Algebra & Number Theory

Volume 11
2017
No. 8

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2017 is US $325/year for the electronic version, and $520/year (+$55, 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 11:8 (2017)
dx.doi.org/10.2140/ant.2017.11.1739

On ℓ-torsion in class groups of number fields

Jordan Ellenberg, Lillian B. Pierce and Melanie Matchett Wood

For each integer $\ell \geq 1$, we prove an unconditional upper bound on the size of the ℓ-torsion subgroup of the class group, which holds for all but a zero-density set of field extensions of $\mathbb{Q}$ of degree d, for any fixed $d \in \{2, 3, 4, 5\}$ (with the additional restriction in the case $d = 4$ that the field be non-D_4). For sufficiently large ℓ (specified explicitly), these results are as strong as a previously known bound that is conditional on GRH. As part of our argument, we develop a probabilistic "Chebyshev sieve," and give uniform, power-saving error terms for the asymptotics of quartic (non-D_4) and quintic fields with chosen splitting types at a finite set of primes.

1. Introduction

The distribution of class groups is a great mystery. The Cohen–Lenstra heuristics [Cohen and Lenstra 1984] (for quadratic fields) and the Cohen–Lenstra–Martinet heuristics [Cohen and Martinet 1990] (for more general number fields) make predictions for the distribution of class groups, including for the average size of the ℓ-torsion subgroups for certain "good" primes ℓ. However, the questions of proving anything towards these predictions are almost entirely open, and mostly apparently inaccessible.

The main goal of the present work is to prove, for each integer $\ell \geq 1$, an unconditional upper bound for the size of the ℓ-torsion subgroup of the class group, which holds for all but a zero-density set of field extensions of $\mathbb{Q}$ of degree d, for any fixed $d \in \{2, 3, 4, 5\}$ (with the additional restriction in the case $d = 4$ that the field be non-D_4). Alternatively, these results may be viewed as the first unconditional upper bounds for the average size of ℓ-torsion in class groups as the field varies over extensions of $\mathbb{Q}$ of fixed degree $d \in \{2, 3, 4, 5\}$ (and non-D_4 in the case $d = 4$).

Let K be a degree d field extension of $\mathbb{Q}$ with absolute discriminant $D_K = |\operatorname{disc} K/\mathbb{Q}|$. We will denote the class group by Cl_K and the ℓ-torsion subgroup by

MSC2010: primary 11R29; secondary 11N36, 11R45.
Keywords: number fields, class groups, Cohen–Lenstra heuristics, sieves.

$\mathrm{Cl}_K[\ell]$. We note the trivial pointwise upper bound (see for example [Narkiewicz 1990, Theorem 4.4])

$$|\mathrm{Cl}_K[\ell]| \le |\mathrm{Cl}_K| \ll_{d,\varepsilon} D_K^{1/2+\varepsilon}, \tag{1-1}$$

for every $\varepsilon > 0$. (Throughout, $\varepsilon > 0$ is allowed to be arbitrarily small (possibly taking a different value in different occurrences), and $A \ll B$ indicates that $|A| \le cB$ for an implied constant c, which we allow in any instance to depend on ℓ, d, ε.)

It is conjectured that

$$|\mathrm{Cl}_K[\ell]| \ll D_K^{\varepsilon} \tag{1-2}$$

for every $\varepsilon > 0$, but improving on the trivial bound (1-1) has proved difficult. (Impetus for this conjecture may be found in [Duke 1998; Zhang 2005, page 10; Brumer and Silverman 1996, "Question CL(ℓ, d)"].) For K quadratic, Gauss's genus theory [1966] implies (1-2) in the case $\ell = 2$. Recently, Bhargava et al. [2017] obtained nontrivial upper bounds for 2-torsion in fields of degree d for all $d \ge 3$, proving $|\mathrm{Cl}_K[2]| \ll D_K^{0.2784\ldots+\varepsilon}$ for $d = 3, 4$ and $|\mathrm{Cl}_K[2]| \ll D_K^{1/2-1/2d+\varepsilon}$ for $d \ge 5$. For $\ell = 3$, after initial incremental improvement in [Helfgott and Venkatesh 2006; Pierce 2005; 2006] over the trivial bound (1-1) for quadratic fields, Ellenberg and Venkatesh [2007, Proposition 3.4, Corollary 3.7] proved that

$$|\mathrm{Cl}_K[3]| \ll D_K^{1/3+\varepsilon} \tag{1-3}$$

holds for both quadratic and cubic fields, and moreover there is a positive constant $\delta > 0$ such that

$$|\mathrm{Cl}_K[3]| \ll D_K^{1/2-\delta+\varepsilon} \tag{1-4}$$

holds for quartic fields. (In particular, one may take $\delta = 1/168$ in (1-4) for quartic fields with Galois closure having Galois group A_4 or S_4.) At this time, these are the best bounds in the literature that are unconditional and hold for all such fields.

In the realm of average results, there is little known, with the exceptions being spectacular successes. For 3-torsion in quadratic fields, Davenport and Heilbronn [1971] proved

$$\sum_{\substack{\deg(K)=2\\ 0<D_K\le X}} |\mathrm{Cl}_K[3]| \sim \Bigl(\frac{2}{3\zeta(2)} + \frac{1}{\zeta(2)}\Bigr)X, \tag{1-5}$$

in which the first contribution is from fields with disc $K/\mathbb{Q} > 0$ and the second is from fields with disc $K/\mathbb{Q} < 0$; this has recently been improved to reflect second order terms by [Bhargava et al. 2013; Taniguchi and Thorne 2013; Hough 2010]. For 2-torsion in cubic fields, Bhargava [2005] proved the asymptotic

$$\sum_{\substack{\deg(K)=3\\ 0<D_K\le X}} |\mathrm{Cl}_K[2]| \sim \Bigl(\frac{5}{48\zeta(3)} + \frac{3}{8\zeta(3)}\Bigr)X, \tag{1-6}$$

in which each isomorphism class of fields is counted once, and the first contribution is from fields with disc $K/\mathbb{Q} > 0$ and the second is from fields with disc $K/\mathbb{Q} < 0$. For 4-torsion in quadratic fields, Fouvry and Klüners [2007] have determined the asymptotics, for each nonnegative integer k,

$$\sum_{\substack{\deg(K)=2 \\ 0<D_K\le X}} |\mathrm{Cl}_K^2[2]|^k \sim (c_k + p^{-k}(c_{k+1} - c_k))X, \tag{1-7}$$

where c_k is the number of vector subspaces of $\mathbb{F}_2^k$. See also the recent work of Klys [2016] giving analogous results on 3-torsion in cyclic cubic fields, and the recent work of Milovic on 16-rank in quadratic fields, e.g., [Milovic 2017].

Turning to conditional results, Klys's results [2016] extend to p-torsion in cyclic degree p fields under GRH and Smith [2016] has results on 8-torsion averages in quadratic fields under GRH as well. In the case of quadratic fields, Wong [1999b] proved that, conditional on the Birch–Swinnerton-Dyer conjecture and the Riemann hypothesis, $|\mathrm{Cl}_K[3]| \ll D_K^{1/4+\varepsilon}$. Before the proof of (1-3), Soundararajan noted (as communicated in [Helfgott and Venkatesh 2006]) that one could prove $|\mathrm{Cl}_K[3]| \ll D_K^{1/3+\varepsilon}$ for K quadratic if one assumed the truth of the Riemann hypothesis for only the L-function $L(s,\chi)$ of the quadratic character χ associated to the quadratic field K. The key idea of the latter bound was the use of many small primes that split in K; the role of the Riemann hypothesis was to guarantee the existence of sufficiently many such primes. This approach has been generalized by Ellenberg and Venkatesh [2007] to number fields of any degree; we recall the key result in the special case of field extensions of $\mathbb{Q}$:

Theorem A [Ellenberg and Venkatesh 2007, Lemma 2.3]. *Let K be a field extension of $\mathbb{Q}$ of degree d, and let ℓ be a positive integer. Let $\delta < \frac{1}{2\ell(d-1)}$. Suppose that $\{\mathfrak{p}_1, \dots, \mathfrak{p}_M\}$ are M prime ideals in $\mathcal{O}_K$ with* $\mathrm{Norm}(\mathfrak{p}_j) \le D_K^{\delta}$ *that are unramified in $K/\mathbb{Q}$ and are not extensions of ideals from any proper subfield $K_0 \subsetneq K$. Then*

$$|\mathrm{Cl}_K[\ell]| \ll_{d,\ell,\varepsilon} D_K^{1/2+\varepsilon} M^{-1}. \tag{1-8}$$

(Here we recall the convention in [Ellenberg and Venkatesh 2007] that an ideal $\mathfrak{p}$ in $\mathcal{O}_K$ is said to be an extension of a prime ideal from a subfield $K_0 \subsetneq K$ if there is a prime ideal $\mathfrak{p}_0$ in $\mathcal{O}_{K_0}$ such that $\mathfrak{p} = \mathfrak{p}_0 \mathcal{O}_K$.)

Upon assuming GRH, an application of the effective Chebotarev theorem of Lagarias and Odlyzko [1977] guarantees, for any fixed $\eta > 0$, the existence of $\gg D_K^{\eta-\varepsilon}$ rational primes of size $\le D_K^{\eta}$ that split completely in K. Upon choosing $\eta = \frac{1}{2\ell(d-1)} - \varepsilon_0$ for arbitrarily small $\varepsilon_0 > 0$, one obtains the following bound, currently the state of the art for conditional pointwise upper bounds for $|\mathrm{Cl}_K[\ell]|$:

Theorem B [Ellenberg and Venkatesh 2007, Proposition 3.1]. *Let K be a field extension of $\mathbb{Q}$ of degree d and ℓ a positive integer. Assuming GRH,*

$$|\mathrm{Cl}_K[\ell]| \ll_{d,\ell,\varepsilon} D_K^{\frac{1}{2}-\frac{1}{2\ell(d-1)}+\varepsilon}, \tag{1-9}$$

for any $\varepsilon > 0$.

One may attempt to remove the conditionality by proving results that hold on average, or for all but a small exceptional family. In this vein, in the case of imaginary quadratic fields, Soundararajan [2000] noted that for all but at most one imaginary quadratic field K with $D_K \in [X, 2X]$, one has the bound $|\mathrm{Cl}_K[\ell]| \ll X^{1/2-1/2\ell+\varepsilon}$, for any prime ℓ. Also in the imaginary quadratic case, a recent result of Heath-Brown and Pierce [2014] provides an upper bound for averages (and in addition higher moments) of $|\mathrm{Cl}_K[\ell]|$, for example proving for any prime $\ell \geq 5$ that

$$\sideset{}{'}\sum_{\substack{\deg(K)=2\\ 0<D_K\leq X}} |\mathrm{Cl}_K[\ell]| \ll X^{\frac{3}{2}-\frac{3}{2\ell+2}+\varepsilon}, \tag{1-10}$$

with the sum restricted to imaginary quadratic fields.

In this paper, we prove *unconditional* results for $|\mathrm{Cl}_K[\ell]|$ that are as strong as (1-9) for all sufficiently large positive integers ℓ, and hold for all but a zero-density family of quadratic, cubic, non-D_4-quartic, or quintic field extensions of $\mathbb{Q}$.

For this we work with families of fields. Let $N_d(X)$ denote the number of degree d extensions of $\mathbb{Q}$ with $0 < D_K \leq X$, in which each isomorphism class is counted once; it is conjectured that for an appropriate constant c_d,

$$N_d(X) \sim c_d X. \tag{1-11}$$

Importantly for our work, this is known to be true for $d = 2$ (classical), $d = 3$ by Davenport and Heilbronn [1971], $d = 4$ by Cohen, Diaz y Diaz, and Olivier [Cohen et al. 2002] and Bhargava [2005], and $d = 5$ by Bhargava [2010]. Throughout our work, in the case of $d = 4$, we restrict our attention to non-D_4-quartic fields (that is, quartic extensions whose Galois closure does not have Galois group D_4); see the remark on page 1758. Thus we let $\widetilde{N}_4(X)$ denote the further restricted count of non-D_4-quartic extensions of $\mathbb{Q}$; then (1-11) is also known to hold for $\widetilde{N}_4(X)$, with a different constant [Bhargava 2005].

As a consequence of the field counts (1-11) combined with the trivial bound (1-1), a trivial average bound for $|\mathrm{Cl}_K[\ell]|$ is

$$\sum_{\substack{\deg(K)=d\\ 0<D_K\leq X}} |\mathrm{Cl}_K[\ell]| \ll_{d,\varepsilon} X^{3/2+\varepsilon}. \tag{1-12}$$

Our approach to improve upon (1-12) is to show that "most" degree d fields K contain sufficiently many small primes that split completely in K for Theorem A to give a good upper bound for $|\mathrm{Cl}_K[\ell]|$. Roughly speaking, we will show that there is some small $\delta_0 > 0$ such that for all but at most $O(X^{1-\delta_0})$ of the degree d fields K with $0 < D_K \le X$, at least a fixed positive proportion of the primes $p \le X^{\delta_0}$ split completely in K. (Under GRH, the small set of exceptional fields is in fact empty.)

Our main results are as follows:

Theorem 1.1. *Let $d \in \{2, 3, 4, 5\}$ and let ℓ be any positive integer with $\ell \ge \ell(d)$ where*

$$\ell(2) = \ell(3) = 1, \quad \ell(4) = 8, \quad \ell(5) = 25.$$

Then for all but $O_{d,\ell,\varepsilon}(X^{1-\frac{1}{2\ell(d-1)}+\varepsilon})$ degree d fields $K/\mathbb{Q}$ with $D_K \le X$ (and non-D_4 in the case $d = 4$),

$$|\mathrm{Cl}_K[\ell]| \ll_{d,\ell,\varepsilon} D_K^{\frac{1}{2}-\frac{1}{2\ell(d-1)}+\varepsilon},$$

for all $\varepsilon > 0$. For $d = 4, 5$, in the remaining cases of positive integers $\ell < \ell(d)$, for all but $O_{d,\varepsilon}(X^{1-\delta_0(d)+\varepsilon})$ degree d fields $K/\mathbb{Q}$ with $0 < D_K \le X$ (and non-D_4 in the case $d = 4$),

$$|\mathrm{Cl}_K[\ell]| \ll_{d,\ell,\varepsilon} D_K^{\frac{1}{2}-\delta_0(d)+\varepsilon},$$

for all $\varepsilon > 0$, where we may take

$$\delta_0(d) = \begin{cases} \frac{1}{48} & \text{if } d = 4, \\ \frac{1}{200} & \text{if } d = 5. \end{cases}$$

Remark. Theorem 2.1 states a version of this result in terms of bounding the number of exceptional fields that fail to have many small split primes. One notes from Theorem 2.1 that for sufficiently large ℓ, the limiting reagent is not the availability of small completely split primes, but the constraint $\delta < \frac{1}{2\ell(d-1)}$ in Theorem A.

As immediate corollaries, we note:

Corollary 1.1.1. *Let $d \in \{2, 3, 4, 5\}$. As K ranges over degree d extensions of $\mathbb{Q}$ with discriminant $0 < D_K \le X$ (and non-D_4 in the case $d = 4$),*

$$\sum_{\substack{\deg(K)=d \\ 0<D_K\le X}} |\mathrm{Cl}_K[\ell]| \ll_{d,\varepsilon} X^{\frac{3}{2}-\frac{1}{2\ell(d-1)}+\varepsilon},$$

for all integers $\ell \ge \ell(d)$, where $\ell(2) = \ell(3) = 1$, $\ell(4) = 8$, $\ell(5) = 25$.

Corollary 1.1.2. *For positive integers $\ell \le 7$, averaging over non-D_4-quartic fields,*

$$\sideset{}{'}\sum_{\substack{\deg(K)=4 \\ 0<D_K\le X}} |\mathrm{Cl}_K[\ell]| \ll_{d,\varepsilon} X^{\frac{3}{2}-\frac{1}{48}+\varepsilon}.$$

For positive integers $\ell \le 24$, *averaging over quintic fields*,

$$\sum_{\substack{\deg(K)=5\\ 0<D_K\le X}} |\mathrm{Cl}_K[\ell]| \ll_{d,\varepsilon} X^{\frac{3}{2}-\frac{1}{200}+\varepsilon}.$$

Our strategy is as follows. Recall that $N_d(X)$ denotes the number of degree d fields K over $\mathbb{Q}$, up to isomorphism, with $0 < D_K \le X$, and let $N_d(X; p)$ denote the number of degree d fields K over $\mathbb{Q}$, up to isomorphism, with $0 < D_K \le X$, such that the rational prime p splits completely in K. (For $d = 4$ we define $\widetilde{N}_4(X; p)$ analogously, restricting to non-D_4-quartic fields.) Suppose we know that for each fixed prime p, $N_d(X; p)$ is a positive proportion of $N_d(X)$, so p splits completely in a positive proportion of the fields. Then one would expect the fields in which the primes split completely to distribute somewhat evenly, so that "most fields" have the property that "near the average number" of primes split completely in them; that is, one would expect that the primes do not conspire to cause many fields to fail the criterion of Theorem A. We will make this argument precise by developing a flexible "Chebyshev sieve" (Lemma 3.1, related to Chebyshev's inequality); the crucial input to the sieve will be asymptotics for $N_d(X; p)$ with power-saving error and explicitly given dependence on p (Lemma 2.2, Theorem C, Theorems 2.3 and 2.4).

Counting quadratic fields may be accomplished by a simple classical argument (given in the Appendix). Power-saving error terms for $N_d(X)$ were first found in the cases $d = 3, 4$ by Belabas, Bhargava, and Pomerance [Belabas et al. 2010], and first found in the case $d = 5$ by Shankar and Tsimerman [2014]. In the case $d = 3$, Bhargava, Shankar, and Tsimerman [Bhargava et al. 2013] and Taniguchi and Thorne [2013] have also proved a second main term and improved the power-saving error term. For the refined estimates that we require on $N_d(X; p)$, we quote the necessary asymptotics for $d = 3$ from [Taniguchi and Thorne 2013], while for $d = 4, 5$ we prove the necessary estimates using the methods and results from [Belabas et al. 2010; Shankar and Tsimerman 2014]. In fact, in Sections 4 and 5, we give the field counting asymptotics for fields with any chosen splitting types at a finite set of primes with the expectation that they could be useful in other applications; see Theorems 4.1 and 5.1.

Our counting theorems improve upon analogous results that appear in four recent papers, three [Yang 2009; Cho and Kim 2015; Shankar et al. 2015] in the area of finding symmetry groups of families of L-functions (see [Sarnak et al. 2016] for a general overview of the area) and one [Lemke Oliver and Thorne 2017] studying the distribution of ramified primes in small-degree number fields. See Sections 4 and 5 for detailed comparisons to these previous works.

2. Anatomy of the proof

2A. ***Reduction to counting bad fields.*** We now outline the strategy in more detail; for the sake of motivation, we focus temporarily on proving upper bounds on average. Let us fix d and define for any degree d field K over $\mathbb{Q}$ and any real parameter $Y \geq 1$,

$$N(K;Y) = \#\{\text{rational primes } p \leq Y \text{ that split completely in } K\}.$$

(Implicitly, in the case $d = 4$ we further restrict to non-D_4-quartic fields.) Let us fix a positive integer ℓ and a parameter $\delta_1 < \frac{1}{2\ell(d-1)}$, to be chosen precisely later. Then by Theorem A, for any $X \geq 1$,

$$\begin{aligned}\sum_{X<D_K\leq 2X} |\mathrm{Cl}_K[\ell]| &\ll \sum_{X<D_K\leq 2X} D_K^{1/2+\varepsilon} N(K;D_K^{\delta_1})^{-1} \\ &\ll X^{1/2+\varepsilon} \sum_{X<D_K\leq 2X} N(K;X^{\delta_1})^{-1}.\end{aligned}$$

Now given real parameters $X \geq 1$ and $1 \leq M \leq Y$, we define $\mathscr{B}_d^0(X;Y,M)$ to be the set

$$\mathscr{B}_d^0(X;Y,M) = \big\{K/\mathbb{Q},\ \deg(K) = d,\ X < D_K \leq 2X : \\ \text{at most } M \text{ primes } p \leq Y \text{ split completely in } K\big\},$$

(with the usual further restriction in the case $d = 4$).

We denote by $\pi(Y)$ the number of rational primes $p \leq Y$, and let us regard $1 \leq M \leq \pi(X^{\delta_1})$ as fixed for the moment, to be specified later. Then we may make the decomposition

$$\begin{aligned}&\sum_{X<D_K\leq 2X} |\mathrm{Cl}_K[\ell]| \\ &\quad\ll X^{1/2+\varepsilon}\Bigg(\sum_{\substack{X<D_K\leq 2X\\ K\notin\mathscr{B}_d^0(X;X^{\delta_1},M)}} N(K;X^{\delta_1})^{-1} + \sum_{K\in\mathscr{B}_d^0(X;X^{\delta_1},M)} N(K;X^{\delta_1})^{-1}\Bigg).\end{aligned}$$

Since $N(K;X^{\delta_1}) \geq M$ if $K \notin \mathscr{B}_d^0(X;X^{\delta_1},M)$, we have

$$\sum_{X<D_K\leq 2X} |\mathrm{Cl}_K[\ell]| \ll X^{1/2+\varepsilon}\Bigg(\sum_{\substack{X<D_K\leq 2X\\ K\notin\mathscr{B}_d^0(X;X^{\delta_1},M)}} M^{-1} + \sum_{K\in\mathscr{B}_d^0(X;X^{\delta_1},M)} 1\Bigg),$$

and we may conclude that

$$\sum_{X<D_K\leq 2X} |\mathrm{Cl}_K[\ell]| \ll X^{3/2+\varepsilon}M^{-1} + \#\mathscr{B}_d^0(X;X^{\delta_1},M)X^{1/2+\varepsilon}. \tag{2-1}$$

Then upon defining the set

$$\mathscr{B}_d(X;Y,M)=\bigl\{K/\mathbb{Q},\ \deg(K)=d,\ 0<D_K\le X:\\ \text{at most } M \text{ primes } p\le Y \text{ split completely in } K\bigr\} \tag{2-2}$$

(with the usual further restriction in the case $d=4$), we may trivially replace the expression $\#\mathscr{B}_d^0(X;X^{\delta_1},M)$ in (2-1) by $\#\mathscr{B}_d(2X;X^{\delta_1},M)$ and only increase the right-hand side.

We now suppose that we can bound from above the cardinality of the "bad set" $\mathscr{B}_d(2X;X^{\delta_1},M)$ for appropriate δ_1 and M. Note that one expects via the Chebotarev density theorem that a positive proportion of the primes up to X^{δ_1} split completely in K, so that a reasonable choice for M will be proportional to $\pi(X^{\delta_1})$. Precisely, we suppose that there is a small fixed $\delta_2>0$ such that for every $X\ge 1$ and an appropriate choice of M with $X^{\delta_1}/\log X \ll M \ll X^{\delta_1}/\log X$ we have

$$\#\mathscr{B}_d(2X;X^{\delta_1},M)\ll X^{1-\delta_2+\varepsilon}, \tag{2-3}$$

for all $\varepsilon>0$. Then upon summing over $O(\log X)$ ranges and applying (2-1) and (2-3) within each range, we see that for any $X\ge 1$,

$$\begin{aligned}\sum_{0<D_K\le X}|\mathrm{Cl}_K[\ell]| &\le \sum_{0\le j\le\lceil\log_2 X\rceil}\ \sum_{2^{j-1}<D_K\le 2^j}|\mathrm{Cl}_K[\ell]|\\ &\ll \sum_{0\le j\le\lceil\log_2 X\rceil}\bigl\{(2^{j-1})^{3/2+\varepsilon}(2^{(j-1)\delta_1})^{-1}\log 2^j\\ &\qquad\qquad + \#\mathscr{B}_d(2^j;2^{(j-1)\delta_1},M)(2^{j-1})^{1/2+\varepsilon}\bigr\}\\ &\ll \log X\sum_{0\le j\le\lceil\log_2 X\rceil}\bigl\{(2^j)^{3/2-\delta_1+\varepsilon}+(2^j)^{3/2-\delta_2+2\varepsilon}\bigr\}\\ &\ll X^{3/2-\delta+3\varepsilon},\end{aligned} \tag{2-4}$$

where $\delta=\min\{\delta_1,\delta_2\}$ and $\varepsilon>0$ is arbitrarily small. Thus we see that an upper bound of the form (2-3) is the key to obtaining an average result in the shape of Corollaries 1.1.1 and 1.1.2; this upper bound plays a similarly crucial role in obtaining the results of Theorem 1.1, as we show in Section 7.

Ultimately, we will prove the following version of (2-3), which controls the number of possible bad fields:

Theorem 2.1. *Let $\mathscr{B}_d(X;Y,M)$ be defined as in* (2-2). *Set*

$$\delta_0(d)=\begin{cases}\frac{1}{6} & \textit{if } d=2,\\ \frac{2}{25} & \textit{if } d=3,\\ \frac{1}{48} & \textit{if } d=4,\\ \frac{1}{200} & \textit{if } d=5.\end{cases} \tag{2-5}$$

For each $d = 2, 3, 4, 5$, *there is a constant* $0 < c_0(d) < 1$ *such that for every* $0 < \delta \le \delta_0(d)$ *and every* $X \ge 1$,

$$\#\mathscr{B}_d\left(X; (X/2)^\delta,\ \tfrac{1}{2}c_0(d)\frac{(X/2)^\delta}{\log(X/2)^\delta}\right) \ll X^{1-\delta+\varepsilon}$$

for every $\varepsilon > 0$.

Remark. The methods of this paper also prove an analogous theorem if the condition "split completely" in the definition (2-2) is replaced by another fixed splitting type.

2B. ***Counting bad fields via a sieve and counts for fields with local conditions.*** We prove Theorem 2.1 via a sieve we develop for this purpose; to describe the strategy, we first recall the simplest classical setting of a sieve. Let $\mathscr{A}$ be a finite set of elements of cardinality N, and let $\mathscr{P}$ denote the set of all rational primes. We assume a certain property of interest has been specified so that each element $a \in \mathscr{A}$ either satisfies it or not, with respect to p, for each $p \in \mathscr{P}$. For each prime $p \in \mathscr{P}$ we let $\mathscr{A}_p$ denote the finite subset of $\mathscr{A}$ that satisfies the fixed property with respect to the prime p. Moreover we assume we know that for each p there exists a real number $0 \le \delta_p < 1$ and a real number R_p with $|R_p| \le N$ such that

$$\#\mathscr{A}_p = \delta_p N + R_p. \tag{2-6}$$

In simplest terms, a classical aim of a sieve is to provide an upper bound for the number of elements in the set $\mathscr{A}$ such that the designated property fails for all primes $p \le z$, for some fixed threshold z. Thus one could use a sieve to provide an upper bound for

$$\#\big(\mathscr{A} \setminus \bigcup_{p \le z} \mathscr{A}_p\big).$$

For example, to sieve for prime numbers, the set $\mathscr{A}$ is a finite set of integers, and the property is that $p|a$. Slightly more generally, one could apply a classical sieve such as the Turán sieve to count

$$\#\big(\mathscr{A} \setminus \bigcup_{p \in P_0} \mathscr{A}_p\big) \tag{2-7}$$

for an arbitrary fixed finite set of primes P_0.

In our application, the set $\mathscr{A}$ is the set of fields $K/\mathbb{Q}$ of degree d with $D_K \in (0, X]$ and the property is that p splits completely in K, so that $\mathscr{A}_p$ is the subset of fields in which the prime p splits completely. In this setting, assuming we possess an appropriate understanding of $\#\mathscr{A}_p$ as in (2-6), then (2-7) would allow us to count those degree d fields K with $D_K \in (0, X]$ in which a *fixed* set of primes fail to split completely. But in order to bound the bad set $\mathscr{B}_d(X; X^{\delta_1}, M)$ we require more flexibility: a field belongs to this set if all the primes in a sufficiently large set fail to split completely in K, but the relevant large set of primes might be different for two different bad fields K. Thus we develop in Section 3 a flexible new sieve that allows us to count elements $a \in \mathscr{A}$ that fail to lie in $\mathscr{A}_p$ for many p, without specifying which p fail for any given a.

The key input to any sieve is an understanding of $\mathscr{A}_p$ that provides the expression (2-6). In our case, this requires an understanding of $N_d(X)$, $N_d(X; p)$, and $N_d(X; pq)$ for two distinct primes p, q; here $N_d(X; pq)$ counts the number of degree d fields $K/\mathbb{Q}$ in which both p and q split completely. In the case of quartic fields, we let

$$\widetilde{N}_4(X), \quad \widetilde{N}_4(X; p) \quad \text{and} \quad \widetilde{N}_4(X; pq)$$

denote the analogous quantities, restricted to non-D_4-quartic fields $K/\mathbb{Q}$.

We now summarize the key results we will require for the sieve. For quadratic fields, we record:

Lemma 2.2. *There exists a constant $c_2 > 0$, such that for $e = e_1$ or $e = e_1e_2$ for distinct primes e_1, e_2,*

$$N_2(X) = c_2 X + O(X^{1/2}), \tag{2-8}$$

$$N_2(X; e) = \delta_e c_2 X + O(eX^{1/2}), \tag{2-9}$$

where δ_e is a multiplicative function defined for any prime e by

$$\delta_e = \frac{1}{2}\frac{1}{(1+e^{-1})}. \tag{2-10}$$

For completeness, we record a simple proof of this classical result in the Appendix; the error terms given here can be improved (see for example the survey [Pappalardi 2005]), but will suffice for our application.

In contrast, the results for cubic, quartic, and quintic fields are deep. For cubic fields, we cite:

Theorem C [Taniguchi and Thorne 2013, Theorems 1.1, 1.3]. *There exist constants $c_3 > 0$, $c_3' < 0$ such that for $e = e_1$ or $e = e_1e_2$ for distinct primes e_1, e_2,*

$$N_3(X) = c_3 X + c_3' X^{5/6} + O(X^{7/9+\varepsilon}), \tag{2-11}$$

$$N_3(X; e) = \delta_e c_3 X + \delta_e' c_3' X^{5/6} + O(e^{8/9} X^{7/9+\varepsilon}), \tag{2-12}$$

where δ_e and δ_e' are multiplicative functions defined for any prime e by

$$\delta_e = \frac{1}{6}\frac{1}{(1+e^{-1}+e^{-2})}, \quad \delta_e' = \frac{1}{6} + O(e^{-1/3}). \tag{2-13}$$

For quartic fields, we have:

Theorem 2.3. *There exists a constant $c_4 > 0$ such that for $e = e_1$ or $e = e_1 e_2$ for distinct primes e_1, e_2,*

$$\widetilde{N}_4(X) = c_4 X + O(X^{23/24+\varepsilon}), \tag{2-14}$$

$$\widetilde{N}_4(X; e) = \delta_e c_4 X + O(e^{1/2+\varepsilon} X^{23/24+\varepsilon}), \tag{2-15}$$

where δ_e is a multiplicative function defined for any prime e by

$$\delta_e = \frac{1}{24} \frac{1}{(1+e^{-1}+2e^{-2}+e^{-3})}. \tag{2-16}$$

We note (2-14) is due to [Belabas et al. 2010, Theorem 1.3]; we deduce (2-15) in Section 4, using the methods of Belabas, Bhargava, and Pomerance [Belabas et al. 2010], which build on the work of Bhargava [2005] that obtained the original count of S_4-quartic fields with an $o(X)$ error term. See Theorem 4.1 for our most general result of this type, of which Theorem 2.3 is a special case.

For quintic fields, we have:

Theorem 2.4. *There exists a constant $c_5 > 0$ such that for $e = e_1$ or $e = e_1 e_2$ for distinct primes e_1, e_2,*

$$N_5(X) = c_5 X + O(X^{199/200+\varepsilon}), \tag{2-17}$$

$$N_5(X; e) = \delta_e c_5 X + O(e^{1/2+\varepsilon} X^{79/80+\varepsilon} + X^{199/200+\varepsilon}), \tag{2-18}$$

where δ_e is a multiplicative function defined for any prime e by

$$\delta_e = \frac{1}{120} \frac{1}{(1+e^{-1}+2e^{-2}+2e^{-3}+e^{-4})}. \tag{2-19}$$

We note (2-17) is due to Shankar and Tsimerman [2014]; we deduce (2-18) in Section 5, using their methods, which build on the work of Bhargava [2010] that obtained the original count of S_5-quintic fields with an $o(X)$ error term. (We also fill in a missing step from [Shankar and Tsimerman 2014].) See Theorem 5.1 for our most general result, of which Theorem 2.4 is a special case.

We remark that the techniques for counting number fields that produced these results for $N_d(X; e)$ continue to be refined, and we may expect that the error terms will continue to be reduced. Thus in our subsequent computations involving $N_d(X; e)$ we have worked more generally with error terms of the form $O(e^{\sigma} X^{\tau})$, so that it will be immediately clear how improvements in counting fields will lead to refinements of our results. (In particular, improved error terms for *smoothed* versions of the counting functions $N_d(X; e)$ would suffice for our application.) We note that the mechanism we employ will apply equally well to higher degree extensions of $\mathbb{Q}$ (or extensions of a fixed number field, using the more general form of Theorem A available in [Ellenberg and Venkatesh 2007, Lemma 2.3]) if suitable results for $N_d(X)$ and $N_d(X; e)$ (or their analogues for extensions of a fixed number field) become available. In addition, one might consider other families of fields for which precise asymptotics are known, such as abelian fields over $\mathbb{Q}$ with a fixed Galois group, ordered either by discriminant [Wright 1989; Frei et al. 2015] or by conductor [Wood 2010]. It would be an interesting question to see whether the existing methods can be refined to produce an appropriate power-saving error term with sufficiently explicit dependence on a finite number of local conditions.

3. The Chebyshev sieve

We now develop in a fully general setting a new sieve that allows us to give an upper bound for the number of elements a belonging to a set $\mathscr{A}$ that satisfy a desired property with respect to p for "few" p (without specifying for which p it is satisfied). We will see that the principal idea is probabilistic, relating to Chebyshev's inequality, thus we dub it the Chebyshev sieve.

As before, let $\mathscr{A}$ be a finite set of cardinality N, let $\mathscr{P}$ denote the set of all rational primes, and let $\mathscr{A}_p$ denote the finite subset of $\mathscr{A}$ that satisfies the fixed property with respect to the prime p. For a fixed real parameter $z \geq 1$, we let

$$P(z) = \prod_{\substack{p \in \mathscr{P} \\ p \leq z}} p$$

and we define for each $a \in \mathscr{A}$ the quantity

$$N(a) = \#\{p : p \mid P(z),\ a \in \mathscr{A}_p\}.$$

Next, we set

$$M(z) = \frac{1}{N} \sum_{a \in \mathscr{A}} N(a) = \frac{1}{N} \sum_{p \mid P(z)} \#\mathscr{A}_p \tag{3-1}$$

to be the mean number of sets $\mathscr{A}_p$ (with $p \leq z$) to which a typical element $a \in \mathscr{A}$ belongs. (In nonvacuous cases, $M(z)$ is nonzero.) We would expect that a typical element $a \in \mathscr{A}$ has $N(a)$ being about size $M(z)$, and we want to bound from above the number of $a \in \mathscr{A}$ which have $N(a)$ being unusually small, that is, less than a fixed small proportion of $M(z)$.

Given $1 \leq M \leq z$, we define $\mathscr{E}(\mathscr{A}; z, M)$ to be the set of elements $a \in \mathscr{A}$ such that at most M primes $p \mid P(z)$ have $a \in \mathscr{A}_p$. (Or in other words, $\mathscr{E}(\mathscr{A}; z, M)$ is the set of elements $a \in \mathscr{A}$ such that $N(a) \leq M$.) Then we set

$$E(\mathscr{A}; z, M) = \#\mathscr{E}(\mathscr{A}; z, M).$$

Our sieve lemma will provide us with an upper bound for $E(\mathscr{A}; z, \frac{1}{2} M(z))$; that is, the number of elements in $\mathscr{A}$ that lie in $\mathscr{A}_p$ for fewer than half the mean number of p.

For the purposes of the lemma, we introduce the following notation. Given distinct primes p, q we let $\mathscr{A}_{pq} = \mathscr{A}_p \cap \mathscr{A}_q$, and let $R_{p,q}$ denote the quantity such that

$$\#\mathscr{A}_{pq} = \delta_p \delta_q N + R_{p,q}.$$

(For notational convenience, we will interpret $R_{p,p}$ as R_p.) Finally, we set

$$U(z) = \sum_{p \mid P(z)} \delta_p.$$

We now state the key sieve lemma.

Lemma 3.1 (Chebyshev sieve). *With the setting described above*,

$$E(\mathscr{A}; z, \tfrac{1}{2}M(z)) \\ \le \frac{4N}{M(z)^2}\left(U(z) + \frac{1}{N}\sum_{p,q \mid P(z)} |R_{p,q}| + \frac{2U(z)}{N}\sum_{p \mid P(z)} |R_p| + \left(\frac{1}{N}\sum_{p \mid P(z)} |R_p|\right)^2\right).$$

3A. ***Proof of the sieve lemma.*** We note that the sieve inequality we prove is related to the classical Turán sieve (see for example Theorem 4.1.1 of [Cojocaru and Murty 2006]), and can be seen as an application of Chebyshev's inequality

$$\mathbb{P}(|X - \mu| \ge \alpha) \le \sigma^2/\alpha^2,$$

for X a random variable with mean μ and variance σ^2, applied to the random variable $N(a)$ when a is drawn uniformly from $\mathscr{A}$.

We prove the lemma directly. We begin by noting that

$$\begin{aligned}\frac{1}{N}E\big(\mathscr{A}; z, \tfrac{1}{2}M(z)\big)\big(\tfrac{1}{2}M(z)\big)^2 &\le \frac{1}{N}\sum_{a \in \mathscr{E}(\mathscr{A}; z, \frac{1}{2}M(z))} (N(a) - M(z))^2 \\ &\le \frac{1}{N}\sum_{a \in \mathscr{A}} (N(a) - M(z))^2.\end{aligned}$$

It then suffices to prove the variance term on the right-hand side satisfies

$$\frac{1}{N}\sum_{a \in \mathscr{A}} (N(a) - M(z))^2 \\ \le U(z) + \frac{1}{N}\sum_{p,q \mid P(z)} |R_{p,q}| + 2U(z)\left(\frac{1}{N}\sum_{p \mid P(z)} |R_p|\right) + \left(\frac{1}{N}\sum_{p \mid P(z)} |R_p|\right)^2. \tag{3-2}$$

We first note from (3-1) that the mean satisfies

$$M(z) = \frac{1}{N}\sum_{p \mid P(z)} \#\mathscr{A}_p = \frac{1}{N}\sum_{p \mid P(z)} (\delta_p N + R_p) = U(z) + \frac{1}{N}\sum_{p \mid P(z)} R_p. \tag{3-3}$$

We now consider the left-hand side of (3-2), which we trivially expand as

$$\frac{1}{N}\sum_{a\in\mathscr{A}} N(a)^2 - \frac{2}{N}\sum_{a\in\mathscr{A}} N(a)M(z) + M(z)^2 = \frac{1}{N}\sum_{a\in\mathscr{A}} N(a)^2 - M(z)^2. \quad (3\text{-}4)$$

The first term on the right-hand side of (3-4) is equal to

$$\begin{aligned}\frac{1}{N}\sum_{p,q\mid P(z)} \#(\mathscr{A}_p\cap\mathscr{A}_q) &= \frac{1}{N}\Bigg(\sum_{p\mid P(z)}\delta_p N + \sum_{\substack{p,q\mid P(z)\\ p\neq q}}\delta_p\delta_q N + \sum_{p,q\mid P(z)} R_{p,q}\Bigg)\\ &= \sum_{p\mid P(z)}\delta_p + \Bigg(\sum_{p\mid P(z)}\delta_p\Bigg)^2 - \sum_{p\mid P(z)}\delta_p^2 + \frac{1}{N}\sum_{p,q\mid P(z)} R_{p,q}\\ &= \sum_{p\mid P(z)}\delta_p(1-\delta_p) + U(z)^2 + \frac{1}{N}\sum_{p,q\mid P(z)} R_{p,q}.\end{aligned}$$

On the other hand, we may expand $M(z)^2$ via (3-3) and see that after cancellation of the $U(z)^2$ factor, the right-hand side of (3-4) is equal to

$$\sum_{p\mid P(z)}\delta_p(1-\delta_p) + \frac{1}{N}\sum_{p,q\mid P(z)} R_{p,q} - 2U(z)\Bigg(\frac{1}{N}\sum_{p\mid P(z)} R_p\Bigg) - \Bigg(\frac{1}{N}\sum_{p\mid P(z)} R_p\Bigg)^2.$$

As R_p may be either positive or negative, we take absolute values; then using the fact that $\delta_p \le 1$ we see the resulting inequality simplifies to (3-2), thus proving the lemma.

