1,p-1}(\mathscr{A}_{\mathbb{R}})(a) = 0$ for $a \neq 2(p-1)$ by Theorem 9. □

5. An axiomatic characterization of the Maillot–Rössler current

We want to prove an axiomatic characterization of $(-1)^g p_{1,*}(\widehat{\mathrm{ch}}(\overline{\mathcal{P}}))^{[g]}$. The result is the following:

Theorem 11. *Let $\hat{\xi}$ be an element of $\widehat{\mathrm{CH}}^g(\mathcal{A})_{\mathbb{Q}}$ satisfying the following two properties*:

- $\zeta(\hat{\xi}) = S_0$ *in* $\mathrm{CH}^g(\mathcal{A})_{\mathbb{Q}}$,
- $([n]^* - n^{2g})^k(\hat{\xi}) = 0$ *in* $\widehat{\mathrm{CH}}^g(\mathcal{A})_{\mathbb{Q}}$ *for some* $n \geq 2$ *and some* $k \geq 1$.

Then $\hat{\xi} = (-1)^g p_{1,*}(\widehat{\mathrm{ch}}(\overline{\mathcal{P}}))^{[g]} = (S_0, \mathfrak{g}_{\mathcal{A}^\vee})$.

Remark. (1) Notice that, even if $(-1)^g p_{1,*}(\widehat{\mathrm{ch}}(\overline{\mathcal{P}}))^{[g]}$ satisfies the second property for every n [Scarponi 2017, Proposition 6], it is sufficient to ask that this property holds for one integer greater than one to uniquely characterize it.

(2) Notice also that the condition $([n]^* - n^{2g})(\hat{\xi}) = 0$ implies $[n]_*(\hat{\xi}) = \hat{\xi}$, thanks to the projection formula $[n]_*[n]^* = n^{2g}$. In the case of an abelian scheme over the ring of integers of a number field, K. Künnemann [1994] showed that there exists a decomposition of the Arakelov Chow groups as a direct sum of eigenspaces for the pullback $[n]^*$. As a consequence, in this particular case the conditions $([n]^* - n^{2g})(\hat{\xi}) = 0$ and $[n]_*(\hat{\xi}) = \hat{\xi}$ are equivalent, if $\hat{\xi}$ belongs to the Arakelov Chow group.

Proof. By definition $\hat{\xi} \in \widehat{\mathrm{CH}}^g(\mathcal{A})_{\mathbb{Q}}(2g)$ and by Corollary 10 one has an injection

$$\widehat{\mathrm{CH}}^g(\mathcal{A})_{\mathbb{Q}}(2g) \hookrightarrow \mathrm{CH}^g(\mathcal{A})_{\mathbb{Q}}(2g).$$

This shows that $\hat{\xi}$ is uniquely determined by its image in $\mathrm{CH}^g(\mathcal{A})_{\mathbb{Q}}$. As this image is the same as that of $(S_0, \mathfrak{g}_{\mathcal{A}^\vee})$, this shows the theorem. $\square$

Theorems 8 and 11 give us the following axiomatic characterization of $\mathfrak{g}_{\mathcal{A}^\vee}$ and therefore of pol^0.

Theorem 12. *The class $\mathfrak{g}_{\mathcal{A}^\vee}$ is the unique element $\mathfrak{g} \in \widetilde{D}^{g-1,g-1}(\mathcal{A}_{\mathbb{R}})$ such that*

(i) $\mathfrak{g}$ *is a Green current for* S_0,

(ii) $([n]^* - n^{2g})^k(S_0, \mathfrak{g}) = 0$ *in* $\widehat{\mathrm{CH}}^g(\mathcal{A})_{\mathbb{Q}}$ *for some* $n \geq 2$ *and some* $k \geq 1$,

(iii) $[m]_*\mathfrak{g} = \mathfrak{g}$ *for some* $m > 1$.

Furthermore, pol^0 *is the unique element in* $H^{2g-1}_{\mathcal{M}}(\mathcal{A} \setminus \mathcal{A}[N], g)^{(0)}$ *such that*

$$-2 \cdot \mathrm{reg}_{\mathrm{an}}(\mathrm{pol}^0) = (2\pi i)^{g-1}([N]^*\mathfrak{g}_{\mathcal{A}^\vee} - N^{2g}\mathfrak{g}_{\mathcal{A}^\vee})|_{\mathcal{A}\setminus\mathcal{A}[N]} \in \mathrm{H}^{2g-1}_{D^{\mathrm{an}}}((\mathcal{A} \setminus \mathcal{A}[N])_{\mathbb{R}}, \mathbb{R}(g)).$$

Proof. By Theorem 11 we know that the first two conditions of our theorem are equivalent to the first two conditions in Theorem 4, so that $\mathfrak{g}_{\mathcal{A}^\vee}$ satisfies the three properties above. Suppose now that $\mathfrak{g} \in \widetilde{D}^{g-1,g-1}(\mathcal{A}_{\mathbb{R}})$ is another element satisfying the three properties of our theorem, and let $m > 1$ be such that $[m]_*\mathfrak{g} = \mathfrak{g}$. We want to show that $\mathfrak{g}_{\mathcal{A}^\vee} = \mathfrak{g}$. Since by Theorem 11

$$(S_0, \mathfrak{g}) = (-1)^g p_{1,*}(\widehat{\mathrm{ch}}(\overline{\mathcal{P}}))^{[g]} = (S_0, \mathfrak{g}_{\mathcal{A}^\vee}),$$

the exact sequence (4) implies that the difference $\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g}$ belongs to the image of the regulator $H^{2g-1}_{\mathcal{M}}(\mathcal{A}, g) \to \widetilde{E}^{g-1,g-1}(\mathcal{A}_{\mathbb{R}})$. Since $H^{2g-1}_{\mathcal{M}}(\mathcal{A}, g)$ is a torsion module over $\mathbb{C}[X]$ (with X acting as $[m]^*$), then $\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g}$ lies in $T\widetilde{E}^{g-1,g-1}(\mathcal{A}_{\mathbb{R}})$. The projection formula and Theorem 9 give

$$m^{2g}(\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g}) = [m]_*[m]^*(\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g}) = m^{2g-2}[m]_*(\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g}),$$

i.e., $[m]_*(\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g}) = m^2(\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g})$, but property (iii) in our theorem implies that $[m]_*(\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g}) = (\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g})$. This is possible only if $\mathfrak{g}_{\mathcal{A}^\vee} - \mathfrak{g}$ is zero.

The second statement is a simple consequence of Theorem 8 and the fact that $\mathrm{reg}_{\mathrm{an}}$ is injective when restricted to $H^{2g-1}_{\mathcal{M}}(\mathcal{A} \setminus \mathcal{A}[N], g)^{(0)}$ [Kings and Rössler 2017, Lemma 4.2.6]. □

Acknowledgements

The authors would like to express their gratitude to José Burgos, who was always ready for encouraging discussions and who explained to us many subtler aspects of his work on arithmetic Chow groups and Deligne cohomology. This work was carried out at the DFG SFB 1085 "Higher invariants" in Regensburg, whose support is gratefully acknowledged.

References

[Beilinson 1985] A. A. Beilinson, "Higher regulators and values of L-functions", *J. Sov. Math.* **30**:2 (1985), 2036–2070. MR Zbl

[Burgos 1997] J. I. Burgos, "Arithmetic Chow rings and Deligne–Beilinson cohomology", *J. Algebraic Geom.* **6**:2 (1997), 335–377. MR Zbl

[Gillet and Soulé 1990] H. Gillet and C. Soulé, "Arithmetic intersection theory", *Inst. Hautes Études Sci. Publ. Math.* **72** (1990), 93–174. MR Zbl

[Huber and Kings 2018] A. Huber and G. Kings, "Polylogarithm for families of commutative group schemes", *J. Algebraic Geom.* **27**:3 (2018), 449–495. MR Zbl

[Kings and Rössler 2017] G. Kings and D. Rössler, "Higher analytic torsion, polylogarithms and norm compatible elements on abelian schemes", pp. 99–126 in *Geometry, analysis and probability*, edited by J.-B. Bost et al., Progr. Math. **310**, Springer, 2017. MR Zbl

[Künnemann 1994] K. Künnemann, "Arakelov Chow groups of abelian schemes, arithmetic Fourier transform, and analogues of the standard conjectures of Lefschetz type", *Math. Ann.* **300**:3 (1994), 365–392. MR Zbl

[Maillot and Rössler 2015] V. Maillot and D. Rössler, "On a canonical class of Green currents for the unit sections of abelian schemes", *Doc. Math.* **20** (2015), 631–668. MR Zbl

[Scarponi 2017] D. Scarponi, "The realization of the degree zero part of the motivic polylogarithm on abelian schemes in Deligne–Beilinson cohomology", *Int. J. Number Theory* **13**:9 (2017), 2471–2485. MR Zbl

[Soulé 1985] C. Soulé, "Opérations en K-théorie algébrique", *Canad. J. Math.* **37**:3 (1985), 488–550. MR Zbl

Communicated by Shou-Wu Zhang
Received 2018-03-20 Revised 2018-09-02 Accepted 2018-11-10

guido.kings@mathematik.uni-regensburg.de *Fakultät für Mathematik, Universität Regensburg, Regensburg, Germany*

danny.scarponi@kcl.ac.uk *King's College London, London, United Kingdom*

ALGEBRA AND NUMBER THEORY 13:2 (2019)
dx.doi.org/10.2140/ant.2019.13.513

Essential dimension of inseparable field extensions

Zinovy Reichstein and Abhishek Kumar Shukla

Let k be a base field, K be a field containing k, and L/K be a field extension of degree n. The essential dimension $\mathrm{ed}(L/K)$ over k is a numerical invariant measuring "the complexity" of L/K. Of particular interest is

$$\tau(n) = \max\left\{\mathrm{ed}(L/K) \mid L/K \text{ is a separable extension of degree } n\right\},$$

also known as the essential dimension of the symmetric group S_n. The exact value of $\tau(n)$ is known only for $n \leqslant 7$. In this paper we assume that k is a field of characteristic $p > 0$ and study the essential dimension of inseparable extensions L/K. Here the degree $n = [L : K]$ is replaced by a pair $(n, \boldsymbol{e})$ which accounts for the size of the separable and the purely inseparable parts of L/K, respectively, and $\tau(n)$ is replaced by

$$\tau(n, \boldsymbol{e}) = \max\left\{\mathrm{ed}(L/K) \mid L/K \text{ is a field extension of type } (n, \boldsymbol{e})\right\}.$$

The symmetric group S_n is replaced by a certain group scheme $G_{n,\boldsymbol{e}}$ over k. This group scheme is neither finite nor smooth; nevertheless, computing its essential dimension turns out to be easier than computing the essential dimension of S_n. Our main result is a simple formula for $\tau(n, \boldsymbol{e})$.