4. Asymptotic count of non-D_4-quartic fields

In this section we will prove the following, of which Theorem 2.3 is a special case.

Theorem 4.1. *Let P be a finite set of primes. For each prime $p \in P$ we choose a splitting type at p and assign a corresponding density as follows*:

$$\begin{aligned}
\delta_p &:= \tfrac{1}{24}(1+p^{-1}+2p^{-2}+p^{-3})^{-1} && \text{for } p = \wp_1\wp_2\wp_3\wp_4,\\
\delta_p &:= \tfrac{1}{4}(1+p^{-1}+2p^{-2}+p^{-3})^{-1} && \text{for } p = \wp_1\wp_2\wp_3,\\
\delta_p &:= \tfrac{1}{3}(1+p^{-1}+2p^{-2}+p^{-3})^{-1} && \text{for } p = \wp_1\wp_2 \text{ with } \wp_2 \text{ inertia degree } 3,\\
\delta_p &:= \tfrac{1}{8}(1+p^{-1}+2p^{-2}+p^{-3})^{-1} && \text{for } p = \wp_1\wp_2 \text{ with } \wp_i \text{ inertia degree } 2,\\
\delta_p &:= \tfrac{1}{4}(1+p^{-1}+2p^{-2}+p^{-3})^{-1} && \text{for } p = \wp_1,\\
\delta_p &:= \frac{p^{-1}+2p^{-2}+p^{-3}}{(1+p^{-1}+2p^{-2}+p^{-3})} && \text{for } p \text{ ramified.}
\end{aligned}$$

Let $\delta_P := \prod_{p\in P}\delta_p$ and let $e = \prod_{p\in P} p$. Let $\widetilde{N}_4(X; P)$ be the number of non-D_4 quartic fields with absolute discriminant at most X such that for each $p \in P$, the *prime p splits in the quartic field in the splitting type chosen for p above. There exists a constant* $c_4 > 0$ *such that*

$$\widetilde{N}_4(X; P) = \delta_P c_4 X + O(e^{1/2+\varepsilon} X^{23/24+\varepsilon}), \tag{4-1}$$

where the implied constant in the O term is absolute (*does not depend on P*). *Moreover, we may choose more than one splitting type at each prime and let* δ_p *be the sum of the corresponding densities and the result still holds.*

Bhargava [2005] first determined the asymptotic count of non-D_4-quartic fields, and Belabas, Bhargava, and Pomerance [Belabas et al. 2010] gave a power-saving asymptotic for this count. We will follow the method of [Belabas et al. 2010], additionally requiring our chosen splitting types. While the main term for such a restricted count appears in [Bhargava 2005, Theorem 3] (at least for one prime, and the same argument would work for more primes), we require a power-saving error term *with explicit dependence* on the primes. In fact, such results have appeared at least four times recently, but we will improve upon the exponents in all of these results and remove various hypotheses that don't hold in the situation in which we need to apply the bound. Yang [2009, Proposition 3.1.7] proved such a power-saving error of the form $\widetilde{N}_4(X; P) = \delta_P c_4 X + O(e^2 X^{143/144+\varepsilon})$. ([Yang 2009, Proposition 3.1.7] only states this for one local condition, but [Cho and Kim 2015, Section 7] remarked it can be extended to finitely many local conditions.) Lemke Oliver and Thorne [2017, Theorem 2.1] proved a power-saving error (in which we may only specify that p is ramified) of $\widetilde{N}_4(X; P) = \delta_P c_4 X + O(e^{9/10} X^{239/240+\varepsilon})$. Shankar, Södergren, and Templier [Shankar et al. 2015] proved $\widetilde{N}_4(X; P) = \delta_P c_4 X + O(e^{12} X^{23/24+\varepsilon})$ when P contains a single prime.

The exposition of the method in [Bhargava 2005; Belabas et al. 2010] is quite clear, so we will focus here on the particular aspects of the computation we need. Instead of directly counting quartic fields, the method, equivalently, counts maximal quartic orders. The parametrization of quartic rings with their cubic resolvents due to Bhargava [2004] (see also [Belabas et al. 2010, Theorem 4.1]) gives an injection from the set of isomorphism classes of maximal quartic orders to the set of $\mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$ classes of pairs of ternary quadratic forms with integral coefficients. Pairs of integral ternary quadratic forms comprise a 12 dimensional lattice $V_{\mathbb{Z}} = \mathbb{Z}^{12}$. Counting $\mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$ classes of lattice points in $\mathbb{Z}^{12}$ is the same as counting lattice points in a fundamental domain for $\mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$ on $\mathbb{R}^{12}$. In this paper, we need to count only these lattice points in particular translates of sublattices of $\mathbb{Z}^{12}$. We collect some basic facts about the lattice translates corresponding to our desired fields, apply the geometry of numbers result from [Belabas et al. 2010] to count the necessary lattice points, and then work to minimize the resulting error terms.

As in [Bhargava 2005, Section 2.2] and [Belabas et al. 2010, Section 4] we use a certain random fundamental domain for the action of $\mathrm{GL}_2(\mathbb{Z})\times\mathrm{SL}_3(\mathbb{Z})$ on $\mathbb{R}^{12}$. For a positive integer m, let L be a translate $v+mV_{\mathbb{Z}}$ ($v\in V_{\mathbb{Z}}$) of the sublattice $mV_{\mathbb{Z}}$ of $V_{\mathbb{Z}}$. Let $N'(L;X)$ denote the expected number of lattice points in L, with first coordinate nonzero and discriminant less than X, in a random fundamental domain. (This notion of expected value for a random fundamental domain is defined as in [Bhargava 2005, Equation (5)], with S the set of points of L with first coordinate nonzero, but *without* the "abs. irr." condition that appears in [Bhargava 2005, Equation (5)]. See also [Belabas et al. 2010, p. 198].) Let $N_{S_4}(q;X)$ be the number of classes in $V_{\mathbb{Z}}$ corresponding to isomorphism classes of S_4-quartic orders and whose index in their maximal order is divisible by q and whose discriminant is less than X. We have the following result that estimates these counts.

Theorem D [Belabas et al. 2010, Theorem 4.11]. *Let L be a translate $v+mV_{\mathbb{Z}}$ ($v\in V_{\mathbb{Z}}$). Let (a,b,c,d) denote the smallest positive first four coordinates of any element of L. Then*

$$N'(L,X)=\frac{N_{S_4}(1;X)}{m^{12}}+O\biggl(\sum_{S}\frac{X^{(|S|+\alpha_S+\beta_S+\gamma_S+\delta_S)/12}}{m^{|S|}a^{\alpha_S}b^{\beta_S}c^{\gamma_S}d^{\delta_S}}+\log X\biggr),$$

where S ranges over the nonempty proper subsets of the set of 12 coordinates on $V_{\mathbb{Z}}$, and $\alpha_S,\beta_S,\gamma_S,\delta_S\in[0,1]$ are real constants that depend only on S and satisfy $|S|+\alpha_S+\beta_S+\gamma_S+\delta_S\le 11$.

Let q be square-free and $(q,e)=1$. First, we will assume that we have chosen unramified splitting types at each prime in P. Now, we will start by counting the expected number $N'(q,e;X)$ of lattice points in a random fundamental domain that satisfy the following conditions: (1) their first coordinate is nonzero, (2) their discriminant is less than X, (3) their corresponding quartic ring is not maximal at each prime dividing q and is maximal and of chosen splitting type at primes in P. We do this by summing Theorem D over the collection T of translates of $eq^2V_{\mathbb{Z}}$ that give quartic rings that are not maximal at each prime dividing q, and are maximal and with chosen local splitting at each $p\in P$. (See [Bhargava 2004, Section 4] for a description of which pairs of ternary quadratic forms correspond to quartic rings that are maximal or split in a certain way at a prime.)

Given $(a,b,c,d)\in[1,eq^2]^4$, we need to bound the number of translates in T that have (a,b,c,d) as the smallest positive first four coordinates of any element. By [Belabas et al. 2010, Corollary 4.8], there are $O(6^{\omega(q)}q^{14})$ translates of $q^2V_{\mathbb{Z}}$ that are congruent to (a,b,c,d) modulo q^2 and whose lattice points correspond to quartic rings that are not maximal at each prime dividing q. Since $V_{\mathbb{Z}}$ is 12 dimensional, there are e^8 translates of $eV_{\mathbb{Z}}$ congruent to (a,b,c,d) modulo e. Thus by the Chinese Remainder Theorem, there are $O(6^{\omega(q)}q^{14}e^8)$ translates in T that have (a,b,c,d) as the smallest positive first four coordinates of any element.

For q square-free, we define $\nu(q)$ to be the multiplicative function defined for a prime p by

$$\nu(p) := p^{-2}+2p^{-3}+2p^{-4}-3p^{-5}-4p^{-6}-p^{-7}+3p^{-8}+3p^{-9}-p^{-10}-p^{-11}.$$

This is the density of lattice points that correspond to quartic rings nonmaximal at p [Belabas et al. 2010, Lemma 4.4]. Then $\#T = \nu(q)q^{24}e^{12}\Gamma_{\boldsymbol{P}}$, where

$$\Gamma_{\boldsymbol{P}} := \prod_{p\in \boldsymbol{P}} \delta_p(1-\nu(p)),$$

and $0 \le \delta_p \le 1$ is the density of lattice points corresponding to quartic rings that are split as we chose at p as a subset of those corresponding to quartic rings that are maximal at p [Bhargava 2004, Lemma 23].

If $q^2 > X$, then all the classes counted by $N'(q,e;X)$ have discriminant 0, and by [Belabas et al. 2010, Lemma 4.10], in this case there are $O(X^{11/12+\varepsilon})$ such classes.

So now we consider the case when $q^2 \le X$, in which case, using the shorthand

$$\boldsymbol{\mu}_S = |S| + \alpha_S + \beta_S + \gamma_S + \delta_S \quad \text{and} \quad \boldsymbol{\varepsilon}_S = a^{\alpha_S} b^{\beta_S} c^{\gamma_S} d^{\delta_S},$$

by Theorem D,

$$N'(q,e;X)$$
$$= \nu(q)\Gamma_{\boldsymbol{P}} N_{S_4}(1;X) + O\biggl(\sum_{(a,b,c,d)\in[1,eq^2]^4} 6^{\omega(q)}q^{14}e^8\biggl(\sum_S \frac{X^{\boldsymbol{\mu}_S/12}}{(eq^2)^{|S|}\boldsymbol{\varepsilon}_S} + \log X\biggr)\biggr).$$

We have

$$\begin{aligned}
&\sum_{(a,b,c,d)\in[1,eq^2]^4} 6^{\omega(q)}q^{14}e^8\biggl(\sum_S \frac{X^{\boldsymbol{\mu}_S/12}}{(eq^2)^{|S|}\boldsymbol{\varepsilon}_S} + \log X\biggr)\\
&\quad = 6^{\omega(q)}q^{14}e^8\biggl(e^4q^8\log X + \sum_S \frac{X^{\boldsymbol{\mu}_S/12}}{(eq^2)^{|S|}} \sum_{(a,b,c,d)\in[1,eq^2]^4} \frac{1}{\boldsymbol{\varepsilon}_S}\biggr)\\
&\quad \le 6^{\omega(q)}q^{14}e^8\biggl(e^4q^8\log X + \sum_S \frac{X^{\boldsymbol{\mu}_S/12}}{(eq^2)^{|S|}}\bigl((eq^2)^{4-\alpha_S-\beta_S-\gamma_S-\delta_S}\log^4(eq^2)\bigr)\biggr)\\
&\quad = 6^{\omega(q)}q^{22}e^{12}\biggl(\log X + \sum_S (X^{1/12}e^{-1}q^{-2})^{\boldsymbol{\mu}_S}\log^4(eq^2)\biggr).
\end{aligned}$$

Since $0 \le \boldsymbol{\mu}_S \le 11$, and recalling that $q^2 \le X$, the above is

$$\begin{aligned} &= O\big(6^{\omega(q)} q^{22} e^{12} \big((X^{1/12} e^{-1} q^{-2})^{11} \log^4(eq^2) + \log^4(eq^2) + \log X\big)\big) \\ &= O\big(e^{1+\varepsilon} X^{11/12+\varepsilon} + q^{22} e^{12+\varepsilon} X^{\varepsilon}\big). \end{aligned}$$

Let $N_{S_4}(q, e; X)$ be the number of classes in $V_{\mathbb{Z}}$, or equivalently lattice points in a fundamental domain, corresponding to isomorphism classes of S_4-quartic orders, whose index in their maximal order is divisible by q and whose discriminant is less than X, and that are maximal and of chosen splitting type at $p \in P$. Now, by inclusion-exclusion, as in the proof of [Belabas et al. 2010, Theorem 4.13], we have that the number of isomorphism classes of maximal S_4-quartic orders splitting as chosen for $p \in P$ and having (absolute) discriminant less than X is given by

$$\sum_{q \ge 1}{}' \mu(q) N_{S_4}(q, e; X)$$

where the sum is restricted to square-free q that are relatively prime to e.

Now we compare $N_{S_4}(q, e; X)$ and $N'(q, e; X)$. Note that the difference is that $N'(q, e; X)$ excludes those lattice points with first coordinate 0, and $N_{S_4}(q, e; X)$ excludes those lattice points that do not correspond to orders in S_4-quartic fields. So by [Belabas et al. 2010, Lemmas 4.9 and 4.10], we have

$$|N_{S_4}(q, e; X) - N'(q, e; X)| = O(X^{11/12+\varepsilon}).$$

Thus by our previous computation for $N'(q, e; X)$,

$$N_{S_4}(q, e; X) = \nu(q) \Gamma_P N_{S_4}(1; X) + O\big(e^{1+\varepsilon} X^{11/12+\varepsilon} + q^{22} e^{12+\varepsilon} X^{\varepsilon}\big). \tag{4-2}$$

So for a fixed Q (to be chosen in terms of X, e later), we sum over square-free q with $(q, e) = 1$ as in (4-2), obtaining

$$\begin{aligned} &\sum_{q \ge 1}{}' \mu(q) N_{S_4}(q, e; X) \\ &\qquad = \sum_{1 \le q \le Q}{}' \mu(q) N_{S_4}(q, e; X) + \sum_{q > Q}{}' \mu(q) N_{S_4}(q, e; X) \\ &\qquad = \sum_{1 \le q \le Q}{}' \mu(q) \nu(q) \Gamma_P N_{S_4}(1; X) + O(E_1) + O(E_2) \\ &\qquad = \sum_{q \ge 1}{}' \mu(q) \nu(q) \Gamma_P N_{S_4}(1; X) + O(E_1) + O(E_2) + O(E_3) \\ &\qquad = \prod_p (1 - \nu(p)) \prod_{p \in P} \delta_p N_{S_4}(1; X) + O(E_1) + O(E_2) + O(E_3), \end{aligned}$$

where

$$E_1 = \sum_{1\le q\le Q}{}' \big(e^{1+\varepsilon}X^{11/12+\varepsilon} + q^{22}e^{12+\varepsilon}X^{\varepsilon}\big),$$
$$E_2 = \sum_{q>Q}{}' \mu(q)N_{S_4}(q,e;X),$$
$$E_3 = \sum_{q>Q}{}' \nu(q)\Gamma_P N_{S_4}(1;X).$$

(Note that we handle the terms slightly differently than in [Belabas et al. 2010], so that E_3 above does not correspond to their E_3 term.)

We have $E_1 = O(e^{1+\varepsilon}QX^{11/12+\varepsilon} + Q^{23}e^{12+\varepsilon}X^{\varepsilon})$. By [Belabas et al. 2010, Lemma 4.3], we have $N_{S_4}(q,e;X) = O(Xq^{-2+\varepsilon})$, and so $E_2 = O(XQ^{-1+\varepsilon})$. We have $E_3 = O(Q^{-1+\varepsilon}X)$, since by [Belabas et al. 2010, Lemma 4.2], we have $N_{S_4}(1;X) = O(X)$, and by definition $\nu(q) = O(q^{-2+\varepsilon})$.

If $e \le X^{1/12}$, then we take $Q = X^{1/24}e^{-1/2}$, and we have

$$\sum_{q\ge 1}{}' \mu(q)N_{S_4}(q,e;X) = \prod_p(1-\nu(p))\prod_{p\in P}\delta_p N_{S_4}(1;X) + O(e^{1/2+\varepsilon}X^{23/24+\varepsilon}).$$

By [Belabas et al. 2010, Lemma 4.2], we have that

$$\prod_p(1-\nu(p))N_{S_4}(1;X) = c_4X + O(X^{23/24+\varepsilon}),$$

for some positive constant c_4. Thus we conclude that the number of isomorphism classes of maximal S_4-quartic orders with our chosen splitting types at $p\in P$ and having (absolute) discriminant less than X is

$$\delta_P c_4 X + O(e^{1/2+\varepsilon}X^{23/24+\varepsilon}).$$

If $e > X^{1/12}$, then the number of isomorphism classes of maximal S_4-quartic orders with chosen splitting types for $p\in P$ and having (absolute) discriminant less than X is $O(X)$ by [Belabas et al. 2010, Lemma 4.2], which we may then also write as

$$\delta_P c_4 X + O(e^{1/2+\varepsilon}X^{23/24+\varepsilon}).$$

There are at most $O(X^{7/8+\varepsilon})$ quartic extensions with $D_K < X$ with Galois closure having Galois group C_4, K_4 or A_4 [Baily 1980; Wong 1999a]. So we can conclude Theorem 4.1 holds for unramified splitting types. This argument shows we can also choose more than one splitting type at each p, and sum the corresponding densities.

Now, given P and choices for local splitting types some of which may be ramified, let P_1 be the subset of P for which we choose only unramified splitting types. We can find $\widetilde{N}_4(X; P_1)$ using the result already proven. For any subset $P_2 \subset P \setminus P_1$, write $\widetilde{N}_4(X; P_1 \cup \bar{P}_2)$ for the number of non-D_4 quartic fields with absolute discriminant at most X such that for each $p \in P_1$ the prime p splits in one of our chosen spitting type, and for each $p \in P_2$ the prime p *does not* split in one of our chosen splitting types. We can also apply the result already proven to find $\widetilde{N}_4(X; P_1 \cup \bar{P}_2)$. Then using inclusion exclusion, we have

$$\begin{aligned}\widetilde{N}_4(X; P) &= \sum_{P_2 \subset P \setminus P_1} (-1)^{|P_2|} \widetilde{N}_4(X; P_1 \cup \bar{P}_2) \\ &= \delta_P c_4 X + \sum_{P_2 \subset P \setminus P_1} (-1)^{|P_2|} O(e^{1/2+\varepsilon} X^{23/24+\varepsilon}).\end{aligned}$$

Since each set P_2 corresponds to a distinct divisor of e there are $O(e^\varepsilon)$ terms in the sum and Theorem 4.1 follows.

Remark. On the other hand, the number of D_4-quartic fields with $D_K < X$ is $\sim cX$ with $c \approx 0.052326$, as initially indicated (as an order of magnitude) by Baily [1980] and refined with an explicit constant by Cohen, Diaz y Diaz and Olivier [Cohen et al. 2002]. It is an interesting open problem to count D_4 fields with local conditions such as certain primes being split completely, and for now we exclude them from our consideration.

5. Asymptotic count of quintic fields

In this section we will prove the following, of which Theorem 2.4 is a special case.

Theorem 5.1. *Let P be a finite set of primes. For each prime $p \in P$ we choose a splitting type at p and assign a corresponding density as follows* (*i.d. = inertia degree*):

$$\begin{aligned}
\delta_p &:= \tfrac{1}{120}(1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4})^{-1} && \text{for } p = \wp_1\wp_2\wp_3\wp_4\wp_5,\\
\delta_p &:= \tfrac{1}{12}(1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4})^{-1} && \text{for } p = \wp_1\wp_2\wp_3\wp_4,\\
\delta_p &:= \tfrac{1}{8}(1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4})^{-1} && \text{for } p = \wp_1\wp_2\wp_3 \text{ with } \wp_2, \wp_3 \text{ i.d. } 2,\\
\delta_p &:= \tfrac{1}{6}(1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4})^{-1} && \text{for } p = \wp_1\wp_2\wp_3 \text{ with } \wp_3 \text{ i.d. } 3,\\
\delta_p &:= \tfrac{1}{6}(1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4})^{-1} && \text{for } p = \wp_1\wp_2 \text{ with } \wp_2 \text{ i.d. } 3,\\
\delta_p &:= \tfrac{1}{4}(1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4})^{-1} && \text{for } p = \wp_1\wp_2 \text{ with } \wp_2 \text{ i.d. } 4,\\
\delta_p &:= \tfrac{1}{5}(1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4})^{-1} && \text{for } p = \wp_1,\\
\delta_p &:= \frac{p^{-1}+2p^{-2}+2p^{-3}+p^{-4}}{(1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4})} && \text{for } p \text{ ramified.}
\end{aligned}$$

Let $\delta_P := \prod_{p \in P} \delta_p$ and let $e = \prod_{p \in P} p$. Let $N_5(X; P)$ be the number of quintic fields with absolute discriminant at most X such that for each $p \in P$, the prime

p splits in the quartic field in the splitting type chosen for p above. There exists a constant $c_5 > 0$ such that

$$N_5(X; P) = \delta_P c_5 X + O(e^{1/2+\varepsilon} X^{79/80+\varepsilon} + X^{199/200+\varepsilon}), \tag{5-1}$$

where the implied constant in the O term is absolute (does not depend on P). Moreover, we may choose more than one splitting type at each prime and let δ_p be the sum of the corresponding densities and the result still holds.

Bhargava [2010] gave the first asymptotic count of quintic number fields, and Shankar and Tsimerman [2014], building on Bhargava's work, gave the first power-saving error term. Both proofs fundamentally rely on Bhargava's [2008] parametrization of quintic rings. We will follow the outline of the argument of [Shankar and Tsimerman 2014], additionally requiring our chosen splitting conditions. While the main term for such a restricted count appears in [Bhargava 2010, Theorem 3] (at least for one prime, and the same argument would work for more primes), we require a power-saving error term *with explicit dependence* on the primes. While such bounds have appeared in at least three recent papers, we will improve on the exponents in all of them, as well as remove hypotheses that do not hold in our cases of interest. Lemke Oliver and Thorne [2017, Theorem 2.1] have shown, assuming that we choose ramification at each prime $p \in P$, that $N_5(X; P) = \delta_P c_5 X + O(e X^{199/200+\varepsilon})$. Cho and Kim [2015, Section 6] have recently proven a bound of the sort we desire; it seems they show $N_5(X; P) = \delta_P c_5 X + O(e^{2-\varepsilon} X^{399/400+\varepsilon})$. Also, Shankar, Södergren, and Templier [Shankar et al. 2015] stated the bound $N_5(X; P) = \delta_P c_5 X + O(e^{40} X^{79/80+\varepsilon} + X^{199/200+\varepsilon})$ when P contains a single prime.

Instead of directly counting quintic fields, the method, equivalently, counts maximal quintic orders. Analogously to the quartic case, we use a parametrization of quintic rings with their sextic resolvents due to Bhargava [2008]. Let $V_{\mathbb{Z}} = \mathbb{Z}^{40}$ denote the space of quadruples of 5×5 skew-symmetric matrices with integer coefficients. Then quintic rings with their sextic resolvents are parametrized by $\mathrm{GL}_4(\mathbb{Z}) \times \mathrm{SL}_5(\mathbb{Z})$ orbits on $V_{\mathbb{Z}}$ [Bhargava 2008, Theorem 1]. These orbits correspond to lattice points in a fundamental domain for $\mathrm{GL}_4(\mathbb{Z}) \times \mathrm{SL}_5(\mathbb{Z})$ on $\mathbb{R}^{40}$. As in [Bhargava 2010, Section 2.2; Shankar and Tsimerman 2014, Section 2.2], we take a certain random fundamental domain for the action of $\mathrm{GL}_4(\mathbb{Z}) \times \mathrm{SL}_5(\mathbb{Z})$ on $\mathbb{R}^{40}$. For a subset $S \subset V_{\mathbb{Z}}$, let $N_{\mathrm{dom}}(S; X)$ denote the expected number of elements of S with absolute discriminant less than X and whose associated quintic ring is an integral domain (i.e., is an order in a quintic field), in a random fundamental domain (as in [Shankar and Tsimerman 2014, Equation (1)], summed over the implicit i there). Let $N^*(S; X)$ denote the expected number of elements of S with absolute discriminant less than X, in a random fundamental domain (as in the equation after (1) in [Shankar and Tsimerman 2014], summed over the implicit i there).

We first consider the case in which only unramified splitting types are chosen. Let a_{12} denote the $(1,2)$ coordinate of the first matrix in a quadruple of 5×5 skew-symmetric matrices. For a square-free integer q relatively prime to e, let $W_{q,e}\subset V_{\mathbb{Z}}$ denote the set of elements corresponding to quintic rings that are not maximal at each prime dividing q and are maximal and of chosen splitting type at primes dividing e. Recall from [Bhargava 2008, Section 12] that $W_{q,e}$ is defined by congruence conditions modulo q^2e (for maximality, an argument analogous to that in [Bhargava 2004, Lemma 22] is necessary). Let $U_e\subset V_{\mathbb{Z}}$ denote the set of elements corresponding to quintic rings that are maximal at all primes and of chosen splitting type at the primes dividing e. Then counting $N_{\text{dom}}(U_e;X)$ will provide us with precisely the count $N_5(X;e)$ we require. We will count lattice points in U_e by using inclusion-exclusion to reduce to counting lattice points in the $W_{q,e}$.

By [Bhargava 2010, Equation (27)] (see also [Shankar and Tsimerman 2014, Equation (4)]), if L is a translate of the lattice $mV_{\mathbb{Z}}$ and $m=O(X^{1/40})$, then

$$N^*(L\cap\{a_{12}\neq 0\};X)=c_0m^{-40}X+O(m^{-39}X^{39/40}), \tag{5-2}$$

for some positive absolute constant c_0.

Bhargava gives the density of lattice points corresponding to rings maximal at a given prime [2008, Equation (48)] and the density of lattice points corresponding to rings maximal and of each splitting type [2008, Lemma 20]. Using these two computed densities, we conclude that of the $(q^2e)^{40}$ quadruples of 5×5 skew-symmetric matrices mod q^2e, we have that $W_{q,e}$ corresponds to $\nu(q)q^{80}\delta_P e^{40}\prod_{p\in P}(1-\nu(p))$ of them, where

$$\nu(p)=\\ 1-\frac{(p-1)^8p^{12}(p+1)^4(p^2+1)^2(p^2+p+1)^2(p^4+p^3+p^2+p+1)(p^4+p^3+2p^2+2p+1)}{p^{40}},$$

and we extend this to a multiplicative function $\nu(q)$ for square-free q. (Here, $\nu(p)$ is the density of lattice points correspond to rings that are *nonmaximal* at p from [Bhargava 2008, Equation (48)].) Note that $\nu(p)=p^{-2}+O(p^{-3})$ and thus $\nu(q)=O(q^{-2+\varepsilon})$.

We have that $\delta_P\le 1$ and $1-\nu(p)\le 1$. So, when $q^2e=O(X^{1/40})$, by summing Equation (5-2) over all the translates of $q^2eV_{\mathbb{Z}}$ that comprise $W_{q,e}$, we find that

$$\begin{aligned}N^*(&W_{q,e}\cap\{a_{12}\neq 0\};X)\\ &=\nu(q)q^{80}\delta_Pe^{40}\prod_{p\in P}(1-\nu(P))c_0q^{-80}e^{-40}X\\ &\qquad\qquad+O\Big(\nu(q)q^{80}\delta_Pe^{40}\prod_{p\in P}(1-\nu(p))q^{-78}e^{-39}X^{39/40}\Big)\end{aligned}$$

$$
\begin{aligned}
&= \nu(q)\delta_P \prod_{p\in P}(1-\nu(p))c_0X + O\Bigl(\nu(q)q^2\delta_P e\prod_{p\in P}(1-\nu(p))X^{39/40}\Bigr)\\
&= \nu(q)\delta_P \prod_{p\in P}(1-\nu(p))c_0X + O(q^{\varepsilon}eX^{39/40}),
\end{aligned}
\tag{5-3}
$$

where in the last identity we have used the fact that $\nu(q)=O(q^{-2+\varepsilon})$.

We then, by inclusion-exclusion as in [Shankar and Tsimerman 2014, Section 4], have for an appropriate Q (to be chosen later in terms of X),

$$
\begin{aligned}
&N_{\mathrm{dom}}(U_e\cap\{a_{12}\neq 0\};X)\\
&= {\sum_{q\geq 1}}' \mu(q)N_{\mathrm{dom}}(W_{q,e}\cap\{a_{12}\neq 0\};X)\\
&= {\sum_{1\leq q\leq Q}}' \mu(q)N_{\mathrm{dom}}(W_{q,e}\cap\{a_{12}\neq 0\};X) + {\sum_{q>Q}}' \mu(q)N_{\mathrm{dom}}(W_{q,e}\cap\{a_{12}\neq 0\};X)\\
&= {\sum_{1\leq q\leq Q}}' \mu(q)N^{*}(W_{q,e}\cap\{a_{12}\neq 0\};X)\\
&\qquad + {\sum_{1\leq q\leq Q}}' \mu(q)\bigl(N_{\mathrm{dom}}(W_{q,e}\cap\{a_{12}\neq 0\};X) - N^{*}(W_{q,e}\cap\{a_{12}\neq 0\};X)\bigr)\\
&\qquad\qquad + {\sum_{q>Q}}' \mu(q)N_{\mathrm{dom}}(W_{q,e}\cap\{a_{12}\neq 0\};X),
\end{aligned}
$$

where the sums are over square-free q relatively prime to e.

By [Shankar and Tsimerman 2014, Lemma 3], we have $N_{\mathrm{dom}}(W_{q,e};X) = O(q^{-2+\varepsilon}X)$ and we use this for the sum for $q>Q$. We will use Equation (5-3) for the first $1\leq q\leq Q$ sum. For the second $1\leq q\leq Q$ sum, note that each lattice point corresponding to a nondomain of discriminant D is counted with coefficient

$$
-\sum_{\substack{1\leq q\leq Q\\ q\mid D}} \mu(q),
$$

which is $O(D^{\varepsilon})=O(X^{\varepsilon})$, and by [Shankar and Tsimerman 2014, Equation (8)] there are at most $O(X^{199/200+\varepsilon})$ lattice points corresponding to nondomains. (This step, or something similar, should be added to the proof in [Shankar and Tsimerman 2014].)

As a result, as long as $Q=O(X^{1/80}e^{-1/2})$,

$$
\begin{aligned}
&N_{\mathrm{dom}}(U_e\cap\{a_{12}\neq 0\};X)\\
&\qquad = {\sum_{q\geq 1}}' \mu(q)\nu(q)\delta_P\prod_{p\in P}(1-\nu(p))c_0X + O(E_1)+O(E_2)+O(E_3),
\end{aligned}
$$

where

$$E_1 = \sum_{1 \le q \le Q} O(q^{\varepsilon} e X^{39/40}), \quad E_2 = O(X^{199/200+\varepsilon}),$$
$$E_3 = \sum_{q > Q} q^{-2+\varepsilon} X, \qquad E_4 = \sum_{q>Q} \mu(q)\nu(q)\delta_P \prod_{p \in P} (1-\nu(p)) c_0 X.$$

These terms trivially admit the estimates

$$E_1 = O(Q^{1+\varepsilon} e X^{39/40}), \quad E_2 = O(X^{199/200+\varepsilon}),$$
$$E_3 = O(Q^{-1+\varepsilon} X), \qquad E_4 = O(Q^{-1+\varepsilon} X),$$

where in the last estimate we have used the fact that $\nu(q) = O(q^{-2+\varepsilon})$.

We take $Q = X^{1/80} e^{-1/2}$, and have

$$\begin{aligned} N_{\mathrm{dom}}&(U_e \cap \{a_{12} \neq 0\}; X) \\ &= \sum_{q \ge 1}{}' \mu(q)\nu(q)\delta_P \prod_{p \in P} (1-\nu(p)) c_0 X + O(e^{1/2} X^{79/80+\varepsilon} + X^{199/200+\varepsilon}) \\ &= \prod_p (1-\nu(p)) \delta_P c_0 X + O(e^{1/2} X^{79/80+\varepsilon} + X^{199/200+\varepsilon}). \end{aligned}$$

From [Bhargava 2010, Lemma 11], we have that $N_{\mathrm{dom}}(\{a_{12} = 0\}; X) = O(X^{39/40})$, and so

$$N_{\mathrm{dom}}\big(U_e \cap \{a_{12} = 0\}; X\big) = O(X^{39/40}).$$

It follows that

$$\begin{aligned} N_5(X; P) &= N_{\mathrm{dom}}(U_e; X) \\ &= \prod_p (1-\nu(p)) \delta_P c_0 X + O\big(e^{1/2} X^{79/80+\varepsilon} + X^{199/200+\varepsilon}\big). \end{aligned}$$

We thus conclude Theorem 5.1 holds with, $c_5 = \prod_p (1-\nu(p)) c_0$ when we only choose unramified splitting types. As at the end of Theorem 4.1, we can apply the result we have just proven and inclusion-exclusion to prove Theorem 5.1 in general.

6. Application of the sieve

6A. ***Summary of the asymptotic inputs to the sieve.*** We now turn to the application of the sieve lemma to degree d field extensions of $\mathbb{Q}$. Note that when applying the sieve, it is crucial to have error terms with explicit dependence on local conditions (such as we have derived in Theorems 2.3 and 2.4): without such an explicit dependence, we would not have quantitative control of the right-hand side of the key sieve inequality in Lemma 3.1, since we would not have an explicit bound for R_p in terms of p.

Let $\mathscr{A}$ and $\mathscr{A}_p$ (for each rational prime p) be the sets such that $\#\mathscr{A} = N_d(X)$ and $\#\mathscr{A}_p = N_d(X; p)$ (or $\widetilde{N}_4(X)$, $\widetilde{N}_4(X; p)$ in the case of $d = 4$). With these definitions, the quantity $E\big(\mathscr{A}; z, \frac{1}{2}M(z)\big)$ treated in the sieve (Lemma 3.1), which we will now denote by $E_d\big(\mathscr{A}; z, \frac{1}{2}M(z)\big)$, is the number of degree d extensions K of $\mathbb{Q}$ with $0 < D_K \le X$ (up to isomorphism, and non-D_4 when $d = 4$) such that there are at most $\frac{1}{2}M(z)$ primes $p \le z$ that split completely in K.

We recall the collection $\mathscr{B}_d(X; Y, M)$ of bad fields, as defined in (2-2). We will think of $Y = z = (X/2)^{\delta_0}$ for $\delta_0 > 0$ to be chosen precisely later, and define $M(z)$ as in (3-1). In particular, the set of bad fields satisfies

$$\#\mathscr{B}_d\big(X; (X/2)^{\delta_0}, \tfrac{1}{2}M((X/2)^{\delta_0})\big) = E_d\big(\mathscr{A}; (X/2)^{\delta_0}, \tfrac{1}{2}M((X/2)^{\delta_0})\big).$$

We will need to apply the sieve separately to fields of each degree, since in several cases the count for $N_d(X; p)$ takes a somewhat different form, but in an effort to unify the presentation, we restate the asymptotics we will assume in more general form. We write the results of Lemma 2.2, Theorem C, Theorems 2.3 and 2.4 as follows.

Quadratic fields: for δ_e as in (2-10), there is some $\sigma_2 > 0$ and $0 < \tau_2 \le 1/2$ such that

$$N_2(X) = c_2 X + O(X^{\tau_2+\varepsilon}), \quad N_2(X; e) = \delta_e c_2 X + O(e^{\sigma_2} X^{\tau_2+\varepsilon}).$$

Cubic fields: for δ_e, δ'_e as in (2-13), there is some $\sigma_3 > 0$ and $0 < \tau_3 < 5/6$ such that

$$N_3(X) = c_3 X + c'_3 X^{5/6} + O(X^{\tau_3+\varepsilon}),$$
$$N_3(X; e) = \delta_e c_3 X + \delta'_e c'_3 X^{5/6} + O(e^{\sigma_3} X^{\tau_3+\varepsilon}).$$

Non-D_4-quartic fields: for δ_e as in (2-16), there is some $\sigma_4 > 0$ and $0 < \tau_4 < 1$ such that

$$\widetilde{N}_4(X) = c_4 X + O(X^{\tau_4+\varepsilon}), \quad \widetilde{N}_4(X; e) = \delta_e c_4 X + O(e^{\sigma_4} X^{\tau_4+\varepsilon}).$$

Quintic fields: for δ_e as in (2-19), there is some $\sigma_5 > 0$ and $0 < \tau_5 < 1$ as well as some $0 < \gamma < 1$ such that

$$N_5(X) = c_5 X + O(X^{\gamma+\varepsilon}), \quad N_5(X; e) = \delta_e c_5 X + O(e^{\sigma_5} X^{\tau_5}) + O(X^{\gamma+\varepsilon}).$$

The main result of the sieve in this context is the following:

Proposition 6.1. *With the notation as above, we have*

$$E_d\big(\mathscr{A}; (X/2)^{\delta_0}, \tfrac{1}{2}M((X/2)^{\delta_0})\big) \ll X^{1-\delta_0+\varepsilon},$$

for any δ_0 such that

$$\delta_0 \le \begin{cases} \dfrac{1-\tau_d}{1+2\sigma_d} & \text{if } d = 2, 4, \\ \min\left\{\dfrac{1-\tau_d}{1+2\sigma_d}, \dfrac{1}{4}\right\} & \text{if } d = 3, \\ \min\left\{\dfrac{1-\tau_d}{1+2\sigma_d}, 1-\gamma\right\} & \text{if } d = 5. \end{cases} \tag{6-1}$$

Moreover, for any such δ_0 there exist positive real constants $c_0(d) < c_1(d) < 1$ and $X_d = X_d(\delta_0) \ge 1$ such that for all $X \ge X_d$,

$$c_0(d)\frac{(X/2)^{\delta_0}}{\log(X/2)^{\delta_0}} \le M((X/2)^{\delta_0}) \le c_1(d)\frac{(X/2)^{\delta_0}}{\log(X/2)^{\delta_0}}. \tag{6-2}$$

The requirement that $X \ge X_d$ simply is a quantification of the requirement that X be sufficiently large, and will be incorporated later simply by enlarging certain implicit constants.

Proposition 6.1 immediately provides the upper bound we require for the bad set $\mathscr{B}_d(X; Y, M)$ defined in (2-2), with an appropriate choice of the parameters Y, M. As there are $\pi(Y) = Y(\log Y)^{-1} + O(Y(\log Y)^{-2})$ primes $p \le Y$, we could of course only expect at most $Y(\log Y)^{-1}$ primes $p \le Y$ to split completely in any given field. Proposition 6.1 shows that, up to a constant factor, this is a reasonable expectation, in that the mean $M((X/2)^{\delta_0})$ is approximately $\mu\pi((X/2)^{\delta_0})$, for some $\mu \in [c_0(d), c_1(d)]$; moreover Proposition 6.1 provides an upper bound for the number of fields with $D_K \le X$ in which at most $\frac{1}{2}M((X/2)^{\delta_0})$ primes $p \le (X/2)^{\delta_0}$ split completely.

We will prove Proposition 6.1 case by case.

6B. ***Sieve for quadratic fields.*** For notational convenience, in this section we write σ, τ for σ_2, τ_2. We compute that for any prime p,

$$R_p = \#\mathscr{A}_p - \delta_p \#\mathscr{A} = N_2(X; p) - \delta_p N_2(X) = O(p^\sigma X^{\tau+\varepsilon}).$$

Similarly, for distinct primes p, q

$$R_{pq} = \#\mathscr{A}_{pq} - \delta_p\delta_q \#\mathscr{A} = N_2(X; pq) - \delta_p\delta_q N_2(X) = O(p^\sigma q^\sigma X^{\tau+\varepsilon}).$$

Thus since $\#\mathscr{A} \gg X$,

$$\frac{1}{\#\mathscr{A}}\sum_{p|P(z)} |R_p| \ll z^{1+\sigma}X^{\tau-1+\varepsilon}, \quad \frac{1}{\#\mathscr{A}}\sum_{p,q|P(z)} |R_{pq}| \ll z^{2+2\sigma}X^{\tau-1+\varepsilon}.$$

We compute

$$U(z) = \sum_{p|P(z)} \delta_p = \frac{1}{2}\sum_{p|P(z)} \frac{1}{1+p^{-1}},$$

from which we deduce that

$$\tfrac{1}{3}z(\log z)^{-1} + O(z(\log z)^{-2}) \le U(z) \le \tfrac{1}{2}z(\log z)^{-1} + O(z(\log z)^{-2}). \quad (6\text{-}3)$$

Indeed, letting $\varepsilon_p = (1 + p^{-1})^{-1}$, the upper bound follows directly from the prime number theorem and the fact that $0 < \varepsilon_p < 1$, while the lower bound only requires noticing

$$U(z) \ge \frac{1}{2} \sum_{p \mid P(z)} \varepsilon_2 = \frac{1}{3} \sum_{p \mid P(z)} 1 = \tfrac{1}{3}z(\log z)^{-1} + O(z(\log z)^{-2}).$$

We may compute the mean as in (3-3):

$$M(z) = U(z) + \frac{1}{\#\mathscr{A}} \sum_{p \mid P(z)} R_p = U(z) + O(z^{1+\sigma} X^{\tau-1+\varepsilon}).$$

Recalling (6-3) and that $z = (X/2)^{\delta_0}$ for a parameter δ_0 to be chosen later, we see the last error term will be $< \frac{1}{2}U(z)$ for sufficiently large X as long as

$$\delta_0 < \frac{1-\tau}{\sigma}. \quad (6\text{-}4)$$

Assuming this, for sufficiently large X we have

$$c_0 z(\log z)^{-1} \le \tfrac{1}{2}U(z) \le M(z) \le \tfrac{3}{2}U(z) \le c_1 z(\log z)^{-1}$$

for absolute constants $0 < c_0 < c_1 \le 1$. We apply Lemma 3.1 to see that

$$\begin{aligned} E_2\big(\mathscr{A}; z, \tfrac{1}{2}M(z)\big) &\ll \frac{X^{1+\varepsilon}}{z^2}\big(z + z^{2+2\sigma} X^{\tau-1} + z(z^{1+\sigma} X^{\tau-1}) + (z^{1+\sigma} X^{\tau-1})^2\big) \\ &\ll X^{\varepsilon}(X z^{-1} + z^{2\sigma} X^{\tau}), \end{aligned}$$

still assuming (6-4). Balancing the terms in the last expression above would set

$$\delta_0 = (1-\tau)/(1+2\sigma), \quad (6\text{-}5)$$

which certainly satisfies (6-4); as a consequence, for any $\delta_0 \le (1-\tau)/(1+2\sigma)$, we obtain

$$E_2\big(\mathscr{A}; (X/2)^{\delta_0}, \tfrac{1}{2}M((X/2)^{\delta_0})\big) \ll X^{1-\delta_0+\varepsilon},$$

which proves Proposition 6.1 in the case of quadratic fields.

6C. ***Sieve for cubic fields.*** For notational convenience, in this section we write σ, τ for σ_3, τ_3. We compute that

$$R_p = \#\mathscr{A}_p - \delta_p \#\mathscr{A} = c_3'(\delta_p' - \delta_p) X^{5/6} + O(p^{\sigma} X^{\tau+\varepsilon}) = O(p^{-1/3} X^{5/6} + p^{\sigma} X^{\tau+\varepsilon}).$$

For distinct primes p, q,

$$\begin{aligned} R_{pq} &= c_3'(\delta_p'\delta_q' - \delta_p\delta_q)X^{5/6} + O(p^\sigma q^\sigma X^{\tau+\varepsilon}) \\ &= O(p^{-1/3}X^{5/6} + q^{-1/3}X^{5/6} + p^\sigma q^\sigma X^{\tau+\varepsilon}). \end{aligned}$$

Since $\#\mathscr{A} \gg X$, we may compute that

$$\frac{1}{\#\mathscr{A}} \sum_{p|P(z)} |R_p| \ll z^{2/3}X^{-1/6} + z^{1+\sigma}X^{\tau-1+\varepsilon},$$
$$\frac{1}{\#\mathscr{A}} \sum_{p,q|P(z)} |R_{p,q}| \ll z^{5/3}X^{-1/6} + z^{2+2\sigma}X^{\tau-1+\varepsilon}.$$

Next, we note that

$$U(z) = \sum_{p|P(z)} \delta_p = \frac{1}{6}\sum_{p|P(z)} \frac{1}{1+p^{-1}+p^{-2}} = \frac{1}{6}\sum_{p|P(z)} e_p,$$

say. From this we can deduce (as in the case of quadratic fields) that

$$\tfrac{2}{21}z(\log z)^{-1} + O(z(\log z)^{-2}) \le U(z) \le \tfrac{1}{6}z(\log z)^{-1} + O(z(\log z)^{-2}). \quad (6\text{-}6)$$

Finally, we compute the mean

$$M(z) = U(z) + \frac{1}{\#\mathscr{A}}\sum_{p|P(z)} R_p = U(z) + O\big(z^{2/3}X^{-1/6} + z^{1+\sigma}X^{\tau-1+\varepsilon}\big).$$

Recalling (6-6) and that $z = (X/2)^{\delta_0}$ for a parameter δ_0 to be chosen later, we see the last error term will be $< \frac{1}{2}U(z)$ for sufficiently large X as long as the analogue of (6-4) holds, in which case

$$c_0 z(\log z)^{-1} \le \tfrac{1}{2}U(z) \le M(z) \le \tfrac{3}{2}U(z) \le c_1 z(\log z)^{-1}.$$

for absolute constants $0 < c_0 < c_1 \le 1$.

We now apply Lemma 3.1, which shows that

$$\begin{aligned} &E_3\big(\mathscr{A}; z, \tfrac{1}{2}M(z)\big) \\ &\quad\ll \frac{X^{1+\varepsilon}}{z^2}\Big(z + \big(z^{5/3}X^{-1/6} + z^{2+2\sigma}X^{\tau-1}\big) \\ &\qquad\qquad + z\big(z^{2/3}X^{-1/6} + z^{1+\sigma}X^{\tau-1}\big) + \big(z^{2/3}X^{-1/6} + z^{1+\sigma}X^{\tau-1}\big)^2\Big). \end{aligned}$$

As long as $\delta_0 \le 1/4$, we have $z^{5/3}X^{-1/6} \ll z$; after further simplification and still assuming the analogue of (6-4), we see that

$$E_3\big(\mathscr{A}; z, \tfrac{1}{2}M(z)\big) \ll X^\varepsilon(Xz^{-1} + z^{2\sigma}X^\tau).$$

This is optimized by choosing δ_0 as in (6-5) as before, which satisfies (6-4). In particular, for any $\delta_0 \le \min\{1/4, (1-\tau)/(1+2\sigma)\}$, we obtain

$$E_3\big(\mathscr{A}; (X/2)^{\delta_0}, \tfrac{1}{2}M((X/2)^{\delta_0})\big) \ll X^{1-\delta_0+\varepsilon},$$

which proves Proposition 6.1 in the case of cubic fields.

6D. ***Sieve for non-D_4-quartic fields.*** The case of non-D_4-quartic fields is very similar to that for real quadratic fields, thus we only mention the highlights, with σ, τ denoting σ_4, τ_4. We have

$$\begin{aligned}
R_p &= \#\mathscr{A}_p - \delta_p \#\mathscr{A} = O(p^\sigma X^{\tau+\varepsilon}),\\
R_{pq} &= \#\mathscr{A}_{pq} - \delta_p\delta_q \#\mathscr{A} = O(p^\sigma q^\sigma X^{\tau+\varepsilon}),\\
U(z) &= \sum_{p \mid P(z)} \delta_p = \frac{1}{24}\sum_{p\mid P(z)} \frac{1}{1+p^{-1}+2p^{-2}+p^{-3}}.
\end{aligned}$$

We deduce that

$$\frac{1}{3\cdot 17} z(\log z)^{-1} + O(z(\log z)^{-2}) \le U(z) \le \frac{1}{24} z(\log z)^{-1} + O(z(\log z)^{-2}). \tag{6-7}$$

Next we compute the mean

$$M(z) = U(z) + \frac{1}{\#\mathscr{A}} \sum_{p\mid P(z)} R_p = U(z) + O(z^{1+\sigma} X^{\tau-1+\varepsilon}).$$

Recalling (6-7) and that $z = (X/2)^{\delta_0}$, we see that as long as the analogous condition to (6-4) holds and X is sufficiently large,

$$c_0 z(\log z)^{-1} \le \tfrac{1}{2}U(z) \le M(z) \le \tfrac{3}{2}U(z) \le c_1 z(\log z)^{-1}$$

for absolute constants $0 < c_0 < c_1 \le 1$.

We apply Lemma 3.1 to see that under the assumption (6-4)

$$\begin{aligned}
E_4\big(\mathscr{A}; z, \tfrac{1}{2}M(z)\big) &\ll \frac{X^{1+\varepsilon}}{z^2}\big(z + z^{2+2\sigma}X^{\tau-1} + z(z^{1+\sigma}X^{\tau-1}) + (z^{1+\sigma}X^{\tau-1})^2\big)\\
&\ll X^{\varepsilon}(Xz^{-1} + z^{2\sigma}X^{\tau}),
\end{aligned}$$

so that

$$E_4\big(\mathscr{A}; (X/2)^{\delta_0}, \tfrac{1}{2}M((X/2)^{\delta_0})\big) \ll X^{1-\delta_0+\varepsilon}$$

for any $\delta_0 \le (1-\tau)/(1+2\sigma)$.

6E. ***Sieve for quintic fields.*** Finally, we apply the sieve to quintic fields, denoting σ_5, τ_5 by σ, τ. We compute that for any $p = O(X^\rho)$,

$$R_p = \#\mathscr{A}_p - \delta_p \#\mathscr{A} = O(X^\varepsilon(p^\sigma X^\tau + X^\gamma)).$$

For distinct primes p, q,

$$R_{pq} = \#\mathscr{A}_{pq} - \delta_p\delta_q \#\mathscr{A} = O(X^\varepsilon(p^\sigma q^\sigma X^\tau + X^\gamma)).$$

We compute

$$U(z) = \sum_{p|P(z)} \delta_p = \frac{1}{120}\sum_{p|P(z)} \frac{1}{1+p^{-1}+2p^{-2}+2p^{-3}+p^{-4}},$$

from which we deduce that

$$\frac{2}{15\cdot 37} z(\log z)^{-1} + O(z(\log z)^{-2}) \le U(z) \le \frac{1}{120} z(\log z)^{-1} + O(z(\log z)^{-2}).$$

The mean may be expressed as

$$M(z) = U(z) + \frac{1}{\#\mathscr{A}} \sum_{p|P(z)} R_p = U(z) + O\big(X^\varepsilon(z^{1+\sigma}X^{\tau-1} + zX^{\gamma-1+\varepsilon})\big).$$

The last term will be $< \frac{1}{2}U(z)$ for sufficiently large X as long as $\gamma < 1$ and the analogous condition to (6-4) holds. Assuming this, we have

$$c_0 z(\log z)^{-1} \le \tfrac{1}{2}U(z) \le M(z) \le \tfrac{3}{2}U(z) \le c_1 z(\log z)^{-1}$$

for absolute constants $0 < c_0 < c_1 \le 1$.

We apply Lemma 3.1 to see that under the assumptions $\tau, \gamma < 1$ and (6-4),

$$E_5\big(\mathscr{A}; z, \tfrac{1}{2}M(z)\big) \ll \frac{X^{1+\varepsilon}}{z^2}\Big(z + z^2X^{-1}(z^{2\sigma}X^\tau + X^\gamma) + zX^{-1}(z^\sigma X^\tau + X^\gamma) + z^2X^{-2}(z^\sigma X^\tau + X^\gamma)^2\Big).$$

After simplification, this shows

$$\begin{aligned} E_5\big(\mathscr{A}; z, \tfrac{1}{2}M(z)\big) &\ll \frac{X^{1+\varepsilon}}{z^2}(z + z^2X^{\gamma-1} + z^{2+2\sigma}X^{\tau-1}) \\ &\ll X^\varepsilon(Xz^{-1} + X^\gamma + z^{2\sigma}X^\tau). \end{aligned}$$

Assuming $z = (X/2)^{\delta_0}$, we may conclude that for any

$$\delta_0 \le \min\big\{(1-\tau)/(1+2\sigma), 1-\gamma\big\},$$

we have

$$E_5\big(\mathscr{A}; (X/2)^{\delta_0}, \tfrac{1}{2}M((X/2)^{\delta_0})\big) \ll X^{1-\delta_0+\varepsilon}.$$

This completes the proof of Proposition 6.1.

7. Proof of the main theorem and corollaries

7A. ***Proof of Theorem 2.1.*** We now derive Theorem 2.1 from Proposition 6.1. By definition, if $M_1 \le M_2$ then $\mathscr{B}_d(X; Y, M_1) \subseteq \mathscr{B}_d(X; Y, M_2)$. If X is sufficiently large that (6-2) holds, say $X \ge X_d(\delta)$, we may apply (6-2) to write

$$\begin{aligned}\#\mathscr{B}_d\left(X; (X/2)^\delta, \frac{1}{2}c_0(d)\frac{(X/2)^\delta}{\log(X/2)^\delta}\right) &\le \#\mathscr{B}_d\big(X; (X/2)^\delta, \tfrac{1}{2}M((X/2)^\delta)\big)\\ &= E_d\big(\mathscr{A}; (X/2)^\delta, \tfrac{1}{2}M((X/2)^\delta)\big).\end{aligned}$$

We then apply Proposition 6.1 and deduce that for $X \ge X_d(\delta)$,

$$\#\mathscr{B}_d\left(X; (X/2)^\delta, \frac{1}{2}c_0(d)\frac{(X/2)^\delta}{\log(X/2)^\delta}\right) \ll X^{1-\delta+\varepsilon}$$

for every $\varepsilon > 0$, and for δ constrained by (6-1). When we make the constraints in (6-1) precise by applying the results of Lemma 2.2, Theorem C, Theorems 2.3 and 2.4, we obtain the parameters defined in (2-5). For any δ satisfying (2-5), we may remove the explicit assumption that $X \ge X_d(\delta)$ by including an appropriate implicit constant, so that

$$\#\mathscr{B}_d\left(X; (X/2)^\delta, \frac{1}{2}c_0(d)\frac{(X/2)^\delta}{\log(X/2)^\delta}\right) \ll_{d,\delta,\varepsilon} X^{1-\delta+\varepsilon} \tag{7-1}$$

for every $X \ge 1$ and every $\varepsilon > 0$.

7B. ***Proof of Theorem 1.1.*** To derive Theorem 1.1 from Theorem 2.1, we proceed via a standard dyadic argument, which we now make precise. Let $\varepsilon > 0$ be fixed and for this ε, let the implied constant in Theorem A be denoted by $C_0 = C_0(d, \varepsilon)$, so that (1-8) becomes

$$|\mathrm{Cl}_K[\ell]| \le C_0 D_K^{\frac{1}{2}+\varepsilon} M^{-1}. \tag{7-2}$$

Fix any $\delta < \frac{1}{2\ell(d-1)}$. Then if K is a degree d extension of $\mathbb{Q}$ with $D_K \in (X, 2X]$ that is not in the bad set $\mathscr{B}_d^0\big(X; X^\delta, \frac{1}{2}c_0(d)X^\delta/\log X^\delta\big)$, we see from (7-2) that

$$|\mathrm{Cl}_K[\ell]| \le C_0\big(\tfrac{1}{2}c_0(d)\big)^{-1} D_K^{\frac{1}{2}+\varepsilon} X^{-\delta}\log(X^\delta) \le C_0' D_K^{\frac{1}{2}-\delta+\varepsilon}\log(D_K^\delta),$$

where it suffices to take $C_0' = C_0 2^{1+\delta}\big(\frac{1}{2}c_0(d)\big)^{-1}$. Now we assume that X is sufficiently large, say $X \ge C(d, \ell, \varepsilon)$, so that for all $\delta < \frac{1}{2\ell(d-1)}$, and for all $D_K \in (X, 2X]$, we have $\log(D_K^\delta) \le D_K^\varepsilon$. Under this assumption we have

$$|\mathrm{Cl}_K[\ell]| \le C_0' D_K^{\frac{1}{2}-\delta+2\varepsilon} \tag{7-3}$$

for all these fields not in $\mathscr{B}_d^0\big(X; X^\delta, \frac{1}{2}c_0(d)X^\delta/\log X^\delta\big)$.

Let $F^0_{d,\ell}(X;\delta,\varepsilon)$ denote the collection of fields $K/\mathbb{Q}$ of degree d with $X < D_K \le 2X$ that fail the bound (7-3); we may conclude that for any $\delta < \frac{1}{2\ell(d-1)}$ and for all $X \ge C(d,\ell,\varepsilon)$,

$$F^0_{d,\ell}(X;\delta,\varepsilon) \subseteq \mathscr{B}^0_d\bigl(X, X^\delta, \tfrac{1}{2}c_0(d)X^\delta/\log X^\delta\bigr). \tag{7-4}$$

Now let $F_{d,\ell}(X;\delta,\varepsilon)$ denote the collection of fields $K/\mathbb{Q}$ of degree d with $0 < D_K \le X$ that fail the bound (7-3); then

$$F_{d,\ell}(X;\delta,\varepsilon) \subseteq \bigcup_{0\le j\le \lceil \log_2 X\rceil} F^0_d(2^j;\delta,\varepsilon).$$

Set j_0 to be the smallest j such that $2^{j_0} \ge C(d,\ell,\varepsilon)$. Then for $j \le j_0$, we apply the trivial bound, $\#F^0_{d,\ell}(2^j,\delta,\varepsilon) \ll 2^j$. (This bound is only "trivial" in the sense that we know by (1-11) how to count fields of degree d with $0 < D_K \le X$, for $d \le 5$.) For $j > j_0$ we apply (7-4) to write

$$\bigcup_{j_0<j\le\lceil\log_2 X\rceil} F^0_{d,\ell}(2^j;\delta,\varepsilon) \subseteq \bigcup_{j_0<j\le\lceil\log_2 X\rceil} \mathscr{B}^0_d\bigl(2^j, 2^{j\delta}, \tfrac{1}{2}c_0(d)2^{j\delta}/\log 2^{j\delta}\bigr).$$

Trivially enlarging each of the last sets to the nondyadic version

$$\mathscr{B}_d\bigl(2^{j+1}, 2^{j\delta}, \tfrac{1}{2}c_0(d)2^{j\delta}/\log 2^{j\delta}\bigr)$$

and applying the result of Theorem 2.1 to each such set, we obtain

$$\#F_{d,\ell}(X;\delta,\varepsilon) \ll C(d,\ell,\varepsilon) + \sum_{j_0<j\le\lceil\log_2 X\rceil} 2^{j(1-\delta+\varepsilon')} \ll_{c,d,\ell,\varepsilon,\varepsilon'} X^{1-\delta+\varepsilon'}, \tag{7-5}$$

which now holds (with a sufficiently large implicit constant) for all $X \ge 1$, for all $\varepsilon' > 0$ arbitrarily small, and for all $\delta < \min\bigl\{\frac{1}{2\ell(d-1)}, \delta_0(d)\bigr\}$ where $\delta_0(d)$ is defined as in (2-5) in Theorem 2.1. For sufficiently large ℓ, the first constraint on δ is a stronger constraint than the second.

To be precise, we now break down into cases depending on d. For $d = 2$, Theorem 1.1 is implied in the case $\ell = 2$ by Gauss genus theory, and in the case $\ell = 3$ by the known asymptotic (1-5). For integers $\ell \ge 4$, Theorem 1.1 follows from (7-5), since $\frac{1}{2\ell} = \min\bigl\{\frac{1}{6}, \frac{1}{2\ell}\bigr\}$ for $\ell \ge 4$. (Of course, for primes $\ell \ge 5$ and imaginary quadratic fields, Theorem 1.1 is implied by the stronger result (1-10), or indeed by an earlier result of Soundararajan [2000] that at most one imaginary quadratic field K with $D_K \in [X, 2X]$ can have $|\mathrm{Cl}_K[\ell]| \gg D_K^{\frac{1}{2}-\frac{1}{2\ell}+\varepsilon}$; see also Corollary 2.2 of [Heath-Brown and Pierce 2014].) For $d = 3$, Theorem 1.1 is implied for $\ell = 2$ by the known asymptotic (1-6), and for $\ell = 3$ by the stronger known result (1-3). The cases $\ell \ge 4$ are implied by (7-5), since $\frac{1}{4\ell} = \min\bigl\{\frac{2}{25}, \frac{1}{4\ell}\bigr\}$ for $\ell \ge 4$. For $d = 4$, Theorem 1.1 follows from (7-5) since $\frac{1}{6\ell} = \min\bigl\{\frac{1}{48}, \frac{1}{6\ell}\bigr\}$ for $\ell \ge 8$; the remaining cases of $\ell \le 7$ follow from the choice $\delta_0 = \frac{1}{48}$. Finally, for $d = 5$, Theorem 1.1

similarly follows from (7-5) since $\frac{1}{8\ell} = \min\{\frac{1}{200}, \frac{1}{8\ell}\}$ for $\ell \geq 25$; the remaining cases of $\ell \leq 24$ follow from the choice $\delta_0 = \frac{1}{200}$.

Corollaries 1.1.1 and 1.1.2 now follow from Theorem 1.1, or can be derived directly from Theorem 2.1, as already demonstrated in Section 2A.

Appendix: Counting quadratic fields

In this appendix we prove the following result, from which Lemma 2.2 may be deduced immediately.

Proposition A.1. *Let P be a finite set of primes. For each prime $p \in P$ we choose a splitting type at p and assign a corresponding density as follows*:

$$\delta_p := \begin{cases} \frac{1}{2}(1+p^{-1})^{-1} & \text{for } p = \mathfrak{p}_1\mathfrak{p}_2, \\ \frac{1}{2}(1+p^{-1})^{-1} & \text{for } p = \mathfrak{p}_1, \\ (p+1)^{-1} & \text{for } p \text{ ramified.} \end{cases}$$

Let $e = \prod_{p \in P} p$ and $\delta_e = \prod_{p \in P} \delta_p$. Let $N_2^{\pm}(X; P)$ denote the number of real (*respectively imaginary*) *quadratic extensions of $\mathbb{Q}$ with fundamental discriminant $|D_K| \leq X$ such that for each $p \in P$, the prime p splits in the quadratic field with splitting type chosen for p above. Then*

$$N_2^{\pm}(X; P) = \delta_e \Big(\frac{1}{3} + \frac{1}{6}\Big) \frac{1}{\zeta(2)} X + O(e\sqrt{X}). \tag{A-1}$$

We remark that in (A-1), the first term is contributed by fundamental discriminants $\equiv 1 \pmod 4$ and the second by fundamental discriminants $\equiv 0 \pmod 4$. We prove the proposition explicitly for $N_2^{+}(X; P)$, and omit the analogous argument for $N_2^{-}(X; P)$. Upon combining the counts for real and imaginary fields, this implies Lemma 2.2 as a special case.

The proof is a simple elaboration on the classical method for counting square-free integers $\leq X$. Recall that, for a fundamental discriminant D, a prime p is ramified in $\mathbb{Q}(\sqrt{D})$ precisely when $p \mid D$; otherwise a prime $p \nmid D$ splits in $\mathbb{Q}(\sqrt{D})$ if the Kronecker symbol $\big(\frac{D}{p}\big)$ evaluates as $+1$, and is nonsplit if $\big(\frac{D}{p}\big) = -1$ (see, e.g., [Hua 1982, Theorem 10.3, Chapter 16]). Thus for each unramified $p \in P$ we assign $\varepsilon_p \in \{-1, +1\}$ according to the specified splitting type of p. Let P_0 be the set of ramified primes in P and set $P' = P \setminus P_0$; define

$$e_0 = \prod_{p \in P_0} p \quad \text{and} \quad e' = \prod_{p \in P'} p.$$

Then we may write

$$N_2^{+}(X; P) = \#\Big\{\text{fundamental discriminants } 1 \leq n \leq X : e_0 | n,\ \Big(\frac{n}{p}\Big) = \varepsilon_p, \forall\, p \in P'\Big\}.$$

We will find a count for this by sieving for fundamental discriminants (that is, elements that are free of odd squares) in the following two sets:

$$\mathcal{A}^{(1)} = \left\{1 \le n \le X : n \equiv 1 \pmod 4,\ e_0|n,\ \left(\tfrac{n}{p}\right) = \varepsilon_p, \forall p \in P'\right\},$$
$$\mathcal{A}^{(0)} = \left\{1 \le n \le X : n \equiv 8, 12 \pmod{16},\ e_0|n,\ \left(\tfrac{n}{p}\right) = \varepsilon_p, \forall p \in P'\right\}.$$

More generally, fix a power g and define for any $b \pmod{2^g}$ the set

$$\mathcal{A} = \left\{1 \le n \le X : n \equiv b \pmod{2^g}, e_0|n, \left(\tfrac{n}{p}\right) = \varepsilon_p, \forall p \in P'\right\}.$$

For each odd prime q let $\mathcal{A}_q = \{n \in \mathcal{A} : q^2|n\}$. Note that certainly $\mathcal{A}_q$ is empty as soon as $q > \sqrt{X}$; we let M be the index of the greatest prime $q_M \le \sqrt{X}$. We will denote by $\bar{\mathcal{A}}_q$ the complement $\mathcal{A} \setminus \mathcal{A}_q$. We will deduce Proposition A.1 from the following lemma:

Lemma A.2. *Let $\mathcal{A}$ be as above, with $P = P_0 \cup P'$ a set of odd primes. Then*

$$\bigcap_{q \text{ odd}} \bar{\mathcal{A}}_q = \frac{X}{3\cdot 2^{g-2}\zeta(2)} \prod_{p\in P'} \delta_p \prod_{p \in P_0} \delta_p + O(e\sqrt{X}),$$

with δ_p as defined in Proposition A.1.

If the set P specified in Proposition A.1 is a set of odd primes, then the proposition follows immediately from this lemma, by applying it to $\mathcal{A}^{(1)}$ with $g = 2$, $b = 1$ and then partitioning $\mathcal{A}^{(0)}$ into two disjoint sets with $g = 4$ and $b = 8$ or 12, respectively, and applying the lemma to each.

If 2 belongs to the set P specified in Proposition A.1, then we consider separately the case when 2 is specified to be ramified or unramified. If $2 \in P_0$ then $\mathcal{A}^{(1)}$ is empty. We already have $2 \mid n$ for every $n \in \mathcal{A}^{(0)}$, so we set $P_{00} = P_0 \setminus \{2\}$ and apply Lemma A.2 to $\mathcal{A}^{(0)}$ with $P = P_{00} \cup P'$ (as before, separating $\mathcal{A}^{(0)}$ into two disjoint sets and applying the lemma to each). We obtain

$$\begin{aligned}\bigcap_{q \text{ odd}} \bar{\mathcal{A}}_q &= 2 \cdot \frac{X}{3\cdot 4\zeta(2)} \prod_{p\in P'} \delta_p \prod_{p \in P_{00}} \delta_p + O(e\sqrt{X}) \\ &= \delta_2 \cdot \frac{X}{2\zeta(2)} \prod_{p\in P'} \delta_p \prod_{p \in P_{00}} \delta_p + O(e\sqrt{X}),\end{aligned}$$

with $\delta_2 = \frac{1}{3}$, as claimed.

If $2 \in P'$ then $\mathcal{A}^{(0)}$ is empty. We recall that for $p = 2$ and $n \equiv 1 \pmod 4$, the Kronecker symbol $\left(\frac{n}{2}\right) = +1$ if $n \equiv 1 \pmod 8$ and -1 if $n \equiv 5 \pmod 8$. Thus if $2 \in P'$, we set $P'' = P' \setminus \{2\}$ and $\mathcal{A}^{(1)}$ becomes

$$\mathcal{A}^{(1)} = \left\{1 \le n \le X : n \equiv b \pmod 8,\ e_0|n,\ \left(\tfrac{n}{p}\right) = \varepsilon_p, \forall p \in P''\right\},$$

with $b=1$ if the original specification was $\varepsilon_2=+1$ and $b=5$ if $\varepsilon_2=-1$. Applying Lemma A.2, we see that

$$\begin{aligned}\bigcap_{q \text{ odd}} \bar{\mathcal{A}}_q &= \frac{X}{3\cdot 2\zeta(2)} \prod_{p\in P''} \delta_p \prod_{p\in P_0} \delta_p + O(e\sqrt{X})\\ &= \delta_2 \cdot \frac{X}{2\zeta(2)} \prod_{p\in P''} \delta_p \prod_{p\in P_0} \delta_p + O(e\sqrt{X}),\end{aligned}$$

with $\delta_2=\frac{1}{3}$, again as claimed. This proves Proposition A.1.

We now prove Lemma A.2. By the inclusion-exclusion principle,

$$\bigcap_{q \text{ odd}} \bar{\mathcal{A}}_q = \sum_{m=0}^{M} (-1)^m \sum_{q_1<\cdots<q_m} |\mathcal{A}_{q_1} \cap \cdots \cap \mathcal{A}_{q_m}|, \tag{A-2}$$

in which for the $m=0$ term we sum the full set $|\mathcal{A}|$. A priori, any fixed term in (A-2) can be written as

$$\begin{aligned}&|\mathcal{A}_{q_1} \cap \cdots \cap \mathcal{A}_{q_m}|\\ &\qquad = \#\bigl\{n \le X : q_1^2 \cdots q_m^2 | n,\ n \equiv b \pmod{2^g},\ e_0 | n,\ \bigl(\tfrac{n}{p}\bigr) = \varepsilon_p, \forall p \in P'\bigr\}.\end{aligned}$$

Denote the set on the right-hand side by S, and let $Q := \{q_1, \ldots, q_m\}$. We first observe that if any $p \in P'$ belongs to Q then the set S must be empty. Thus we may reduce to considering the case in which P' and Q are disjoint, in which case we will prove that

$$\#S = \frac{1}{2^g} \frac{X}{q_1^2 \cdots q_m^2} \frac{\gcd(q_1\cdots q_m, e_0)}{e_0} \prod_{p\in P'} \frac{1}{2}\Bigl(\frac{p-1}{p}\Bigr) + O(e'). \tag{A-3}$$

First note that if a prime p in P_0 belongs to Q as well, then the condition $q_1^2\cdots q_m^2|n$ already specifies that p is ramified. Thus upon defining $e_{00} = \prod_{p\in P_0\setminus Q} p$, we may deduce that

$$\begin{aligned}S = \Bigl\{&k \le X(q_1^2\cdots q_m^2 e_{00})^{-1} :\\ &\qquad k \equiv b(q_1^2\cdots q_m^2 e_{00})^{-1} \pmod{2^g},\ \bigl(\tfrac{k}{p}\bigr) = \varepsilon_p', \forall p \in P'\Bigr\},\end{aligned} \tag{A-4}$$

where for each $p \in P'$ we have defined $\varepsilon_p' = \varepsilon_p\bigl(\frac{e_{00}}{p}\bigr)$.

We note that for any integer $K \ge 1$ and any residue class b modulo 2^g, the quantity

$$\#\bigl\{k \le K : k \equiv b \pmod{2^g},\ \bigl(\tfrac{k}{p}\bigr) = \varepsilon_p', \forall p \in P'\bigr\}$$

may be expressed as

$$\begin{aligned}\sum_{\substack{a \pmod{e'}\\(a,e')=1}} \biggl(\prod_{p\in P'} \tfrac{1}{2}\bigl(1+\varepsilon_p'\bigl(\tfrac{a}{p}\bigr)\bigr)\biggr) \sum_{\substack{k\le K\\ k\equiv a \pmod{e'}\\ k\equiv b \pmod{2^g}}} 1 &= \frac{1}{2^{|P'|}} \sum_{\substack{a \pmod{e'}\\(a,e')=1}} \prod_{p\in P'} \bigl(1+\varepsilon_p'\bigl(\tfrac{a}{p}\bigr)\bigr)\Bigl(\frac{K}{2^g e'}+O(1)\Bigr)\\ &= \biggl(\prod_{p\in P'} \frac{p-1}{p}\biggr)\frac{K}{2^{g+|P'|}} + 0 + O(e');\end{aligned}$$

all the intermediate terms vanish by orthogonality of characters. Applying this to S in (A-4), we obtain

$$\#S = \frac{X}{q_1^2\cdots q_m^2 e_{00}} \frac{1}{2^{g+|P'|}} \prod_{p\in P'} \Bigl(\frac{p-1}{p}\Bigr) + O(e'),$$

proving (A-3).

Applying (A-3) to the inclusion-exclusion in (A-2) shows that

$$\bigcap_{q \text{ odd}} \bar{\mathcal{A}}_q = \sum_{\substack{d\le\sqrt{X}\\(d,2e')=1}} \mu(d)\biggl(\frac{X}{2^g d^2}\frac{\gcd(d,e_0)}{e_0}\prod_{p\in P'}\frac{1}{2}\Bigl(\frac{p-1}{p}\Bigr)+O(e')\biggr).$$

The error term contributes $O(e\sqrt{X})$, while the main term contributes

$$X\frac{1}{2^g}\biggl(\prod_{p\in P'}\frac{1}{2}\Bigl(\frac{p-1}{p}\Bigr)\biggr)\sum_{\substack{d=1\\(d,2e')=1}}^{\infty}\frac{\mu(d)}{d^2}\frac{\gcd(d,e_0)}{e_0} + O\biggl(X\sum_{d>\sqrt{X}}\frac{1}{d^2}\biggr).$$

Here the error term is $O(\sqrt{X})$, with an implied constant which may be taken to be independent of P.

We now simplify the main term. We note that since P consists of odd primes and $(e_0,e')=1$, upon setting $d=\delta f$ with $\delta=\gcd(d,e_0)$, we have

$$\begin{aligned}\sum_{\substack{d=1\\(d,2e')=1}}^{\infty}\frac{\mu(d)}{d^2}\frac{\gcd(d,e_0)}{e_0} &= \sum_{\delta\mid e_0}\sum_{\substack{f=1\\(\delta f,2e')=(f,e_0)=1}}^{\infty}\frac{\mu(\delta f)}{\delta^2 f^2}\frac{\delta}{e_0}\\ &= \frac{1}{e_0}\biggl(\sum_{\delta\mid e_0}\frac{\mu(\delta)}{\delta}\biggr)\biggl(\sum_{\substack{f=1\\(f,2e'e_0)=1}}^{\infty}\frac{\mu(f)}{f^2}\biggr).\end{aligned}$$

The sum over $\delta | e_0$ is a multiplicative function with respect to e_0. For p prime we have

$$\sum_{\delta | p} \frac{\mu(\delta)}{\delta} = 1 - \frac{1}{p}$$

and thus for e_0 square-free we may compute by multiplicativity that

$$\frac{1}{e_0} \sum_{\delta | e_0} \frac{\mu(\delta)}{\delta} = \prod_{p \in P_0} \frac{p-1}{p^2}.$$

We next recall that for any $\Re(s) > 1$ and any distinct primes $q_1, \ldots, q_r$,

$$\left(\prod_{i=1}^{r} \Big(1 - \frac{1}{q_i^s}\Big) \zeta(s) \right)^{-1} = \prod_{p \notin \{q_1, \ldots, q_r\}} \Big(1 - \frac{1}{p^s}\Big) = \sum_{\substack{d=1 \\ (d, \prod q_i)=1}}^{\infty} \frac{\mu(d)}{d^s}.$$

Thus

$$\sum_{\substack{f=1 \\ (f, 2e' e_0)=1}}^{\infty} \frac{\mu(f)}{f^2} = \Big(1 - \frac{1}{2^2}\Big)^{-1} \prod_{p \in P} \Big(1 - \frac{1}{p^2}\Big)^{-1} \frac{1}{\zeta(2)}.$$

Assembling this all together, we see that

$$\bigcap_{q \text{ odd}} \bar{\mathcal{A}}_q$$

$$= \frac{X}{2^g} \Big(1 - \frac{1}{2^2}\Big)^{-1} \frac{1}{\zeta(2)} \prod_{p \in P'} \left(\frac{p-1}{2p} \frac{1}{1 - \frac{1}{p^2}} \right) \prod_{p \in P_0} \left(\frac{p-1}{p^2} \frac{1}{1 - \frac{1}{p^2}} \right) + O(e\sqrt{X}).$$

This reduces to

$$\bigcap_{q \text{ odd}} \bar{\mathcal{A}}_q = \frac{X}{3 \cdot 2^{g-2} \zeta(2)} \prod_{p \in P'} \frac{1}{2} \frac{1}{(1 + p^{-1})} \prod_{p \in P_0} \frac{1}{1+p} + O(e\sqrt{X}),$$

proving Lemma A.2, with δ_p as in Proposition A.1.

Acknowledgements

The authors thank Paul Pollack, Arul Shankar, Frank Thorne, and Jacob Tsimerman for very helpful conversations, and the referee for helpful comments on exposition. We thank Frank Thorne in particular for a suggestion that helped improved the exponent in Theorem 5.1. Ellenberg is partially supported by National Science Foundation Grant DMS-1402620 and a Guggenheim Fellowship. Pierce has been partially supported by National Science Foundation Grant DMS-1402121 and CAREER grant DMS-1652173. Wood is partially supported by an American Institute of Mathematics Five-Year Fellowship, a Packard Fellowship for Science and Engineering, a Sloan Research Fellowship, and National Science Foundation Grant DMS-1301690.