1. Introduction

Throughout this paper k will denote a base field. All other fields will be assumed to contain k. A field extension L/K of finite degree is said to descend to a subfield $K_0 \subset K$ if there exists an intermediate field $K_0 \subset L_0 \subset L$ such that L_0 and K generate L and $[L_0 : K_0] = [L : K]$. Equivalently, L is isomorphic to $L_0 \otimes_{K_0} K$ over K, as is shown in the diagram

$$\begin{tikzcd}
 & L \arrow[dl, dash] \arrow[d, dash] \\
L_0 \arrow[d, dash] & K \arrow[dl, dash] \\
K_0 &
\end{tikzcd}$$

The essential dimension of L/K (over k) is defined as

$$\mathrm{ed}(L/K) = \min\left\{\mathrm{trdeg}(K_0/k) \mid L/K \text{ descends to } K_0 \text{ and } k \subset K_0\right\}.$$

Reichstein was partially supported by National Sciences and Engineering Research Council of Canada Discovery grant 253424-2017. Shukla was partially supported by a graduate fellowship from the Science and Engineering Research Board of India.
MSC2010: 12F05, 12F15, 12F20, 20G10.
Keywords: inseparable field extension, essential dimension, group scheme in prime characteristic.

Essential dimension of separable field extensions was studied in [Buhler and Reichstein 1997]. Of particular interest is

$$\tau(n) = \max\left\{\mathrm{ed}(L/K) \mid L/K \text{ is a separable extension of degree } n \text{ and } k \subset K\right\}, \tag{1-1}$$

otherwise known as the essential dimension of the symmetric group S_n. It is shown in [Buhler and Reichstein 1997] that if $\mathrm{char}(k) = 0$, then $\lfloor n/2 \rfloor \leqslant \tau(n) \leqslant n-3$ for every $n \geqslant 5$.[1] A. Duncan [2010] later strengthened the lower bound as follows.

Theorem 1.1. *If* $\mathrm{char}(k) = 0$, *then* $\lfloor (n+1)/2 \rfloor \leqslant \tau(n) \leqslant n-3$ *for every* $n \geqslant 6$.

This paper is a sequel to [Buhler and Reichstein 1997]. Here we will assume that $\mathrm{char}(k) = p > 0$ and study inseparable field extensions L/K. The role of the degree, $n = [L:K]$ in the separable case, will be played by a pair $(n, \boldsymbol{e})$. The first component of this pair is the separable degree, $n = [S:K]$, where S is the separable closure of K in L. The second component is the so-called type $\boldsymbol{e} = (e_1, \dots, e_r)$ of the purely inseparable extension $[L:S]$, where $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_r \geqslant 1$ are integers; see Section 4 for the definition. Note that the type $\boldsymbol{e} = (e_1, \dots, e_r)$ uniquely determines the inseparable degree $[L:S] = p^{e_1+\cdots+e_r}$ of L/K but not conversely. By analogy with (1-1) it is natural to define

$$\tau(n, \boldsymbol{e}) = \max\left\{\mathrm{ed}(L/K) \mid L/K \text{ is a field extension of type } (n, \boldsymbol{e}) \text{ and } k \subset K\right\}. \tag{1-2}$$

Our main result is the following:

Theorem 1.2. *Let k be a base field of characteristic $p > 0$, $n \geqslant 1$ and $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_r \geqslant 1$ be integers, $\boldsymbol{e} = (e_1, \dots, e_r)$, and $s_i = e_1 + \cdots + e_i$ for $i = 1, \dots, r$. Then*

$$\tau(n, \boldsymbol{e}) = n \sum_{i=1}^{r} p^{s_i - i e_i}.$$

Some remarks are in order.

(1) Theorem 1.2 gives the exact value for $\tau(n, \boldsymbol{e})$. This is in contrast to the separable case, where Theorem 1.1 only gives estimates and the exact value of $\tau(n)$ is unknown for any $n \geqslant 8$.

(2) A priori, the integers $\mathrm{ed}(L/K)$, $\tau(n)$, and $\tau(n, \boldsymbol{e})$ all depend on the base field k. However, Theorem 1.2 shows that for a fixed $p = \mathrm{char}(k)$, $\tau(n, \boldsymbol{e})$ is independent of the choice of k.

(3) Theorem 1.2 implies that for any inseparable extension L/K of finite degree,

$$\mathrm{ed}(L/K) \leqslant \frac{1}{p}[L:K];$$

see Remark 5.3. This is again in contrast to the separable case, where Theorem 1.1 tells us that there exists an extension L/K of degree n such that $\mathrm{ed}(L/K) > \frac{1}{2}[L:K]$ for every odd $n \geqslant 7$ (assuming $\mathrm{char}(k) = 0$).

[1] These inequalities hold for any base field k of characteristic $\neq 2$. On the other hand, the stronger lower bound of Theorem 1.1, due to Duncan, is only known in characteristic 0.

(4) We will also show that the formula for $\tau(n, \boldsymbol{e})$ remains valid if we replace the essential dimension $\mathrm{ed}(L/K)$ in the definition (1-2) by the essential dimension at p, $\mathrm{ed}_p(L/K)$; see Theorem 7.1. For the definition of the essential dimension at a prime, see Section 5 in [Reichstein 2010] or Section 3 below.

The number $\tau(n)$ has two natural interpretations. On the one hand, $\tau(n)$ is the essential dimension of the functor Et_n which associates to a field K the set of isomorphism classes of étale algebras of degree n over K. On the other hand, $\tau(n)$ is the essential dimension of the symmetric group S_n. Recall that an étale algebra L/K is a direct product $L = L_1 \times \cdots \times L_m$ of separable field extensions L_i/K. Equivalently, an étale algebra of degree n over K can be thought of as a twisted K-form of the split algebra $k^n = k \times \cdots \times k$ (n times). The symmetric group S_n arises as the automorphism group of this split algebra, so that $\mathrm{Et}_n = H^1(K, \mathrm{S}_n)$; see Example 3.5.

Our proof of Theorem 1.2 relies on interpreting $\tau(n, \boldsymbol{e})$ in a similar manner. Here the role of the split étale algebra k^n will be played by the algebra $\Lambda_{n,\boldsymbol{e}}$, which is the direct product of n copies of the truncated polynomial algebra

$$\Lambda_{\boldsymbol{e}} = k[x_1, \dots, x_r]/\big(x_1^{p^{e_1}}, \dots, x_r^{p^{e_r}}\big).$$

Note that the k-algebra $\Lambda_{n,\boldsymbol{e}}$ is finite-dimensional, associative, and commutative, but not semisimple. Étale algebras over K will get replaced by K-forms of $\Lambda_{n,\boldsymbol{e}}$. The role of the symmetric group S_n will be played by the algebraic group scheme $G_{n,\boldsymbol{e}} = \mathrm{Aut}_k(\Lambda_{n,\boldsymbol{e}})$ over k. We will show that $\tau(n, \boldsymbol{e})$ is the essential dimension of $G_{n,\boldsymbol{e}}$, just like $\tau(n)$ is the essential dimension of S_n in the separable case. The group scheme $G_{n,\boldsymbol{e}}$ is neither finite nor smooth; however, much to our surprise, computing its essential dimension turned out to be easier than computing the essential dimension of S_n.

The remainder of this paper is structured as follows. Sections 2 and 3 contain preliminary results on finite-dimensional algebras, their automorphism groups, and essential dimension. In Section 4 we recall the structure theory of inseparable field extensions. Section 6 is devoted to versal algebras. The upper bound of Theorem 1.2 is proved in Section 5; alternative proofs are outlined in Section 8. The lower bound of Theorem 1.2 is established in Section 7; our proof relies on the inequality (7-2) due to D. Tossici and A. Vistoli [2013]. Finally, in Section 9 we prove a stronger version of Theorem 1.2 in the special case where $n = 1$, $e_1 = \cdots = e_r$, and k is perfect.

2. Finite-dimensional algebras and their automorphisms

Recall that in the introduction we defined the essential dimension of a field extension L/K of finite degree, where K contains k. The same definition is valid for any finite-dimensional algebra A/K. That is, we say that A descends to a subfield K_0 if there exists a K_0-algebra A_0 such that $A_0 \otimes_{K_0} K$ is isomorphic to A (as a K-algebra). The essential dimension $\mathrm{ed}(A)$ is then the minimal value of $\mathrm{trdeg}(K_0/k)$, where the minimum is taken over the intermediate fields $k \subset K_0 \subset K$ such that A descends to K_0.

Here by a K-algebra A we mean a K-vector space with a bilinear "multiplication" map $m : A \times A \to A$. Later on we will primarily be interested in commutative associative algebras with 1, but at this stage m can be arbitrary: we will not assume that A is commutative or associative or has an identity element. (For example, one can talk of the essential dimension of a finite-dimensional Lie algebra A/K.) Recall that to each basis $x_1, \dots, x_n$ of A one can associate a set of n^3 structure constants $c_{ij}^h \in K$, where

$$x_i \cdot x_j = \sum_{h=1}^{n} c_{ij}^h x_h. \tag{2-1}$$

Lemma 2.1. *Let A be an n-dimensional K-algebra with structure constants c_{ij}^h (relative to some K-basis of A). Suppose a subfield $K_0 \subset K$ contains c_{ij}^h for every $i, j, h = 1, \dots, n$. Then A descends to K_0. In particular*, $\mathrm{ed}(A) \leqslant \mathrm{trdeg}(K_0/k)$.

Proof. Let A_0 be the K_0-vector space with basis $b_1, \dots, b_n$. Define the K_0-algebra structure on A_0 by (2-1). Clearly $A_0 \otimes_{K_0} K = A$, and the lemma follows. □

The following lemma will be helpful to us in the sequel.

Lemma 2.2. *Suppose $k \subset K \subset S$ are field extensions, such that S/K is separable of degree n. Let A be a finite-dimensional algebra over S. If A descends to a subfield S_0 of S such that $K(S_0) = S$, then*

$$\mathrm{ed}(A/K) \leqslant n\,\mathrm{trdeg}(S_0/k).$$

Here $\mathrm{ed}(A/K)$ *is the essential dimension of A, viewed as a K-algebra.*

Proof. By our assumption there exists an S_0-algebra A_0 such that $A = A_0 \otimes_{S_0} S$.

Denote the normal closure of S over K by S^{norm}, and the associated Galois groups by $G = \mathrm{Gal}(S^{\mathrm{norm}}/K)$ and $H = \mathrm{Gal}(S^{\mathrm{norm}}/S) \subset G$. Now define $S_1 = k(g(s) \mid s \in S_0,\ g \in G)$. Choose a transcendence basis $t_1, \dots, t_d$ for S_0 over k, where $d = \mathrm{trdeg}(S_0/k)$. Clearly S_1 is algebraic over $k(g(t_i) \mid g \in G,\ i = 1, \dots, d)$. Since H fixes every element of S, each t_i has at most $[G : H] = n$ distinct translates of the form $g(t_i)$, $g \in G$. This shows that $\mathrm{trdeg}(S_1/k) \leqslant nd$.

Now let $K_1 = S_1^G \subset K$ and $A_1 = A_0 \otimes_{K_0} K_1$. Since S_1 is algebraic over K_1, we have

$$\mathrm{trdeg}(K_1/k) = \mathrm{trdeg}(S_1/k) \leqslant nd.$$

Examining the diagram

$$\begin{tikzcd}
A_0 \arrow[r, dash] \arrow[d, dash] & A_1 \arrow[r, dash] \arrow[d, dash] & A \arrow[d, dash] \\
S_0 \arrow[r, dash] & S_1 \arrow[r, dash] \arrow[d, dash] & S \arrow[d, dash] \\
 & K_1 \arrow[r, dash] & K
\end{tikzcd}$$

we see that A/K descends to K_1, and the lemma follows. □

Now let Λ be a finite-dimensional k-algebra with multiplication map $m : \Lambda \times \Lambda \to \Lambda$. The general linear group $\mathrm{GL}_k(\Lambda)$ acts on the vector space $\Lambda^* \otimes_k \Lambda^* \otimes_k \Lambda$ of bilinear maps $\Lambda \times \Lambda \to \Lambda$. The automorphism group scheme $G = \mathrm{Aut}_k(\Lambda)$ of Λ is defined as the stabilizer of m under this action. It is a closed subgroup scheme of $\mathrm{GL}_k(\Lambda)$ defined over k. The reason we use the term "group scheme" here, rather than "algebraic group", is that G may not be smooth; see the Remark after Lemma III.1.1 in [Serre 1997].

Proposition 2.3. *Let Λ be a commutative finite-dimensional local k-algebra with residue field k, and $G = \mathrm{Aut}_k(\Lambda)$ be its automorphism group scheme. Then the natural map*

$$f : G^n \rtimes \mathrm{S}_n \to \mathrm{Aut}_k(\Lambda^n)$$

is an isomorphism. Here $G^n = G \times \cdots \times G$ (n times) acts on $\Lambda^n = \Lambda \times \cdots \times \Lambda$ (n times) componentwise and S_n acts by permuting the factors.

Before proceeding with the proof of the proposition, recall that an element α of a ring R is called an idempotent if $\alpha^2 = \alpha$.

Lemma 2.4. *Let Λ be a commutative finite-dimensional local k-algebra with residue field k and R be an arbitrary commutative k-algebra with 1. Then the only idempotents of $\Lambda_R = \Lambda \otimes_k R$ are those in R (more precisely in $1 \otimes R$).*

Proof. By Lemma 6.2 in [Waterhouse 1979], the maximal ideal M of Λ consists of nilpotent elements. Tensoring the natural projection $\Lambda \to \Lambda/M \simeq k$ with R, we obtain a surjective homomorphism $\Lambda_R \to R$ whose kernel again consists of nilpotent elements. By Proposition 7.14 in [Jacobson 1980], every idempotent in R lifts to a unique idempotent in Λ_R, and the lemma follows. □

Proof of Proposition 2.3. Let $\alpha_i = (0, \dots, 1, \dots, 0)$ where 1 appears in the i-th position. Then $\bigoplus_{i=1}^n R\alpha_i$ is an R-subalgebra of Λ_R^n.

Let $f \in \mathrm{Aut}_R(\Lambda_R^n)$. Since each α_i is an idempotent in Λ_R^n, so is each $f(\alpha_i)$. The components of each $f(\alpha_i)$ are idempotents in Λ_R. By Lemma 2.4, they lie in R. Thus, $f(\alpha_i) \in \bigoplus_{i=1}^n R\alpha_i$ for every $i = 1, \dots, n$. As a result, we obtain a morphism

$$\mathrm{Aut}_R(\Lambda_R^n) \xrightarrow{\tau_R} \mathrm{Aut}_R\left(\bigoplus_{i=1}^n R\alpha_i\right) = S_n(R).$$

For the second equality, see, e.g., p. 59 in [Waterhouse 1979]. These maps are functorial in R and thus give rise to a morphism $\tau : \mathrm{Aut}(\Lambda^n) \to \mathrm{S}_n$ of group schemes over k. The kernel of τ is $\mathrm{Aut}(\Lambda)^n$, and τ clearly has a section. The proposition follows. □

Remark 2.5. The assumption that Λ is commutative in Proposition 2.3 can be dropped, as long as we assume that the center of Λ is a finite-dimensional local k-algebra with residue field k. The proof proceeds along similar lines, except that we restrict f to an automorphism of the center $Z(\Lambda^n) = Z(\Lambda)^n$ and apply Lemma 2.4 to $Z(\Lambda)$, rather than Λ itself. This more general variant of Proposition 2.3 will not be needed in the sequel.

Remark 2.6. On the other hand, the assumption that the residue field of Λ is k cannot be dropped. For example, if Λ is a separable field extension of k of degree d, then $\mathrm{Aut}_k(\Lambda^n)$ is a twisted form of

$$\mathrm{Aut}_{\bar{k}}(\Lambda^n \otimes_k \bar{k}) = \mathrm{Aut}_{\bar{k}}(\bar{k}^{dn}) = \mathrm{S}_{nd}.$$

Here $\bar{k}$ denotes the separable closure of k. Similarly, $\mathrm{Aut}_k(\Lambda)^n \rtimes \mathrm{S}_d$ is a twisted form of $(\mathrm{S}_d)^n \rtimes \mathrm{S}_n$. For $d, n > 1$, these groups have different orders, so they cannot be isomorphic.

3. Essential dimension of a functor

In the sequel we will need the following general notion of essential dimension, due to A. Merkurjev [Berhuy and Favi 2003]. Let $\mathcal{F} : \mathrm{Fields}_k \to \mathrm{Sets}$ be a covariant functor from the category of field extensions K/k to the category of sets. Here k is assumed to be fixed throughout, and K ranges over all fields containing k. We say that an object $a \in \mathcal{F}(K)$ descends to a subfield $K_0 \subset K$ if a lies in the image of the natural restriction map $\mathcal{F}(K_0) \to \mathcal{F}(K)$. The essential dimension $\mathrm{ed}(a)$ of a is defined as the minimal value of $\mathrm{trdeg}(K_0/k)$, where $k \subset K_0$ and a descends to K_0. The essential dimension of the functor $\mathcal{F}$, denoted by $\mathrm{ed}(\mathcal{F})$, is the supremum of $\mathrm{ed}(a)$ for all $a \in F(K)$, and all fields K in Fields_k.

If l is a prime, there is also a related notion of essential dimension at l, which we denote by ed_l. For an object $a \in \mathcal{F}$, we define $\mathrm{ed}_l(a)$ as the minimal value of $\mathrm{ed}(a')$, where a' is the image of a in $\mathcal{F}(K')$, and the minimum is taken over all field extensions K'/K such that the degree $[K' : K]$ is finite and prime to l. The essential dimension $\mathrm{ed}_l(\mathcal{F})$ of the functor $\mathcal{F}$ at l is defined as the supremum of $\mathrm{ed}_l(a)$ for all $a \in F(K)$ and all fields K in Fields_k. Note that the prime l in this definition is unrelated to $p = \mathrm{char}(k)$; we allow both $l = p$ and $l \neq p$.

Example 3.1. Let G be a group scheme over a base field k and $\mathcal{F}_G : K \to H^1(K, G)$ be the functor defined by

$$\mathcal{F}_G(K) = \{\text{isomorphism classes of } G\text{-torsors } T \to \mathrm{Spec}(K)\}.$$

Here by a torsor we mean a torsor in the flat (fppf) topology. If G is smooth, then $H^1(K, G)$ is the first Galois cohomology set, as in [Serre 1997]; see Section II.1. The essential dimension $\mathrm{ed}(G)$ is, by definition, $\mathrm{ed}(\mathcal{F}_G)$, and similarly for the essential dimension $\mathrm{ed}_l(G)$ of G at prime l. These numerical invariants of G have been extensively studied; see, e.g., [Merkurjev 2009] or [Reichstein 2010] for a survey.

Example 3.2. Define the functor $\mathrm{Alg}_n : K \to H^1(K, G)$ by

$$\mathrm{Alg}_n(K) = \{\text{isomorphism classes of } n\text{-dimensional } K\text{-algebras}\}.$$

If A is an n-dimensional algebra, and $[A]$ is its class in $\mathrm{Alg}_n(K)$, then $\mathrm{ed}([A])$ coincides with $\mathrm{ed}(A)$ defined at the beginning of Section 2. By Lemma 2.1, $\mathrm{ed}(\mathrm{Alg}_n) \leqslant n^3$; the exact value is unknown (except for very small n).

We will now restrict our attention to certain subfunctors of Alg_n which are better understood.

Definition 3.3. Let Λ/k be a finite-dimensional algebra and K/k be a field extension (not necessarily finite or separable). We say that an algebra A/K is a *K-form* of Λ if there exists a field L containing K such that $\Lambda \otimes_k L$ is isomorphic to $A \otimes_K L$ as an L-algebra. We will write

$$\mathrm{Alg}_\Lambda : \mathrm{Fields}_k \to \mathrm{Sets}$$

for the functor which sends a field K/k to the set of K-isomorphism classes of K-forms of Λ.