References

[Baily 1980] A. M. Baily, “On the density of discriminants of quartic fields”, *J. Reine Angew. Math.* **315** (1980), 190–210. MR Zbl

[Belabas et al. 2010] K. Belabas, M. Bhargava, and C. Pomerance, “Error estimates for the Davenport–Heilbronn theorems”, *Duke Math. J.* **153**:1 (2010), 173–210. MR Zbl

[Bhargava 2004] M. Bhargava, “Higher composition laws, III: The parametrization of quartic rings”, *Ann. of Math.* (2) **159**:3 (2004), 1329–1360. MR Zbl

[Bhargava 2005] M. Bhargava, “The density of discriminants of quartic rings and fields”, *Ann. of Math.* (2) **162**:2 (2005), 1031–1063. MR Zbl

[Bhargava 2008] M. Bhargava, “Higher composition laws, IV: The parametrization of quintic rings”, *Ann. of Math.* (2) **167**:1 (2008), 53–94. MR Zbl

[Bhargava 2010] M. Bhargava, “The density of discriminants of quintic rings and fields”, *Ann. of Math.* (2) **172**:3 (2010), 1559–1591. MR Zbl

[Bhargava et al. 2013] M. Bhargava, A. Shankar, and J. Tsimerman, “On the Davenport–Heilbronn theorems and second order terms”, *Invent. Math.* **193**:2 (2013), 439–499. MR Zbl

[Bhargava et al. 2017] M. Bhargava, A. Shankar, T. Taniguchi, F. Thorne, J. Tsimerman, and Y. Zhao, “Bounds on 2-torsion in class groups of number fields and integral points on elliptic curves”, preprint, 2017. arXiv

[Brumer and Silverman 1996] A. Brumer and J. H. Silverman, “The number of elliptic curves over $\mathbb{Q}$ with conductor N”, *Manuscripta Math.* **91**:1 (1996), 95–102. MR Zbl

[Cho and Kim 2015] P. J. Cho and H. H. Kim, “Central limit theorem for Artin L-functions”, preprint, 2015. arXiv

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

[Cohen and Martinet 1990] H. Cohen and J. Martinet, “Étude heuristique des groupes de classes des corps de nombres”, *J. Reine Angew. Math.* **404** (1990), 39–76. MR Zbl

[Cohen et al. 2002] H. Cohen, F. Diaz y Diaz, and M. Olivier, “Enumerating quartic dihedral extensions of $\mathbb{Q}$”, *Compositio Math.* **133**:1 (2002), 65–93. MR Zbl

[Cojocaru and Murty 2006] A. C. Cojocaru and M. R. Murty, *An introduction to sieve methods and their applications*, London Mathematical Society Student Texts **66**, Cambridge University Press, 2006. MR Zbl

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

[Duke 1998] W. Duke, “Bounds for arithmetic multiplicities”, *Doc. Math.* Extra Vol. II (1998), 163–172. MR Zbl

[Ellenberg and Venkatesh 2007] J. S. Ellenberg and A. Venkatesh, “Reflection principles and bounds for class group torsion”, *Int. Math. Res. Not.* **2007**:1 (2007), Art. ID rnm002, 18. MR Zbl

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

[Frei et al. 2015] C. Frei, D. Loughran, and R. Newton, "The Hasse norm principle for abelian extensions", preprint, 2015. arXiv

[Gauss 1966] C. F. Gauss, *Disquisitiones arithmeticae*, Yale University Press, New Haven, CT, 1966. MR Zbl

[Heath-Brown and Pierce 2014] D. Heath-Brown and L. B. Pierce, "Averages and moments associated to class numbers of imaginary quadratic fields", preprint, 2014. to appear in *Compositio Math.* arXiv

[Helfgott and Venkatesh 2006] H. A. Helfgott and A. Venkatesh, "Integral points on elliptic curves and 3-torsion in class groups", *J. Amer. Math. Soc.* **19**:3 (2006), 527–550. MR Zbl

[Hough 2010] B. Hough, "Average equidistribution of Heegner points associated to the 3-part of the class group of imaginary quadratic fields", preprint, 2010. arXiv

[Hua 1982] L. K. Hua, *Introduction to number theory*, Springer, 1982. MR Zbl

[Klys 2016] J. Klys, "The Distribution of p-Torsion in Degree p Cyclic Fields", preprint, 2016. arXiv

[Lagarias and Odlyzko 1977] J. C. Lagarias and A. M. Odlyzko, "Effective versions of the Chebotarev density theorem", pp. 409–464 in *Algebraic number fields: L-functions and Galois properties* (Durham, United Kingdom, 1975), edited by A. Fröhlich, Academic Press, London, 1977. MR Zbl

[Lemke Oliver and Thorne 2017] R. J. Lemke Oliver and F. Thorne, "The number of ramified primes in number fields of small degree", *Proc. Amer. Math. Soc.* **145**:8 (2017), 3201–3210. MR Zbl

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

[Narkiewicz 1990] W. Narkiewicz, *Elementary and analytic theory of algebraic numbers*, 2nd ed., Springer, 1990. MR Zbl

[Pappalardi 2005] F. Pappalardi, "A survey on k-freeness", pp. 71–88 in *Number theory*, edited by S. D. Adhikari et al., Ramanujan Math. Soc. Lect. Notes Ser. **1**, Ramanujan Math. Soc., Mysore, India, 2005. MR Zbl

[Pierce 2005] L. B. Pierce, "The 3-part of class numbers of quadratic fields", *J. London Math. Soc.* (2) **71**:3 (2005), 579–598. MR Zbl

[Pierce 2006] L. B. Pierce, "A bound for the 3-part of class numbers of quadratic fields by means of the square sieve", *Forum Math.* **18**:4 (2006), 677–698. MR Zbl

[Sarnak et al. 2016] P. Sarnak, S. W. Shin, and N. Templier, "Families of L-functions and their symmetry", pp. 531–578 in *Families of automorphic forms and the trace formula* (Puerto Rico, 2014), edited by W. Müller et al., Springer, 2016. Zbl

[Shankar and Tsimerman 2014] A. Shankar and J. Tsimerman, "Counting S_5-fields with a power saving error term", *Forum Math. Sigma* **2** (2014), e13, 8. MR Zbl

[Shankar et al. 2015] A. Shankar, A. Södergren, and N. Templier, "Low-lying zeros of certain families of Artin L-functions", preprint, 2015. arXiv

[Smith 2016] A. Smith, "Governing fields and statistics for 4-Selmer groups and 8-class groups", preprint, 2016. arXiv

[Soundararajan 2000] K. Soundararajan, "Divisibility of class numbers of imaginary quadratic fields", *J. London Math. Soc.* (2) **61**:3 (2000), 681–690. MR Zbl

[Taniguchi and Thorne 2013] T. Taniguchi and F. Thorne, "Secondary terms in counting functions for cubic fields", *Duke Math. J.* **162**:13 (2013), 2451–2508. MR Zbl

[Wong 1999a] S. Wong, "Automorphic forms on GL(2) and the rank of class groups", *J. Reine Angew. Math.* **515** (1999), 125–153. MR Zbl

[Wong 1999b] S. Wong, "On the rank of ideal class groups", pp. 377–383 in *Number theory* (Ottawa, ON, 1996), edited by R. Gupta and K. S. Williams, CRM Proc. Lecture Notes **19**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Wood 2010] M. M. Wood, "On the probabilities of local behaviors in abelian field extensions", *Compos. Math.* **146**:1 (2010), 102–128. MR Zbl

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

[Yang 2009] A. Yang, *Distribution problems associated to zeta functions and invariant theory*, Ph.D. thesis, Princeton University, 2009, Available at https://search.proquest.com/docview/304982142. MR

[Zhang 2005] S.-W. Zhang, "Equidistribution of CM-points on quaternion Shimura varieties", *Int. Math. Res. Not.* **2005**:59 (2005), 3657–3689. MR Zbl

Communicated by Philippe Michel
Received 2016-04-01 Revised 2017-06-10 Accepted 2017-07-10

ellenber@math.wisc.edu *Department of Mathematics, University of Wisconsin, Madison, WI 53706, United States*

pierce@math.duke.edu *Mathematics Department, Duke University, Durham, NC 27708, United States*

mmwood@math.wisc.edu *Department of Mathematics, University of Wisconsin, Van Vleck Hall, Madison, WI 53711, United States*

American Institute of Mathematics, San Jose, CA 95112, United States

ALGEBRA AND NUMBER THEORY 11:8 (2017)
dx.doi.org/10.2140/ant.2017.11.1779

Torsion orders of complete intersections

Andre Chatzistamatiou and Marc Levine

By a classical method due to Roitman, a complete intersection X of sufficiently small degree admits a rational decomposition of the diagonal. This means that some multiple of the diagonal by a positive integer N, when viewed as a cycle in the Chow group, has support in $X \times D \cup F \times X$, for some divisor D and a finite set of closed points F. The minimal such N is called the torsion order. We study lower bounds for the torsion order following the specialization method of Voisin, Colliot-Thélène, and Pirutka. We give a lower bound for the generic complete intersection with and without point. Moreover, we use methods of Kollár and Totaro to exhibit lower bounds for the very general complete intersection.

Introduction

Decomposition of the diagonal has played a prominent role in recent progress on stable rationality questions. For a rationally connected variety over a field k, there is a minimal integer $\mathrm{Tor}_k(X) \geq 1$ such that the multiple of the diagonal $\mathrm{Tor}_k(X) \cdot \Delta_X$, when viewed in the Chow group of $X \times X$, is supported in $X \times D \cup F \times X$, for some divisor D and some finite set of closed points F. We will call $\mathrm{Tor}_k(X)$ the *torsion order* of X; it is a stable birational invariant which equals 1 if X is stably

Chatzistamatiou was supported by the Heisenberg Program (DFG); Levine was supported by the DFG through the SFB Transregio 45 and the SPP 1786.
MSC2010: 14C25.
Keywords: algebraic cycles, decomposition of the diagonal.

rational and in general gives an upper bound on the exponent of the unramified cohomology of X. This invariant is also studied by Kahn [2016]. In a proper flat family the torsion order of a fiber divides the torsion order of the generic fiber (see Lemma 1.5 for the precise statement). One can thus deduce a nontrivial torsion order from a nontrivial torsion order of a cleverly chosen degeneration. In all current implementations of this strategy divisors of the torsion order of the degeneration are computed by finding a good resolution of singularities. On the resolution, the action of algebraic correspondences on a suitable cohomology can be used to produce divisors of the torsion order.

This method was pioneered by Voisin [2015]. It was significantly simplified and applied by Colliot-Thélène and Pirutka [2016b] to show the nonrationality of a very general quartic threefold by using a degeneration to a classical example of Artin and Mumford (after a "universally CH_0-trivial" resolution of singularities [Colliot-Thélène and Pirutka 2016b, Définitions 1.1 and 1.2]), which is a unirational but nonrational variety. The nontrivial 2-torsion in its Brauer group forces nontriviality of the torsion order (in fact, it implies that the torsion order is even). The degeneration method is also used in the recent work of Hasset, Pirutka, and Tschinkel [Hassett et al. 2016] exhibiting a family of smooth projective fourfolds containing both rational and nonstably rational members. Totaro [2016] used [Colliot-Thélène and Pirutka 2016b] and Voisin's method combined with work of Kollár [1995] to improve Kollár's nonrationality results for hypersurfaces in [loc. cit.]. Roughly speaking, Totaro showed how, for large enough degree, a general hypersurface of even degree degenerates to an inseparable degree-2 cover in characteristic 2 with a universally CH_0-trivial resolution of singularities that supports nonvanishing differential forms. An action of correspondences on differentials shows that the torsion order is even.

In this paper we study the torsion order of complete intersections in projective space. The method used by Roitman to show that a degree-zero 0-cycle on a hypersurface of degree $d \le n$ in $\mathbb{P}^n$ over an algebraically closed field is d-torsion is applied in Proposition 5.2 to establish an upper bound for the torsion index, more precisely, that a complete intersection X of multidegree $d_1, \dots, d_r$ in $\mathbb{P}_k^{n+r}$ (over any field k) with $\sum_{i=1}^r d_i \le n+r$ satisfies $\mathrm{Tor}_k(X) \mid \prod_{i=1}^r (d_i!)$. Our first result is a lower bound for a generic complete intersection.

Theorem (Theorem 6.5 and Corollary 6.6). *Let* $\mathcal{Y} := \prod_{i=1}^r \mathbb{P}(H^0(\mathbb{P}_k^{n+r}, \mathcal{O}(d_i))^\vee)$, *and let* $\mathcal{X} \subset \mathcal{Y} \times_k \mathbb{P}_k^{n+r}$ *be the incidence variety*

$$\mathcal{X} = \{(f_1, \dots, f_r, x) \in \mathcal{Y} \times_k \mathbb{P}_k^{n+r} \mid f_1(x) = \cdots = f_r(x) = 0\}.$$

We denote by K the quotient field of $\mathcal{Y}$, and let X/K be the generic fiber of the family $\mathcal{X} \to \mathcal{Y}$. For an integer $d \ge 1$, let $d!^*$ be the l.c.m. of the integers $1, \dots, d$. Then:

(i) $\mathrm{Tor}_K(X)$ *is divisible by* $\prod_{i=1}^r d_i!^*$.

(ii) $\mathrm{Tor}_{K(X)}(X \otimes_K K(X))$ *is divisible by* $\prod_{i=1}^r d_i!^*/(d_1 \cdots d_r)$.

The invariant which detects divisors of the torsion order in the first part of the theorem is the index of a variety, that is, the image of the Chow group of zero cycles via the degree map. The index of X/K is given by $d_1 \cdots d_r$. Divisibility of the torsion order by other integers of the form $i_1 \cdots i_r$ with $1 \leq i_j \leq d_j$ is shown by degeneration to a union of complete intersections with lower degrees and using induction.

We also consider the generic cubic hypersurface with a line, and use Theorem 6.5 to show that this has torsion order exactly 2 (Example 6.8). We show the existence of a cubic threefold over $K = \mathbb{Q}_p((x))$ or $K = \mathbb{F}_p((t))((x))$, having a K-point and torsion order divisible by 2 (Example 6.9); more generally, we construct examples of cubic hypersurfaces of dimension n over a field $K = k((x))$, where k is a field of characteristic zero and u-invariant at least $n+1$, which have a K-point and for which 2 divides the torsion order. This last series of examples is taken over from [Colliot-Thélène 2016] (without the assumption that the u-invariant is a power of 2), with the kind permission of the author, and it gives an improvement over a construction in an earlier version of this paper, which relied on Rost's degree formula. We should mention that other examples of this kind already exist in the literature; see for example [Colliot-Thélène and Pirutka 2016b, Théorème 1.21], where cubic threefolds over a p-adic field with nonzero torsion order are constructed, as well as examples over $\mathbb{F}_p((x))$ [Colliot-Thélène and Pirutka 2016b, Remarque 1.23]; both examples have a rational point.

Our second result concerns the torsion order of very general complete intersections over algebraically closed fields of characteristic zero. The idea of the proof is as in the papers of Kollár and Totaro. We are able to generalize the results on the Hodge cohomology of the degeneration in characteristic p to Hodge–Witt cohomology. In this way we can establish results on divisibility by powers of p.

Theorem (Theorem 8.2). *Let k be an algebraically closed field of characteristic zero. Let $X \subset \mathbb{P}_k^{n+r}$ be a very general complete intersection of multidegree $d_1, d_2, \ldots, d_r$ such that $d' := \sum_{i=1}^r d_i \leq n+r$ and $n \geq 3$. Let p be a prime*, *let $m \geq 1$, and suppose*

$$d_i \geq p^m \cdot \left\lceil \frac{n+r+1-d'+d_i}{p^m+1} \right\rceil$$

for some i, where $\lceil \cdot \rceil$ denotes the ceiling function. Assume that p is odd or n is even. Then $p^m \mid \mathrm{Tor}_k(X)$.

For example, it is easy to see that if $\sum_{i=1}^r d_i = n+r$ and $n \geq 3$, which is the extreme case, then $d_i \mid \mathrm{Tor}_k(X)$ if d_i is odd or n is even. For hypersurfaces and

$m = 1$, the theorem is due to Totaro, and we give a short proof of the straightforward generalization to complete intersections and the case $m = 1$ in Theorem 7.1. We should mention that our Theorems 7.1 and 8.2 are actually a bit stronger, in that we prove the same divisibility result for the torsion orders of level $n-2$ (see below), which automatically divide the torsion orders described above.

The paper is divided into seven sections. Section 1 contains the definition and basic properties of the torsion order. Following a suggestion of Claire Voisin, we consider decompositions of the diagonal of higher level and the associated torsion invariants; we also describe some elementary specialization results. In Section 3 we recall from Colliot-Thélène and Pirutka the notion of a universally CH_0-trivial morphism and a related notion, that of a totally CH_0-trivial morphism. Behavior under a combination of degeneration and modification by a birational totally CH_0-trivial morphism, which is the basic tool used for divisibility results, is the focus of Section 4; in this section we follow Colliot-Thélène and Pirutka [2016b] and extend their specialization results to cover decompositions of higher level. We recall Roitman's theorem in Section 5 and discuss the case of the generic complete intersection in Section 6. We recall Totaro's arguments leading to the divisibility results for the torsion order of a very general complete intersection in Section 7 and conclude by proving our refined version in Section 8.

1. Torsion orders

For a noetherian scheme Y, we let $\mathscr{Z}(Y)$ denote the group of algebraic cycles on Y, that is, the free abelian group on the integral closed subschemes of Y. If Y is a scheme of finite type over a field k, we grade $\mathscr{Z}(Y)$ by dimension over k. For such a scheme, we have the n-th Chow group $\mathrm{CH}_n(Y) := \mathscr{Z}_n(Y)/\sim$, where $\sim$ is the relation of rational equivalence (see [Fulton 1984, §1.3], where this group is denoted $A_n(Y)$). By an *integral component* of Y, we mean an irreducible component of Y, endowed with the reduced scheme structure.

Let k be a field and X a k-scheme of finite type. If A is a presheaf on X_{Zar}, we let

$$A(X(i)) := \mathrm{colim}_F\, A(X \setminus F)$$

where F runs over all closed subsets of X with $\dim_k F \leq i$. We extend this notation to products, defining for a presheaf A on $(X \times_k Y)_{\mathrm{Zar}}$

$$A(X(i) \times Y(j)) = \mathrm{colim}_{F,G}\, A((X \setminus F) \times_k (Y \setminus G)).$$

For example, the contravariant functoriality of the classical Chow groups for open immersions [Fulton 1984, §1.7] allows us to apply this notation to $A(X) := \mathrm{CH}_n(X)$ for some n.

Let k be a field with algebraic closure $\bar{k}$. Let P be a property of k-schemes, such as “reduced” or “smooth over k”. We say that a finite type k-scheme X is *generically* P if there exists an open Zariski dense subset $U \subset X$ having property P. The property of being generically smooth over k will be used frequently in this paper. Recall that X is generically smooth over k if and only if $X \times_k \bar{k}$ is generically smooth over $\bar{k}$. Moreover, this notion is stable under taking products; that is, if X and Y are generically smooth over k, then so is $X \times_k Y$.

A closed subset D of a finite type k-scheme X is called *nowhere dense* if the complement $X \setminus D$ is Zariski dense. We denote by $k(X)$ the product over the residue fields at the generic points of X, that is,

$$k(X) := \prod_{\eta \in X} \mathcal{O}_{X,\eta}/\mathfrak{m}_\eta,$$

where η runs over the generic points of the irreducible components of X (we note that $\mathcal{O}_{X,\eta}$ is a field if X is generically reduced). We have an evident morphism of schemes $\operatorname{Spec} k(X) \to X$. If X is equidimensional of dimension d, then we can see from the definition of Chow groups that

$$\begin{aligned}\varinjlim_{\substack{D \subset X \\ D \text{ nowhere dense}}} \mathrm{CH}_d(Y \times_k (X \setminus D)) &\xrightarrow{\cong} \mathrm{CH}_0(Y \times_k \operatorname{Spec} k(X)) \\ &= \bigoplus_{\eta \in X} \mathrm{CH}_0(Y \times_k \operatorname{Spec} \mathcal{O}_{X,\eta}/\mathfrak{m}_\eta),\end{aligned}$$

for any Y. For any class $\alpha \in \mathrm{CH}_d(Y \times_k X)$, we will call its image in $\bigoplus_{\eta \in X} \mathrm{CH}_0(Y \times_k \operatorname{Spec} \mathcal{O}_{X,\eta}/\mathfrak{m}_\eta)$ under this composition the pullback under the morphism $Y \times_k \operatorname{Spec} k(X) \to Y \times_k X$.

Definition 1.1. Let k be a field, and let X be a reduced proper k-scheme that is equidimensional of dimension d.

(1) For $i = 0, 1, 2, \ldots$, the i-th torsion order of X, $\mathrm{Tor}_k^{(i)}(X) \in \mathbb{N}_+ \cup \{\infty\}$, is the order of the image of the diagonal $\Delta_X \subset X \times_k X$ in $\mathrm{CH}_d(X(i) \times X(d-1))$. We write $\mathrm{Tor}_k(X)$ for $\mathrm{Tor}_k^{(0)}(X)$ and call this the torsion order of X.

(2) Suppose X is generically smooth over k. For $1 \le i < j \le 3$, let $p_{ij} : X \times_k X \times_k X \to X \times_k X$ denote the projection on the i-th and j-th factors, and let $\Delta_{ij} \subset X \times_k X \times_k X$ denote the pullback $p_{ij}^{-1}(\Delta_X)$. Consider the Cartesian diagram

$$\begin{tikzcd} X_{k(X \times_k X)} \arrow[r, "\tilde{j}"] \arrow[d] & X \times_k X \times_k X \arrow[d, "p_{23}"] \\ \operatorname{Spec} k(X \times_k X) \arrow[r, "i"'] & X \times_k X \end{tikzcd}$$

Let $\eta_1 - \eta_2 \in \mathrm{CH}_0(X_{k(X\times_k X)})$ denote the class of the pullback $\tilde{j}^*(\Delta_{12} - \Delta_{13})$, via the flat morphism $\tilde{j}$. The *generic torsion order* of X, $\mathrm{gTor}_k(X) \in \mathbb{N}_+ \cup \{\infty\}$, is the order of $\eta_1 - \eta_2$ in $\mathrm{CH}_0(X_{k(X\times_k X)})$.

(3) We say that X admits a decomposition of the diagonal of order N and level i if there is a nowhere dense closed subset D, a closed subset Z of X with $\dim_k Z \le i$, and cycles γ, γ' on $X \times_k X$, with γ supported in $X \times_k D$, with γ' supported in $Z \times_k X$, and with

$$N \cdot [\Delta_X] = \gamma' + \gamma$$

in $\mathrm{CH}_d(X \times_k X)$.

(4) Suppose X is geometrically integral. For an integer $N \ge 1$, we say that X admits a decomposition of the diagonal of order N if there is a 0-cycle x on X, a proper closed subset D of X, and a dimension-d cycle γ on $X \times_k X$, supported in $X \times_k D$, such that

$$N \cdot [\Delta_X] = x \times X + \gamma$$

in $\mathrm{CH}_d(X \times_k X)$. Then x has degree N over k, which can be seen by pushing forward along the second projection. We say that X admits a $\mathbb{Q}$-decomposition of the diagonal if X admits a decomposition of order N for some N, and that X admits a $\mathbb{Z}$-decomposition of the diagonal if X admits a decomposition of the diagonal of order 1.

(5) Let $\deg_k : \mathrm{CH}_0(X) \to \mathbb{Z}$ be the degree map. For X smooth and integral, the *index* of X is the positive generator I_X of the subgroup $\deg_k \mathrm{CH}_0(X) \subset \mathbb{Z}$. Equivalently, I_X is the g.c.d. of all degrees $[k(x) : k]$ as x runs over closed points of X. We extend the definition of the index to proper, integral, generically smooth k-schemes Y by defining I_Y to be the g.c.d. of all degrees $[k(y) : k]$ as y runs over closed points of the smooth locus Y_{sm} of Y (which is dense in Y).

Remarks 1.2. (1) Suppose X is equidimensional of dimension d and is geometrically integral. Since the only dimension-d cycles $\gamma_{(0)}$ on $X \times_k X$, supported on $Z_{(0)} \times_k X$ with $Z_{(0)} \subset X$ a dimension-zero closed subset, are of the form $\gamma_{(0)} = x \times X$ for some 0-cycle x on X, a decomposition of the diagonal of order N and level 0 is the same as decomposition of the diagonal of order N.

(2) We extend the definition of $\mathrm{Tor}_k^{(i)}(X)$ to all proper, equidimensional k-schemes by setting $\mathrm{Tor}_k^{(i)}(X) := \mathrm{Tor}_k^{(i)}(X_{\mathrm{red}})$.

(3) We will often use an equivalent formulation of Definition 1.1(3), namely, that X admits a decomposition of the diagonal of order N and level i if there is a closed subset D containing no generic point of X and a closed subset Z of X with $\dim_k Z \le i$ such that

$$N \cdot j^*[\Delta_X] = 0$$

in $\mathrm{CH}_d((X \setminus Z) \times_k (X \setminus D))$, where $j : (X \setminus Z) \times_k (X \setminus D) \to X \times_k X$ is the inclusion. This equivalence follows from the localization sequence

$$\mathrm{CH}_d(Z \times_k X \cup X \times_k D) \xrightarrow{i_*} \mathrm{CH}_d(X \times_k X) \xrightarrow{j^*} \mathrm{CH}_d((X \setminus Z) \times_k (X \setminus D)) \to 0$$

and the surjection

$$\mathrm{CH}_d(Z \times_k X) \oplus \mathrm{CH}_d(X \times_k D) \to \mathrm{CH}_d(Z \times_k X \cup X \times_k D).$$

(4) Decompositions of the diagonal for *smooth* proper k-varieties have been considered in [Bloch and Srinivas 1983; Colliot-Thélène and Pirutka 2016b; Totaro 2016] and by many others. Here we have extended the definition to proper, equidimensional, but not necessarily smooth k-schemes.

(5) In the same way as in [Colliot-Thélène and Pirutka 2016b, Proposition 1.4], one can prove the following equivalence for a smooth, proper, and equidimensional k-scheme X, namely, X admits a decomposition of the diagonal of order N and level i if and only if there exists a closed subvariety $\iota : Z \subset X$ with $\dim Z \leq i$ and

$$\mathrm{image}((\iota \times_k K)_* : \mathrm{CH}_0(Z \times_k K) \to \mathrm{CH}_0(X \times_k K)) \supset N \cdot \mathrm{CH}_0(X \times_k K)$$

for all field extensions $k \subset K$.

Lemma 1.3. *Let X be a k-scheme that is proper and equidimensional of dimension d over k.*

(1) *If $\mathrm{Tor}_k^{(i)}(X)$ is finite, then so is $\mathrm{Tor}_k^{(i+1)}(X)$ and in this case, $\mathrm{Tor}_k^{(i+1)}(X)$ divides $\mathrm{Tor}_k^{(i)}(X)$.*

(2) *X admits a decomposition of the diagonal of order N and level i if and only if $\mathrm{Tor}_k^{(i)}(X)$ divides N; if X is geometrically integral, then X admits a decomposition of the diagonal of order N if and only if $\mathrm{Tor}_k(X)$ divides N and X does not admit a $\mathbb{Q}$-decomposition of the diagonal if and only if $\mathrm{Tor}_k(X) = \infty$.*

(3) *Suppose X is smooth over k and geometrically integral. If $\mathrm{Tor}_k(X)$ is finite, then so is $\mathrm{gTor}_k(X)$ and $\mathrm{gTor}_k(X)$ divides $\mathrm{Tor}_k(X)$.*

(4) *Suppose X is generically smooth over k, and let $L \supset k$ be a field extension. If $\mathrm{Tor}_k^{(i)}(X)$ is finite, then so is $\mathrm{Tor}_L^{(i)}(X_L)$ and in this case $\mathrm{Tor}_L^{(i)}(X_L)$ divides $\mathrm{Tor}_k^{(i)}(X)$. If L is finite over k, then $\mathrm{Tor}_k^{(i)}(X)$ is finite if and only if $\mathrm{Tor}_L^{(i)}(X_L)$ is finite and in this case $\mathrm{Tor}_k^{(i)}(X)$ divides $[L:k] \cdot \mathrm{Tor}_L^{(i)}(X_L)$. The corresponding statements hold replacing $\mathrm{Tor}^{(i)}$ with gTor.*

(5) *X admits a decomposition of the diagonal of level i and order N if and only if there is a closed subset $Z \subset X$ of dimension $\leq i$ such that the pullback of Δ_X*

to $(X \setminus Z) \times_k \operatorname{Spec} k(X)$ *via the inclusion*

$$(X \setminus Z) \times_k \operatorname{Spec} k(X) \to X \times_k X$$

has order dividing N *in* $\mathrm{CH}_0((X \setminus Z) \times_k \operatorname{Spec} k(X))$.

Proof. Statement (1) follows from the existence of the restriction homomorphism

$$\mathrm{CH}_d((X \setminus F) \times_k (X \setminus D)) \to \mathrm{CH}_d((X \setminus F') \times_k (X \setminus D))$$

for $F \subset F'$. Statement (2) follows from the localization sequence for $\mathrm{CH}_*(\,\cdot\,)$, as in Remarks 1.2(3).

For (3), suppose

$$N \cdot [\Delta_X] = x \times X + \gamma$$

in $\mathrm{CH}_d(X \times_k X)$ for x and γ as in Definition 1.1. Since X is smooth and proper, we have for every field extension F of k the action of $\mathrm{CH}_d(X_F \times_F X_F)$ on $\mathrm{CH}_n(X_F)$ as correspondences [Fulton 1984, Chapter 16]; that is, for $\alpha \in \mathrm{CH}_d(X_F \times_F X_F)$ and $\rho \in \mathrm{CH}_n(X_F)$, one has the well defined element

$$\alpha^*(\rho) := p_{1*}(p_2^*\rho \cdot \alpha).$$

Acting by the correspondence $N \cdot \Delta^*_{X_{k(X \times_k X)}}$ on $\mathrm{CH}_0(X_{k(X \times_k X)})$ gives

$$N \cdot (\eta_1 - \eta_2) = x - x = 0$$

and thus $\mathrm{gTor}_k(X)$ divides N. Applying (2) gives (3).

For (4), the first assertion follows by applying the pullback in CH_d for $X_L \times_L X_L \to X \times_k X$ and using (2). The second part follows by applying the push-forward map $\mathrm{CH}_d(X_L \times_L X_L) \to \mathrm{CH}_d(X \times_k X)$ and using (2), and the assertion for $\mathrm{gTor}_k(X)$ follows similarly by applying the pushforward map $\mathrm{CH}_d(X_{L(X_L \times_L X_L)}) \to \mathrm{CH}_d(X_{k(X \times_k X)})$.

The last assertion (5) follows from the identity

$$\mathrm{CH}_0((X \setminus Z) \times_k \operatorname{Spec} k(X)) = \varinjlim_{D \subset X} \mathrm{CH}_d((X \setminus Z) \times_k (X \setminus D))$$

where the limit is over all closed nowhere dense $D \subset X$. □

Remark 1.4. We have restricted our attention to proper k-schemes for the definitions of torsion orders and decompositions of the diagonal. Even though the definitions would make sense for nonproper equidimensional k-schemes, a naive extension is probably not useful. Possibly replacing Chow groups with Suslin homology would make more sense: following Lemma 1.3, one could define $\mathrm{Tor}^{(i)}(X)$ for an equidimensional finite type k-scheme as the order of the restriction of Δ_X

to $X \times_k \operatorname{Spec} k(X)$ in the quotient group

$$\varinjlim_{Z \subset X} H_0^{\mathrm{Sus}}(X \times_k \operatorname{Spec} k(X))/im(H_0^{\mathrm{Sus}}(Z \times_k \operatorname{Spec} k(X)))$$

where $Z \subset X$ runs over all closed subsets of dimension at most i. We will not investigate properties of these torsion orders for nonproper k-schemes here.

Here is the first in a series of elementary but useful specialization lemmas.

Lemma 1.5. *Let $\mathcal{O}$ be a noetherian regular local ring and $f : \mathcal{X} \to \operatorname{Spec} \mathcal{O}$ a proper flat morphism, with $\mathcal{X}$ equidimensional over $\operatorname{Spec} \mathcal{O}$ of relative dimension d, $X \to \operatorname{Spec} K$ the generic fiber and $Y \to \operatorname{Spec} k$ the special fiber. Fix an integer i. Suppose that, for each $z \in \operatorname{Spec} \mathcal{O}$, the geometric fiber $\mathcal{X}_{\bar{z}}$ is generically reduced over $\overline{k(z)}$.*

(1) *If $\operatorname{Tor}_K^{(i)}(X)$ is finite, then so is $\operatorname{Tor}_k^{(i)}(Y)$, and $\operatorname{Tor}_k^{(i)}(Y)$ divides $\operatorname{Tor}_K^{(i)}(X)$.*

(2) *Suppose that, for each $z \in \operatorname{Spec} \mathcal{O}$, the fiber $\mathcal{X}_z$ is generically smooth over $k(z)$. If $\operatorname{gTor}_K(X)$ is finite, then so is $\operatorname{gTor}_k(Y)$, and $\operatorname{gTor}_k(Y)$ divides $\operatorname{gTor}_K(X)$.*

(3) *Let $\bar{k}$ and $\bar{K}$ be the algebraic closures of k and K, respectively, and suppose either K has characteristic zero, or that $\mathcal{O}$ is excellent. If $\operatorname{Tor}_{\bar{K}}^{(i)}(X_{\bar{K}})$ is finite, then so is $\operatorname{Tor}_{\bar{k}}^{(i)}(Y_{\bar{k}})$, and $\operatorname{Tor}_{\bar{k}}^{(i)}(Y_{\bar{k}})$ divides $\operatorname{Tor}_{\bar{K}}^{(i)}(X_{\bar{K}})$.*

Proof. We use the definition of $\mathrm{CH}_d(X(i) \times X(d-1))$ as a limit to reduce to making computations in groups of the form $\mathrm{CH}_d((X \setminus Z) \times_K (X \setminus D))$ where Z, D are closed subsets of X with $\dim Z \le i$ and $\dim D \le d-1$. We can find a chain of regular closed subschemes $Z_0 \subset \cdots \subset Z_r = \operatorname{Spec} \mathcal{O}$, with Z_i of Krull dimension i. This gives us the DVRs $\mathcal{O}_i := \mathcal{O}_{Z_i, Z_{i-1}}$ and the restriction of $\mathcal{X}$ to $\mathcal{X}_i \to \operatorname{Spec} \mathcal{O}_i$. Regarding the proof of (3), if the original local ring $\mathcal{O}$ has characteristic-zero quotient field, we can find a chain as above such that each DVR $\mathcal{O}_i$ has characteristic-zero quotient field, and if $\mathcal{O}$ is excellent, so are each of the $\mathcal{O}_i$. Proving the result for each of the families $\mathcal{X}_i$ gives the result for $\mathcal{X}$, which reduces us to the case of a DVR $\mathcal{O}$.

In this case, suppose we have a relation

$$N \cdot \Delta_X = 0 \tag{1-1}$$

in $\mathrm{CH}_d((X \setminus Z) \times_K (X \setminus D))$, with $\dim_K Z \le i$ and D nowhere dense. Taking the closures $\bar{Z}$ and $\bar{D}$ in $\mathcal{X}$, and letting $Z_0 = Y \cap \bar{Z}$ and $D_0 = Y \cap \bar{D}$ (as intersections of closed subsets of $\mathcal{X}$), we have the specialization homomorphism (see for example [Fulton 1984, §20.3])

$$\mathrm{sp} : \mathrm{CH}_d((X \setminus Z) \times_K (X \setminus D)) \to \mathrm{CH}_d((Y \setminus Z_0) \times_k (Y \setminus D_0))$$

associated to the family

$$\mathcal{X} \times_{\mathcal{O}} \mathcal{X} \setminus \bar{Z} \times_{\mathcal{O}} \mathcal{X} \cup \mathcal{X} \times_{\mathcal{O}} \bar{D} \to \operatorname{Spec} \mathcal{O}.$$

Since $\mathbb{O}$ is a DVR, the closure $\bar{Z}$ is automatically flat over $\operatorname{Spec}\mathbb{O}$, and thus $\dim_k Z_0 \le i$; similarly, D_0 is nowhere dense in Y. We have two cartesian diagrams

$$\begin{tikzcd} Y \times_k Y \arrow[r] & \mathscr{X} \times_{\mathbb{O}} \mathscr{X} & X \times_K X \arrow[l] \\ Y \arrow[u, "\Delta_Y"] \arrow[r] & \mathscr{X} \arrow[u, "\Delta_{\mathscr{X}}"] & X \arrow[l] \arrow[u, "\Delta_X"'] \end{tikzcd}$$

From [Fulton 1984, Proposition 20.3(a)] applied to $f = \Delta_{\mathscr{X}}$, [Fulton 1984, Example 6.2.1], and our assumption that X and Y are generically reduced, we conclude

$$\begin{aligned}\mathrm{sp}([\Delta_{X_{\mathrm{red}}}]) = \mathrm{sp}([\Delta_X]) = \mathrm{sp}(\Delta_{X*}([X])) = \Delta_{Y*}(\mathrm{sp}([X])) \\ = \Delta_{Y*}([Y]) = [\Delta_Y] = [\Delta_{Y_{\mathrm{red}}}]\end{aligned}$$

in $\mathrm{CH}_d(Y \times_k Y)$. We used [Fulton 1984, Example 6.2.1] in order to obtain $\iota^*([\mathscr{X}]) = [Y]$, where ι is the evident (regular) closed immersion, which implies $\mathrm{sp}([X]) = [Y]$ by definition of the specialization map. By using compatibility of sp with pullback along open immersions and applying sp to (1-1), we have proved (1).

The proof of (2) is a similar specialization argument. Indeed, we reduce as before to the case of a DVR $\mathbb{O}$. Due to the generic smoothness assumption, there is a dense open subscheme $\mathscr{U}$ of $\mathscr{X} \times_{\mathbb{O}} \mathscr{X}$ that is smooth over $\operatorname{Spec}\mathbb{O}$, with special fiber dense in $Y \times_k Y$. If now τ is a generic point of $Y \times_k Y$, let $\mathscr{R}$ be the local ring $\mathbb{O}_{\mathscr{U},\tau}$. Then $\mathscr{R}$ is a DVR and we may consider the $\mathscr{R}$-scheme $\mathscr{X} \otimes_{\mathbb{O}} \mathscr{R} \to \operatorname{Spec}\mathscr{R}$. The quotient field F of $\mathscr{R}$ is one of the field factors of $k(X \times_K X)$, and the residue field $\mathfrak{f}$ of $\mathscr{R}$ is the factor of $k(Y \times_k Y)$ corresponding to τ. Let η_i^X, η_i^Y, $i = 1, 2$, denote the images of the "generic" points used to define $\mathrm{gTor}_K(X)$ and $\mathrm{gTor}_k(Y)$ in $\mathrm{CH}_0(\mathscr{X}_F)$ and $\mathrm{CH}_0(Y_{\mathfrak{f}})$, respectively. Applying the specialization homomorphism

$$\mathrm{sp} : \mathrm{CH}_0(\mathscr{X}_F) \to \mathrm{CH}_0(Y_{\mathfrak{f}})$$

to a relation $N \cdot (\eta_1^X - \eta_2^X)$ in $\mathrm{CH}_0(\mathscr{X}_F)$ shows that $N \cdot (\eta_1^Y - \eta_2^Y) = 0$ in $\mathrm{CH}_0(Y_{\mathfrak{f}})$ for each generic point τ, and thus $\mathrm{gTor}_k(Y)$ divides N.