Proposition 3.4. *Let Λ be a finite-dimensional k-algebra and $G = \mathrm{Aut}_k(\Lambda) \subset \mathrm{GL}(\Lambda)$ be its automorphism group scheme. Then the functors Alg_Λ and $\mathcal{F}_G = H^1(*, G)$ are isomorphic. In particular*, $\mathrm{ed}(\mathrm{Alg}_\Lambda) = \mathrm{ed}(G)$ *and* $\mathrm{ed}_l(\mathrm{Alg}_\Lambda) = \mathrm{ed}_l(G)$ *for every prime l.*

Proof. For the proof of the first assertion, see Proposition X.2.4 in [Serre 1979] or Proposition III.2.2.2 in [Knus 1991]. The second assertion is an immediate consequence of the first, since isomorphic functors have the same essential dimension. □

Example 3.5. The K-forms of $\Lambda_n = k \times \cdots \times k$ (n times) are called étale algebras of degree n. An étale algebra L/K of degree n is a direct products of separable field extensions,

$$L = L_1 \times \cdots \times L_r, \quad \text{where } \sum_{i=1}^{r} [L_i : K] = n.$$

The functor Alg_{Λ_n} is usually denoted by Et_n. The automorphism group $\mathrm{Aut}_k(\Lambda_n)$ is the symmetric group S_n, acting on Λ_n by permuting the n factors of k; see Proposition 2.3. Thus, $\mathrm{Et}_n = H^1(K, \mathrm{S}_n)$; see, e.g., Examples 2.1 and 3.2 in [Serre 2003].

4. Field extensions of type (n, e)

Let L/S be a purely inseparable extension of finite degree. For $x \in L$ we define the exponent of x over S as the smallest integer e such that $x^{p^e} \in S$. We will denote this number by $e(x, S)$. We will say that $x \in L$ is *normal* in L/S if $e(x, S) = \max\{e(y, S) \mid y \in L\}$. A sequence $x_1, \ldots, x_r$ in L is called normal if each x_i is normal in L_i/L_{i-1} and $x_i \notin L_{i-1}$. Here $L_i = S(x_1, \ldots, x_{i-1})$ and $L_0 = S$. If $L = S(x_1, \ldots, x_r)$, where $x_1, \ldots, x_r$ is a normal sequence in L/S, then we call $x_1, \ldots, x_r$ a *normal generating sequence* of L/S. We will say that this sequence is *of type* $\boldsymbol{e} = (e_1, \ldots, e_r)$ if $e_i := e(x_i, L_{i-1})$ for each i. Here $L_i = S(x_1, \ldots, x_i)$, as above. It is clear that $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_r$.

Proposition 4.1 (G. Pickert [1949]). *Let L/S be a purely inseparable field extension of finite degree.*

(a) *For any generating set Λ of L/S there exists a normal generating sequence $x_1, \ldots, x_r$ with each $x_i \in \Lambda$.*

(b) *If $x_1, \ldots, x_r$ and $y_1, \ldots, y_s$ are two normal generating sequences for L/S, of types $(e_1, \ldots, e_r)$ and $(f_1, \ldots, f_s)$, respectively, then $r = s$ and $e_i = f_i$ for each $i = 1, \ldots, r$.*

Proof. For modern proofs of both parts, see Propositions 6 and 8 in [Rasala 1971] or Lemma 1.2 and Corollary 1.5 in [Karpilovsky 1989]. □

Proposition 4.1 allows us to talk about the *type* of a purely inseparable extension L/S. We say that L/S is of type $\boldsymbol{e} = (e_1, \ldots, e_r)$ if it admits a normal generating sequence $x_1, \ldots, x_r$ of type $\boldsymbol{e}$.

Now suppose L/K is an arbitrary inseparable (but not necessarily purely inseparable) field extension L/K of finite degree. Denote the separable closure of K in L by S. We will say that L/K is of type $(n, \boldsymbol{e})$ if $[S : K] = n$ and the purely inseparable extension L/S is of type $\boldsymbol{e}$.

Remark 4.2. Note that we will assume throughout that $r \geqslant 1$, i.e., that L/K is not separable. In particular, a finite field K does not admit an extension of type $(n, \boldsymbol{e})$ for any n and $\boldsymbol{e}$.

Remark 4.3. It follows from Proposition 4.1 that L/K cannot be generated by fewer than r elements. Note also that the integer r can be determined directly, without constructing a normal generating sequence. Indeed, by Theorem 6 in [Becker and MacLane 1940], $[L : K(L^p)] = p^r$. Here $K(L^p)$ denotes the subfield of L generated by L^p and K.

Lemma 4.4. *Let $n \geqslant 1$ and $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_r \geqslant 1$ be integers. Then there exist*

(a) *a separable field extension E/F of degree n with $k \subset F$ and*

(b) *a field extension L/K of type $(n, \boldsymbol{e})$ with $k \subset K$ and $\boldsymbol{e} = (e_1, \dots, e_r)$.*

In particular, this lemma shows that the maxima in definitions (1-1) and (1-2) are taken over a nonempty set of integers.

Proof. (a) Let $x_1, \dots, x_n$ be independent variables over k. Set $E = k(x_1, \dots, x_n)$ and $F = E^C$, where C is the cyclic group of order n acting on E by permuting the variables. Clearly E/F is a Galois (and hence, separable) extension of degree n.

(b) Let E/F be as in part (a) and $y_1, \dots, y_r$ be independent variables over F. Set $L = E(y_1, \dots, y_r)$ and $K = F(z_1, \dots, z_r)$, where $z_i = y_i^{p^{e_i}}$. One readily checks that $S = E(z_1, \dots, z_n)$ is the separable closure of K in L and L/S is a purely inseparable extension of type $\boldsymbol{e}$. □

Now suppose $n \geqslant 1$ and $\boldsymbol{e} = (e_1, \dots, e_r)$ are as above, with $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_r \geqslant 1$. The following finite-dimensional commutative k-algebras will play an important role in the sequel:

$$\Lambda_{n,\boldsymbol{e}} = \Lambda_{\boldsymbol{e}} \times \cdots \times \Lambda_{\boldsymbol{e}} \quad (n \text{ times}), \quad \text{where } \Lambda_{\boldsymbol{e}} = k[x_1, \dots, x_r]/(x_1^{p^{e_1}}, \dots, x_r^{p^{e_r}}) \tag{4-1}$$

is a truncated polynomial algebra.

Lemma 4.5. *$\Lambda_{n,\boldsymbol{e}}$ is isomorphic to $\Lambda_{m,\boldsymbol{f}}$ if and only if $m = n$ and $\boldsymbol{e} = \boldsymbol{f}$.*

Proof. One direction is obvious: if $m = n$ and $\boldsymbol{e} = \boldsymbol{f}$, then $\Lambda_{n,\boldsymbol{e}}$ is isomorphic to $\Lambda_{m,\boldsymbol{f}}$

To prove the converse, note that $\Lambda_{\boldsymbol{e}}$ is a finite-dimensional local k-algebra with residue field k. By Lemma 2.4, the only idempotents in $\Lambda_{\boldsymbol{e}}$ are 0 and 1. This readily implies that the only idempotents in $\Lambda_{n,\boldsymbol{e}}$ are of the form $(\epsilon_1, \dots, \epsilon_n)$, where each ϵ_i is 0 or 1, and the only minimal idempotents are

$$\alpha_1 = (1, 0, \dots, 0), \quad \dots, \quad \alpha_n = (0, \dots, 0, 1).$$

(Recall that idempotents α and β are called *orthogonal* if $\alpha\beta = \beta\alpha = 0$. If α and β are orthogonal, then one readily checks that $\alpha + \beta$ is also an idempotent. An idempotent is *minimal* if it cannot be written as a sum of two orthogonal idempotents.)

If $\Lambda_{n,\boldsymbol{e}}$ and $\Lambda_{m,\boldsymbol{f}}$ are isomorphic, then they have the same number of minimal idempotents; hence, $m = n$. Denote the minimal idempotents of $\Lambda_{m,\boldsymbol{f}}$ by

$$\beta_1 = (1, 0, \dots, 0), \quad \dots, \quad \beta_m = (0, \dots, 0, 1).$$

A k-algebra isomorphism $\Lambda_{n,\boldsymbol{e}} \to \Lambda_{m,\boldsymbol{f}}$ takes α_1 to β_j for some $j = 1, \ldots, n$ and, hence, induces a k-algebra isomorphism between $\alpha_1 \Lambda_{n,\boldsymbol{e}} \simeq \Lambda_{\boldsymbol{e}}$ and $\beta_j \Lambda_{m,\boldsymbol{f}} \simeq \Lambda_{\boldsymbol{f}}$. To complete the proof, we appeal to Proposition 8 in [Rasala 1971], which asserts that $\Lambda_{\mathbf{e}}$ and $\Lambda_{\boldsymbol{f}}$ are isomorphic if and only if $\boldsymbol{e} = \boldsymbol{f}$. $\square$

Lemma 4.6. *Let L/K be a field extension of finite degree. Then the following are equivalent.*

(a) *L/K is of type $(n, \boldsymbol{e})$.*

(b) *L is a K-form of $\Lambda_{n,\boldsymbol{e}}$. In other words, $L \otimes_K K'$ is isomorphic to $\Lambda_{n,\boldsymbol{e}} \otimes_k K'$ as a K'-algebra for some field extension K'/K.*

Proof. (a) $\Longrightarrow$ (b). Assume L/K is a field extension of type $(n, \boldsymbol{e})$. Let S be the separable closure of K in L and K' be an algebraic closure of S (which is also an algebraic closure of K). Then

$$L \otimes_K K' = L \otimes_S (S \otimes_K K') = (L \otimes_S K') \times \cdots \times (L \otimes_S K') \quad (n \text{ times}).$$

On the other hand, by [Rasala 1971, Theorem 3], $L \otimes_S K'$ is isomorphic to $\Lambda_{\boldsymbol{e}}$ as a K'-algebra, and part (b) follows.

(b) $\Longrightarrow$ (a). Assume $L \otimes_K K'$ is isomorphic to $\Lambda_{n,\boldsymbol{e}} \otimes_k K'$ as a K'-algebra for some field extension K'/K. After replacing K' by a larger field, we may assume that K' contains the normal closure of S over K. Since $\Lambda_{n,\boldsymbol{e}} \otimes_k K'$ is not separable over K', L is not separable over K. Thus, L/K is of type $(m, \boldsymbol{f})$ for some $m \geqslant 1$ and $\boldsymbol{f} = (f_1, \ldots, f_s)$ with $f_1 \geqslant f_2 \geqslant \cdots \geqslant f_s \geqslant 1$. As shown above, this implies that $L \otimes_K K''$ is isomorphic to $\Lambda_{m,\boldsymbol{f}} \otimes_k K''$ for a suitable field extension K''/K. After enlarging K'', we may assume without loss of generality that $K' \subset K''$. We conclude that $\Lambda_{n,\boldsymbol{e}} \otimes_k K''$ is isomorphic to $\Lambda_{m,\boldsymbol{f}} \otimes_k K''$ as a K''-algebra. By Lemma 4.5, with k replaced by K'', this is only possible if $(n, \boldsymbol{e}) = (m, \boldsymbol{f})$. $\square$

5. Proof of the upper bound of Theorem 1.2

In this section we will prove the following proposition.

Proposition 5.1. *Let $n \geqslant 1$, $\boldsymbol{e} = (e_1, \ldots, e_r)$, where $e_1 \geqslant \cdots \geqslant e_r \geqslant 1$, and $s_i = e_1 + \cdots + e_i$ for $i = 1, \ldots, r$. Then*

$$\tau(n, \boldsymbol{e}) \leqslant n \sum_{i=1}^{r} p^{s_i - i e_i}.$$

Our proof of Proposition 5.1 will be facilitated by the following lemma.

Lemma 5.2. *Let K be an infinite field of characteristic p, q be a power of p, S/K be a separable field extension of finite degree, and $0 \neq a \in S$. Then there exists an $s \in S$ such that as^q is a primitive element for S/K.*

Proof. Assume the contrary. It is well known that there are only finitely many intermediate fields between K and S; see, e.g., [Lang 1984, Theorem V.4.6]. Denote the intermediate fields properly contained in S by $S_1, \ldots, S_n \subsetneq S$, and let $\mathbb{A}_K(S)$ be the affine space associated to S. (Here we view S as a K-vector space.) The nongenerators of S/K may now be viewed as K-points of the finite union

$$Z = \bigcup_{i=1}^{n} \mathbb{A}_K(S_i).$$

Since we are assuming that every element of S of the form as^q is a nongenerator, and K is an infinite field, the image of the K-morphism $f : \mathbb{A}(S) \to \mathbb{A}(S)$ given by $s \mapsto as^q$ lies in $Z = \bigcup_{i=1}^{n} \mathbb{A}_K(S_i)$. Since $\mathbb{A}_K(S)$ is irreducible, we conclude that the image of f lies in one of the affine subspaces $\mathbb{A}_K(S_i)$, say in $\mathbb{A}_K(S_1)$. Equivalently, $as^q \in S_1$ for every $s \in S$. Setting $s = 1$, we see that $a \in S_1$. Dividing $as^q \in S_1$ by $0 \neq a \in S_1$, we conclude that $s^q \in S_1$ for every $s \in S$. Thus, S is purely inseparable over S_1, contradicting our assumption that S/K is separable. □

Proof of Proposition 5.1. Let L/K be a field extension of type $(n, \boldsymbol{e})$. Our goal is to show that $\operatorname{ed}(L/K) \leqslant n \sum_{j=1}^{r} p^{s_j - je_j}$. By Remark 4.2, K is infinite.

Let S be the separable closure of K in L and $x_1, \dots, x_r$ be a normal generating sequence for the purely inseparable extension L/S of type $\boldsymbol{e}$. Set $q_i = p^{e_i}$. Recall that by the definition of normal sequence, $x_1^{q_1} \in S$. We are free to replace x_1 by $x_1 s$ for any $0 \neq s \in S$; clearly $x_1 s, x_2, \dots, x_r$ is another normal generating sequence. By Lemma 5.2, we may choose $s \in S$ so that $(x_1 s)^{q_1}$ is a primitive element for S/K. In other words, we may assume without loss of generality that $x_1^{q_1}$ is a primitive element for S/K.

By the structure theorem of Pickert, each $x_i^{q_i}$ lies in $S[x_1^{q_i}, \dots, x_{i-1}^{q_i}]$, where $q_i = p^{e_i}$ [Rasala 1971, Theorem 1]. In other words, for each $i = 1, \dots, r$,

$$x_i^{q_i} = \sum a_{d_1, \dots, d_{i-1}} x_1^{q_i d_1} \cdots x_{i-1}^{q_i d_{i-1}} \tag{5-1}$$

for some $a_{d_1,\dots,d_{i-1}} \in S$. Here the sum is taken over all integers $d_1, \dots, d_{i-1}$, where each $0 \leqslant d_j < p^{e_j - e_i}$. Note that for $i = 1$ (5-1) reduces to

$$x_1^{q_1} = a_\varnothing,$$

for some $a_\varnothing \in S$. By Lemma 2.1, L (viewed as an S-algebra), descends to

$$S_0 = k(a_{d_1,\dots,d_{i-1}} \mid i = 1, \dots, r \text{ and } 0 \leqslant d_j < p^{e_j - e_i}).$$

Note that for each $i = 1, \dots, r$, there are exactly

$$p^{e_1 - e_i} \cdot p^{e_2 - e_i} \cdot \dots \cdot p^{e_{i-1} - e_i} = p^{s_i - ie_i}$$

choices of the subscripts $d_1, \dots, d_{i-1}$. Hence, S_0 is generated over k by $\sum_{i=1}^{r} p^{s_i - ie_i}$ elements and consequently,

$$\operatorname{trdeg}(S_0/k) \leqslant \sum_{i=1}^{r} p^{s_i - ie_i}.$$

Moreover, since S_0 contains $a_\varnothing = x_1^q$, which is a primitive element for S/K, we conclude that $K(S_0) = S$. Thus, Lemma 2.2 can be applied to $A = L$; it yields $\operatorname{ed}(L/K) \leqslant n \operatorname{trdeg}(S_0/k)$, and the proposition follows. □

Remark 5.3. Suppose L/K is an extension of type $(n, \boldsymbol{e})$, where $e = (e_1, \dots, e_r)$. Here, as usual, K is assumed to contain the base field k of characteristic $p > 0$. Dividing both sides of the inequality in Proposition 5.1 by $[L : K] = np^{e_1+\cdots+e_r}$, we readily deduce that

$$\frac{\operatorname{ed}(L/K)}{[L:K]} \leqslant \frac{\tau(n)}{[L:K]} \leqslant \sum_{i=1}^{r} p^{-ie_i - e_{i+1} - \cdots - e_r} \leqslant \frac{r}{p^r} \leqslant \frac{1}{p}.$$

In particular, $\operatorname{ed}(L/K) \leqslant \frac{1}{2}[L : K]$ for any inseparable extension $[L : K]$ of finite degree, in any (positive) characteristic. As we pointed out in the introduction, this inequality fails in characteristic 0 (even for $k = \mathbb{C}$).

6. Versal algebras

Let K be a field and A be a finite-dimensional associative K-algebra with 1. Every $a \in A$ gives rise to the K-linear map $l_a : A \to A$ given by $l_a(x) = ax$ (left multiplication by a). Note that $l_{ab} = l_a \cdot l_b$. It readily follows from this that a has a multiplicative inverse in A if and only if l_a is nonsingular.

Proposition 6.1. *Let l be a prime integer and Λ be a finite-dimensional associative k-algebra with* 1. *Assume that there exists a field extension K/k and a K-form A of Λ such that A is a division algebra. Then*:

(a) *There exists a field K_{ver} containing k and a K_{ver}-form A_{ver} of Λ such that*

$$\operatorname{ed}(A_{\text{ver}}) = \operatorname{ed}(\operatorname{Alg}_\Lambda), \quad \operatorname{ed}_l(A_{\text{ver}}) = \operatorname{ed}_l(\operatorname{Alg}_\Lambda) \quad \textit{for every prime integer } l, \quad \textit{and}$$

A_{ver} is a division algebra.

(b) *If G is the automorphism group scheme of Λ, then*

$$\operatorname{ed}(G) = \operatorname{ed}(\operatorname{Alg}_\Lambda) \ = \max\{\operatorname{ed}(A/K) \mid A \textit{ is a } K\textit{-form of } \Lambda \textit{ and a division algebra}\},$$
$$\operatorname{ed}_l(G) = \operatorname{ed}_l(\operatorname{Alg}_\Lambda) = \max\{\operatorname{ed}_l(A/K) \mid A \textit{ is a } K\textit{-form of } \Lambda \textit{ and a division algebra}\}.$$

Here the subscript "ver" is meant to indicate that $A_{\text{ver}}/K_{\text{ver}}$ is a versal object for $\operatorname{Alg}_\Lambda = H^1(*, G)$. For a discussion of versal torsors, see Section I.5 in [Serre 2003] or [Duncan and Reichstein 2015].

Proof. (a) We begin by constructing a versal G-torsor $T_{\text{ver}} \to \operatorname{Spec}(K_{\text{ver}})$. Recall that $G = \operatorname{Aut}_k(\Lambda)$ is defined as a closed subgroup of the general linear group $\operatorname{GL}_k(\Lambda)$. This general linear group admits a generically free linear action on some vector space V (e.g., we can take $V = \operatorname{End}_k(\Lambda)$, with the natural left G-action). Restricting to G we obtain a generically free representation $G \to \operatorname{GL}(V)$. We can now choose a dense open G-invariant subscheme $U \subset V$ over k which is the total space of a G-torsor $\pi : U \to B$; see, e.g., Example 5.4 in [Serre 2003]. Passing to the generic point of B, we obtain a G-torsor $T_{\text{ver}} \to \operatorname{Spec}(K_{\text{ver}})$, where K_{ver} is the function field of B over k. Then $\operatorname{ed}(T_{\text{ver}}/K_{\text{ver}}) = \operatorname{ed}(G)$ (see, e.g., Section 4 in [Berhuy and Favi 2003]) and $\operatorname{ed}_l(T_{\text{ver}}/K_{\text{ver}}) = \operatorname{ed}_l(G)$ (see Lemma 6.6 in [Reichstein and Youssin 2000] or Theorem 4.1 in [Merkurjev 2009]).