For (3), we note that there is a finite extension L of K so that

$$\mathrm{Tor}^{(i)}_{\bar{K}}(X_{\bar{K}}) = \mathrm{Tor}^{(i)}_L(X_L) = \mathrm{Tor}^{(i)}_F(X_F)$$

for all finite extensions F of L. Since either K has characteristic zero or $\mathbb{O}$ is excellent, the normalization $\mathbb{O}^N$ of $\mathbb{O}$ in L is a semilocal principal ideal ring, finite over $\mathbb{O}$ (the characteristic-zero case follows from [Zariski and Samuel 1975, Chapter V, Theorem 7], and the excellent case follows from [Matsumura 1980, Theorem 78]). Thus, after replacing $\mathbb{O}$ with the localization $\mathbb{O}'$ of $\mathbb{O}^N$ at a maximal ideal, and replacing $\mathscr{X}$ with $\mathscr{X}' := \mathscr{X} \otimes_{\mathbb{O}} \mathbb{O}'$, we may assume that $\mathrm{Tor}^{(i)}_K(X) = \mathrm{Tor}^{(i)}_{\bar{K}}(X_{\bar{K}})$. Since $\mathrm{Tor}^{(i)}_{\bar{k}}(Y_{\bar{k}})$ divides $\mathrm{Tor}^{(i)}_k(Y)$ by Lemma 1.3(4), (3) follows from (1). □

Remark 1.6. We did not use the properness of $\mathscr{X} \to \operatorname{Spec} \mathbb{O}$ in the proof of Lemma 1.5, but we have defined $\operatorname{Tor}^{(i)}$ for proper k-schemes only.

Next, we prove a modification of the specialization Lemma 1.5. A related result may be found in [Totaro 2016, Lemma 2.4].

Lemma 1.7. *Let $\mathbb{O}$ be a discrete valuation ring with quotient field K and residue field k. Let $f : \mathscr{X} \to \operatorname{Spec} \mathbb{O}$ be a flat morphism of dimension d over $\operatorname{Spec} \mathbb{O}$ with generic fiber X and special fiber Y. We suppose Y is a union of closed subschemes, $Y = Y_1 \cup Y_2$, with Y_1 and Y_2 having no common components. Suppose in addition that X admits a decomposition of the diagonal of order N and level i. Then there is an identity in $\mathrm{CH}_d(Y_1 \times_k Y_1)$*

$$N\Delta_{Y_1} = \gamma + \gamma_1 + \gamma_2$$

with γ supported in $Z_1 \times_k Y_1$ for some closed subset $Z_1 \subset Y_1$ of dimension $\le i$, γ_1 supported on $Y_1 \times_k D_1$ for some nowhere dense closed subset $D_1 \subset Y_1$, and γ_2 supported in $(Y_1 \cap Y_2) \times_k Y_1$.

Proof. We consider the (nonproper) $\mathbb{O}$-scheme $(\mathscr{X} \setminus Y_2) \times_{\mathbb{O}} (\mathscr{X} \setminus Y_2) \to \operatorname{Spec} \mathbb{O}$, closed subsets Z, D of X with $\dim_K Z \le i$, D nowhere dense, and a relation

$$N \cdot [\Delta_X] = 0$$

in $\mathrm{CH}_d((X \setminus Z) \times_K (X \setminus D))$, where $[\Delta_X]$ denotes the cycle class represented by the restriction of the diagonal.

As in the proof of Lemma 1.5(1), we have closed subsets Z_0, D_0 of $Y_1^0 := Y_1 \setminus Y_2$ with $\dim_k Z_0 \le i$, D_0 nowhere dense, and a specialization homomorphism

$$\mathrm{sp} : \mathrm{CH}_d((X \setminus Z) \times_K (X \setminus D)) \to \mathrm{CH}_d((Y_1^0 \setminus Z_0) \times_k (Y_1^0 \setminus D_0)), \tag{1-2}$$

which is induced by

$$\mathrm{sp} : \mathrm{CH}_d(X \times_K X) \to \mathrm{CH}_d(Y \times_k Y).$$

As in the proof of Lemma 1.5(1), we have $\mathrm{sp}([\Delta_X]) = [\Delta_Y]$ in $\mathrm{CH}_d(Y \times_k Y)$. It follows immediately that $\mathrm{sp}([\Delta_X]) = [\Delta_{Y_1^0}]$ in $\mathrm{CH}_d(Y_1^0 \times_k Y_1^0)$, where $[\Delta_{Y_1^0}]$ is the cycle class of the restriction of the diagonal on Y_1^0. Applying (1-2) thus gives the relation

$$N \cdot [\Delta_{Y_1^0}] = 0$$

in $\mathrm{CH}_d((Y_1^0 \setminus Z_0) \times_k (Y_1^0 \setminus D_0))$.

Let $Z_1 := \bar{Z}_0$ be the closures of Z_0 in Y_1, let $\bar{D}_0$ be the closure of D_0 in Y_1, and let $D_1 = \bar{D}_0 \cup (Y_1 \cap Y_2)$. Using the localization sequence

$$\begin{aligned}\mathrm{CH}_d(Z_1 \times_k Y_1 \cup Y_1 \times_k D_1 \cup (Y_1 \cap Y_2) \times_k Y_1)&\\ \to \mathrm{CH}_d(Y_1 \times_k Y_1) \to \mathrm{CH}_d((Y_1^0 \setminus Z_0) \times_k (Y_1^0 \setminus D_0)) \to 0&\end{aligned}$$

and the surjection

$$\mathrm{CH}_d(Z_1 \times_k Y_1) \oplus \mathrm{CH}_d(Y_1 \times_k D_1) \oplus \mathrm{CH}_d((Y_1 \cap Y_2) \times_k Y_1) \\ \to \mathrm{CH}_d(Z_1 \times_k Y_1 \cup Y_1 \times_k D_1 \cup (Y_1 \cap Y_2) \times_k Y_1),$$

the relation $N \cdot [\Delta_{Y_1^0}] = 0$ in $\mathrm{CH}_d((Y_1^0 \setminus Z_0) \times_k (Y_1^0 \setminus D_0))$ lifts to a relation of the desired form in $\mathrm{CH}_d(Y_1 \times_k Y_1)$. □

We conclude this series of specialization results with the following variation on Lemma 1.7; a similar result may be found in [Colliot-Thélène 2016, Lemme 2.2].

Lemma 1.8. *Let $\mathcal{O}$ be a discrete valuation ring with quotient field K and residue field k. Let $f : \mathcal{X} \to \operatorname{Spec} \mathcal{O}$ be a flat and proper morphism of dimension d over $\operatorname{Spec} \mathcal{O}$ with generic fiber X and special fiber Y. We suppose Y is a union of closed subschemes, $Y = Y_1 \cup Y_2$, with X and Y_1 geometrically irreducible, X generically smooth over K, and Y_1 generically smooth over k. Suppose that X admits a decomposition of the diagonal of order N. Let $Z = (Y_1 \cap Y_2)_{\mathrm{red}}$ with inclusion $i_Z : Z \to Y_1$. Suppose further that $Y_{2k(Y_1)}$ admits a zero-cycle y_2 of degree r supported in the smooth locus of $Y_{2k(Y_1)}$.*

Then there is an identity in $\mathrm{CH}_d(Y_1 \times_k Y_1)$

$$Nr\Delta_{Y_1} = \gamma_1 + \gamma_2$$

with γ_1 supported on $Y_1 \times_k D_1$, for some divisor $D_1 \subset Y_1$, and γ_2 supported in $Z \times_k Y_1$.

Proof. Let η_1 be the generic point of Y_1, let $\mathcal{O}_1 = \mathcal{O}_{\mathcal{X},\eta_1}$, and let $\mathfrak{D}$ be the henselization of $\mathcal{O}_1$. Let L be the quotient field of $\mathfrak{D}$; clearly $\mathfrak{D}$ has residue field $k(Y_1)$. Then as $\operatorname{Spec} \mathcal{O}_1 \to \operatorname{Spec} \mathcal{O}$ is essentially smooth, the base-change $\mathcal{X}_{\mathfrak{D}} := \mathcal{X} \otimes_{\mathcal{O}} \mathfrak{D} \to \operatorname{Spec} \mathfrak{D}$ has generic fiber $\mathcal{X}_L$ and special fiber $Y_{k(Y_1)} = Y_{1k(Y_1)} \cup Y_{2k(Y_1)}$. Let $\mathcal{X}_{\mathfrak{D}}^{\mathrm{sm}} \subset \mathcal{X}_{\mathfrak{D}}$ be the maximal open subscheme of $\mathcal{X}_{\mathfrak{D}}$ that is smooth over $\mathfrak{D}$.

Fix a rational equivalence

$$N \cdot \Delta_X \sim x \times X + \gamma$$

with x a 0-cycle on X and γ supported on $X \times_K E$ for some divisor E. Pulling this back to X_L gives the rational equivalence

$$N \cdot \Delta_{X_L} \sim x_L \times X_L + \gamma_L$$

with γ_L supported on $X_L \times_L E_L$. Let $\mathcal{E}$ be the closure of E_L in $\mathcal{X}_{\mathfrak{D}}$, and let $E_0 = \mathcal{E} \cap Y_{k(Y_1)}$; E_0 contains no generic point of $Y_{k(Y_1)}$. Furthermore, since the 0-cycle y_2 on $Y_{2k(Y_1)}$ is contained in the smooth locus of $Y_{2k(Y_1)}$, we may find a 0-cycle y_2' on $Y_{2k(Y_1)}$, rationally equivalent to y_2, and with support in the smooth locus of $Y_{2k(Y_1)} \setminus (E_0 \cup Z_{k(Y_1)})$. Changing notation, we may assume that y_2 is supported in the smooth locus of $Y_{2k(Y_1)} \setminus (E_0 \cup Z_{k(Y_1)})$.

Since $\mathscr{D}$ is hensel, we may lift $\eta_1 \in Y_1(k(Y_1))$ to a section $s_1 : \operatorname{Spec}\mathscr{D} \to \mathscr{X}_{\mathscr{D}}$. Since y_2 is supported in the smooth locus of $Y_{k(Y_1)}$, we may similarly lift the 0-cycle y_2 on $Y_{2k(Y_1)}$ to a cycle $\mathfrak{y}_2$ on $\mathscr{X}_{\mathscr{D}}$ of relative dimension zero and relative degree r over $\mathscr{D}$. This gives us the 0-cycle of degree zero $\rho_L := r \cdot s_1(\operatorname{Spec} L) - \mathfrak{y}_{2L}$ on X_L. Since $\mathscr{D}$ is local, $\mathscr{X}_{\mathscr{D}}$ is flat over $\mathscr{D}$ and both y_2 and η_1 are supported in the smooth locus of $Y \setminus E_0$, it follows that both $s_1(\operatorname{Spec}\mathscr{D})$ and $\mathfrak{y}_2$ are supported in $\mathscr{X}^{\mathrm{sm}}_{\mathscr{D}} \setminus \mathscr{E}$, and thus ρ_L is supported in the smooth locus of $X_L \setminus E$.

Let p be a closed point in the smooth locus of X_L, inducing the inclusion $i_p : X_L \times_L p \to X_L \times_L X_L$. Since i_p is a regular codimension-$d = \dim X$ embedding, we have the pullback map (see [Fulton 1984, §6.2, pp. 97–98], where this map is called the *Gysin homomorphism*)

$$i_p^* : \mathrm{CH}_d(X_L \times_L X_L) \to \mathrm{CH}_0(X_L \times_L p).$$

If $\mathfrak{z}$ is a 0-cycle supported in the smooth locus of X_L, $\mathfrak{z} = \sum_j n_j p_j$, we have the map

$$\mathfrak{z}^* : \mathrm{CH}_d(X_L \times_L X_L) \to \mathrm{CH}_0(X_L)$$

defined as the sum $\sum_j n_j p_{1*} \circ i_{p_j}^*$. If γ is a d-cycle on $X_L \times_L X_L$ such that each component of γ intersects each subvariety $X_L \times p_j$ properly, then $\gamma^*(\mathfrak{z})$ is well defined and

$$\mathfrak{z}^*(\gamma) = \gamma^*(\mathfrak{z}).$$

We apply these comments to the 0-cycle ρ_L and the cycles $N \cdot \Delta_{X_L}$, $x_L \times_L X_L$, and γ_L. We get the identities in $\mathrm{CH}_0(X_L)$

$$\begin{aligned} N \cdot \rho_L &= \rho_L^*(N \cdot \Delta_{X_L}) \\ &= \rho_L^*(x_L \times_L X_L) + \rho_L^*(\gamma_L). \end{aligned}$$

Both terms in this last line are zero: the first since, as X_L is irreducible, we have $\rho_L^*(x_L \times_L X_L) = \deg_L(\rho_L) \cdot x_L = 0$, and the second since $X_L \times_L \operatorname{supp}(\rho_L) \cap \operatorname{supp}(\gamma_L) = \varnothing$. In other words, $N \cdot \rho_L = 0$ in $\mathrm{CH}_0(X_L)$.

We apply the specialization map

$$\mathrm{sp} : \mathrm{CH}_0(X_L) \to \mathrm{CH}_0(Y_{k(Y_1)})$$

and find $N(r \cdot \eta_1 - y_2) = 0$ in $\mathrm{CH}_0(Y_{k(Y_1)})$. Thus, $Nr \cdot \eta_1 = 0$ in $\mathrm{CH}_0(Y_{1k(Y_1)} \setminus Z_{k(Y_1)})$, and by using the localization sequence for the inclusion $Z_{k(Y_1)} \to Y_{k(Y_1)}$, there is a 0-cycle $\gamma_{2k(Y_1)}$ on $Z_{k(Y_1)}$ with

$$Nr \cdot \eta_1 = i_{Z*}(\gamma_{2k(Y_1)})$$

in $\mathrm{CH}_0(Y_{1k(Y_1)})$. Spreading this relation out over Y_1 as in previous proofs gives the desired decomposition of $Nr \cdot \Delta_{Y_1}$. □

Remark 1.9. Suppose we have $\mathscr{X}$, $Y = Y_1 \cup Y_2$, and $Z = Y_1 \cap Y_2$ satisfying the hypotheses of Lemma 1.8; suppose in addition that Y_1 is smooth over k. Then for all fields $F \supset k$, the quotient group $\mathrm{CH}_0(Y_{1F})/i_{Z*}(\mathrm{CH}_0(Z_F))$ is Nr-torsion. Indeed, since Y_1 is smooth, we have an operation of correspondences on $\mathrm{CH}_0(Y_{1F})$, the correspondence γ_1^* of Lemma 1.8 acts trivially on $\mathrm{CH}_0(Y_{1F})$, γ_2^* maps $\mathrm{CH}_0(Y_{1F})$ to $i_{Z*}(\mathrm{CH}_0(Z_F))$, and the sum acts by multiplication by Nr.

The torsion orders behave well with respect to base-change.

Lemma 1.10. *Let X and Y be proper generically smooth k-schemes, with Y integral and with X equidimensional over k. Let K be the function field $k(Y)$ and I_Y the index of Y.*

(1) *For all i, $\mathrm{Tor}_k^{(i)}(X)$ is finite if and only if $\mathrm{Tor}_K^{(i)}(X_K)$ is finite and in this case, $\mathrm{Tor}_k^{(i)}(X)$ divides $I_Y\,\mathrm{Tor}_K^{(i)}(X_K)$.*

(2) *Suppose X is geometrically integral. If $\mathrm{gTor}_k(X)$ is finite, then so is $\mathrm{Tor}_k(X)$ and $\mathrm{Tor}_k(X)$ divides $I_X \cdot \mathrm{gTor}_k(X)$.*

Proof. For (1), if $\mathrm{Tor}_k^{(i)}(X)$ is finite, then so is $\mathrm{Tor}_K^{(i)}(X_K)$ by Lemma 1.3(4). Suppose $\mathrm{Tor}_K^{(i)}(X_K)$ is finite. Let y be a closed point of Y, contained in the smooth locus of Y over k, and let $\mathcal{O} := \mathcal{O}_{Y,y}$. Applying Lemma 1.5 to the constant family $\mathscr{X} := X \times_k \mathcal{O}$, we see that $\mathrm{Tor}_{k(y)}^{(i)}(X_{k(y)})$ is finite and $\mathrm{Tor}_{k(y)}^{(i)}(X_{k(y)})$ divides $\mathrm{Tor}_K^{(i)}(X_K)$. Applying Lemma 1.3(4) again, $\mathrm{Tor}_k^{(i)}(X)$ is finite and divides $[k(y):k]\cdot \mathrm{Tor}_{k(y)}^{(i)}(X_{k(y)})$. This proves the first assertion.

For (2), let y be a closed point of X, contained in the smooth locus of X over k, let $\mathcal{O} := \mathcal{O}_{X,y}$, and let $\eta \in X(k(X))$ be the canonical point, that is, the restriction of the diagonal section $X \to X \times_k X$ to $\operatorname{Spec} k(X)$. As in the proof of Lemma 1.5, we may find a sequence of regular closed subschemes $y = Z_0 \subset \cdots \subset Z_d = \operatorname{Spec} \mathcal{O}$, $d = \dim_k X$, and thereby define specialization homomorphisms

$$\mathrm{sp}_i : \mathrm{CH}_0(X_{k(Z_i)(X)}) \to \mathrm{CH}_0(X_{k(Z_{i-1})(X)}), \quad i = 1, \dots, d.$$

Letting $\mathrm{sp}_y : \mathrm{CH}_0(X_{k(X\times_k X)}) \to \mathrm{CH}_0(X_{k(y)(X)})$ be the composition of the sp_i, we have $\mathrm{sp}_y(\eta_1 - \eta_2) = \eta_y - y_{\mathrm{gen}}$, where $\eta_y \in X(k(y)(X))$ is the base-change of $y \in X(k(y))$ and $y_{\mathrm{gen}} \in X(k(y)(X))$ is the base-change of $\eta \in X(k(X))$. Thus, $\mathrm{gTor}_k(X)\cdot(\eta_y - y_{\mathrm{gen}}) = 0$ in $\mathrm{CH}_0(X_{k(y)(X)})$; pushing forward to $\mathrm{CH}_0(X_{k(X)})$ gives $[k(y):k]\cdot \mathrm{gTor}_k(X)\cdot \eta - \mathrm{gTor}_k(X)\cdot y \times_k k(X) = 0$ in $\mathrm{CH}_0(X_{k(X)})$. Applying localization gives us the decomposition of the diagonal Δ_X of order $[k(y):k]\cdot \mathrm{gTor}_k(X)$; doing this for each closed point y gives us the decomposition of the diagonal of order $I_X \cdot \mathrm{gTor}_k(X)$. Hence, $\mathrm{Tor}_k(X)$ is finite and divides $I_X \cdot \mathrm{gTor}_k(X)$. □

For example, $\mathrm{Tor}_k^{(i)}(X) = \mathrm{Tor}_L^{(i)}(X_L)$ if L is a pure transcendental extension of a field k.

Lemma 1.11. *Let X be a proper k-scheme. Let $k \subset L$ be an extension of fields with k algebraically closed. The following hold*:

(1) *For all i*, $\mathrm{Tor}_k^{(i)}(X) = \mathrm{Tor}_L^{(i)}(X_L)$.

(2) *Suppose in addition X is smooth and integral. Then* $\mathrm{gTor}_k(X) = \mathrm{gTor}_L(X_L)$ *and* $\mathrm{Tor}_k(X) = \mathrm{gTor}_k(X)$.

Proof. We may assume that L is finitely generated over k. Using openness of the regular locus for finite type k-schemes and k algebraically closed, we can find a noetherian local regular k-algebra $\mathbb{O}$ with quotient field L and residue field k. Applying Lemma 1.5(1) to $\mathscr{X} := X \times_k \mathbb{O} \to \operatorname{Spec} \mathbb{O}$ implies (1).

The assertions about gTor follow from (1), Lemma 1.10(2), and Lemma 1.3. □

Definition 1.12. Let X be a proper, generically smooth k-scheme. Let $\bar{k}$ be the algebraic closure of k, and define $\mathrm{Tor}^{(i)}(X) := \mathrm{Tor}_{\bar{k}}^{(i)}(X_{\bar{k}})$. We call $\mathrm{Tor}^{(i)}(X)$ the i-th *geometric torsion order* of X. We write $\mathrm{Tor}(X)$ for $\mathrm{Tor}^{(0)}(X)$.

Note that $\mathrm{Tor}^{(i)}(X)$ is invariant under base-extension $X \rightsquigarrow X_L$ for a field extension $L \supset k$. Also, assuming X to be smooth and geometrically integral, $\mathrm{Tor}(X)$ is equal to $\mathrm{gTor}_{\bar{k}}(X_{\bar{k}})$.

In much the same vein as Lemma 1.3, we show that the generic torsion order measures the torsion order after adjoining a "generic" rational point, that is:

Lemma 1.13. *Let X be a smooth proper geometrically integral k-scheme*, *and let $K = k(X)$. Then* $\mathrm{gTor}_k(X) = \mathrm{Tor}_K(X_K)$.

Proof. If $N \cdot (\eta_1 - \eta_2) = 0$ in $\mathrm{CH}_0(X_{k(X \times_k X)})$, then we have a decomposition of the diagonal of order N for $X_{k(X)}$:

$$N \cdot \Delta_{X_K} = N \cdot [\eta] \times_K X_K + \gamma$$

with γ supported in $X_K \times_K D$, with $D \subsetneq X_K$, and with η the restriction of Δ_X to $X \times_k k(X) \subset X \times_k X$. In other words, η is the K-rational point of X_K induced by the generic point of X. Thus, $\mathrm{Tor}_K(X_K)$ divides $\mathrm{gTor}_k X$. Conversely, if X_K admits a decomposition of the diagonal of order n,

$$n \cdot \Delta_{X_K} = x \times X_K + \gamma \tag{1-3}$$

with x a 0-cycle on X_K and γ supported on $X_K \times_K D$ for some divisor $D \subset X_K$, then pulling (1-3) back along $(\mathrm{id}_{X_K}, \eta) : X_K \to X_K \times_K X_K$ gives us $x = n \cdot [\eta]$ in $\mathrm{CH}_0(X_K)$, so $n \cdot \Delta_{X_K} = n \cdot [\eta] \times X_K + \gamma$ in $\mathrm{CH}_d(X_K \times_K X_K)$. Restriction to $X_K \times_K K(X_K)$ gives $n \cdot \eta_1 = n \cdot \eta_2$ in $\mathrm{CH}_0(X_{k(X \times_k X)})$, so $\mathrm{gTor}_k(X)$ divides $\mathrm{Tor}_K(X_K)$. □

One last elementary property of the torsion indices concerns the behavior with respect to morphisms.

Lemma 1.14. *Let $f : Y \to X$ be a surjective morphism of integral reduced proper k-schemes of the same dimension d. Then $\mathrm{Tor}_k^{(i)} X$ divides $\deg f \cdot \mathrm{Tor}_k^{(i)} Y$ for all i. If X and Y are generically smooth over k, then $\mathrm{gTor}_k X$ divides $(\deg f)^2 \cdot \mathrm{gTor}_k Y$.*

Proof. Suppose the diagonal for Y admits a decomposition of order N and level i:

$$N \cdot \Delta_Y = \gamma_i + \gamma'$$

with γ' supported on $Y \times_k D$ for some divisor D and γ_i supported on $Z \times_k Y$ for some closed subset Z of Y with $\dim_k Z \leq i$. Pushing forward by $f \times f$ gives

$$\deg f \cdot N \cdot \Delta_X = (f \times f)_* \gamma_i + (f \times f)_* \gamma',$$

and thus $\mathrm{Tor}_k^{(i)} X$ divides $\deg f \cdot \mathrm{Tor}_k^{(i)} Y$. Similarly, we have $(f \times f \times f)_*(\Delta_{Y,ij}) = (\deg f)^2 \cdot \Delta_{X,ij}$ for $ij = 12, 13$, which shows $\mathrm{gTor}_k X$ divides $(\deg f)^2 \cdot \mathrm{gTor}_k Y$. □

The behavior of the torsion indices with respect to rational and birational maps will be discussed in Section 3.

2. Torsion orders for very general fibers

The following global version of Lemma 1.5(3) follows by an argument using Hilbert schemes. See [Voisin 2015, Theorem 1.1 and Proposition 1.4] or [Colliot-Thélène and Pirutka 2016b, Appendice B] for similar statements. This result will only be used in Sections 7 and 8.

Proposition 2.1. *Let $p : \mathscr{X} \to B$ be a flat, equidimensional, and projective family over a scheme B of finite type over a field k, and let b_0 be a point of B. We suppose that each geometric fiber of p is generically reduced. Fix an integer $i \geq 0$. Then there is a countable union of closed subsets $F = \bigcup_{j=1}^{\infty} F_j$ with $b_0 \notin F$ such that for all $b \in B \setminus F$, the geometric fiber $\mathscr{X}_{\overline{k(b)}}$ satisfies $\mathrm{Tor}^{(i)}(\mathscr{X}_{\overline{k(b_0)}}) \mid \mathrm{Tor}^{(i)}(\mathscr{X}_{\overline{k(b)}})$. Here we use the convention that $N \mid \infty$ for all $N \in \mathbb{N}_+ \cup \{\infty\}$ and $\infty \mid N \implies N = \infty$.*

The proof uses the following elementary lemma, which we were not able to find in the literature.

Let X be a noetherian equidimensional scheme with integral components $X_1, \dots, X_t$. Let $x_i \in X_i$ be the generic point. The *associated cycle* [Fulton 1984, §1.5] of X is the cycle $\mathrm{cyc}(X) := \sum_{i=1}^{s} e_i X_i \in \mathscr{Z}(X)$ with e_i defined as

$$e_i := \mathrm{lng}_{\mathscr{O}_{X_i,x_i}} \mathscr{O}_{X_i,x_i}.$$

Lemma 2.2. *Let B be a noetherian scheme, let $p : \mathscr{Y} \to B$ be a flat morphism, and let $\mathscr{W}_0, \mathscr{W}_1, \dots, \mathscr{W}_s$ be closed subschemes of $\mathscr{Y}$, flat and equidimensional of dimension r over B. For each $b \in B$, let $\mathscr{W}_{ib} \subset \mathscr{Y}_b$ be the respective fibers over b. Fix integers $m_0, \dots, m_s$.*

(1) *The subset*

$$\mathcal{T}_{\mathcal{Y}}(B) := \left\{ b \in B \;\middle|\; \sum_{i=0}^{s} m_i \cdot \operatorname{cyc}(\mathcal{W}_{ib}) = 0 \textit{ in } \mathcal{Z}(\mathcal{Y}_b) \right\}$$

is a constructible subset of B.

(2) *For each $b \in B$, let $\bar{b}$ be a geometric point mapping to b, and $\mathcal{W}_{i\bar{b}} \subset \mathcal{Y}_{\bar{b}}$ be the respective fibers over $\bar{b}$. Fix integers $m_0, \dots, m_s$. Then, the equality*

$$\mathcal{T}_{\mathcal{Y}}(B) = \left\{ b \in B \;\middle|\; \sum_{i=0}^{s} m_i \cdot \operatorname{cyc}(\mathcal{W}_{i\bar{b}}) = 0 \textit{ in } \mathcal{Z}(\mathcal{Y}_{\bar{b}}) \right\} \tag{2-1}$$

holds.

Proof. We first prove (1); we proceed by a series of reductions. Firstly, we may assume that B is integral and separated. As the assertion is obvious if B is a point, we may use noetherian induction and replace B with any dense open subscheme. Moreover, if $\mathcal{Y} = \bigcup_i \mathcal{U}_i$ is a finite open covering, then $\mathcal{T}_{\mathcal{Y}}(B) = \bigcap_i \mathcal{T}_{\mathcal{U}_i}(B)$; hence, it suffices to prove the assertion for each i.

Let S be the set of all integral components of the subschemes $\mathcal{W}_0, \dots, \mathcal{W}_s$. Let us consider the elements of S as integral schemes. We claim that there is an open dense subset U of B such that for all $b \in U$ and $\mathcal{W}, \mathcal{V} \in S$ with $\mathcal{W} \neq \mathcal{V}$, $\mathcal{W}_b$ and $\mathcal{V}_b$ have no common integral component. Indeed, for $\mathcal{W}, \mathcal{V} \in S$ with $\mathcal{W} \neq \mathcal{V}$, there is an open dense $U_{\mathcal{V},\mathcal{W}} \subset B$ such that $\mathcal{W} \times_{\mathcal{Y}} \mathcal{V} \times_{\mathcal{Y}} p^{-1}(U_{\mathcal{V},\mathcal{W}}) \to U_{\mathcal{V},\mathcal{W}}$ is flat of relative dimension $\leq r-1$ (it need not be equidimensional). All integral components of $\mathcal{W}_b$ and $\mathcal{V}_b$ have dimension r, and therefore $\mathcal{W}_b$ and $\mathcal{V}_b$ do not have a common integral component if $b \in U_{\mathcal{V},\mathcal{W}}$. Taking the intersection of the $U_{\mathcal{V},\mathcal{W}}$ for all $\mathcal{V} \neq \mathcal{W}$ in S gives the desired open dense subset U.

After passing from B to U, we get $\mathcal{T}_{\mathcal{Y}}(B) = \mathcal{T}_{\mathcal{Y}'}(B)$ with

$$\mathcal{Y}' = \mathcal{Y} \setminus \bigcup_{\substack{\mathcal{W},\mathcal{V} \in S \\ \mathcal{W} \neq \mathcal{V}}} \mathcal{W} \cap \mathcal{V} = \bigcup_{\mathcal{W} \in S} \mathcal{Y} \setminus \Bigg(\bigcup_{\mathcal{V} \in S \setminus \{\mathcal{W}\}} \mathcal{V} \Bigg).$$

Therefore, we may suppose $S = \{\mathcal{W}\}$, in other words, there is only one integral component.

For each i, define n_i by $\operatorname{cyc}(\mathcal{W}_i) = n_i \cdot \mathcal{W}$. We claim that there is an open dense $U \subset B$ such that $\operatorname{cyc}(\mathcal{W}_{ib}) = n_i \cdot \operatorname{cyc}(\mathcal{W}_b)$ for all $b \in U$. This will imply the assertion, because after shrinking U further so that $\mathcal{W}_b \neq \varnothing$ holds, either $\mathcal{T}_{\mathcal{Y} \cap p^{-1}(U)}(U) = U$ or $\mathcal{T}_{\mathcal{Y} \cap p^{-1}(U)}(U) = \varnothing$, depending on whether $0 = \sum_i m_i \cdot n_i$ holds.

In order to prove our claim, let η be the generic point of $\mathcal{W}$ and $\mathcal{W}_i$. Since $\mathcal{O}_{\mathcal{W},\eta}$ is a field, there is a filtration

$$\mathcal{O}_{\mathcal{W}_i,\eta} = F^0 \subset F^1 \supset \cdots \supset F^r = 0$$

by $\mathbb{O}_{\mathcal{W}_i,\eta}$ submodules such that the quotients F^{j+1}/F^j are free $\mathbb{O}_{\mathcal{W},\eta}$ modules of rank r_j. By definition, the equality $n_i = \sum_j r_j$ holds. We can extend this filtration to a nonempty open subset $\mathcal{W}'_i$ of $\mathcal{W}_i$ such that the quotients are free $\mathbb{O}_{\mathcal{W}'}$ modules of rank r_j, where $\mathcal{W}' = \mathcal{W} \cap \mathcal{W}'_i$. We denote it by

$$\mathbb{O}_{\mathcal{W}'_i} = \tilde{F}^0 \subset \tilde{F}^1 \supset \cdots \supset \tilde{F}^r = 0.$$

Define $U \subset B$ to be a nonempty open subset such that $(\mathcal{W} \setminus \mathcal{W}') \cap p^{-1}(U) \to U$ is flat of relative dimension $\leq r-1$ and $\mathcal{W}' \cap p^{-1}(U) \to U$ is flat. For $b \in U$, the generic points of the integral components of $\mathcal{W}_{ib}$ (= integral components of $\mathcal{W}_b$) are contained in $\mathcal{W}'_i$. Flatness of $\mathcal{W}' \cap p^{-1}(U) \to U$ implies that we get an induced filtration

$$\mathbb{O}_{\mathcal{W}'_{ib}} = \tilde{F}^0_b \subset \tilde{F}^1_b \supset \cdots \supset \tilde{F}^r_b = 0$$

with quotients $\tilde{F}^{j+1}_b/\tilde{F}^j_b$ free of rank r_j as $\mathbb{O}_{\mathcal{W}'_b}$-modules. For every generic point ϵ of $\mathcal{W}_{ib}$ (hence $\mathcal{W}'_{ib}$) we get

$$\operatorname{lng}_{\mathbb{O}_{\mathcal{W}_{ib},\epsilon}} \mathbb{O}_{\mathcal{W}_{ib},\epsilon} = \Big(\sum_j r_j\Big) \cdot \operatorname{lng}_{\mathbb{O}_{\mathcal{W}_b,\epsilon}} \mathbb{O}_{\mathcal{W}_b,\epsilon} = n_i \cdot \operatorname{lng}_{\mathbb{O}_{\mathcal{W}_b,\epsilon}} \mathbb{O}_{\mathcal{W}_b,\epsilon},$$

which proves the claim.

For (2), we note that for each $\bar{b} \to b \in B$, the map $\bar{b} \to b$ is flat and the pullback map

$$\mathfrak{X}(Y_b) \to \mathfrak{X}(Y_{\bar{b}})$$

is injective; (2) follows directly from this and (1). $\square$

Proof of Proposition 2.1. Let d be the relative dimension of $\mathcal{X}$ over B. For a positive integer M, let $\mathcal{S}(M)$ be the set of $b \in B$ such that M does not divide $\operatorname{Tor}^{(i)}(\mathcal{X}_{\overline{k(b)}})$. Taking $M = \operatorname{Tor}^{(i)}(\mathcal{X}_{\overline{k(b_0)}})$ and $F = \mathcal{S}(M)$, it suffices to show that $\mathcal{S}(M)$ is a countable union of closed subsets of B.

We first show that $\mathcal{S}(M)$ is closed under specialization. Indeed, if we have a specialization $b \rightsquigarrow \tilde{b}$ with $b \in \mathcal{S}(M)$, then there is an excellent DVR $\mathbb{O}$ and a morphism $\operatorname{Spec} \mathbb{O} \to B$ with b the image of the generic point of $\operatorname{Spec} \mathbb{O}$ and $\tilde{b}$ the image of the closed point. Indeed, let C be the closure of b in B, blow-up $\operatorname{Spec} \mathbb{O}_{C,\tilde{b}}$ along $\tilde{b}$, normalize to obtain a normal scheme $\pi : T \to \operatorname{Spec} \mathbb{O}_{C,\tilde{b}}$ of finite type over $\mathbb{O}_{C,\tilde{b}}$, choose a generic point t of the Cartier divisor $\pi^{-1}(\tilde{b})$ on T, and take $\mathbb{O} := \mathbb{O}_{T,t}$. The local ring $\mathbb{O}_{C,\tilde{b}}$ is excellent since C is of finite type over a field, and the operations used in constructing $\mathbb{O}$ from $\mathbb{O}_{C,\tilde{b}}$ all preserve excellence [Matsumura 1980, Chapters 12 and 13]. Pulling back $\mathcal{X}$ to $\operatorname{Spec} \mathbb{O}$, it follows from Lemmas 1.5(3) and 1.11(1) that $\tilde{b}$ is also in $\mathcal{S}(M)$.

Since $\mathcal{S}(M)$ is closed under specialization, it suffices to show that, for each affine open subscheme U of B, $\mathcal{S}(M) \cap U$ is a countable union of constructible subsets of U. Thus, we may assume that B is affine, and that $\mathscr{X}$ is a closed subscheme of $B \times_k \mathbb{P}^n_k$ for some n, with $p : \mathscr{X} \to B$ the restriction of the projection.

By standard Hilbert scheme arguments, there is a projective B-scheme $q : \mathscr{Y}_{\alpha,\beta} \to B$ such that the geometric points of $\mathscr{Y}_{\alpha,\beta}$ consist of triples (b, Z_b, D_b), with b a geometric point of B, $Z_b \subset \mathscr{X}_b$ a closed subscheme of dimension $j \leq i$, and $D_b \subset \mathscr{X}_b$ a closed subscheme of dimension $< d$, and with Z_b and D_b having fixed Hilbert polynomials α, β. Let $\mathscr{Z} \subset \mathscr{X} \times_B \mathscr{Y}_{\alpha,\beta}$ and $\mathscr{D} \subset \mathscr{X} \times_B \mathscr{Y}_{\alpha,\beta}$ be the universal subschemes. We set

$$\mathscr{U} := \mathscr{X} \times_B \mathscr{X} \times_B \mathscr{Y}_{\alpha,\beta} \setminus (\mathscr{X} \times_B \mathscr{D} \cup \mathscr{Z} \times_B \mathscr{X}).$$

Similarly, there is a finite type B-scheme $g : \mathscr{W}_\phi \to B$ whose geometric points consist of pairs (b, W) with $W \subset \mathscr{X}_b \times_b \mathscr{X}_b \times_b \mathbb{P}^1_b$ a closed subscheme of dimension $d+1$, having Hilbert polynomial ϕ and being flat over $\mathbb{P}^1_b$. Indeed, denoting by

$$H := \mathrm{Hilb}_\phi(\mathscr{X} \times_B \mathscr{X} \times_B \mathbb{P}^1_B)$$

the Hilbert scheme, we can consider the subfunctor F of H defined by

$$F(T) = \{W \in H(T) \mid W \to T \times_B \mathbb{P}^1_B \text{ is flat}\}.$$

If $W_{\mathrm{uni}} \subset H \times_B \mathscr{X} \times_B \mathscr{X} \times_B \mathbb{P}^1_B$ denotes the universal subscheme, then we let $W'_{\mathrm{uni}} \subset W_{\mathrm{uni}}$ be the closed subset where $W_{\mathrm{uni}} \to H \times_B \mathbb{P}^1_B$ is not flat. We define $\mathscr{W}_\phi$ as the complement of the image of W'_{uni} in H. By using critère de platitude par fibres [EGA IV$_3$ 1966, Théorème 11.3.10], we conclude that $\mathscr{W}_\phi$ represents F.

For each integer $r \geq 1$, and each choice of Hilbert polynomials α, β and $\phi_1, \ldots, \phi_r$, we obtain subschemes $\mathscr{W}^0_1, \mathscr{W}^\infty_1, \ldots, \mathscr{W}^0_r, \mathscr{W}^\infty_r$ of $\mathscr{U} \times_B \mathscr{W}_{\phi_1} \times_B \cdots \times_B \mathscr{W}_{\phi_r}$ that are flat of relative dimension d over $\mathscr{Y}_{\alpha,\beta} \times_B \mathscr{W}_{\phi_1} \times_B \cdots \times_B \mathscr{W}_{\phi_r}$ as follows. Let $\mathscr{V}_i \subset \mathscr{X} \times_B \mathscr{X} \times_B \mathbb{P}^1_B \times_B \mathscr{W}_{\phi_i}$ be the universal subscheme. Since $\mathscr{V}_i \to \mathbb{P}^1_B \times_B \mathscr{W}_{\phi_i}$ is flat, the base-change $\mathscr{V}^\epsilon_i$ to $\mathscr{W}_{\phi_i}$ via $B \xrightarrow{\epsilon} \mathbb{P}^1_B$, for $\epsilon = 0$ and $\epsilon = \infty$, is flat. We define $\mathscr{W}^\epsilon_1$ to be the restriction of $\mathscr{V}^\epsilon_1 \times_B \mathscr{Y}_{\alpha,\beta} \times_B \mathscr{W}_{\phi_2} \times_B \cdots \times_B \mathscr{W}_{\phi_2}$ to $\mathscr{U} \times_B \mathscr{W}_{\phi_1} \times_B \cdots \times_B \mathscr{W}_{\phi_r}$, and similarly for $\mathscr{W}^\epsilon_i$.