Let $T \to \operatorname{Spec}(K)$ be the torsor associated to the K-algebra A and A_{ver} be the K_{ver}-algebra associated to $T_{\text{ver}} \to \operatorname{Spec}(K_{\text{ver}})$ under the isomorphism between the functors $\operatorname{Alg}_\Lambda$ and $H^1(*, G)$ of Proposition 3.4.

By the characteristic-free version of the no-name lemma, proved in [Reichstein and Vistoli 2006, §2], $T \times V$ is G-equivariantly birationally isomorphic to $T \times \mathbb{A}^d_k$, where $d = \dim V$ and G acts trivially on $\mathbb{A}^d_k$. In other words, we have a Cartesian diagram of rational maps defined over k:

$$\begin{tikzcd} & T \times \mathbb{A}^d \arrow[r, dashed, "\simeq"] \arrow[d] & T \times V \arrow[r, dashed, "\mathrm{pr}_2"] & U \arrow[d] \\ \mathbb{A}^d_K \arrow[r, equal] & \mathrm{Spec}(K) \times \mathbb{A}^d \arrow[rr, dashed] & & B \end{tikzcd}$$

Here all direct products are over $\mathrm{Spec}(k)$, and pr_2 denotes the rational G-equivariant projection map taking $(t, v) \in T \times V$ to $v \in V$ for $v \in U$. The map $\mathrm{Spec}(K) \times \mathbb{A}^d \dashrightarrow B$ in the bottom row is induced from the dominant G-equivariant map $T \times \mathbb{A}^d \dashrightarrow U$ on top. Passing to generic points, we obtain an inclusion of field $K_{\mathrm{ver}} \hookrightarrow K(x_1, \ldots, x_d)$ such that the induced map $H^1(K_{\mathrm{ver}}, G) \to H^1(K(x_1, \ldots, x_d), G)$ sends the class of $T_{\mathrm{ver}} \to \mathrm{Spec}(K_{\mathrm{ver}})$ to the class associated to $T \times \mathbb{A}^d \to \mathbb{A}^d_K$. Under the isomorphism of Proposition 3.4 between the functors Alg_Λ and $\mathcal{F}_G = H^1(*, G)$, this translates to

$$A_{\mathrm{ver}} \otimes_{K_{\mathrm{ver}}} K(x_1, \ldots, x_d) \simeq A \otimes_K K(x_1, \ldots, x_d)$$

as $K(x_1, \ldots, x_d)$-algebras.

For simplicity we will write $A(x_1, \ldots, x_d)$ in place of $A \otimes_K K(x_1, \ldots, x_d)$. Since A is a division algebra, so is $A(x_1, \ldots, x_d)$. Thus, the linear map $l_a : A(x_1, \ldots, x_d) \to A(x_1, \ldots, x_d)$ is nonsingular (i.e., has trivial kernel) for every $a \in A_{\mathrm{ver}}$. Hence, the same is true for the restriction of l_a to A_{ver}. We conclude that A_{ver} is a division algebra. Remembering that A_{ver} corresponds to T_{ver} under the isomorphism of functors between Alg_Λ and $\mathcal{F}_G$, we see that

$$\begin{aligned} \mathrm{ed}(A_{\mathrm{ver}}) &= \mathrm{ed}(T_{\mathrm{ver}}/K_{\mathrm{ver}}) = \mathrm{ed}(G) = \mathrm{ed}(\mathrm{Alg}_\Lambda), \\ \mathrm{ed}_l(A_{\mathrm{ver}}) &= \mathrm{ed}_l(T_{\mathrm{ver}}/K_{\mathrm{ver}}) = \mathrm{ed}_l(G) = \mathrm{ed}_l(\mathrm{Alg}_\Lambda), \end{aligned}$$

as desired.

(b) The first equality in both formulas follows from Proposition 3.4, and the second from part (a). □

We will now revisit the finite-dimensional k-algebras $\Lambda_{\boldsymbol{e}}$ and $\Lambda_{n,\boldsymbol{e}} = \Lambda_{\boldsymbol{e}} \times \cdots \times \Lambda_{\boldsymbol{e}}$ (n times) defined in Section 4; see (4-1). We will write

$$G_{n,\boldsymbol{e}} = \mathrm{Aut}(\Lambda_{n,\boldsymbol{e}}) \subset \mathrm{GL}_k(\Lambda_{n,\boldsymbol{e}})$$

for the automorphism group scheme of $\Lambda_{n,\boldsymbol{e}}$ and $\mathrm{Alg}_{n,\boldsymbol{e}}$ for the functor $\mathrm{Alg}_{\Lambda_{n,\boldsymbol{e}}} : \mathrm{Fields}_k \to \mathrm{Sets}$. Recall that this functor associates to a field K/k the set of isomorphism classes of K-forms of $\Lambda_{n,\boldsymbol{e}}$.

Replacing essential dimension by essential dimension at a prime l in the definitions (1-1) and (1-2), we set

$$\begin{aligned} \tau_l(n) &= \max\big\{\mathrm{ed}_l(L/K) \mid L/K \text{ is a separable field extension of degree } n \text{ and } k \subset K\big\}, \\ \tau_l(n, \boldsymbol{e}) &= \max\big\{\mathrm{ed}_l(L/K) \mid L/K \text{ is a field extension of type } (n, \boldsymbol{e}) \text{ and } k \subset K\big\}. \end{aligned}$$

Corollary 6.2. *Let l be a prime integer. Then*:

(a) $\operatorname{ed}(\mathrm{S}_n) = \operatorname{ed}(\mathrm{Et}_n) = \tau(n)$ *and* $\operatorname{ed}_l(\mathrm{S}_n) = \operatorname{ed}_l(\mathrm{Et}_n) = \tau_l(n)$. *Here* Et_n *is the functor of* n*-dimensional étale algebras*, *as in Example 3.5.*

(b) $\operatorname{ed}(G_{n,\boldsymbol{e}}) = \operatorname{ed}(\mathrm{Alg}_{n,\boldsymbol{e}}) = \tau(n, \boldsymbol{e})$ *and* $\operatorname{ed}_l(G_{n,\boldsymbol{e}}) = \operatorname{ed}_l(\mathrm{Alg}_{n,\boldsymbol{e}}) = \tau_l(n, \boldsymbol{e})$.

Proof. (a) Recall that étale algebras are, by definition, commutative and associative with identity. For such algebras "division algebra" is the same as "field". By Lemma 4.4(a) there exists a separable field extension E/F of degree n with $k \subset F$. The desired equality follows from Proposition 6.1(b).

(b) The same argument as in part (a) goes through, with part (a) of Lemma 4.4 replaced by part (b). □

Remark 6.3. The value of $\operatorname{ed}_l(\mathrm{S}_n)$ is known for every integer $n \geqslant$ and every prime $l \geqslant 2$:

$$\operatorname{ed}_l(\mathrm{S}_n) = \begin{cases} \lfloor n/l \rfloor & \text{if } \operatorname{char}(k) \neq l, \\ 1 & \text{if } \operatorname{char}(k) = l \leqslant n, \\ 0 & \text{if } \operatorname{char}(k) = l > n. \end{cases}$$

See respectively [Meyer and Reichstein 2009, Corollary 4.2], [Reichstein and Vistoli 2018, Theorem 1], and either [Meyer and Reichstein 2009, Lemma 4.1] or [Reichstein and Vistoli 2018, Theorem 1].

7. Conclusion of the proof of Theorem 1.2

In this section we will prove Theorem 1.2 in the following strengthened form.

Theorem 7.1. *Let k be a base field of characteristic $p > 0$, $n \geqslant 1$ and $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_r \geqslant 1$ be integers*, $\boldsymbol{e} = (e_1, \ldots, e_r)$, *and $s_i = e_1 + \cdots + e_i$ for $i = 1, \ldots, r$. Then*

$$\tau_p(n, \boldsymbol{e}) = \tau(n, \boldsymbol{e}) = n \sum_{i=1}^{r} p^{s_i - i e_i}.$$

By definition $\tau_p(n, \boldsymbol{e}) \leqslant \tau(n, \boldsymbol{e})$ and by Proposition 5.1, $\tau(n, \boldsymbol{e}) \leqslant n \sum_{i=1}^{r} p^{s_i - i e_i}$. Moreover, by Corollary 6.2(b), $\tau_p(n, \boldsymbol{e}) = \operatorname{ed}_p(G_{n,\boldsymbol{e}})$. It thus remains to show that

$$\operatorname{ed}_p(G_{n,\boldsymbol{e}}) \geqslant n \sum_{i=1}^{r} p^{s_i - i e_i}. \tag{7-1}$$

Our proof of (7-1) will be based on the general inequality, due to Tossici and Vistoli [2013],

$$\operatorname{ed}_p(G) \geqslant \dim \operatorname{Lie}(G) - \dim G \tag{7-2}$$

for any group scheme G of finite type over a field k of characteristic p. Now recall that $G_{\boldsymbol{e}} = \operatorname{Aut}_k(\Lambda_{\boldsymbol{e}})$, and $G_{n,\boldsymbol{e}} = \operatorname{Aut}_k(\Lambda_{n,\boldsymbol{e}})$, where $\Lambda_{n,\boldsymbol{e}} = \Lambda_{\boldsymbol{e}}^n$. Since $\Lambda_{\boldsymbol{e}}$ is a commutative local k-algebra with residue field k, Proposition 2.3 tells us that $G_{n,\boldsymbol{e}} = G_{\boldsymbol{e}}^n \rtimes \mathrm{S}_n$ (see also Proposition 5.1 in [Sancho de Salas 2000]). We conclude that

$$\dim G_{n,\boldsymbol{e}} = n \dim G_{\boldsymbol{e}} \quad \text{and} \quad \dim \operatorname{Lie}(G_{n,\boldsymbol{e}}) = n \dim \operatorname{Lie}(G_{\boldsymbol{e}}).$$

Substituting these formulas into (7-2), we see that the proof of the inequality (7-1) (and thus of Theorem 7.1) reduces to the following:

Proposition 7.2. *Let $\boldsymbol{e} = (e_1, \dots, e_r)$, where $e_1 \geqslant \cdots \geqslant e_r \geqslant 1$ are integers. Then*

(a) $\dim \operatorname{Lie}(G_{\boldsymbol{e}}) = rp^{e_1+\cdots+e_r}$, *and*

(b) $\dim G_{\boldsymbol{e}} = rp^{e_1+\cdots+e_r} - \sum_{i=1}^{r} p^{s_i - ie_i}$.