Fix a sequence of integers $m_1, \ldots, m_r$ and an integer $N > 0$. By Lemma 2.2, the image of all geometric points $(b, Z_b, D_b, W_1, \ldots, W_r)$ satisfying

$$N \cdot \Delta_{\mathscr{X}_b \times_b \mathscr{X}_b}|_{\mathscr{X}_b \times_b \mathscr{X}_b \setminus (\mathscr{X}_b \times_b D_b \cup \mathscr{Z}_b \times_b \mathscr{X}_b)} = \sum_i m_i \cdot (\mathrm{cyc}(W^0_i) - \mathrm{cyc}(W^\infty_i)),$$

where W^ϵ_i is the scheme theoretic intersection of W_i with $\mathscr{X}_b \times_b \mathscr{X}_b \times \epsilon$ in $\mathscr{X}_b \times_b \mathscr{X}_b \times_b \mathbb{P}^1_b$, forms a constructible subset $\mathscr{T}_{r,\phi_*,\alpha,\beta,m_*,N}$ of $\mathscr{Y}_{\alpha,\beta} \times_B \mathscr{W}_{\phi_1} \times_B \cdots \times_B \mathscr{W}_{\phi_r}$. Let $\mathscr{R}_{r,\phi_*,\alpha,\beta,m_*,N}$ be the image of $\mathscr{T}_{r,\phi_*,\alpha,\beta,m_*,N}$ in B.

If b is a geometric point of B with image in $\mathscr{R}_{r,\phi_*,\alpha,\beta,m_*,N}$, and if we choose a geometric point $(b, Z_b, D_b, W_1, \ldots, W_r)$ of $\mathscr{Y}_{\alpha,\beta} \times_B \mathscr{W}_{\phi_1} \times_B \cdots \times_B \mathscr{W}_{\phi_r}$ lying over b, then the cycle $\sum_i m_i \cdot \mathrm{cyc}(W_i)$ on $\mathscr{X}_b \times_b \mathscr{X}_b \times \mathbb{P}^1$ gives a rational equivalence showing that $\mathrm{Tor}^{(i)}(\mathscr{X}_b) \mid N$. Conversely, as each integral closed subscheme $W \subset$

$\mathscr{X}_b \times_b \mathscr{X}_b \times \mathbb{P}^1$ that dominates $\mathbb{P}^1$ is flat over $\mathbb{P}^1$, each geometric point $b \in B$ such that $\mathrm{Tor}^{(i)}(\mathscr{X}_b) \mid N$ is in $\mathscr{R}_{r,\phi_*,\alpha,\beta,m_*,N}$ for some choice of Hilbert polynomials α, β, ϕ_* and integers r and $m_1, \dots, m_r$. Thus, $\mathscr{S}(M)$ is the union of the subsets $\mathscr{R}_{r,\phi_*,\alpha,\beta,m_*,N}$ over all α, β, r, ϕ_*, m_*, and all $N > 0$ not divisible by M. As this set of choices is countable, it follows that $\mathscr{S}(M)$ is a countable union of constructible subsets of B. Since $\mathscr{S}(M)$ is closed under specialization, the proof is complete. □

3. Universally and totally CH_0-trivial morphisms

We recall the notion of a universally CH_0-trivial morphism and a related notion, that of a totally CH_0-trivial morphism.

Definition 3.1 [Colliot-Thélène and Pirutka 2016b, Définitions 1.1 and 1.2]. Let $p : Z \to Y$ be a proper morphism of finite type k-schemes for some field k. The morphism p is *universally* CH_0*-trivial* if for all field extensions $F \supset k$, the map $p_* : \mathrm{CH}_0(Z_F) \to \mathrm{CH}_0(Y_F)$ is an isomorphism. A proper k-scheme $\pi_Y : Y \to \operatorname{Spec} k$ is called a universally CH_0-trivial k-scheme if π_Y is a universally CH_0-trivial morphism.

Definition 3.2. A proper morphism $p : Z \to Y$ of k-schemes is *totally* CH_0*-trivial* if for each point $y \in Y$, the fiber $p^{-1}(y)$ is a universally CH_0-trivial $k(y)$-scheme.

It follows directly from the definition that the property of a proper morphism being totally CH_0-trivial is stable under arbitrary base-change.

We rephrase a result of Colliot-Thélène and Pirutka.

Proposition 3.3 [Colliot-Thélène and Pirutka 2016b, Proposition 1.8]. *Let $p : Z \to Y$ be a totally* CH_0*-trivial morphism. Then p is universally* CH_0*-trivial.*

Remarks 3.4. (1) By the base-change property of totally CH_0-trivial morphisms, we see that for $p : Z \to Y$ a totally CH_0-trivial morphism and $W \to Y$ a morphism of k-schemes, the projection $Z \times_Y W \to W$ is universally CH_0-trivial.

(2) There are examples of universally CH_0-trivial morphisms that are not totally CH_0-trivial;[1] in particular, the property of a morphism being universally CH_0-trivial is not stable under base-change.

Corollary 3.5. (1) *Universally* CH_0*-trivial morphisms and totally* CH_0*-trivial morphisms are closed under composition.*

[1] For example, let k be an algebraically closed field of characteristic $\neq 2$, let S be the cone in $\mathbb{P}^3_k$ over a smooth plane curve C of degree ≥ 3, let $Y \to S$ be the double cover branched over the transverse intersection of S with a quadric, and let $y_1, y_2 \in Y$ be the points lying over the vertex of S. Let $p : Z \to Y$ be the blow-up of Y at y_1, and let $z = p^{-1}(y_2)$. Then for all fields $L \supset k$, $\mathrm{CH}_0(z_L) \xrightarrow{i_{z*}} \mathrm{CH}_0(Z_L)$ and $\mathrm{CH}_0(y_{2L}) \xrightarrow{i_{y_2*}} \mathrm{CH}_0(Y_L)$ are isomorphisms, and thus p is universally CH_0-trivial. However, $p^{-1}(y_1) \cong C$, so p is not totally CH_0-trivial.

(2) *Let $p : Z \to Y$ be a morphism of smooth k-schemes that is a sequence of blow-ups with smooth centers. Then p is a totally* CH_0*-trivial morphism.*

(3) *Suppose that the field k admits resolution of singularities of birational morphisms for smooth k-schemes of dimension $\leq d$; that is, if $p : Z \to Y$ is a proper birational morphism of smooth k-schemes of dimension $\leq d$, there is a sequence of blow-ups of Y with smooth centers, $q : W \to Y$, such that the resulting birational map $r : W \to Z$ is a morphism. Then each proper birational morphism $p : Z \to Y$ of smooth k-schemes of dimension $\leq d$ is totally* CH_0*-trivial. In particular, this holds for k of characteristic zero, or for $d \leq 3$ and k algebraically closed* [Abhyankar 1966].

Proof. Statement (1) for universally CH_0-trivial morphisms is obvious from the definition, and for totally CH_0-trivial morphisms this follows with the help of Proposition 3.3.

For (2), we use (1) to reduce to checking for the blow-up of Y along a smooth closed subscheme F, for which the assertion is clear.

For (3), let y be a point of Y and $L \supset k(y)$ a field extension. Dominating Z by a $q : W \to Y$ as above, we have the maps

$$\mathrm{CH}_0(q^{-1}(y)_L) \xrightarrow{r_*} \mathrm{CH}_0(p^{-1}(y)_L) \xrightarrow{p_*} \mathrm{CH}_0(\mathrm{Spec}\, L) = \mathbb{Z}$$

which, as $\mathrm{CH}_0(q^{-1}(y)_L) \to \mathrm{CH}_0(\mathrm{Spec}\, L)$ is an isomorphism, gives us a splitting to p_*. Applying resolution of singularities to $r : W \to Z$ gives a sequence of blow-ups with smooth centers $s : X \to Z$ such that $t := r^{-1}s : X \to W$ is a morphism. Since $X \to Z$ is totally CH_0-trivial, the sequence

$$\mathrm{CH}_0(t^{-1}(q^{-1}(y))_L) \xrightarrow{t_*} \mathrm{CH}_0(q^{-1}(y)_L) \xrightarrow{r_*} \mathrm{CH}_0(p^{-1}(y)_L)$$

gives a splitting to r_*, so p_* is an isomorphism. □

Lemma 3.6. (1) *Let $q : Z \to Y$ be a birational totally* CH_0*-trivial morphism of integral, generically smooth k-schemes. Let $N > 0$ be an integer, let $Y_i, W, D \subset Y$ be proper closed subsets with* $\dim Y_i \leq i$, *and suppose we have a decomposition of Δ_Y as*

$$N \cdot \Delta_Y = \gamma + \gamma_1 + \gamma_2,$$

with γ supported on $Y_i \times_k Y$, γ_1 supported on $Y \times_k D$, and γ_2 supported on $W \times_k Y$. Then there are proper closed subsets $Z_i, D' \subset Z$ with $\dim Z_i \leq i$ *and a decomposition of Δ_Z as*

$$N \cdot \Delta_Z = \gamma' + \gamma'_1 + \gamma'_2,$$

with γ' supported on $Z_i \times_k Z$, γ'_1 supported on $Z \times_k D'$, and γ'_2 supported on $q^{-1}(W) \times_k Z$.

(2) *Let $q : Z \to Y$ be a birational totally* CH_0*-trivial morphism of integral, generically smooth, proper k-schemes. Then* $\mathrm{Tor}_k^{(i)}(Z) = \mathrm{Tor}_k^{(i)}(Y)$ *for all i.*

(3) *Let $q : Z \to Y$ be a birational universally* CH_0*-trivial morphism of integral proper k-schemes. Then* $\mathrm{Tor}_k(Z) = \mathrm{Tor}_k(Y)$. *If moreover Z and Y are geometrically integral, then* $\mathrm{gTor}_k(Z) = \mathrm{gTor}_k(Y)$.

Proof. We note that (2) follows easily from (1). Indeed, (1) with $W = \varnothing$ shows that $\mathrm{Tor}_k^{(i)}(Z)$ divides $\mathrm{Tor}_k^{(i)}(Y)$ for all i; as $(q \times q)_*(\Delta_Z) = \Delta_Y$, it follows that a decomposition of Δ_Z of order N and level i gives a similar decomposition of Δ_Y by applying $(q \times q)_*$.

We now prove (1). We may assume that $W = \varnothing$. Indeed, if we replace Y with $Y' := Y \setminus W$ and Z with $Z' := Z \setminus q^{-1}(W)$, the result for $q|_{Z'} : Z' \to Y'$ and the decomposition

$$N \cdot \Delta_{Y'} = \gamma|_{Y' \times_k Y'} + \gamma_{1|Y' \times_k Y'},$$

together with localization gives (1) for the original data.

Suppose then we have

$$N \cdot \Delta_Y = \gamma + \gamma_1$$

with γ supported on $Y_i \times_k Y$ and γ_1 supported on $Y \times_k D$. Let $K = k(Y)$, and let $\eta_Y \in Y$ be the generic point. We have a rational equivalence of 0-cycles on $Y \times_k \eta_Y$

$$N \cdot \eta_Y \times \eta_Y \sim \gamma_{\eta_Y}$$

with γ_{η_Y} a 0-cycle supported on $Y_i \times_k \eta_Y$. Thus, $N \cdot \eta_Y \times \eta_Y \sim 0$ on $(Y \setminus Y_i) \times_k \eta_Y$.

Since $Z \setminus q^{-1}(Y_i) \to Y \setminus Y_i$ is birational and universally CH_0-trivial (Remarks 3.4), there is a rational equivalence of 0-cycles

$$N \cdot \eta_Z \times \eta_Z \sim 0$$

on $(Z \setminus q^{-1}(Y_i)) \times_k \eta_Z$, where $\eta_Z \in Z$ is the generic point. We claim that there is a dimension $\leq i$ closed subset Z' of Z and a rational equivalence of 0-cycles on $Z \times_k \eta_Z$

$$N \cdot \eta_Z \times \eta_Z \sim \rho_Z$$

with ρ_Z a 0-cycle supported on $Z' \times_k \eta_Z$. We proceed by a noetherian induction. We assume there is a closed subset $Y^j \subset Y_i$, a dimension $\leq i$ closed subset Z_j of $q^{-1}(Y_i)$, and a rational equivalence of 0-cycles on $(Z \setminus q^{-1}(Y^j)) \times_k \eta_Z$

$$N \cdot \eta_Z \times \eta_Z \sim \rho_j$$

with ρ_j a 0-cycle supported on $Z_j \times_k \eta_Z$, and we show the parallel statement for a proper closed subset Y^{j+1} of Y^j. The induction starts with $Y^0 = Y_i$.

Choose an integral component Y_0^j of Y^j, and let ν be its generic point. Let Y' be the union of the components of Y^j different from Y_0^j. We have the exact localization sequence

$$\begin{aligned}\mathrm{CH}_0((q^{-1}(Y_0^j \setminus Y')) \times_k \eta_Z) &\xrightarrow{i_*} \mathrm{CH}_0((Z \setminus q^{-1}(Y')) \times_k \eta_Z)\\ &\to \mathrm{CH}_0((Z \setminus q^{-1}(Y^j)) \times_k \eta_Z) \to 0,\end{aligned}$$

and thus there is a 0-cycle ρ' on $q^{-1}(Y_0^j \setminus Y') \times_k \eta_Z$ and a rational equivalence

$$N \cdot \eta_Z \times \eta_Z \sim \rho_j + i_*(\rho')$$

on $(Z \setminus q^{-1}(Y')) \times_k \eta_Z$.

Write

$$\rho' = \sum_i m_i x_i + \sum_j n_j x'_j,$$

where the x_i, x'_j are closed points of $q^{-1}(Y_0^j \setminus Y') \times_k \eta_Z$, such that $q \circ p_1(x_i) = \nu$ for all i and $q \circ p_1(x'_j)$ is contained in some proper closed subset (say Y'') of Y_0^j for all j. Replacing Y' with $Y' \cup Y''$ and changing notation, we may assume that $\rho' = \sum_i m_i x_i$.

By assumption, the map $q^{-1}(\nu) \to \nu$ is universally CH_0-trivial, so there is a degree-one 0-cycle ϵ on $q^{-1}(\nu)$ so that ϵ_L generates $\mathrm{CH}_0(q^{-1}(\nu)_L)$ for all field extensions $L \supset k(\nu)$; in particular, $\epsilon \times \eta_Z$ generates $\mathrm{CH}_0(q^{-1}(\nu) \times_k \eta_Z)$. Enlarging Y' again by a proper closed subset of Y_0^j, we may assume that

$$\rho' = m \cdot \epsilon \times \eta_Z$$

in $\mathrm{CH}_0(q^{-1}(Y_0^j \setminus Y') \times_k \eta_Z)$, for some $m \in \mathbb{Z}$. Since ϵ is a 0-cycle on $q^{-1}(\nu)$, the dimension of the closure Z' of the support of ϵ in $q^{-1}(Y_0^j)$ is bounded by the transcendence dimension of $k(\nu)$ over k, that is, by $\dim_k Y_0^j$; since $Y_0^j \subset Y_i$,

$$\dim_k Z' \leq i.$$

Taking $Y^{j+1} = Y'$, $Z_{j+1} = Z_j \cup Z'$, and $\rho_{j+1} = \rho_j + m \cdot \epsilon \times \eta_Z$, the 0-cycle ρ_{j+1} is supported on $Z_{j+1} \times_k \eta_Z$, $\dim_k Z_{j+1} \leq i$, and we have

$$N \cdot \eta_Z \times \eta_Z = \rho_{j+1}$$

in $\mathrm{CH}_0((Z \setminus q^{-1}(Y^{j+1})) \times_k \eta_Z)$. The induction thus goes through, proving the result.

The proof of (3) is similar but easier. We have already seen that if Z has a decomposition of the diagonal of order N, then so does Y. If conversely Y has a decomposition of the diagonal of order N, then there is a 0-cycle y on Y with

$$N \cdot \eta_Y \times \eta_Y = y \times \eta_Y$$

in $\mathrm{CH}_0(Y \times \eta_Y)$. As $q : Z \to Y$ is universally CH_0 trivial, there is a 0-cycle z on Z with $q_* z = y$ in $\mathrm{CH}_0(Y)$ and since $(q \times q)_* : \mathrm{CH}_0(Z \times_k \eta_Z) \to \mathrm{CH}_0(Y \times_k \eta_Y)$ is an isomorphism, we have

$$N \cdot \eta_Z \times \eta_Z = z \times \eta_Z$$

in $\mathrm{CH}_0(Z \times_k \eta_Z)$. The proof for gTor is the same. □

We note some consequences of Lemma 3.6.

Proposition 3.7. *Let $f : Y \to X$ be a dominant rational map of smooth integral proper k-schemes of the same dimension d.*

(1) *Suppose k admits resolution of singularities for rational maps of varieties of dimension $\leq d$; that is, if $p : Y \to X$ is a rational morphism of smooth k-schemes of dimension $\leq d$, there is a sequence of blow-ups of Y with smooth center, $q : W \to Y$, such that the resulting rational map $r : W \to X$ is a morphism. Then* $\mathrm{Tor}_k^{(i)} X$ *divides* $\deg f \cdot \mathrm{Tor}_k^{(i)} Y$ *for all i.*

(2) *Without assumption on k,* $\mathrm{Tor}_k X$ *divides* $\deg f \cdot \mathrm{Tor}_k Y$ *and* $\mathrm{gTor}_k X$ *divides* $(\deg f)^2 \cdot \mathrm{gTor}_k Y$.

Proof. For (1) we may find a sequence of blow-ups with smooth centers, $g : Z \to Y$, so that the induced rational map $h : Z \to X$ is a morphism. Since g is a totally CH_0-trivial morphism, $\mathrm{Tor}_k^{(i)} Z = \mathrm{Tor}_k^{(i)} Y$ by Lemma 3.6(2), so we may assume that g is a morphism; the result then follows from Lemma 1.14.

For (2), let $Z \subset Y \times_k X$ be the graph of f, that is, the closure of the graph of $f : V \to X$ for a nonempty open subset $V \subset Y$ on which f is defined. The map $p_1 : Z \to Y$ is birational and there is a nonempty open $X_0 \subset X$ such that $p_1 : p_2^{-1}(X_0) \cap Z \to Y$ is an open immersion; set $Y_0 := p_1(p_2^{-1}(X_0) \cap Z)$. The correspondence $Z \times_k Z$ yields a homomorphism

$$g : \mathrm{CH}_d(Y \times_k Y) \to \mathrm{CH}_d(X \times_k X).$$

We claim that $g(\Delta_Y) = \deg(f) \cdot \Delta_X + \gamma$ where γ is a cycle supported on $X \times_k (X \setminus X_0)$, which implies the assertion for $\mathrm{Tor}_k X$. Keeping track of supports and using localization, we have an identity in $\mathrm{CH}_d(Z \times_k Z)$ of the form

$$[Z \times_k Z] \cdot (p_1 \times p_1)^*(\Delta_Y) = \Delta_Z + \gamma', \tag{3-1}$$

where γ' has support in $(p_1^{-1}(Y \setminus Y_0) \cap Z) \times_{Y \setminus Y_0} (p_1^{-1}(Y \setminus Y_0) \cap Z)$. Therefore, $(p_2 \times p_2)_*(\gamma')$ has support in $X \times_k (X \setminus X_0)$. Applying $(p_2 \times p_2)_*$ to (3-1) we prove our claim.

The proof for gTor_k is similar. □

In particular, if we have resolution of singularities of birational maps, $\mathrm{Tor}_k^{(i)}$ is a birational invariant and in general Tor_k is a birational invariant; from this it follows easily that $\mathrm{Tor}_k^{(i)}$ is a stable birational invariant if we have resolution of singularities of birational maps and in general Tor_k is a stable birational invariant.

4. Specialization and degeneration

The next result, in a somewhat different form, is proven in [Colliot-Thélène and Pirutka 2016b, Théorème 1.12]. In a less general setting, a similar result may be found in [Voisin 2015, Theorem 1.1].

Proposition 4.1. *Let $\mathbb{O}$ be a regular local ring with quotient field K and residue field k. Let $f:\mathscr{X}\to \operatorname{Spec}\mathbb{O}$ be a flat and proper morphism with geometrically integral fibers, and let X be the generic fiber $\mathscr{X}_K$ and Y the special fiber $\mathscr{X}_k$. We suppose that Y admit a resolution of singularities $q: Z\to Y$ such that q is a universally CH_0-trivial morphism. Suppose in addition that X admits a decomposition of the diagonal of order N. Then Z also admits a decomposition of the diagonal of order N. In particular, if $\mathrm{Tor}_K(X)$ is finite, then so is $\mathrm{Tor}_k(Z)$, and in this case $\mathrm{Tor}_k(Z)\mid \mathrm{Tor}_K(X)$.*

In [Colliot-Thélène and Pirutka 2016b] it is assumed that X has a resolution of singularities $\tilde{X}\to X$ such that $\tilde{X}_K$ admits a decomposition of the diagonal of order N, which implies the same condition on X by pushing forward; there is also an assumption that Z has a 0-cycle of degree 1. This resolution of singularities in [Colliot-Thélène and Pirutka 2016b] arises because they consider decompositions of the diagonal only on smooth proper varieties; the existence of a degree-1 0-cycle comes from considering only the case $N=1$. The modified version stated above is proved exactly as in [loc. cit.].

We prove an extension of this specialization result which takes the decompositions of higher level into account.

Proposition 4.2. *Let $\mathbb{O}$ be a regular local ring with quotient field K and residue field k. Let $f:\mathscr{X}\to \operatorname{Spec}\mathbb{O}$ be a flat and proper morphism with geometrically integral fibers, and let X be the generic fiber $\mathscr{X}_K$ and Y the special fiber $\mathscr{X}_k$. Suppose that there is a birational totally CH_0-trivial morphism $q: Z\to Y$ of geometrically integral proper k-schemes.*

(1) *Suppose X admits a decomposition of the diagonal of order N and level i. Then Z also admits a decomposition of the diagonal of order N and level i. If $\mathrm{Tor}_K^{(i)}(X)$ is finite, then so is $\mathrm{Tor}_k^{(i)}(Z)$ and in this case $\mathrm{Tor}_k^{(i)}(Z)\mid \mathrm{Tor}_K^{(i)}(X)$.*

(2) *Let $\bar{K}$ and $\bar{k}$ be the respective algebraic closures of K and k, and suppose that $X_{\bar{K}}$ admits a decomposition of the diagonal of order N and level i. Suppose that K has characteristic zero, or that $\mathbb{O}$ is excellent. Then $Z_{\bar{k}}$ also admits a decomposition of the diagonal of order N and level i. If $\mathrm{Tor}^{(i)}(X)$ is finite, then so is $\mathrm{Tor}^{(i)}(Z)$ and in this case $\mathrm{Tor}^{(i)}(Z)\mid \mathrm{Tor}^{(i)}(X)$.*

Proof. The assertion (2) follows from (1) by first stratifying $\operatorname{Spec}\mathcal{O}$ as in the proof of Lemma 1.5 to reduce to the case of a DVR. We then take a finite extension L of K so that $\mathrm{Tor}^{(i)}(X) = \mathrm{Tor}^{(i)}_L(X_L)$, take the normalization $\mathcal{O} \to \mathcal{O}^N$ of $\mathcal{O}$ in L, and replace $\mathcal{O}$ with the localization $\mathcal{O}'$ of $\mathcal{O}^N$ at some maximal ideal. Letting k' be the residue field of $\mathcal{O}'$, $\mathrm{Tor}^{(i)}(Z)$ divides $\mathrm{Tor}^{(i)}_{k'}(Z_{k'})$, so (1) implies (2). We now prove (1).

By Lemma 1.5, Y admits a decomposition of the diagonal of order N and level i. By Lemma 3.6, Z also admits a decomposition of the diagonal of order N and level i, proving (1). □

We also have a version that incorporates Totaro's extended specialization Lemma 1.7.

Proposition 4.3. *Let $\mathcal{O}$ be a discrete valuation ring with quotient field K and residue field k. Let $f:\mathcal{X}\to \operatorname{Spec}\mathcal{O}$ be a flat and proper morphism of dimension d over* $\operatorname{Spec}\mathcal{O}$ *with generic fiber X and special fiber Y. We suppose Y is a union of closed subschemes, $Y = Y_1 \cup Y_2$, and that X and Y_1 are geometrically integral. Suppose there is a birational totally* CH_0*-trivial morphism $q : Z \to Y_1$ of geometrically integral proper k-schemes and that X admits a decomposition of the diagonal of order N and level i. Then there are proper closed subsets Z_i, $D \subset Z$ with* $\dim Z_i \le i$ *and a decomposition*

$$N \cdot \Delta_Z = \gamma + \gamma_1 + \gamma_2$$

with γ supported in $Z_i \times_k Z$, γ_1 supported in $Z \times_k D$, and γ_2 supported in $q^{-1}(Y_1 \cap Y_2) \times_k Z$.

Proof. This follows directly from Lemmas 1.7 and 3.6. □

Remark 4.4. As in the second part of Proposition 4.2, we may take the N in Proposition 4.3 to be $\mathrm{Tor}^{(i)}_{\bar{K}}(X_{\bar{K}})$ if $\mathcal{O}$ is excellent or if K has characteristic zero, by replacing $\mathcal{O}$ with its normalization $\mathcal{O}'$ in a finite extension L of K so that $\mathrm{Tor}^{(i)}_{\bar{K}}(X_{\bar{K}}) = \mathrm{Tor}^{(i)}_L(X_L)$, replacing $\mathcal{X}$ with $\mathcal{X} \times_{\mathcal{O}} \mathcal{O}'$, replacing k with a residue field k' of $\mathcal{O}'$, and replacing Z with $Z \otimes_k k'$.

5. Torsion order for complete intersections in a projective space: an upper bound

We concentrate on the 0-th torsion order of a (reduced, generically smooth) complete intersection $X = X^n_{d_1,\dots,d_r}$ in $\mathbb{P}^{n+r}$ of dimension n and multidegree $d_1, d_2, \dots, d_r$. In this section, we recall the construction of Roitman [1980], which when suitably refined gives an upper bound for $\mathrm{Tor}_k(X)$; by Lemma 1.3(1), this gives an upper bound for $\mathrm{Tor}^{(i)}_k(X)$ for all i.

Remark 5.1. Roitman [1980] considered 0-cycles modulo rational equivalence on a smooth hypersurface X of degree $d \le n$ in $\mathbb{P}^n_k$ for k an algebraically closed field.

His argument (in part) consisted in showing that through each point x of X, there is a line ℓ in $\mathbb{P}^n$ containing x with either $\ell \subset X$ or $\ell \cap X = \{x\}$ (set-theoretically). To do this, he showed how the defining equation for X gives equations for the set of all ℓ with the above properties, as a closed subset of the projective space $\mathbb{P}^{n-1}(x)$ of lines through x. For his purpose, it is enough to show that the closed subset of all ℓ containing x, with $\ell \subset X$ or with $\ell \cap X = \{x\}$, is nonempty; for our purposes, we need the degree of this closed subscheme. More concretely, if x, x' are points of $X(k)$ with k algebraically closed, Roitman's argument shows that $d \cdot x \sim d \cdot x'$ by finding lines ℓ, ℓ' as described above, whereas we need to consider points x, x' in $X(k(X))$, so the factor d becomes multiplied by the degree of the closed subscheme of lines through x or x'. Finally, Roitman eventually shows that $x \sim x'$ for all $x \in X(k)$, k algebraically closed, by applying his famous theorem on the torsion in the group of 0-cycles modulo rational equivalence.

We often shorten the notation by writing d_* for a sequence $d_1, d_2, \dots, d_r$.

Proposition 5.2. *Let k be a field, and let $X = X^n_{d_1,\dots,d_r}$ in $\mathbb{P}^{n+r}_k$ with $\sum_i d_i \le n + r$ be a reduced, generically smooth complete intersection of multidegree $d_1, \dots, d_r$, with $n \ge 1$. Then $\mathrm{Tor}_k(X)$ is finite and divides $\prod_{i=1}^r d_i!$.*

Proof. The reduced, generically smooth complete intersections in $\mathbb{P}^{n+r}$ and of multidegree $d_1, \dots, d_r$ are parametrized by an open subscheme $\mathcal{U}_{d_*;n}$ of a product of projective spaces; by Lemma 1.5 it suffices to prove the result for the subscheme $X := X_{d_*,\mathrm{gen}}$ of $\mathbb{P}^{n+r}_K$ defined over the field $K := k(\mathcal{U}_{d_*;n})$ corresponding to the generic point of $\mathcal{U}_{d_*;n}$. For such an X, there is an open subset $V \subset X$, such that, for $x \in V$, the set of lines $\ell \subset \mathbb{P}^{n+r}$ such that $x \in \ell$ and $(\ell \cap X)_{\mathrm{red}}$ is either $\{x\}$ or is ℓ is defined by a complete intersection W_x of multidegree

$$d_1 - 1, d_1 - 2, \dots, 2, 1, d_2 - 1, d_2 - 2, \dots 2, 1, \dots, d_r - 1, \dots, 2, 1$$

in the projective space $\mathbb{P}^{n+r-1}_{K(x)}$ of lines through x. Indeed, we may choose a standard affine open U in $\mathbb{P}^{n+r}_{K(x)}$ containing x and choose affine coordinates $t_0, \dots, t_{n+r-1}$ for U so that x is the origin, and $X \cap U$ is defined by inhomogeneous equations $F_1 = \dots = F_r = 0$. Writing each F_i as a sum of homogeneous terms $F_i^{(j)}$ of degree j,

$$F_i = \sum_{j=1}^{d_i} F_i^{(j)},$$

W_x is defined by ideal $(\cdots F_i^{(j)} \cdots)$, $i = 1, \dots, r$ and $j = 1, \dots, d_i - 1$. Since we are choosing X to be the generic hypersurface, and as we may also choose x to lie outside any proper closed subset of X, the homogeneous terms $F_i^{(j)} \in K(x)[t_0, \dots, t_{n+r-1}]_j$ will define a complete intersection in $\mathbb{P}^{n+r-1}_{K(x)}$. In particular W_x has codimension $\sum_{i=1}^r (d_i - 1) \le n + r - 1$ in $\mathbb{P}^{n+r-1}_{K(x)}$, is nonempty, and has degree $\prod_{i=1}^r (d_i - 1)!$.

Let $W_x^0 \subset W_x$ be the closed subset of lines ℓ containing x with $\ell \subset X$; this is defined by the r additional equations $F_i^{(d_i)} = 0$. Thus, for general (X, x), W_x^0 has codimension r on W_x (or is empty).

Since $n + r - 1 - \sum_{i=1}^r (d_i - 1) \geq r - 1$, we may intersect W_x with a suitably general linear space $L \subset \mathbb{P}^{n+r-1}_{K(x)}$ to form a closed subscheme $\overline{W}_x \subset W_x$ of dimension $r-1$ and degree $\prod_{i=1}^r (d_i - 1)!$ and we may choose L with $L \cap W_x^0 = \emptyset$. The cone over $\overline{W}_x$ with vertex x, $\mathscr{C}_x \subset \mathbb{P}^{n+r}_{K(x)}$, is thus a dimension-$r$ closed subscheme of degree $\prod_{i=1}^r (d_i - 1)!$ with intersection (set) $\mathscr{C}_x \cap X = \{x\}$. Thus, as cycles

$$\mathscr{C}_x \cdot X = \left(\prod_{i=1}^r d_i!\right) \cdot x.$$

Let η be the generic point of X. Taking $x = \eta$ in the above discussion gives

$$\prod_{i=1}^r d_i! \cdot \eta = \mathscr{C}_\eta \cdot X.$$

But $\mathscr{C}_\eta$ is an r-cycle on $\mathbb{P}^{n+r}_{K(\eta)}$ of degree $\prod_{i=1}^r (d_i - 1)!$, so $\mathscr{C}_\eta = \prod_{i=1}^r (d_i - 1)! \cdot L_r$ in $\mathrm{CH}_r(\mathbb{P}^{n+r}_{K(\eta)})$, where $L_r \subset \mathbb{P}^{n+r}_K$ is any dimension-r linear subspace. Since K is infinite, we may choose L_r so that the intersection $L_r \cap X$ has dimension zero. Thus, letting $z = \prod_{i=1}^r (d_i - 1)! \cdot (L_r \cdot X)$, we have

$$\prod_{i=1}^r d_i! \cdot \eta - z_{K(\eta)} = 0$$

in $\mathrm{CH}_0(X_{K(\eta)})$, which gives a decomposition of the diagonal in X of order $\prod_{i=1}^r d_i!$. Thus, $\mathrm{Tor}_K(X)$ is finite and divides $\prod_{i=1}^r d_i!$, as desired. $\square$

Corollary 5.3. *Let $X = X^n_{d_1,\ldots,d_r}$ in $\mathbb{P}^{n+r}_k$ be a smooth complete intersection of multidegree $d_1, \ldots, d_r$ and of dimension $n \geq 1$ with $\sum_i d_i \leq n + r$. Then* $\mathrm{gTor}_k(X)$ *and* $\mathrm{Tor}(X)$ *are both finite and both divide $\prod_{i=1}^r d_i!$.*

Proof. Both $\mathrm{gTor}_k(X)$ and $\mathrm{Tor}(X) := \mathrm{Tor}_{\bar{k}}(X_{\bar{k}})$ divide $\mathrm{Tor}_k(X)$ (Lemma 1.3), so the result follows from Proposition 5.2. $\square$

6. A lower bound in the generic case

In this section we discuss the case of the generic complete intersection. Let k denote a fixed base-field, for instance the prime field. The bounds we find for the generic case are independent of k, so one could equally well take k to be the reader's favorite field, even an algebraically closed one.

Before going into details, we outline the case of hypersurfaces, which uses all the main ideas.

Let $d!^*$ denote the l.c.m. of the integers $2, \dots, d$. Note that $d!^*$ is inductively the l.c.m. of d and $(d-1)!^*$ (Lemma 6.4). Our main result in the case of hypersurfaces is that the torsion order of level 0 of the generic hypersurface of degree $d \le n+1$ in $\mathbb{P}^{n+1}$ is divisible by $d!^*$, in other words, if the generic hypersurface admits a decomposition of the diagonal of degree N, then $d!^*$ divides N.

The hypersurfaces of degree $d \le n+1$ in $\mathbb{P}_k^{n+1}$ are parametrized by a projective space $\mathbb{P}^{N_{n,d}}$, and it is not hard to show that the index over $k(\mathbb{P}^{N_{n,d}})$ of the generic degree-d hypersurface X is d. In fact, we have a much stronger statement, namely $\mathrm{CH}_0(X) = \mathbb{Z}$, generated by $X \cdot \ell$ for $\ell \subset \mathbb{P}^{n+1}$ a line (Lemma 6.1(1)). In particular, for any zero cycle x on X, we have $d \mid \deg_{k(\mathbb{P}^{N_{n,d}})} x$.

If we have a decomposition of order N of the diagonal on X,

$$N \cdot \Delta_X \sim x \times X + \gamma,$$

then, as projecting this identity on the second factor shows that $N = \deg_{k(\mathbb{P}^{N_{n,d}})} x$, it follows that $d \mid N$. Now degenerate X to the generic degree-$(d-1)$ hypersurface Y in $\mathbb{P}^{n+1}$ plus the hyperplane H given by $x_{n+1} = 0$, and let $Z = Y \cap H$. Here Y and Z are defined over $L := k(\mathbb{P}^{N_{n,d-1}})$. Specializing the above rational equivalence using Lemma 1.7 gives a rational equivalence on $Y \times_L Y$ of the form

$$N \cdot \Delta_Y \sim \bar{x} \times Y + \gamma_1 + \gamma_2$$

with $\bar{x}$ a zero-cycle on Y, γ_1 a dimension-n cycle on $Z \times_L Y$, and γ_2 supported in $Y \times_L D$ for some divisor D on Y. Passing to the generic point of Y, γ_1 gives a 0-cycle on $Z \times_L L(Y)$. The main point is to show that $\mathrm{CH}_0(Z \times_L L(Y))$ is also $\mathbb{Z}$, generated by intersections from $\mathbb{P}^{n-1}$ (Lemma 6.1(3)), so we can replace γ_1 with $y \times Y + \gamma_3$, where y is a 0-cycle on Z and γ_3 is supported on $Z \times_L D'$ for some divisor D' on Y (Lemma 6.2). In other words,

$$N \cdot \Delta_Y \sim (\bar{x} + y) \times Y + \gamma_2 + \gamma_3,$$

so Y admits a decomposition of the diagonal of degree N. Now use induction on d to conclude that $(d-1)!^* \mid N$. As we already know that $d \mid N$, we find $d!^* \mid N$.

Now we address the details and the case of a general complete intersection. Fix integers $n, r \ge 1$. For an integer d, let $\mathcal{S}_{d,n+r}$ be the set of indices $I = (i_0, \dots, i_{n+r})$ with $0 \le i_j$ and $\sum_j i_j = d$. We let $\mathcal{S}_i = \mathcal{S}_{d_i,n+r}$ and let $N_i := \#\mathcal{S}_i - 1$. Let $\{u_i^{(I)} \mid I \in \mathcal{S}_i\}$ be homogeneous coordinates for $\mathbb{P}^{N_i}$, and let $x_0, \dots, x_{n+r}$ be homogeneous coordinates for $\mathbb{P}^{n+r}$. The universal family of intersections of multidegree $d_1, \dots, d_r$ in $\mathbb{P}^{n+r}$, $\mathcal{X}^{d_*,n}$, is the subscheme of $\mathbb{P}^{N_1} \times \cdots \times \mathbb{P}^{N_r} \times \mathbb{P}^{n+r}$ defined by the multihomogeneous ideal in the polynomial ring $k[\{u_i^{(I)}\}_{I \in \mathcal{S}_i,\, i=1,\dots,r}, x_0, \dots, x_{n+r}]$ generated by the elements

$$\sum_{I\in\mathcal{S}_i} u_i^{(I)} x^I, \quad i = 1, \ldots, r,$$

where as usual $x^I = x_0^{i_0} \cdots x_{n+r}^{i_{n+r}}$ for $I = (i_0, \ldots, i_{n+r})$. We let $\eta := \eta_{d_*;n}$ denote the generic point of $\mathbb{P}^{N_1} \times \ldots \times \mathbb{P}^{N_r}$ and let $\mathscr{X}_\eta^{d_*,n}$ denote the fiber product

$$\mathscr{X}_\eta^{d_*,n} := \mathscr{X}^{d_*,n} \times_{\mathbb{P}^{N_1}\times\cdots\times\mathbb{P}^{N_r}} \eta \subset \mathbb{P}_\eta^{n+r}.$$

By Proposition 5.2, we know that if $\sum_{i=1}^r d_i \le n+r$, then $\mathrm{Tor}_{k(\eta)}(\mathscr{X}_\eta^{d_*,n})$ is finite and divides $\prod_i d_i!$. We turn to a computation of a lower bound.

Let $H \subset \mathbb{P}^{N_1} \times \cdots \times \mathbb{P}^{N_r} \times \mathbb{P}^{n+r}$ be the subscheme defined by $(x_{n+r} = 0)$, let $\mathscr{X}_H^{d_*,n} := \mathscr{X}^{d_*,n} \cap H$, and let $\mathscr{X}_{H,\eta}^{d_*,n} := \mathscr{X}_\eta^{d_*,n} \cap H$. Let $\eta' := \eta_{d_*,n-1}$.