The remainder of this section will be devoted to proving Proposition 7.2. We will use the following notations.

(1) We fix the type $\boldsymbol{e} = (e_1, \dots, e_r)$ and set $q_i = p^{e_i}$.

(2) The infinitesimal group scheme α_{p^j} over a commutative ring S of characteristic p is defined as the kernel of the j-th power of the Frobenius map, $\mathbb{G}_a \to \mathbb{G}_a$, $x \mapsto x^{p^j}$, viewed as a homomorphism of group schemes over S. We will be particularly interested in the case where $S = \Lambda_{\boldsymbol{e}}$.

(3) Suppose X is a scheme over Λ, where Λ is a finite-dimensional commutative k-algebra. We will denote the Weil restriction of the Λ-scheme X to k by $R_{\Lambda/k}(X)$. For generalities on Weil restriction, see Chapter 2 and the Appendix in [Milne 2017].

(4) We will denote by $\operatorname{End}(\Lambda_{\boldsymbol{e}})$ the functor

$$\mathrm{Comm}_k \to \mathrm{Sets}, \qquad R \to \operatorname{End}_{R\text{-alg}}(\Lambda_{\boldsymbol{e}} \otimes_k R)$$

of algebra endomorphisms of $\Lambda_{\boldsymbol{e}}$. Here Comm_k denotes the category of commutative associative k-algebras with 1 and Sets denotes the category of sets.

Lemma 7.3. (a) *The functor* $\operatorname{End}(\Lambda_{\boldsymbol{e}})$ *is represented by an irreducible, nonreduced, affine k-scheme $X_{\boldsymbol{e}}$.*

(b) $\dim X_{\boldsymbol{e}} = rp^{e_1+\cdots+e_r} - \sum_{i=1}^{r} p^{s_i - ie_i}$.

(c) $\dim T_\gamma(X_{\boldsymbol{e}}) = rp^{e_1+\cdots+e_r}$ *for any k-point γ of $X_{\boldsymbol{e}}$. Here $T_\gamma(X_{\boldsymbol{e}})$ denotes the tangent space to $X_{\boldsymbol{e}}$ at γ.*

Proof. An endomorphism F in $\operatorname{End}(\Lambda_{\boldsymbol{e}})(R)$ is uniquely determined by the images

$$F(x_1), F(x_2), \dots, F(x_r) \in \Lambda_{\boldsymbol{e}}(R)$$

of the generators $x_1, \dots, x_r$ of $\Lambda_{\boldsymbol{e}}$. These elements of $\Lambda_{\boldsymbol{e}}$ satisfy $F(x_i)^{q_i} = 0$. Conversely, any r elements $F_1, \dots, F_r$ in $\Lambda_{\boldsymbol{e}} \otimes R$ satisfying $F_i^{q_i} = 0$ give rise to an algebra endomorphism F in $\operatorname{End}(\Lambda_{\boldsymbol{e}})(R)$. We thus have

$$\begin{aligned}
\operatorname{End}(\Lambda_{\boldsymbol{e}})(R) &= \operatorname{Hom}_{R\text{-alg}}(\Lambda_{\boldsymbol{e}} \otimes_k R, \Lambda_{\boldsymbol{e}} \otimes R) \\
&\cong \alpha_{q_1}(\Lambda_{\boldsymbol{e}} \otimes R) \times \cdots \times \alpha_{q_r}(\Lambda_{\boldsymbol{e}} \otimes R) \\
&\cong R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_1})(R) \times \cdots \times R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_r})(R) \\
&\cong \prod_{i=1}^{r} R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_i})(R).
\end{aligned}$$

We conclude that $\mathrm{End}(\Lambda_{\boldsymbol{e}})$ is represented by an affine k-scheme $X_{\boldsymbol{e}} = \prod_{i=1}^{r} R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_i})$. Note that $X_{\boldsymbol{e}}$ is isomorphic to $\prod_{i=1}^{r} R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_i})$ as a k-scheme only, not as a group scheme. To complete the proof of the lemma it remains to establish the following assertions, claimed for all $q_j \in \{q_1, \dots, q_r\}$:

(a′) $R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_j})$ is irreducible.

(b′) $\dim R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_j}) = p^{e_1+\cdots+e_r} - p^{s_j - je_j}$.

(c′) $\dim T_\gamma(R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_j})) = p^{e_1+\cdots+e_r}$ for any k-point γ of $R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_j})$.

To prove (a′), (b′), and (c′), we will write out explicit equations for $R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_j})$ in $R_{\Lambda_{\boldsymbol{e}}/k}(\mathbb{A}^1) \simeq \mathbb{A}_k(\Lambda_{\boldsymbol{e}})$. We will work in the basis $\{x_1^{i_1} x_2^{i_2} \cdots x_r^{i_r}\}$ of monomials in $\Lambda_{\boldsymbol{e}}$, where $0 \leqslant i_1 < q_1, 0 \leqslant i_2 < q_2, \dots, 0 \leqslant i_r < q_r$. Over $\Lambda_{\boldsymbol{e}}$, α_{q_j} is cut out (scheme-theoretically) in $\mathbb{A}^1$ by the single equation $t^{q_j} = 0$, where t is a coordinate function on $\mathbb{A}^1$. Since $x_i^{q_i} = 0$ for every i, writing

$$t = \sum y_{i_1,\dots,i_r} x_1^{i_1} x_2^{i_2} \cdots x_r^{i_r}$$

and expanding

$$t^{q_j} = \sum y_{i_1,\dots,i_r}^{q_j} x_1^{q_j i_1} x_2^{q_j i_2} \cdots x_r^{q_j i_r}$$

we see that the only monomials appearing in the above sum are those for which

$$q_j i_1 < q_1, \quad q_j i_2 < q_2, \quad \dots, \quad q_j i_r < q_r.$$

Thus, $R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_j})$ is cut out (again, scheme-theoretically) in $R_{\Lambda_{\boldsymbol{e}}/k}(\mathbb{A}^1) \simeq \mathbb{A}(\Lambda_{\boldsymbol{e}})$ by

$$y_{i_1,\dots,i_{j-1},0,\dots,0}^{q_j} = 0 \quad \text{for } 0 \leqslant i_1 < \frac{q_1}{q_j}, \quad \dots, \quad 0 \leqslant i_{j-1} < \frac{q_{j-1}}{q_j},$$

where $y_{i_1,\dots,i_r}$ are the coordinates in $\mathbb{A}(\Lambda_{\boldsymbol{e}})$. In other words, $R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_j})$ is the subscheme of $R_{\Lambda_{\boldsymbol{e}}/k}(\mathbb{A}^1) \simeq \mathbb{A}_k(\Lambda_{\boldsymbol{e}}) \simeq \mathbb{A}_k^{p^{e_1+\cdots+e_r}}$ cut out (again, scheme-theoretically) by q_j-th powers of

$$\frac{q_1}{q_j} \frac{q_2}{q_j} \cdots \frac{q_{j-1}}{q_j} = p^{s_j - je_j}$$

distinct coordinate functions. The reduced scheme $R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_j})_{\mathrm{red}}$ is thus isomorphic to an affine space of dimension $p^{e_1+\cdots+e_r} - \sum_{j=1}^{r} p^{s_j - je_j}$. On the other hand, since q_j is a power of p, the Jacobian criterion tells us that the tangent space to $R_{\Lambda_{\boldsymbol{e}}/k}(\alpha_{q_l})$ at any k-point is the same as the tangent space to $\mathbb{A}(\Lambda_{\boldsymbol{e}}) = \mathbb{A}^{p^{e_1+\cdots+e_r}}$, and (a′), (b′), and (c′) follow. □

Conclusion of the proof of Proposition 7.2. The automorphism group scheme $G_{\boldsymbol{e}}$ is the group of invertible elements in $\mathrm{End}(\Lambda_{\boldsymbol{e}})$. In other words, the natural diagram

$$\begin{array}{ccc} G_{\boldsymbol{e}} & \longrightarrow & \mathrm{GL}_N \\ \downarrow & & \downarrow \\ \mathrm{End}(\Lambda_{\boldsymbol{e}}) & \longrightarrow & \mathrm{Mat}_{N\times N} \end{array}$$

where $N = \dim \Lambda_{\boldsymbol{e}} = p^{e_1+\dots+e_r}$, is Cartesian. Hence, $G_{\boldsymbol{e}}$ is an open subscheme of $X_{\boldsymbol{e}}$. Since $X_{\boldsymbol{e}}$ is irreducible, Proposition 7.2 follows from Lemma 7.3. This completes the proof of Proposition 7.2 and thus of Theorem 7.1. □

8. Alternative proofs of Theorem 1.2

The proof of the lower bound of Theorem 1.2 given in Section 7 is the only one we know. However, we have two other proofs for the upper bound (Proposition 5.1), in addition to the one given in Section 5. In this section we will briefly outline these arguments for the interested reader.

Our first alternative proof of Proposition 5.1 is based on an explicit construction of the versal algebra A_{ver} of type $(n, \boldsymbol{e})$ whose existence is asserted by Proposition 6.1. This construction is via generators and relations, by taking "the most general" structure constants in (5-1). Versality of A_{ver} constructed this way takes some work to prove; however, once versality is established, it is easy to see directly that A_{ver} is a field and thus

$$\tau(n, \boldsymbol{e}) = \mathrm{ed}(A_{\mathrm{ver}}) \leqslant \mathrm{trdeg}(K_{\mathrm{ver}}/k) = n \sum_{i=1}^{r} p^{s_i - i e_i}.$$

Our second alternative proof of Proposition 5.1 is based on showing that the natural representation of $G_{n,\boldsymbol{e}}$ on $V = \Lambda_{n,\boldsymbol{e}}^r$ is generically free. Intuitively speaking, this is clear: $\Lambda_{n,e}$ is generated by r elements as a k-algebra, so r-tuples of generators of $\Lambda_{n,\boldsymbol{e}}$ are dense in V and have trivial stabilizer in $G_{n,\boldsymbol{e}}$. The actual proof involves checking that the stabilizer in general position is trivial scheme-theoretically and not just on the level of points. Once generic freeness of this linear action is established, the upper bound of Proposition 5.1 follows from the inequality

$$\mathrm{ed}(G_{n,\boldsymbol{e}}) \leqslant \dim V - \dim G_{n,\boldsymbol{e}};$$

see, e.g., Proposition 4.11 in [Berhuy and Favi 2003]. To deduce the upper bound of Proposition 5.1 from this inequality, recall that

- $\tau(n, \boldsymbol{e}) = \mathrm{ed}(G_{n,\boldsymbol{e}})$ (see Corollary 6.2(b)),
- $\dim V = r \dim \Lambda_{n,\boldsymbol{e}} = nr \dim \Lambda_{\mathbf{e}} = nrp^{e_1+\dots+e_r}$ (clear from the definition), and
- $\dim G_{n,\boldsymbol{e}} = n \dim G_{\boldsymbol{e}} = nrp^{e_1+\dots+e_r} - n \sum_{i=1}^{r} p^{s_i - i e_i}$ (see Proposition 7.2(b)).