We separate the indices $\mathcal{S}_i$ into two disjoint subsets $\mathcal{S}_i^0$ and $\mathcal{S}_i^1$, with $\mathcal{S}_i^0$ the set of $(i_0, \ldots, i_{n+r})$ with $i_{n+r} = 0$ and $\mathcal{S}_i^1$ those with $i_{r+n} > 0$. We set $v_i^{(I)} = u_i^{(I)}$ for $I \in \mathcal{S}_i^0$ and $w_i^{(I)} = u_i^{(I)}$ for $I \in \mathcal{S}_i^1$. We write $k(\{u_i^{(I)}\}_0)$ for the field extension of k generated by the ratios $u_i^{(I)}/u_i^{(I')}$, $I \neq I'$, and similarly for $k(\{v_i^{(I)}\}_0)$, giving us the field extension $k(\{v_i^{(I)}\}_0) \subset k(\{u_i^{(I)}\}_0)$. We note that $k(\{u_i^{(I)}\}_0) = k(\eta)$, $k(\{v_i^{(I)}\}_0) = k(\eta')$, and the $k(\eta)$-scheme $\mathscr{X}_{H,\eta}^{d_*,n}$ is canonically isomorphic to the base-change of the $k(\eta')$-scheme $\mathscr{X}_{\eta'}^{d_*,n-1}$ via the base-extension $k(\eta') \subset k(\eta)$:

$$\mathscr{X}_{H,\eta}^{d_*,n} \cong \mathscr{X}_{\eta'}^{d_*,n-1} \otimes_{k(\eta')} k(\eta).$$

This defines for us the projection $q_1 : \mathscr{X}_{H,\eta}^{d_*,n} \to \mathscr{X}_{\eta'}^{d_*,n-1}$.

Let $K = k(\eta)(\mathscr{X}_\eta^{d_*;n}) = k(\mathscr{X}^{d_*,n})$. We have the morphism of $k(\eta')$-schemes

$$\pi : \mathscr{X}_{H,\eta}^{d_*,n} \otimes_{k(\eta_{d_*;n})} K \to \mathscr{X}_{\eta'}^{d_*,n-1}$$

formed by the composition

$$\mathscr{X}_{H,\eta}^{d_*,n} \otimes_{k(\eta)} K \xrightarrow{p_1} \mathscr{X}_{H,\eta}^{d_*,n} \xrightarrow{q_1} \mathscr{X}_{\eta'}^{d_*,n-1}.$$

Lemma 6.1. (1) *For $i = 0, \ldots, n$, the intersection map*

$$\mathrm{CH}_{r+i}(\mathbb{P}_{k(\eta)}^{n+r}) \to \mathrm{CH}_i(\mathscr{X}_\eta^{d_*;n})$$

is an isomorphism.

(2) *For $i = 0, \ldots, n-1$, the pullback*

$$\pi^* : \mathrm{CH}_i(\mathscr{X}_{\eta'}^{d_*,n-1}) \to \mathrm{CH}_i(\mathscr{X}_{H,\eta}^{d_*,n} \otimes_{k(\eta)} K)$$

is an isomorphism.

(3) *For $i = 0, \ldots, n-1$, the intersection map*

$$\mathrm{CH}_{r+i+1}(\mathbb{P}_K^{n+r}) \to \mathrm{CH}_i(\mathscr{X}_{H,\eta}^{d_*,n} \otimes_{k(\eta)} K)$$

is an isomorphism.

Proof. Noting that the base-extension $\mathrm{CH}_*(\mathbb{P}^{n+r}_{k(\eta)}) \to \mathrm{CH}_*(\mathbb{P}^{n+r}_K)$ is an isomorphism, the assertion (3) follows from (1) (for $n-1$) and (2). For (1), the projection

$$p_2 : \mathscr{X}^{d_*,n} \to \mathbb{P}^{n+r}$$

expresses $\mathscr{X}^{d_*,n}$ as a $\mathbb{P}^{N_1-1} \times \cdots \times \mathbb{P}^{N_r-1}$-bundle over $\mathbb{P}^{n+r}$, with fibers embedded in $\mathbb{P}^{N_1} \times \cdots \times \mathbb{P}^{N_r}$ linearly in each factor. Thus, $\mathrm{CH}_*(\mathscr{X}^{d_*;n})$ is generated by $\mathrm{CH}_*(\mathbb{P}^{N_1} \times \cdots \times \mathbb{P}^{N_r} \times \mathbb{P}^{n+r})$ via restriction. After localization at η, this shows that $\mathrm{CH}_*(\mathscr{X}^{d_*;n}_\eta)$ is generated by $\mathrm{CH}_*(\mathbb{P}^{n+r}_{k(\eta)})$ via restriction. The fact that the surjective map $\mathrm{CH}_{r+i}(\mathbb{P}^{n+r}_{k(\eta)}) \to \mathrm{CH}_i(\mathscr{X}^{d_*;n}_\eta)$ is also injective in the stated range follows by noting that the intersection pairing on $\mathscr{X}^{d_*;n}_\eta$ is nondegenerate when restricted to these cycles. This proves (1).

For (2), fix for each i the index $I_i^0 := (d_i, 0, \ldots, 0)$, and the index $I_i^1 := (0, \ldots, 0, d_i)$, and for each homogeneous variable $w_i^{(I)}$, let $w_i^{(I)0}$ be the corresponding affine coordinate $w_i^{(I)}/v_i^{(I_i^0)}$. Similarly, we let $v_i^{(I)0} = v_i^{(I)}/v_i^{(I_i^0)}$. Let $y_i = x_i/x_{n+r}$, $i = 0, \ldots, n+r-1$ and $y_{n+r} = 1$. The field extension $k(\eta') \to K$ is isomorphic to the field extension given by including the constants $k(\{v_i^{(I)}\}_0)$ of the $k(\{v_i^{(I)}\}_0)$-algebra A,

$$A := k(\{v_i^{(I)}\}_0, y_0, \ldots, y_{n+r})[\{w_i^{(I)0}\}]/\Bigl(\ldots, \sum_{I \in \mathscr{S}_i^0} v_i^{(I)0} \cdot y^I + \sum_{I' \in \mathscr{S}_i^1} w_i^{(I')0} \cdot y^{I'}, \ldots\Bigr)$$

into the quotient field L of A. In each defining relation for A, we can solve for $w_i^{(I_i^1)0}$ in terms of the y_i and the other $w_i^{(I')0}$. After eliminating each $w_i^{(I_i^1)0}$ in this way, we see that A is a polynomial algebra over $k(\{v^{(I)}{}_i\}_0, y_0, \ldots, y_{n+r-1})$. The y_i and the $w_i^{(I')0}$, after removing $w_i^{(I_i^1)0}$ for each i, therefore form an algebraically independent set of generators for L over $k(\{v_i^{(I)}\}_0)$, and thus K is a pure transcendental extension of $k(\eta')$. As Chow groups are invariant under base-change by purely transcendental field extensions, this proves (2). $\square$

Lemma 6.2. *Take γ in $\mathrm{CH}_n(\mathscr{X}^{d_*,n}_{H,\eta} \times_{k(\eta)} \mathscr{X}^{d_*,n}_\eta)$. Then there is a zero cycle y on $\mathscr{X}^{d_*,n}_{H,\eta}$, a proper closed subset D' of $\mathscr{X}^{d_*,n}_\eta$, and a cycle γ' supported on $\mathscr{X}^{d_*,n}_{H,\eta} \times_{k(\eta)} D'$ such that*

$$\gamma = y \times \mathscr{X}^{d_*,n}_\eta + \gamma'$$

in $\mathrm{CH}_n(\mathscr{X}^{d_*,n}_{H,\eta} \times_{k(\eta)} \mathscr{X}^{d_*,n}_\eta)$. Furthermore, the degree of y is divisible by $\prod_{i=1}^r d_i$.

Proof. Let ξ denote the generic point of $\mathscr{X}^{d_*,n}_\eta$. By Lemma 6.1(3), the class of the restriction $j^*\gamma$ of γ to $\mathscr{X}^{d_*,n}_{H,\eta} \times_{k(\eta)} \xi$ is of the form

$$j^*\gamma = M \cdot L \cdot \mathscr{X}^{d_*,n}_{H,\eta} \times_{k(\eta)} \xi,$$

where L is a linear subspace of $H \subset \mathbb{P}^{n+r}$, M an integer. Letting $y \in \mathrm{CH}_0(\mathscr{X}^{d_*,n}_{H,\eta})$ be the 0-cycle $M \cdot L \cdot \mathscr{X}^{d_*,n}_{H,\eta}$, the result follows from the localization theorem for the Chow groups; the assertion on the degree follows from the fact that $\mathscr{X}_{H,\eta}^{d_*,n}$ has degree $\prod_{i=1}^r d_i$ and hence y has degree $M \cdot \prod_{i=1}^r d_i$. $\square$

Definition 6.3. For a natural number $n \geq 1$, we let $n!^*$ denote the l.c.m. of the numbers $1, 2, \ldots, n$.

Lemma 6.4. *Let $d_1, \ldots, d_r$ be a sequence of positive natural numbers. Then the product $\prod_{i=1}^r (d_i!^*)$ is equal to the l.c.m. M of all products $i_1 \cdots i_r$ with $1 \leq i_j \leq d_j$, $j = 1, \ldots, r$.*

Proof. Fix a prime number p. For each $j = 1, \ldots, r$, let i_j^* be an integer with $1 \leq i_j^* \leq d_j$ and with p-adic valuation $\nu_p(i_j^*)$ equal to $\nu_p(d_j!^*)$. Then

$$\nu_p\left(\prod_{j=1}^r i_j^*\right) = \nu_p\left(\prod_{i=1}^r (d_i!^*)\right)$$

and $\nu_p\left(\prod_{j=1}^r i_j\right) \leq \nu_p\left(\prod_{j=1}^r i_j^*\right)$ for all sequences $i_1, \ldots, i_r$ with $1 \leq i_j \leq d_j$. Thus, $\nu_p(M) = \nu_p\left(\prod_{i=1}^r i_j^*\right) = \nu_p\left(\prod_{i=1}^r (d_i!^*)\right)$. Since p was arbitrary, this gives $M = \prod_{i=1}^r (d_i!^*)$. $\square$

Theorem 6.5. *For integers $d_1, \ldots, d_r$ with $\sum_i d_i \leq n+r$, $\prod_{i=1}^r d_i!^* \mid \mathrm{Tor}_{k(\eta)}(\mathscr{X}_\eta^{d_*,n})$.*

Proof. We may suppose that $d_1 > 1$. Let $d_*' = (d_1 - 1, d_2, \ldots, d_r)$. Let $\mathbb{O}$ be the local ring of the origin in $\mathbb{A}^1_{k(\eta)} = \operatorname{Spec} k(\eta)[t]$, and let $\widetilde{\mathscr{X}}$ be the subscheme of $\mathbb{P}^{N_1} \times \cdots \times \mathbb{P}^{N_r}_{\mathbb{O}}$ defined by the homogeneous ideal $(f_1, \ldots, f_r)$, with

$$f_j = \begin{cases} \sum_{I \in \mathscr{S}_{d_j,n+r}} u_j^{(I)} x^I & \text{for } j \neq 1, \\ t \cdot \sum_{I \in \mathscr{S}_{d_1,n+r}} u_1^{(I)} x^I + (1-t) \cdot x_{n+r} \cdot \sum_{J \in \mathscr{S}_{d_1-1,n+r}} u_1^{(J)} x^J & \text{for } j = 1. \end{cases}$$

The generic fiber of $\mathscr{X}$ is thus isomorphic to $\mathscr{X}_\eta^{d_*,n} \times_{k(\eta)} k(\eta, t)$, and the special fiber is $\mathscr{X}_\eta^{d_*',n} \cup H$.

Suppose that $\mathscr{X}_\eta^{d_*,n}$ admits a decomposition of the diagonal of order N:

$$N \cdot \Delta_{\mathscr{X}_\eta^{d_*,n}} = x \times \mathscr{X}_\eta^{d_*,n} + \gamma$$

with γ supported on $\mathscr{X}_\eta^{d_*,n} \times D$ for some divisor D. By Lemma 6.1, $\deg x$ is divisible by $\prod_{i=1}^r d_i$, and thus $\prod_{i=1}^r d_i$ divides N.

By applying Totaro's specialization lemma (Lemma 1.7) to the family $\mathscr{X} \to \operatorname{Spec} \mathbb{O}$, the diagonal for $\mathscr{X}_\eta^{d_*',n}$ admits a decomposition of the form

$$N \cdot \Delta_{\mathscr{X}_\eta^{d_*',n}} = \bar{x} \times \mathscr{X}_\eta^{d_*',n} + \gamma_1 + \gamma_2$$

with γ_1 supported in $\mathscr{X}_{H,\eta}^{d_*',n} \times \mathscr{X}_\eta^{d_*',n}$ and γ_2 supported in $\mathscr{X}_\eta^{d_*',n} \times D_2$ for some divisor D_2 on $\mathscr{X}_\eta^{d_*',n}$. By Lemma 6.2, we have the identity

$$\gamma_1 = y \times \mathscr{X}_\eta^{d_*',n} + \gamma_3$$

with y a zero-cycle on $\mathscr{X}_\eta^{d'_*,n}$ and γ_3 supported on $\mathscr{X}_\eta^{d'_*,n} \times D_3$ for some divisor D_3. Thus, the diagonal on $\mathscr{X}_\eta^{d'_*,n}$ admits a decomposition of order N as well. By induction $(d_1-1)!^* \cdot \prod_{i=2}^r (d_i!^*)$ divides N; by symmetry $(d_j-1)!^* \cdot \prod_{i=1, i\neq j}^r (d_i!^*)$ divides N for all j with $d_j > 1$. As we have already seen that $\prod_i d_i$ divides N, Lemma 6.4 completes the proof. □

We also have a lower bound for the generic complete intersection with a rational point.

Corollary 6.6. *For integers $d_1, \dots, d_r$ with $\sum_i d_i \le n+r$, let K be the function field of the generic complete intersection of multidegree $d_1, \dots, d_r$, $K := k(\eta)(\mathscr{X}_\eta^{d_*,n})$. Then $\big(1/\prod_{i=1}^r d_i\big) \prod_{i=1}^r (d_i!^*)$ divides $\mathrm{Tor}_K(\mathscr{X}_\eta^{d_*,n} \times_{k(\eta)} K)$.*

Proof. Let $X = \mathscr{X}_\eta^{d_*,n}$. By Lemma 6.1, $I_X = \prod_{i=1}^r d_i$ and thus by Lemma 1.10, $\mathrm{Tor}_{k(\eta)}(\mathscr{X}_\eta^{d_*,n})$ divides $I_X \cdot \mathrm{Tor}_K(\mathscr{X}_\eta^{d_*,n} \times_{k(\eta)} K)$. Clearly $\prod_{i=1}^r d_i$ divides $\prod_{i=1}^r (d_i!^*)$, whence the result. □

Example 6.7 (generic cubic hypersurfaces). For the generic cubic hypersurface $X := \mathscr{X}_\eta^{3,n}$, $n \ge 2$, we thus have $\mathrm{Tor}_{k(\eta)} X = 6$ and the generic cubic hypersurface with a rational point X_K, $K = k(\eta)(X)$, has $2 \mid \mathrm{Tor}_K X_K \mid 6$.

It follows from [Colliot-Thélène 2016, Théorème 4.1] that the generic cubic hypersurface with a rational point does have $\mathrm{Tor}_K X_K = 6$, at least if k has characteristic not equal to 3. Indeed, in view of Lemma 1.3(4), we may enlarge k. First we may suppose that k contains a primitive third root of unity. Then we pass from k to $k(\lambda_0, \dots, \lambda_{n-2})$. The smooth cubic $Y \subset \mathbb{P}^{n+1}_{k(\lambda_0,\dots,\lambda_{n-2})}$ given by $x_0^3 + x_1^3 + x_2^3 + \sum_{i=0}^{n-2} \lambda_i x_{i+3}^3 = 0$ has a rational point but also has nontrivial higher unramified cohomology with $\mathbb{Z}/3$-coefficients. We apply Lemma 1.5(1) to conclude that $3 \mid \mathrm{Tor}_K X_K$.

In particular, the generic dimension-n cubic hypersurface with a rational point does not admit a rational map $\mathbb{P}^n \dashrightarrow X_K$ of degree not divisible by 6 by Proposition 3.7(2).

Example 6.8 (generic cubic hypersurfaces with a line). Take $n \ge 2$. For X a cubic hypersurface in $\mathbb{P}^{n+1}_L$ (defined over some field $L \supset k$), we have the Fano variety of lines on X, F_X, a closed subscheme of the Grassmann variety $\mathrm{Gr}(2, n+2)_L$. In fact, if $U \to \mathrm{Gr}(2, n+2)$ is the universal rank-two bundle, and f is the defining equation for X, then F_X is the closed subscheme defined by the vanishing of the section of the rank-four bundle $\mathrm{Sym}^3 U$ determined by f. In particular, the class of F_X in $\mathrm{CH}^4(\mathrm{Gr}(2, n+2)_L)$ is given by the Chern class $c_4(\mathrm{Sym}^3 U)$. One computes this easily as $c_4 = 9c_2^2(U) + 18c_1(U)^2 c_2(U)$. As $c_2(U)^n$ and $c_2(U)^{n-2}c_1(U)^2$ both have degree one, we see that $F_X \cdot c_2(U)^{n-2}$ has degree 27, and thus I_{F_X} divides 27. This 27 is of course the famous 27 lines on a cubic surface, as intersecting F_X with $c_2(U)^{n-2}$ in $\mathrm{Gr}(2, n+2)$ is the same as taking the Fano variety of the intersection of X with a general $\mathbb{P}^3$ in $\mathbb{P}^{n+1}$. See for example [Fulton 1984, Example 14.7.13] for details of the Chern class computation.

Taking $X = \mathscr{X}_\eta^{3,n}$, and letting $K = k(\eta)(F_X)$, it follows from Lemma 1.10(1) that $6 = \mathrm{Tor}_{k(\eta)}(\mathscr{X}_\eta^{3,n})$ divides $27 \cdot \mathrm{Tor}_K(\mathscr{X}_\eta^{3,n} \times_{k(\eta)} K)$; since we have the degree-two rational map $\mathbb{P}^n_K \dashrightarrow \mathscr{X}_\eta^{3,n} \times_{k(\eta)} K$, we have $\mathrm{Tor}_K(\mathscr{X}_\eta^{3,n} \times_{k(\eta)} K) = 2$. In particular, the generic cubic with a line is not stably rational over its natural field of definition $k(\eta)(F_X)$.

We are indebted to J.-L. Colliot-Thélène [2016, Théorème 3.2] for the next example, which improves the bounds and simplifies the argument of an example in an earlier version of this paper.

Example 6.9 (cubics over a "small" field). Take $n \geq 2$. We consider a DVR $\mathscr{O}$ with quotient field K and residue field k (of characteristic $\neq 2$), and a degree-3 hypersurface $\mathscr{X} \subset \mathbb{P}^{n+1}_{\mathscr{O}}$. Let $X = \mathscr{X}_K$ and $Y = \mathscr{X}_k$. We suppose that X is smooth and $Y = Q \cup H$, with Q a smooth quadric and H a hyperplane. Furthermore, we assume

(1) $I_Q = 1$,

(2) Q and H intersect transversely, and

(3) $I_{Q \cap H} = 2$.

From Proposition 5.2, we know that $\mathrm{Tor}_K(X)$ is finite and divides 6. We will show that 2 divides $\mathrm{Tor}_K(X)$.

For this, suppose we have a decomposition of the diagonal of X of order N. We note that our family $\mathscr{X}$ satisfies the hypotheses of Lemma 1.8, with $Y_1 = Q$, $Y_2 = H$, and $r = 1$. By Remark 1.9, $N \cdot (\mathrm{CH}_0(Q)/i_{Q\cap H*}(\mathrm{CH}_0(Q \cap H))) = 0$; considering degrees, we see that $2 \mid N$.

To construct an explicit example, recall [Lam 1980, Chapter 11, Definition 4.1] that the *u-invariant* $u(k)$ of a field k is the maximum r such that there exists an anisotropic quadratic form over k of dimension r, or is ∞ if no maximum exists. For example, for p odd, $\mathbb{F}_p$ has u-invariant 2, and $\mathbb{Q}_p$ has u-invariant 4; more generally, for a field k of characteristic different from 2, $k((t))$ has u-invariant $2 \cdot u(k)$ [Lam 1980, Chapter 4, Examples 4.2].

The above construction gives us a cubic hypersurface X of dimension $n \geq 2$ over $K := k((x))$ with $2 \mid \mathrm{Tor}_K(X)$ and $X(K) \neq \varnothing$ if k is an infinite field of characteristic $\neq 2$ with u-invariant $\geq n+1$. Indeed, take an anisotropic quadratic form q_0 in $(n+1)$-variables $X_0, \ldots, X_n$, choose $\alpha \in k^\times$ represented by q_0, and let $q = q_0 - \alpha \cdot X_{n+1}^2$, so q is nondegenerate. Let $Q \subset \mathbb{P}^{n+1}_k$ be the quadric defined by q, and let H be the hyperplane $X_{n+1} = 0$. Take a cubic form $c_0 \in k[X_0, \ldots, X_{n+1}]$, and let $c = xc_0 + q \cdot X_{n+1} \in k[[x]][X_0, \ldots, X_{n+1}]$. Since k is infinite, we can choose c_0 so that the subscheme X of $\mathbb{P}^{n+1}_{k((x))}$ defined by c is smooth (and hence geometrically integral); it suffices to choose c_0 so that $c_0 = 0$ is smooth and intersects Q and H transversely. Clearly $I_Q = 1$, Q and H intersect transversely, and $I_{Q\cap H} = 2$, giving us the desired example.

Thus, there are cubic threefolds X over $K := \mathbb{Q}_p((x))$ with $2 \mid \mathrm{Tor}_K(X)$ and with $X(K) \neq \varnothing$. Similarly, there are examples of such cubic threefolds over $K = \mathbb{F}_p((t))((x))$ for $p \neq 2$. Over $K = \mathbb{Q}((x))$ or even over $K = \mathbb{R}((x))$ there are cubic hypersurfaces X of dimension n over K for arbitrary $n \geq 2$, with $2 \mid \mathrm{Tor}_K(X)$ and $X(K) \neq \varnothing$. As in the previous example, we may pass to an odd-degree field extension L of K to find a cubic hypersurface X_L with a line, and with $\mathrm{Tor}_L(X_L) = 2$; all these cubics are thus not stably rational over their corresponding field of definition.

Remark 6.10. As mentioned in the introduction, Colliot-Thélène and Pirutka have constructed cubic threefolds over a p-adic field [2016b, Théorème 1.21] and over $\mathbb{F}_p((x))$ [2016b, Remarque 1.23] with nonzero torsion order and having a rational point.

7. Torsion order for very general complete intersections in a projective space: a lower bound

As in the previous sections, we consider smooth complete intersection subschemes X of $\mathbb{P}^{n+r}$ of multidegree $d_1, \dots, d_r$.

By saying a property holds for a very general complete intersection in $\mathbb{P}^{n+r}_k$ of multidegree $d_1, \dots, d_r$, we mean that there is a countable union F of proper closed subsets of the parameter scheme of such complete intersections (an open subset in a product of projective spaces over k) such that the property holds for X_b if $b \notin F$.

Recall that for X a proper, generically smooth L-scheme for some field L, and $\bar{L}$ the algebraic closure of L, we have defined $\mathrm{Tor}^{(i)}(X) := \mathrm{Tor}^{(i)}_{\bar{L}}(X_{\bar{L}})$ (Definition 1.12).

Theorem 7.1. *Let k be a field of characteristic zero. Let $d_1, \dots, d_r$ and $n \geq 3$ be integers with $d' := \sum_{j=1}^{r} d_j \leq n + r$. Let p be a prime number. Suppose that*

$$d_i \geq p \cdot \left\lceil \frac{n + r + 1 - d' + d_i}{p + 1} \right\rceil \tag{7-1}$$

for some i, $1 \leq i \leq r$. Then $p \mid \mathrm{Tor}^{(n-2)}(X)$ for all very general $X = X_{d_1,\dots,d_r} \subset \mathbb{P}^{n+r}_k$.

Corollary 7.2. *Let k, $d_1, \dots, d_r$, n, and p be as in Theorem 7.1, and suppose that d_i satisfies* (7-1). *Then $p \mid \mathrm{Tor}(X)$ for all very general $X = X_{d_1,\dots,d_r} \subset \mathbb{P}^{n+r}_k$.*

Proof. $\mathrm{Tor}^{(n-2)}(X)$ divides $\mathrm{Tor}(X) := \mathrm{Tor}^{(0)}(X)$ by Lemma 1.3(1). □

Remarks 7.3. (1) We know that $\mathrm{Tor}(X)$ is finite for all $X = X_{d_1,\dots,d_r} \subset \mathbb{P}^{n+r}$ with $\sum_j d_j \leq n + r$ by Proposition 5.2 and hence $\mathrm{Tor}^{(n-2)}(X)$ is also finite.

(2) For $p = 2$ and for hypersurfaces, the corollary follows directly from the results in [Totaro 2016].

(3) We only use the hypothesis of characteristic zero to allow for a specialization to characteristic p, where p is the prime number in the statement. For k a field of positive characteristic, the analogous result holds, but only for $p = \operatorname{char} k$.

(4) There are two interesting cases of complete intersection threefolds we would like to mention: that of a multidegree-(3, 2) complete intersection in $\mathbb{P}^5$ and a multidegree-(2, 2, 2) complete intersection in $\mathbb{P}^6$ (see the recent results of Hassett and Tschinkel [2016]). In both cases we take $d_i = 2$ and get a divisibility by 2. Notice that in the (2, 3) case taking $d_i = 3$ and $p = 3$ works.

Proof of Theorem 7.1. This is another application of the argument of Kollár [1995], as used for example by Totaro [2016], Colliot-Thélène and Pirutka [2016a], or Okada [2016]. We may reorder the d_j so that $d_i = d_1$. We first assume that p divides d_1, $d_1 = q \cdot p$. Take f and g suitably general homogeneous polynomials of degree d_1 and q, respectively, and let $f_2, \dots, f_r$ be suitably general homogeneous polynomials, with f_j of degree d_j, $j = 2, \dots, r$. We take these to be in the polynomial ring $\mathbb{O}[X_0, \dots, X_{n+r}]$, where $\mathbb{O}$ is a complete (hence excellent) discrete valuation ring with maximal ideal (t), with residue field $k = \overline{\mathbb{F}}_p$, the algebraic closure of $\mathbb{F}_p$, and with quotient field K a field of characteristic zero. We let $\mathscr{X} \to \operatorname{Spec} \mathbb{O}$ be the closed subscheme of a weighted projective space $\mathbb{P} = \operatorname{Proj} \mathbb{O}[X_0, \dots, X_{n+r}, Y]$, with the X_i having weight 1 and Y having weight q, defined by the homogeneous ideal

$$(f_2, \dots, f_r, Y^p - f, g - tY).$$

The generic fiber $X := \mathscr{X}_K$ is isomorphic to the complete intersection subscheme of $\mathbb{P}^{n+r}_K$ defined by $g^p - t^p f = f_2 = \dots = f_r = 0$, and the special fiber $Y := \mathscr{X}_k$ is the cyclic p to 1 cover $Y \to W$, with $W \subset \mathbb{P}^{n+r}_k$ the complete intersection defined by $\bar{g} = \bar{f}_2 = \dots = \bar{f}_r = 0$, and $y^p = f|_W$.

For general $f, g, f_2, \dots, f_r$, X and W are smooth, and Y has only finitely many singularities, which may be resolved by an explicit iterated blow-up $q : Z \to Y$ which is totally CH_0-trivial: for details, see Proposition 8.5 if $p \geq 3$. If $p = d_1 = 2$, then we use Lemma 8.7 and Proposition 8.8 for the construction of the resolution of singularities and the proof that the resolution morphism q is totally CH_0-trivial. Kollár shows in addition that, under the assumption (7-1), one has $H^0(Z, \Omega^{n-1}_{Z/k}) \neq \{0\}$. In somewhat more detail, Kollár [1995, §15, Lemma 16] defines an invertible sheaf Q (denoted $\pi^* Q(L, s)$ in [loc. cit.]) with an injection $Q \to (\Omega^{n-1}_{Y/k})^{**}$, where ** denotes the double dual. A local computation (see [Colliot-Thélène and Pirutka 2016a], [Okada 2016], or Remark 8.18 for details) in a neighborhood of the finitely many singularities of Y shows that this injection extends to an injection $q^*Q \to \Omega^{n-1}_{Z/k}$; here is where the condition $n \geq 3$ is used. In addition, q^*Q is isomorphic to the pullback to Z of $\omega_W \otimes \mathcal{O}_W(d_1)$, where ω_W is the canonical sheaf on W. As $\omega_W = \mathcal{O}_W\big(d_1/p + \sum_{j\geq 2} d_j - n - r - 1\big)$, we have a nonzero section of $\Omega^{n-1}_{Z/k}$ if $d_1(p+1)/p \geq n + r + 1 - \sum_{i=2}^{r} d_i$, which is exactly the condition in the statement of the theorem.

By Proposition 5.2, we know that $\mathrm{Tor}(X_{\bar{K}})$ is finite and thus $\mathrm{Tor}^{(n-2)}(X_{\bar{K}})$ is finite as well. The specialization result Proposition 4.2 thus implies that $\mathrm{Tor}^{(n-2)}(Z_{\bar{k}})$ is finite and divides $\mathrm{Tor}^{(n-2)}(X_{\bar{K}})$. By [Gros 1985, Chapitre II, Proposition 4.2.33; Chatzistamatiou and Rülling 2011, Theorem 3.1.8; El Zein 1978, §3.3, Proposition 4], correspondences on $Z \times_k Z$ act on $H^0(Z, \Omega^{n-1}_{Z/k})$, and if γ is a correspondence on $Z \times_k Z$ supported in some $Z' \times_k Z$ with $\dim_k Z' \leq n-2$, then by [Chatzistamatiou and Rülling 2011, Proposition 3.2.2(2)], γ_* acts by zero on $H^0(Z, \Omega^{n-1}_{Z/k})$. Similarly, if γ is a correspondence on $Z \times_k Z$, supported in $Z \times_k D$ for some divisor $D \subset Z$, then $\gamma_*(\omega)|_{Z\setminus D} = 0$ for each $\omega \in H^0(Z, \Omega^{n-1}_{Z/k})$; as $\Omega^{n-1}_{Z/k}$ is locally free, it follows that $\gamma_*(\omega) = 0$. Thus, if Δ_Z admits a decomposition of order N and level $n-2$, this implies that $N \cdot \omega = 0$ for all $\omega \in H^0(Z, \Omega^{n-1}_{Z/k})$, and since $H^0(Z, \Omega^{n-1}_{Z/k})$ is a nonzero k-vector space, this implies that $p \mid N$. Since $\mathrm{Tor}^{(n-2)}(Z_{\bar{k}})$ divides $\mathrm{Tor}^{(n-2)}(X_{\bar{K}})$, it follows that $p \mid \mathrm{Tor}^{(n-2)}(X_{\bar{K}})$ and Proposition 2.1 finishes the proof in this case.

In the case of a general d_1, write $d_1 = q \cdot p + c$, $0 < c < p$, and consider a family $\mathcal{X} \to \operatorname{Spec} \mathcal{O}$ defined by a homogeneous ideal of the form

$$(f_2, \ldots, f_r, (Y^p - h)s + tu, g - tY),$$

with $u, h, g, s \in \mathcal{O}[X_0, \ldots, X_{n+r}]$, u of degree d_1, h of degree pq, g of degree q, and s of degree c, suitably general, and with Y as above of weight q. The generic fiber X is the complete intersection $f_1 = f_2 = \cdots = f_r = 0$, with $f_1 = (g^p - t^p h)s + t^{p+1}u$; the special fiber Y has two components Y_1, Y_2, with Y_1 the p to 1 cyclic cover of $W := (\bar{f}_2 = \cdots = \bar{f}_r = \bar{g} = 0)$, branched along $W \cap (h = 0)$. We take $q : Z \to Y_1$ to be the resolution as in the previous case. Having chosen h, g, s, we may take u sufficiently general so that X is a smooth complete intersection.

Since $\mathcal{O}$ is excellent, we are free to make a finite extension L of K, take the integral closure $\mathcal{O}_L$ of $\mathcal{O}$ in L, replace $\mathcal{O}$ with the localization $\mathcal{O}'$ at a maximal ideal of $\mathcal{O}_L$, and replace $\mathcal{X}$ with $\mathcal{X} \otimes_{\mathcal{O}} \mathcal{O}'$; changing notation, we may assume that $\mathrm{Tor}^{(n-2)}_K(X)$ is the geometric torsion order $\mathrm{Tor}^{(n-2)}(X)$. By Proposition 4.3, the smooth proper k-scheme Z admits a decomposition of the diagonal as

$$N \cdot \Delta_Z = \gamma + \gamma_1 + \gamma_2,$$

with $N = \mathrm{Tor}^{(n-2)}(X)$, γ supported in $Z_{n-2} \times_k Z$ with $\dim Z_{n-2} \leq n-2$, γ_1 supported in $q^{-1}(Y_1 \cap Y_2) \times_k Z$, and γ_2 supported in $Z \times_k D$ for some divisor D on Z.

We may take the degree-c part s as general as we like. In particular, we may assume that $Y_1 \cap Y_2$ is contained in the smooth locus of Y_1 and is thus isomorphic to a closed subscheme Z' of Z.

Our decomposition of the diagonal on Z gives the relation

$$N \cdot \omega = \gamma_{1*}\omega$$

for each $\omega \in H^0(Z, \Omega_Z^{n-1})$. Indeed,

$$N \cdot \omega = N \cdot \Delta_{Z*}\omega = \gamma_{1*}\omega + \gamma_{2*}\omega + \gamma_*\omega.$$

But γ_* factors through the restriction to Z_{n-2}, so $\gamma_*\omega = 0$. Similarly, $\gamma_{2*}\omega$ is a global section of Ω_Z^{n-1} supported in D, which is zero, since Ω_Z^{n-1} is a locally free sheaf.

One computes that the canonical class of $Y_1 \cap Y_2$ is antiample, and thus the canonical line bundle on the dimension-$(n-1)$ subscheme Z' has no sections. Note that Z' is a cyclic p to 1 cover of the complete intersection $W \cap V(\bar{s})$. If s is general, then there is a rational resolution of singularities $\widetilde{Z}'$ (Proposition 8.8 and Lemma 8.9), hence the canonical line bundle of $\widetilde{Z}'$ has no nonvanishing sections. But $\gamma_{1*}\omega$ factors through the restriction of ω to $\widetilde{Z}'$; hence, $\gamma_{1*}\omega = 0$. Since h has degree $q \cdot p$ in the range needed to give the existence of a nonzero ω in $H^0(Z, \Omega_Z^{n-1})$, we conclude as before that $p \mid N$. $\square$

Example 7.4. We consider the case of hypersurfaces of degree d in $\mathbb{P}^{n+1}$, $n \geq 3$. The theorem says that p divides $\mathrm{Tor}^{(n-2)}(X)$ for very general degree $d \leq n+1$ hypersurfaces X in $\mathbb{P}^{n+1}$ if

$$d \geq p \cdot \left\lceil \frac{n+2}{p+1} \right\rceil.$$

For $p = 2$, this is the range considered by Totaro; for $p = 3$, the first case is degree 6 in $\mathbb{P}^6$. For the extreme case of degree $d = n+1$ in $\mathbb{P}^{n+1}$, we have $p \mid \mathrm{Tor}^{(n-2)}(X)$ for all p dividing $n+1$.

8. An improved lower bound for the very general complete intersection

In this section we extend Theorem 7.1 to cover prime powers. The basic idea is to replace the differential forms with Hodge–Witt cohomology. We are grateful to Kay Rülling for providing the argument for the next lemma which shows that a cycle on $Z \times_k Z$, supported on $Z' \times_k Z$ with $\dim Z' \leq n-2$, acts trivially on $H^0(Z, W_m\Omega_Z^{n-1})$.

Lemma 8.1. *Let k be a perfect field of positive characteristic p, and let X, Y be smooth, equidimensional, and quasiprojective k-schemes. Set $n = \dim X$ and $\mathrm{CH}^n_{\mathrm{prop}/Y}(X \times_k Y) = \varinjlim_Z \mathrm{CH}_{\dim Y}(Z)$, where the limit is over all closed subsets*

$Z \subset X \times_k Y$ that are proper over Y. For $\alpha \in \mathrm{CH}^n_{\mathrm{prop}/Y}(X \times_k Y)$ denote by

$$\alpha_* : \bigoplus_{i,j} H^i(X, W_m\Omega^j) \to \bigoplus_{i,j} H^i(Y, W_m\Omega^j)$$

the map induced by α via the cycle action from [Chatzistamatiou and Rülling 2012, §3.5]. *Assume α is supported on $A \times_k Y$, where $A \subset X$ is a closed subset of codimension $\geq r$. Then α_* vanishes on $\bigoplus_{i,j+r>n} H^i(X, W_m\Omega^j)$.*

Proof. We may assume $\alpha = [Z]$, with $Z \subset X \times_k Y$ an integral closed subscheme of codimension n supported on $A \times_k Y$. Denote by p_X, p_Y the respective projections from $X \times_k Y$. It suffices to show for $i \geq 0$, $j + r > n$, and $b \in H^i(X, W_m\Omega^j)$ that

$$p_X^*(b) \cup \mathrm{cl}[Z] = 0 \quad \text{in } H_Z^{i+n}(X \times_k Y, W_m\Omega^{j+n}_{X\times_k Y}). \tag{8-1}$$

Then $\alpha_*(b) = p_{Y*}(p_X^*(b) \cup \mathrm{cl}[Z])$ will also vanish.

We first prove (8-1) for $i = 0$. Denote by $\eta \in X \times_k Y$ the generic point of Z. Since $W_m\Omega^{j+n}_{X\times_k Y}$ is Cohen–Macaulay, the natural map $H^n_Z(X \times_k Y, W_m\Omega^{j+n}_{X\times_k Y}) \to H^n_\eta(X \times_k Y, W_m\Omega^{j+n}_{X\times_k Y})$ is injective. Set $B = \mathcal{O}_{X\times_k Y,\eta}$ and $C = \mathcal{O}_{X,p_X(\eta)}$; by assumption we have $\dim C \geq r$. Since B is formally smooth over C we find $t_1, \dots, t_r \in C$ and $s_{r+1}, \dots, s_n \in B$ such that $p_X^*(t_1), \dots, p_X^*(t_r), s_{r+1}, \dots, s_n$ form a regular sequence of parameters of B. Hence, by [Gros 1985, Chapitre II, §3.5] (see also [Chatzistamatiou and Rülling 2012, Proposition 2.4.1]) and [Chatzistamatiou and Rülling 2012, Lemma 3.1.5] and in the notation of [Chatzistamatiou and Rülling 2012, §1.11.1] the image of $p_X^*(b) \cup \mathrm{cl}[Z] = \Delta^*(p_X^*(b) \times \mathrm{cl}[Z])$ in $H^n_\eta(X \times_k Y, W_m\Omega^{j+n}_{X\times_k Y})$ is up to a sign given by

$$\begin{bmatrix} p_X^*(b \cdot d[t_1] \cdots d[t_r]) \cdot d[s_{r+1}] \cdots d[s_n] \\ p_X^*([t_1]), \dots, p_X^*([t_r]), [s_{r+1}], \dots, [s_n] \end{bmatrix}.$$

Hence, the vanishing follows from $b \cdot d[t_1] \cdots d[t_r] \in W_m\Omega_X^{j+r} = 0$.