9. The case, where $e_1 = \dots = e_r$

In the special case where $n = 1$ and $e_1 = \dots = e_r$, Theorem 1.2 tells us that $\tau(n, \boldsymbol{e}) = r$. In this section, we will give a short proof of the following stronger assertion under the assumption that k is perfect.

Proposition 9.1. *Let $\boldsymbol{e} = (e, \dots, e)$ (r times) and L/K be purely inseparable extension of type $\boldsymbol{e}$, with $k \subset K$. Assume that the base field k is perfect. Then $\mathrm{ed}_p(L/K) = \mathrm{ed}(L/K) = r$.*

The assumption that k is perfect is crucial here. Indeed, by Lemma 4.4(b), there exists a field extension L/K of type $\boldsymbol{e}$. If we do not require k to be perfect, then we may set $k = K$. In this case $\mathrm{ed}(L/K) = 0$, and the proposition fails.

The remainder of this section will be devoted to proving Proposition 9.1. We begin with two reductions.

(1) It suffices to show that

$$\mathrm{ed}(L/K) = r \quad \text{for every field extension } L/K \text{ of type } \boldsymbol{e}; \tag{9-1}$$

the identity $\mathrm{ed}_p(L/K)$ will then follow. Indeed, $\mathrm{ed}_p(L/K)$ is defined as the minimal value of $\mathrm{ed}(L'/K')$ taken over all finite extensions K'/K of degree prime to p. Here $L' = L \otimes_K K'$. Since $[L:K]$ is a power of p, L' is a field, so (9-1) tells us that $\mathrm{ed}(L'/K') = r$.

(2) The proof of the upper bound,

$$\mathrm{ed}(L/K) \leqslant r, \tag{9-2}$$

is the same as in Section 5, but in this special case the argument is much simplified. For the sake of completeness we reproduce it here. Let $x_1, \dots, x_r$ be a normal generating sequence for L/K. By a theorem of Pickert [Rasala 1971, Theorem 1], $x_1^q, \dots, x_r^q \in K$, where $q = p^e$. Set $a_i = x_i^q$ and $K_0 = k(a_1, \dots, a_r)$. The structure constants of L relative to the K-basis $x_1^{d_1} \cdots x_r^{d_r}$ of L, with $0 \leqslant d_1, \dots, d_r \leqslant q-1$ all lie in K_0. Clearly $\mathrm{trdeg}(K_0/k) \leqslant r$; the inequality (9-2) now follows from Lemma 2.1.

It remains to prove the lower bound, $\mathrm{ed}(L/K) \geqslant r$. Assume the contrary: L/K descends to L_0/K_0 with $\mathrm{trdeg}(K_0/k) < r$. By Lemma 2.1, L_0/K_0 further descends to L_1/K_1, where K_1 is finitely generated over k. By Lemma 4.6, L_1/K_1 is a purely inseparable extension of type $\boldsymbol{e}$. After replacing L/K by L_1/K_1, it remains to prove the following:

Lemma 9.2. *Let k be a perfect field and K/k be a finitely generated field extension of transcendence degree $< r$. There does not exist a purely inseparable field extension L/K of type $\boldsymbol{e} = (e_1, \dots, e_r)$, where $e_1 \geqslant \cdots \geqslant e_r \geqslant 1$.*

Proof. Assume the contrary. Let $a_1, \dots, a_s$ be a transcendence basis for K/k. That is, $a_1, \dots, a_s$ are algebraically independent over k, K is algebraic and finitely generated (hence, finite) over $k(a_1, \dots, a_s)$, and $s \leqslant r-1$. By Remark 4.3,

$$[L:L^p] \geqslant [L:(L^p \cdot K)] = p^r. \tag{9-3}$$

On the other hand, since $[L:k(a_1, \dots, a_s)] < \infty$, Theorem 3 in [Becker and MacLane 1940] tells us that

$$[L:L^p] = [k(a_1, \dots, a_s) : k(a_1, \dots, a_s)^p] = [k(a_1, \dots, a_s) : k(a_1^p, \dots, a_s^p)] = p^s < p^r. \tag{9-4}$$

Note that the second equality relies on our assumption that k is perfect. The contradiction between (9-3) and (9-4) completes the proof of Lemma 9.2 and thus of Proposition 9.1. □

Acknowledgements

We are grateful to Madhav Nori, Julia Pevtsova, Federico Scavia, and Angelo Vistoli for stimulating discussions.

References

[Becker and MacLane 1940] M. F. Becker and S. MacLane, "The minimum number of generators for inseparable algebraic extensions", *Bull. Amer. Math. Soc.* **46** (1940), 182–186. MR Zbl

[Berhuy and Favi 2003] G. Berhuy and G. Favi, "Essential dimension: a functorial point of view (after A. Merkurjev)", *Doc. Math.* **8** (2003), 279–330. MR Zbl

[Buhler and Reichstein 1997] J. Buhler and Z. Reichstein, "On the essential dimension of a finite group", *Compositio Math.* **106**:2 (1997), 159–179. MR Zbl

[Duncan 2010] A. Duncan, "Essential dimensions of A_7 and S_7", *Math. Res. Lett.* **17**:2 (2010), 263–266. MR Zbl

[Duncan and Reichstein 2015] A. Duncan and Z. Reichstein, "Versality of algebraic group actions and rational points on twisted varieties", *J. Algebraic Geom.* **24**:3 (2015), 499–530. MR Zbl

[Jacobson 1980] N. Jacobson, *Basic algebra, II*, W. H. Freeman, San Francisco, 1980. MR Zbl

[Karpilovsky 1989] G. Karpilovsky, *Topics in field theory*, North-Holland Math. Studies **155**, North-Holland, Amsterdam, 1989. MR Zbl

[Knus 1991] M.-A. Knus, *Quadratic and Hermitian forms over rings*, Grundlehren der Math. Wissenschaften **294**, Springer, 1991. MR Zbl

[Lang 1984] S. Lang, *Algebra*, 2nd ed., Addison-Wesley, Reading, MA, 1984. MR Zbl

[Merkurjev 2009] A. S. Merkurjev, "Essential dimension", pp. 299–325 in *Quadratic forms: algebra, arithmetic, and geometry*, edited by R. Baeza et al., Contemp. Math. **493**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

[Meyer and Reichstein 2009] A. Meyer and Z. Reichstein, "The essential dimension of the normalizer of a maximal torus in the projective linear group", *Algebra Number Theory* **3**:4 (2009), 467–487. MR Zbl

[Milne 2017] J. S. Milne, *Algebraic groups*, Cambridge Studies in Adv. Math. **170**, Cambridge Univ. Press, 2017. MR Zbl

[Pickert 1949] G. Pickert, "Inseparable Körpererweiterungen", *Math. Z.* **52** (1949), 81–136. MR Zbl

[Rasala 1971] R. Rasala, "Inseparable splitting theory", *Trans. Amer. Math. Soc.* **162** (1971), 411–448. MR Zbl

[Reichstein 2010] Z. Reichstein, "Essential dimension", pp. 162–188 in *Proceedings of the International Congress of Mathematicians, II* (Hyderabad, 2010), edited by R. Bhatia et al., Hindustan Book Agency, New Delhi, 2010. MR Zbl

[Reichstein and Vistoli 2006] Z. Reichstein and A. Vistoli, "Birational isomorphisms between twisted group actions", *J. Lie Theory* **16**:4 (2006), 791–802. MR Zbl

[Reichstein and Vistoli 2018] Z. Reichstein and A. Vistoli, "Essential dimension of finite groups in prime characteristic", *C. R. Math. Acad. Sci. Paris* **356**:5 (2018), 463–467. MR Zbl

[Reichstein and Youssin 2000] Z. Reichstein and B. Youssin, "Essential dimensions of algebraic groups and a resolution theorem for G-varieties", *Canad. J. Math.* **52**:5 (2000), 1018–1056. MR Zbl

[Sancho de Salas 2000] P. J. Sancho de Salas, "Automorphism scheme of a finite field extension", *Trans. Amer. Math. Soc.* **352**:2 (2000), 595–608. MR Zbl

[Serre 1979] J.-P. Serre, *Local fields*, Graduate Texts in Math. **67**, Springer, 1979. MR Zbl

[Serre 1997] J.-P. Serre, *Galois cohomology*, Springer, 1997. MR Zbl

[Serre 2003] J.-P. Serre, "Cohomological invariants, Witt invariants, and trace forms", pp. 1–100 in *Cohomological invariants in Galois cohomology*, Univ. Lecture Ser. **28**, Amer. Math. Soc., Providence, RI, 2003. MR

[Tossici and Vistoli 2013] D. Tossici and A. Vistoli, "On the essential dimension of infinitesimal group schemes", *Amer. J. Math.* **135**:1 (2013), 103–114. MR Zbl

[Waterhouse 1979] W. C. Waterhouse, *Introduction to affine group schemes*, Graduate Texts in Math. **66**, Springer, 1979. MR Zbl

Communicated by Susan Montgomery
Received 2018-06-28 Accepted 2018-12-24

reichst@math.ubc.ca *Department of Mathematics, University of British Columbia, Vancouver, BC, Canada*

abhisheks@math.ubc.ca *Department of Mathematics, University of British Columbia, Vancouver, BC, Canada*

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 13 No. 2 2019