For the general case $i \geq 0$, we first observe that the CM property of $W_m\Omega^{j+n}_{X\times_k Y}$ implies $R\underline{\Gamma}_Z(W_m\Omega^{j+n}_{X\times_k Y}) \cong \mathcal{H}^n_Z(W_m\Omega^{j+n}_{X\times_k Y})[-n]$. Therefore,

$$H_Z^{i+n}(X \times_k Y, W_m\Omega^{j+n}_{X\times_k Y}) = H^i(X \times Y, \mathcal{H}^n_Z(W_m\Omega^{j+n}_{X\times_k Y})).$$

Let $\mathcal{U}$ be an open affine cover of X, and denote by $\mathcal{U} \times_k Y$ the open (not necessarily affine) cover of $X \times_k Y$. We can consider the Cech cohomology with respect to $\mathcal{U} \times_k Y$ and obtain a natural map

$$\check{H}^i(\mathcal{U} \times_k Y, \mathcal{H}^n_Z(W_m\Omega^{j+n}_{X\times_k Y})) \to H^i(X \times_k Y, \mathcal{H}^n_Z(W_m\Omega^{j+n}_{X\times_k Y})). \tag{8-2}$$

Since $\check{H}^i(\mathcal{U}, W_m\Omega^j_X) = H^i(X, W_m\Omega^j_X)$ and pullback and cup product are compatible with restriction to open subsets, we see that $p_X^*(\,\cdot\,) \cup \mathrm{cl}[Z] : H^i(X, W_m\Omega^j_X) \to$

$H_Z^{i+n}(X \times_k Y, W_m\Omega^{j+n}_{X\times_k Y})$ naturally factors via (8-2). Therefore, the case $i \geq 0$ follows from the case $i = 0$. □

Theorem 8.2. *Let k be a field of characteristic zero. Let $X \subset \mathbb{P}^{n+r}_k$ be a very general complete intersection of multidegree $d_1, d_2, \ldots, d_r$ such that $d' := \sum_{i=1}^r d_i \leq n+r$ and $n \geq 3$. Let p be a prime and $m \geq 1$, and suppose*

$$d_i \geq p^m \cdot \left\lceil \frac{n+r+1-d'+d_i}{p^m+1} \right\rceil \tag{8-3}$$

for some i. Furthermore, suppose that p is odd or n is even. Then $p^m \mid \mathrm{Tor}^{(n-2)}(X)$.

Remark 8.3. Just as for Theorem 7.1, the same result holds for k a field of positive characteristic, but only for $p = \operatorname{char} k$.

Proof. The proof relies on Theorem 8.17, which we prove later in this section.

By Proposition 2.1, we need to find only one smooth complete intersection $X \subset \mathbb{P}^{n+r}_k$ such that $p^m \mid \mathrm{Tor}^{(n-2)}(X)$.

For a scheme X with locally free sheaf $\mathscr{E}$ and a section $s : \mathbb{O}_X \to \mathscr{E}$, we let $V(s)$ denote the closed subscheme of X defined by s.

We set $d = d_i$, $a = \lceil (n+r+1-d'+d)/(p^m+1) \rceil$, and $c = d - p^m \cdot a$. Let $\mathbb{O} = W(\bar{\mathbb{F}}_p)$ and $K = \mathrm{Frac}(\mathbb{O})$; we take r, f, g, l, and $f_2, \ldots, f_r$ suitably general (we will make this precise) homogeneous polynomials in $\mathbb{O}[X_0, \ldots, X_{n+r}]$ of degrees d, $d-c$, a, 1, and $d_2, \ldots, d_r$, respectively. We let $\mathscr{X} \to \operatorname{Spec} \mathbb{O}$ be the closed subscheme of the weighted projective space $\mathbb{P} = \operatorname{Proj} \mathbb{O}[X_0, \ldots, X_{n+r}, Y]$, with the X_i having weight 1, and Y having weight a, defined by the homogeneous ideal

$$l^c \cdot (Y^{p^m} - f) + p \cdot r,\, g - p \cdot Y,\, f_2, \ldots, f_r. \tag{8-4}$$

The generic fiber $X := \mathscr{X}_K$ is isomorphic to the complete intersection of $\mathbb{P}^{n+r}_K$ defined by $l^c \cdot (g^{p^m} - p^{p^m} \cdot f) + p^{p^m+1} \cdot r$, $f_2, \ldots, f_r$. For r, $f_2, \ldots, f_r$ general, it is smooth. By replacing $\mathbb{O}$ with its normalization in a suitable finite extension of K and changing notation, we may assume that $\mathrm{Tor}^{(n-2)}_K(X)$ is equal to the geometric torsion order $\mathrm{Tor}^{(n-2)}(X)$.

The special fiber $Y := \mathscr{X}_{\bar{\mathbb{F}}_p}$ is $Y = Y_1 + c \cdot Y_2$. Here, Y_1 is the cyclic p^m cover $Y_1 \to W$ defined by $f \in H^0(W, \mathbb{O}(a)^{\otimes p^m})$, with $W \subset \mathbb{P}^{n+r}_{\bar{\mathbb{F}}_p}$ the complete intersection defined by $g, f_2, \ldots, f_r$. We will take $f, g, f_2, \ldots, f_r$ general enough so that

(1) W is smooth,

(2) Y_1 has nondegenerate singularities (see Section 8A), and

(3) the assumption (3) of Theorem 8.17 is satisfied for Y_1.

For (2) we use Proposition 8.5 if $d - c \geq 3$. If $d - c = 2$ and hence $p = 2$, then we use Lemma 8.7. For (3) we use the theorem of Illusie [1990, Théorème 2.2] about ordinarity of a general complete intersection. Let us check that all other assumptions of Theorem 8.17 are satisfied. Assumption (1) is evident, and (2) is equivalent to $(p^m+1)\cdot a-n-r-1+d'-d\geq 0$, which follows immediately from the definition of a. Assumption (4) is equivalent to $i\cdot a+a-n-r-1+d'-d<0$, for all $i=0,\dots,p^m-1$, which follows from $d'<n+r+1$; (5) is obvious.

The variety Y_2 is defined by $l, g, f_2, \dots, f_r$, and only exists if $c\neq 0$. We take l general so that Y_2 does not contain the singular points of Y_1, $W\cap V(l)$ is smooth, and the p^m cyclic covering of $W\cap V(l)$ corresponding to $f|_{W\cap V(l)}$ has nondegenerate singularities.

Let $r:\tilde{Y}_1\to Y_1$ be the resolution of singularities constructed in Proposition 8.8; the map $r:\tilde{Y}_1\to Y_1$ is totally CH_0-trivial. By Proposition 4.3,

$$\mathrm{Tor}^{(n-2)}(X)\cdot\Delta_{\tilde{Y}_1}=\gamma+Z+Z_2,$$

where γ is a cycle with support in $A\times_{\bar{\mathbb{F}}_p}\tilde{Y}_1$ with $\dim A\leq n-2$, Z has support in $\tilde{Y}_1\times_{\bar{\mathbb{F}}_p}D$ with D a divisor, and Z_2 has support in $(Y_1\cap Y_2)\times_{\bar{\mathbb{F}}_p}\tilde{Y}_1$.

In view of Theorem 8.17, we have $\mathbb{Z}/p^m\subset H^0(\tilde{Y}_1, W_m\Omega^{n-1})$. By the work [Chatzistamatiou and Rülling 2012] on Hodge–Witt cohomology, we have an action of algebraic correspondences on $H^0(\tilde{Y}_1, W_m\Omega^{n-1})$ (relying on the cycle class of Gros [1985, Chapitre II, §3.4]; see [Chatzistamatiou and Rülling 2012, Proposition 2.4.1]). Let us show that Z_2 acts trivially. Note that $T:=Y_1\cap Y_2$ is the p^m cyclic covering of $W\cap V(l)$ corresponding to $f|_{W\cap V(l)}$. An easy computation shows $H^{>0}(Y_1\cap Y_2,\mathbb{O})=0$; hence, $H^{>0}(\tilde{T},\mathbb{O})=0$ by Lemma 8.9, where $\tilde{T}$ is the resolution constructed in Proposition 8.8, and $H^{>0}(\tilde{T}, W_m(\mathbb{O}))=0$. By Ekedahl duality [Ekedahl 1984, Chapter I, Theorem 4.1, Chapter II, Theorem 2.2, and Chapter III, Proposition 2.4] (see [Chatzistamatiou and Rülling 2012, Theorems 1.10.1 and 1.10.3]), we get $H^{<n-1}(\tilde{T}, W_m\Omega^{n-1})=0$. Let $\tilde{Z}_2$ be a lift of Z_2 to $\tilde{T}\times_{\bar{\mathbb{F}}_p}\tilde{Y}_1$. The action of Z_2 factors as

$$H^0(\tilde{Y}_1, W_m\Omega^{n-1})\to H^0(\tilde{T}, W_m\Omega^{n-1})\xrightarrow{\tilde{Z}_2}H^0(\tilde{Y}_1, W_m\Omega^{n-1}),$$

the first map being the pullback for the map $\tilde{T}\to\tilde{Y}_1$; thus, it is zero.

Lemma 8.1 implies that the action of γ on $H^0(\tilde{Y}_1, W_m\Omega^{n-1})$ vanishes. Therefore,

$$H^0(\tilde{Y}_1, W_m\Omega^{n-1})\xrightarrow{\mathrm{Tor}^{(n-2)}(X)\cdot}H^0(\tilde{Y}_1, W_m\Omega^{n-1})\xrightarrow{\text{restriction}}H^0(\tilde{Y}_1\setminus D, W_m\Omega^{n-1})$$

is zero. Since the restriction map is injective, we get $p^m\mid\mathrm{Tor}^{(n-2)}(X)$. □

Corollary 8.4. *Let k be a field of characteristic zero. Let $X\subset\mathbb{P}_k^{n+r}$ be a very general complete intersection of multidegree $d_1,\dots,d_r$ with $\sum_i d_i=n+r$ and $n\geq 3$. For each i, $d_i\mid\mathrm{Tor}^{(n-2)}(X)$ if d_i is odd or if n is even.*

8A. Let X be a smooth variety over an algebraically closed field k of characteristic p. Suppose that $n:=\dim X\geq 2$. Let L be a line bundle on X, and let $s\in H^0(X, L^{\otimes p^m})$.

We denote by $\pi : Y \to X$ the p^m cyclic covering corresponding to s. It is an inseparable morphism and induces a homeomorphism on the underlying topological spaces.

There is a tautological connection $d : L^{\otimes p^m} \to L^{\otimes p^m} \otimes \Omega^1_X$ which satisfies $d(t^{p^m}) = 0$ for all sections $t \in L$. In particular, we have $d(s) \in H^0(X, L^{\otimes p^m} \otimes \Omega^1_X)$. Note that $Y_{\text{sing}} = \pi^{-1}(V(d(s)))$.

We say that Y has *nondegenerate* singularities if the following conditions hold:

(1) Y has at most isolated singularities, or equivalently, $\dim(V(d(s))) = 0$ or $V(d(s)) = \varnothing$.

(2) For all $x \in V(d(s))$, $\text{length}(\mathcal{O}_{V(d(s)),x}) \le 1$, if p is odd or $p = 2$ and n is even. If $p = 2$ and n is odd, then we require $\text{length}(\mathcal{O}_{V(d(s)),x}) \le 2$ and the blow-up $\text{Bl}_x Y$ of x has an exceptional divisor that is a cone over a smooth quadric.

Around a nondegenerate singularity of Y, we can find local coordinates $x_1, \ldots, x_n$ of X such that Y is defined by

$$y^{p^m} + x_1^2 + \cdots + x_n^2 + f_3 \qquad \text{if } p \text{ is odd,} \tag{8-5}$$

$$y^{p^m} + x_1x_2 + \cdots + x_{n-1}x_n + f_3 \qquad \text{if } p = 2 \text{ and } n \text{ is even,} \tag{8-6}$$

$$y^{p^m} + x_1^2 + x_2x_3 + \cdots + x_{n-1}x_n + b \cdot x_1^3 + f_3 \qquad \text{if } p = 2 \text{ and } n \text{ is odd,} \tag{8-7}$$

where $f_3 \in (x_1, \ldots, x_n)^3$, $b \in k^\times$, and f_3 has no x_1^3 term in the last case.

An easy dimension counting argument yields the following proposition (cf. [Kollár 1995, §18]).

Proposition 8.5. *Let $W \subset H^0(X, L^{\otimes p^m})$ be such that for every closed point $x \in X$ the restriction map*

$$W \to \mathcal{O}_{X,x}/m_x^4 \otimes L^{\otimes p^m}$$

is surjective. For a general section $s \in W$ the corresponding p^m cyclic covering has nondegenerate singularities.

Remark 8.6. If $p \ne 2$ or $\dim X$ even, then the following surjectivity is sufficient to conclude the assertion of the proposition:

$$W \to \mathcal{O}_{X,x}/m_x^3 \otimes L^{\otimes p^m} \quad \text{for every closed point } x \in X.$$

In order to handle the case $d_i = 2 = p$, $m = 1$, and $n + r + 1 - d' + 2 \le 3$ in Theorem 8.2 we need the following lemma.

Lemma 8.7. *For a general complete intersection X in $\mathbb{P}^{n+r}$ with $n \ge 2$ and multidegree $d_1, d_2, \ldots, d_r$ such that $d_1 \ge 2$, and a general $s \in H^0(\mathbb{P}^{n+r}, \mathcal{O}(2))$, the double covering corresponding to $s|_X$ has nondegenerate singularities.*

Proof. Only the case $p = 2$ and n odd has to be proved. Consider the variety A consisting of points $(x, f_1, \dots, f_r, s)$ where $x \in \mathbb{P}^{n+r}$, $(f_1, \dots, f_r, s)$ are homogeneous of degree $d_1, \dots, d_r, 2$, $X = V(f_1) \cap \dots \cap V(f_r)$ is smooth at x, and $d(s)|_X$ is vanishing at x. Those points for which the double covering corresponding to $s|_X$ has nondegenerate singularities at x form an open set B. It is not difficult to show that it is nonempty. Indeed, take $x = [1 : 0 : 0 : \dots : 0]$, and (in coordinates $x_1, \dots, x_{n+r}$ around x) $s = 1 + x_1^2 + x_2x_3 + \dots + x_{n-1}x_n + x_1x_{n+1}$, $f_1 = x_{n+1} + x_1^2$, and $f_i = x_{n+i} +$ terms of degree ≥ 2.

Let $V \subset A$ be the open set consisting of points such that $V(f_1) \cap \dots \cap V(f_r)$ is smooth. Since $B \cap V \neq \varnothing$, we conclude that for a general complete intersection X there is an open nonempty set $U \subset X$ such that for any $x \in U$ the set

$$\big\{s \in H^0(\mathbb{P}^{n+r}, \mathscr{O}(2)) \;\big|\; d(s)|_X(x) = 0 \text{ and } s \text{ does not yield a nondegenerate double covering at } x\big\}$$

has codimension $\geq n + 1$. Counting dimensions yields the claim. $\square$

The following proposition has been proved for the case $m = 1$ in [Colliot-Thélène and Pirutka 2016a], and for the general case in [Okada 2016].

Proposition 8.8. *Suppose Y has nondegenerate singularities. Then by successively blowing up singular points, we can construct a resolution of singularities $r : \tilde{Y} \to Y$ such that the exceptional divisor is a normal crossings divisor* (cf. [Kollár 1995]). *Over every singular point $y \in Y$ the fiber $r^{-1}(y)$ is a chain of smooth irreducible divisors, each component of which is either a projective space, a smooth quadric, or a projective bundle over a smooth quadric. The intersection of two irreducible components is a smooth quadric or is empty. In particular, since k is algebraically closed, the morphism r is totally* CH_0 *trivial.*

Proof. We distinguish three cases:

(1) p is odd,

(2) $p = 2$, and n is even, and

(3) $p = 2$, and n is odd.

In any case we will only blow up singular points, and over any singular s there will be at most one singular point appearing in the exceptional divisor of the blow-up of s.

We may assume that Y has only one singular point. In case (1), note that we have a singularity of the form (8-5). We need $(p^m - 1)/2 + 1$ blow-ups:

$$\tilde{Y} := Y_{(p^m-1)/2+1} \to Y_{(p^m-1)/2} \to \dots \to Y_1 \to Y_0 := Y.$$

Around the singularity of Y_i, for $0 \leq i < (p^m - 1)/2$, Y_i is defined by

$$y^{p^m-2\cdot i} + x_1'^2 + \dots + x_n'^2 + f_3', \tag{8-8}$$

where $x'_i = x_i/y^i$ and $f'_3 \in y^i \cdot (x'_1, \dots, x'_n)^3$. Therefore, the exceptional divisor of $Y_{i+1} \to Y_i$ is the cone C defined by $x'^2_1 + \dots + x'^2_n$ in the projective space with homogeneous variables $y, x'_1, \dots, x'_n$. For $i = (p^m - 1)/2$, Y_i is also given by (8-8) around the vertex of the exceptional divisor; hence, $p^m - 2i = 1$ implies that it is smooth and the exceptional divisor of $Y_{(p^m-1)/2+1} \to Y_{(p^m-1)/2}$ is $\mathbb{P}^{n-1}$. Denoting by $\tilde{E}_i$ the strict transform in $\tilde{Y}$ of the exceptional divisor of $Y_i \to Y_{i-1}$, we conclude that $\tilde{E}_i$ is the blow-up of C in its vertex if $i \le (p^m - 1)/2$, and $\tilde{E}_{(p^m-1)/2+1} = \mathbb{P}^{n-1}$. Every $\tilde{E}_i$ has only nonempty intersection with $\tilde{E}_{i+1}$ (if $i \le (p^m - 1)/2$) and $\tilde{E}_{i-1}$ (if $i > 1$); the intersection is the smooth quadric given by $x'^2_1 + \dots + x'^2_n$ in the projective space with homogeneous variables $x'_1, \dots, x'_n$.

For case (2), this case is similar to (1). We need 2^{m-1} blow-ups to arrive at $\tilde{Y}$. Around the singularity of Y_i, for $0 \le i < 2^{m-1}$, Y_i is defined by

$$y^{2^m-2\cdot i} + x'_1x'_2 + \dots + x'_{n-1}x'_n + f'_3, \tag{8-9}$$

and the exceptional divisor of $Y_i \to Y_{i-1}$ is the cone C defined by $x'_1x'_2 + \dots + x'_{n-1}x'_n$ in the projective space P with homogeneous variables $y, x'_1, \dots, x'_n$. The exceptional divisor of $\tilde{Y} := Y_{2^{m-1}} \to Y_{2^{m-1}-1}$ is the smooth quadric defined by $y^2 + x'_1x'_2 + \dots + x'_{n-1}x'_n$ in P. Again, the intersection of $\tilde{E}_i$ with $\tilde{E}_{i-1}$ is the smooth quadric given by $x'_1x'_2 + \dots + x'_{n-1}x'_n$ in the projective space with homogeneous variables $x'_1, \dots, x'_n$.

For case (3), we need 2^m blow-ups to arrive at $\tilde{Y}$. The case $m = 1$ is easy to check; we will assume $m > 1$. We start with Y and the singularity (8-7). After $2^{m-1} - 1$ blow-ups the singularity is of the form

$$b \cdot y^{2^{m-1}+2} + x_1^{[1]^2} + x'_2x'_3 + \dots + x'_{n-1}x'_n + b \cdot x_1^{[1]} \cdot y^{2^{m-1}+1} + \text{h.o.t.},$$

where $x'_i = x_i/y^{2^{m-1}-1}$, $x_1^{[1]} = x'_1 + y$, and the higher order terms h.o.t. can be ignored. After 2^{m-2} more blow-ups we introduce $x_1^{[2]} = x_1^{[1]}/y^{2^{m-2}} + \sqrt{b} \cdot y$, after 2^{m-3} more blow-ups we introduce $x_1^{[3]} = x_1^{[2]}/y^{2^{m-3}} + \sqrt{\sqrt{b} \cdot b} \cdot y$, etc. The singularity is after $2^{m-1} - 1 + 2^{m-2} + 2^{m-3} + \dots + 2^{m-i}$ blow-ups of the form

$$b_i \cdot y^{2^{m-i}+2} + x_1^{[i]^2} + x'_2x'_3 + \dots + x'_{n-1}x'_n + b \cdot x_1^{[i]} \cdot y^{2^{m-i}+1} + \text{h.o.t.}, \tag{8-10}$$

where $x'_i = x_i/y^{-1+\sum_{j=1}^{i} 2^{m-j}}$ and $b_i = b \cdot \sqrt{b_{i-1}}$ with $b_1 = b$. After $2^m - 2$ blow-ups we get a singularity (8-10) with $i = m$. After one more blow-up the variety becomes smooth, and we need one more blow-up to obtain an exceptional divisor with strict normal crossings.

The exceptional divisor E_i of $Y_i \to Y_{i-1}$ is a cone defined by $x_1^{[j]^2} + x'_2x'_3 + \dots + x'_{n-1}x'_n$ in the projective space with homogeneous variables $y, x_1^{[j]}, x'_2, \dots, x'_n$, except for the last blow-up where it is a projective space. The strict transform $\tilde{E}_i$ is the blow-up of the vertex. $\square$

For p odd or n odd, we get a projective space as exceptional divisor in the last step. Denoting by E the sum over all components of the exceptional divisor of r, we set

$$E' := \begin{cases} E + (\text{exc. div. from last step}) & \text{if } p \text{ is odd or } n \text{ is odd,} \\ E & \text{if } p = 2 \text{ and } n \text{ is even.} \end{cases} \tag{8-11}$$

Thus, the exceptional divisor of the last blow-up (a projective space) has multiplicity 2 in E' in the first case. If the singularity is of the form (8-5), (8-6), or (8-7), then E' is the restriction of $\operatorname{div}(y)$ to the exceptional divisor of the resolution r.

Lemma 8.9. *The resolution $r : \tilde{Y} \to Y$ is rational*; *that is*, $Rr_*\mathcal{O}_{\tilde{Y}} = \mathcal{O}_Y$.

Proof. We may suppose that Y has only one singularity. We will show that for each $r_i : Y_i \to Y_{i-1}$, we have $Rr_{i*}\mathcal{O}_{Y_i} = \mathcal{O}_{Y_{i-1}}$. Since Y_{i-1} is normal, it suffices to prove $R^j r_{i*}\mathcal{O}_{Y_i} = 0$ for $j > 0$. We know that r_i is the blow-up of a point and the exceptional divisor D is a cone over a smooth quadric, a smooth quadric, or a projective space, and comes with a given embedding into projective space; we call the corresponding ample line bundle $\mathcal{O}_D(1)$. In any case, $H^{>0}(D, \mathcal{O}(-s \cdot D)) \cong H^{>0}(D, \mathcal{O}(s)) = 0$ for all $s \geq 0$, where $\mathcal{O}_D(s) = \mathcal{O}_D(1)^{\otimes s}$. This implies the claim. □

Lemma 8.10. *Let E' be as defined in* (8-11). *For all $i \geq 2$ we have*

$$H^i(E', \mathcal{O}(E')) = 0.$$

Proof. We may suppose that Y has only one singular point. The exceptional divisor is $\sum_{i=1}^{s} \tilde{E}_i$, and $\tilde{E}_i$ has nonempty intersection only with $\tilde{E}_{i+1}$ and $\tilde{E}_{i-1}$. Recall that all intersections are smooth quadrics. If $i \neq s$, then $\tilde{E}_i$ is the blow-up at the vertex of a cone $C_i \subset \mathbb{P}^n$ over a smooth quadric $Q_i \subset \mathbb{P}^{n-1}$; let $r_i : \tilde{E}_i \to C_i$ denote the blow-up.

For $i = 1, \ldots, s-2$, we have $\mathcal{O}_{\tilde{E}_i}(\tilde{E}_i + \tilde{E}_{i+1}) \cong r_i^*\mathcal{O}_{C_i}(-1)$, hence

$$\mathcal{O}_{\tilde{E}_i \cap \tilde{E}_{i+1}}(E') \cong \mathcal{O}_{\tilde{E}_i \cap \tilde{E}_{i+1}}. \tag{8-12}$$

For $i = 2, \ldots, s-2$, we obtain $\mathcal{O}_{\tilde{E}_i}(E') \cong \mathcal{O}_{\tilde{E}_i}$.

If p or n is odd, then

$$\mathcal{O}_{\tilde{E}_{s-1}}(\tilde{E}_{s-1} + 2 \cdot \tilde{E}_s) \cong r_{s-1}^*\mathcal{O}_{C_{s-1}}(-1);$$

hence $\mathcal{O}_{\tilde{E}_{s-1}}(E') \cong \mathcal{O}_{\tilde{E}_{s-1}}$, and $\mathcal{O}_{\tilde{E}_s}(\tilde{E}_{s-1} + 2 \cdot \tilde{E}_s) \cong \mathcal{O}_{\mathbb{P}^{n-1}}$; thus (8-12) holds for $i = s-1$. If p and n are even, then $\mathcal{O}_{\tilde{E}_{s-1}}(\tilde{E}_{s-1} + \tilde{E}_s) \cong r_{s-1}^*\mathcal{O}_{C_{s-1}}(-1)$, hence $\mathcal{O}_{\tilde{E}_{s-1}}(E') \cong \mathcal{O}_{\tilde{E}_{s-1}}$. Moreover, $\mathcal{O}_{\tilde{E}_s}(E') \cong \mathcal{O}_{\tilde{E}_s}$. This implies the assertion easily. □

8B. Again, we assume that Y has nondegenerate singularities. We denote by $U \subset X$

the complement of the critical points, $Y_{\mathrm{sm}} = \pi^{-1}(U)$; we have

$$W_l(\pi)^* W_l\Omega^1_{U/k} \to W_l\Omega^1_{Y_{\mathrm{sm}}/k},$$

but there is no Verschiebung on $W_l(\pi)^* W_l\Omega^1_{U/k}$. Therefore, we define

$$\mathrm{Im}_V(W_l\Omega^1_{U/k}) \subset W_l\Omega^1_{Y_{\mathrm{sm}}/k}$$

inductively on l by

$$\mathrm{Im}_V(W_l\Omega^1_{U/k}) = \mathrm{image}(W_l(\pi)^* W_l\Omega^1_{U/k} \to W_l\Omega^1_{Y_{\mathrm{sm}}/k}) + V(\mathrm{Im}_V(W_{l-1}\Omega^1_{U/k})).$$

We have an R, V, F calculus for $\mathrm{Im}_V(W_l\Omega^1_{U/k})$, that is, morphisms

$$\begin{aligned} R &: \mathrm{Im}_V(W_l\Omega^1_{U/k}) \to \mathrm{Im}_V(W_{l-1}\Omega^1_{U/k}), \\ V &: \mathrm{Im}_V(W_{l-1}\Omega^1_{U/k}) \to \mathrm{Im}_V(W_l\Omega^1_{U/k}), \\ F &: \mathrm{Im}_V(W_l\Omega^1_{U/k}) \to \mathrm{Im}_V(W_{l-1}\Omega^1_{U/k}) \end{aligned}$$

satisfying the relations induced by $W_*\Omega^1_{Y_{\mathrm{sm}}/k}$ [Illusie 1979, p. 541]. By abuse of notation, any composition of maps R will be also denoted by R.

We are going to need several statements on $\mathrm{Im}_V(W_l\Omega^1_{U/k})$ in Theorem 8.17 which we provide in the following:

Lemma 8.11. *The evident map*

$$\begin{aligned} &\ker(R : W_l(\pi)^* W_l\Omega^1_{U/k} \to \pi^*\Omega^1_U) \\ &\qquad \to \ker(R : \mathrm{Im}_V(W_l\Omega^1_{U/k}) \to \mathrm{Im}_V(W_1\Omega^1_{U/k}))/V(\mathrm{Im}_V(W_{l-1}\Omega^1_{U/k})) \end{aligned} \tag{8-13}$$

is surjective if $l \le m$.

Proof. The target is the image of $R^{-1}(\ker(\pi^*\Omega^1_U \to \Omega^1_{Y_{\mathrm{sm}}})) \subset W_l(\pi)^* W_l\Omega^1_{U/k}$ via the evident map $W_l(\pi)^* W_l\Omega^1_{U/k} \to \mathrm{Im}_V(W_l\Omega^1_{U/k})/V(\mathrm{Im}_V(W_{l-1}\Omega^1_{U/k}))$. Locally, Y_{sm} is defined by $y^{p^m} - f$, for $f \in \mathbb{O}_U$, and $\ker(\pi^*\Omega^1_U \to \Omega^1_{Y_{\mathrm{sm}}})$ is generated by $d(f)$. Since $d([f]) \in W_l(\pi)^* W_l\Omega^1_{U/k}$ is a lifting of $d(f)$ whose image vanishes in $\mathrm{Im}_V(W_l\Omega^1_{U/k})$ (here we use $l \le m$), the claim follows. $\square$

Recall the subsheaves $B_n\Omega^1_{U/k}$ of $\Omega^1_{U/k}$, $n = 1, 2, \ldots$ (see for example [Illusie 1979, Chapitre I, §2.2]). We have a short exact sequence

$$W_{l-1}\Omega^1_{U/k} \xrightarrow{V} \ker(R : W_l\Omega^1_{U/k} \to \Omega^1_U) \xrightarrow{F^{l-1}} B_{l-1}\Omega^1_U \to 0.$$

With the appropriate $W_l(\mathbb{O}_U)$-module structures this becomes a short exact sequence of $W_l(\mathbb{O}_U)$-modules. We obtain the diagram

$$\begin{tikzcd} W_l(\pi)^* \ker(R : W_l\Omega^1_{U/k} \to \Omega^1_U)/W_l(\pi)^*(V)(W_l(\pi)^* W_{l-1}\Omega^1_{U/k}) \arrow[r, "\cong"] \arrow[d, two heads, "\text{surjective by Lemma 8.11}"] & \pi^* B_{l-1}\Omega^1_U \\ \ker(R : \mathrm{Im}_V(W_l\Omega^1_{U/k}) \to \mathrm{Im}_V(W_1\Omega^1_{U/k}))/V(\mathrm{Im}_V(W_{l-1}\Omega^1_{U/k})) \arrow[d, "(*)"] & \\ \ker(R : W_l\Omega^1_{Y_{\mathrm{sm}}/k} \to \Omega^1_{Y_{\mathrm{sm}}})/V(W_{l-1}\Omega^1_{Y_{\mathrm{sm}}/k}) \arrow[r, "\cong"] & B_{l-1}\Omega^1_{Y_{\mathrm{sm}}} \end{tikzcd} \tag{8-14}$$

The induced map

$$\pi^* B_{l-1}\Omega^1_U \to B_{l-1}\Omega^1_{Y_{\mathrm{sm}}} \tag{8-15}$$

is the natural one, that is, given by $a \otimes \pi^{-1}(\omega) \mapsto \mathrm{Frob}^{l-1}(a) \cdot \pi^{-1}(\omega)$. We would like to show that $(*)$ is injective, which we prove by computing the kernel of (8-15) and showing that it is killed in $\mathrm{Im}_V(W_l\Omega^1_{U/k})$.

It is convenient to use the isomorphism [Illusie 1979, Chapitre I, (3.11.4)]

$$F^{l-2}d : W_{l-1}(\mathbb{O}_U)/F(W_{l-1}(\mathbb{O}_U)) \xrightarrow{\cong} B_{l-1}\Omega^1_U. \tag{8-16}$$

The $W_l(\mathbb{O}_U)$-module structure on the left is via the Frobenius $F : W_l(\mathbb{O}_U) \to W_{l-1}(\mathbb{O}_U)$. We give $W_{l-1}(\mathbb{O}_{Y_{\mathrm{sm}}})/F(W_{l-1}(\mathbb{O}_{Y_{\mathrm{sm}}}))$ the analogous $W_l(\mathbb{O}_{Y_{\mathrm{sm}}})$-module structure.

Lemma 8.12. *Suppose Y_{sm} is defined by $y^{p^m} - f$ for $f \in \mathbb{O}_U$ (this is the local picture). The kernel of*

$$W_l(\pi)^*(W_{l-1}(\mathbb{O}_U)/F(W_{l-1}(\mathbb{O}_U))) \to W_{l-1}(\mathbb{O}_{Y_{\mathrm{sm}}})/F(W_{l-1}(\mathbb{O}_{Y_{\mathrm{sm}}}))$$

is generated by $V(W_{l-1}(\mathbb{O}_{Y_{\mathrm{sm}}})) \otimes W_l(\pi)^{-1}(W_{l-1}(\mathbb{O}_U))$, and elements of the form

$$[y^i] \otimes \pi^{-1}(V^j(b)) - [y^{i\% p^{m-1-j}}] \otimes \pi^{-1}(V^j([f^{(i:p^{m-1-j})}] \cdot b)), \tag{8-17}$$

for all $0 \le j \le l-2$, $i \ge p^{m-1-j}$, and $b \in W_{l-1-j}(\mathbb{O}_U)$. Here, $i\% p^{m-1-j}$ means the remainder of i in the division by p^{m-1-j}, and $i = (i : p^{m-1-j}) \cdot p^{m-1-j} + i\% p^{m-1-j}$.

Proof. The kernel contains $V(W_{l-1}(\mathbb{O}_{Y_{\mathrm{sm}}})) \otimes W_l(\pi)^{-1}(W_{l-1}(\mathbb{O}_U))$, because $V(a) \otimes \pi^{-1}(b)$ maps to $F(V(a)) \cdot \pi^{-1}(b) = pa \cdot \pi^{-1}(b) = F(V(a \cdot \pi^{-1}(b)))$. Moreover,

$$\begin{aligned} &[y^i] \otimes \pi^{-1}(V^j(b)) - [y^{i\% p^{m-1-j}}] \otimes \pi^{-1}(V^j([f^{(i:p^{m-1-j})}] \cdot b)) \\ &\qquad \mapsto [y^{pi}] \cdot V^j(b) - [y^{(i\% p^{m-1-j}) \cdot p}] \cdot V^j([f^{(i:p^{m-1-j})}] \cdot b) \\ &\qquad\qquad = V^j(([y^{p^{1+j} \cdot i}] - [y^{(i\% p^{m-1-j}) \cdot p^{1+j}} \cdot f^{(i:p^{m-1-j})}]) \cdot b) = 0. \end{aligned}$$

To show that these are all elements in the kernel, we proceed by induction on l. First, we assume $l = 2$. Without loss of generality, we need only consider elements in the kernel that are of the form $\sum_i [y^i] \otimes \pi^{-1}(b_i)$. By étale base-change, we may assume that $U = \mathrm{Spec}(k[x_1, \dots, x_n])$ and $x_1 = f$; hence, $Y_{\mathrm{sm}} = \mathrm{Spec}(k[y, x_2, \dots, x_n])$. By using elements of the form (8-17), we may suppose that $b_i = b_i(x_2, \dots, x_n)$. Since $\sum_i y^{ip} b_i \in k[y^p, x_2^p, \dots, x_n^p]$ implies $b_i \in k[x_2^p, \dots, x_n^p]$, we are done.

Suppose now that $l > 2$. By induction, we need only consider elements in the kernel that are of the form

$$\sum_i [y^i] \otimes \pi^{-1}(V^{l-2}(b_i)),$$

and we may use the same argument as for the $l = 2$ case. □

Proposition 8.13. *Suppose $l \le m$. The map*

$$\ker(R : \mathrm{Im}_V(W_l \Omega^1_{U/k}) \to \mathrm{Im}_V(W_1 \Omega^1_{U/k}))/V(\mathrm{Im}_V(W_{l-1}\Omega^1_{U/k})) \\ \to \ker(R : W_l \Omega^1_{Y_{\mathrm{sm}}/k} \to \Omega^1_{Y_{\mathrm{sm}}})/V(W_{l-1}\Omega^1_{Y_{\mathrm{sm}}/k})$$

is injective.

Proof. In view of diagram (8-14) and Lemma 8.12, we need to prove that the following elements vanish in $\mathrm{Im}_V(W_l \Omega^1_{U/k})/V(\mathrm{Im}_V(W_{l-1}\Omega^1_{U/k}))$:

(1) $V(a) \cdot dV(b)$ for $a \in W_l(\mathbb{O}_{Y_{\mathrm{sm}}})$ and $b \in W_{l-1}(\mathbb{O}_U)$ and

(2) $[y^i] \cdot dV^{j+1}(b) - [y^{i \% p^{m-1-j}}] \cdot dV^{j+1}([f^{(i : p^{m-1-j})}] \cdot b)$ for $b \in W_{l-1-j}(\mathbb{O}_U)$.

For (1), we have

$$V(a) \cdot dV(b) = V(a \cdot d(b)) \in V(\mathrm{Im}_V(W_{l-1}\Omega^1_{U/k})).$$

For (2), we compute

$$\begin{aligned}
[y^i] \cdot dV^{j+1}(b) &= d([y^i] \cdot V^{j+1}(b)) - V^{j+1}(b) \cdot d([y^i]) \\
&= dV^{j+1}([y^{i \cdot p^{1+j}}] \cdot b) - V^{j+1}(b) \cdot d([y^i]) \\
&= dV^{j+1}([y^{(i \% p^{m-1-j}) \cdot p^{1+j}}] \cdot [f^{(i : p^{m-1-j})}] \cdot b) - V^{j+1}(b) \cdot d([y^i]) \\
&= d([y^{i \% p^{m-1-j}}] \cdot V^{j+1}([f^{(i : p^{m-1-j})}] \cdot b)) - V^{j+1}(b) \cdot d([y^i]) \\
&= V^{j+1}([f^{(i : p^{m-1-j})}] \cdot b) \cdot d([y^{i \% p^{m-1-j}}]) \\
&\qquad + [y^{i \% p^{m-1-j}}] \cdot dV^{j+1}([f^{(i : p^{m-1-j})}] \cdot b) - V^{j+1}(b) \cdot d([y^i]),
\end{aligned}$$

which together with

$$\begin{aligned}
&V^{j+1}([f^{(i : p^{m-1-j})}] \cdot b) \cdot d([y^{i \% p^{m-1-j}}]) - V^{j+1}(b) \cdot d([y^i]) \\
&\quad = V^{j+1}(b \cdot ([f^{(i : p^{m-1-j})}] \cdot F^{j+1}(d([y^{i \% p^{m-1-j}}])) - F^{j+1}(d([y^i])))) \\
&\qquad = V^{j+1}(b \cdot F^{j+1}(d([y^{(i : p^{m-1-j}) \cdot p^{m-1-j}}][y^{i \% p^{m-1-j}}] - [y^i]))) = 0
\end{aligned}$$

(note that $F^{j+1}(d([y^{(i : p^{m-1-j}) \cdot p^{m-1-j}}])) = 0$) implies the claim. □

8C. We denote by $\jmath : r^{-1}(Y_{\mathrm{sm}}) \to \tilde{Y}$ the open immersion. We will work with the logarithmic de Rham–Witt complex

$$W_l\Omega^1_{\tilde{Y}/k}(\log E) \subset \jmath_* W_l\Omega^1_{Y_{\mathrm{sm}}/k}.$$

Locally, when $E = \bigcup_{i=1}^r V(f_i)$ with $V(f_i)$ smooth, $W_l\Omega^1_{\tilde{Y}/k}(\log E)$ is generated as a $W_l(\mathbb{O}_{\tilde{Y}})$ submodule of $\jmath_* W_l\Omega^1_{Y_{\mathrm{sm}}/k}$ by $W_l\Omega^1_{\tilde{Y}/k}$ and $\langle d[f_i]/[f_i] \mid i = 1, \dots, r\rangle$. As for the de Rham complex there is an exact sequence

$$0 \to W_l\Omega^1_{\tilde{Y}/k} \to W_l\Omega^1_{\tilde{Y}/k}(\log E) \to \bigoplus_{i=0}^{r} W_l(\mathbb{O}_{V(f_i)}) \to 0. \tag{8-18}$$

We have the usual F, V, R calculus for $W_*\Omega^1_{\tilde{Y}/k}(\log E)$.

We define

$$K_l := \jmath_* \operatorname{Im}_V(W_l\Omega^1_{U/k}) \cap W_l\Omega^1_{\tilde{Y}/k}(\log E) \subset \jmath_* W_l\Omega^1_{Y_{\mathrm{sm}}/k}.$$

We have an F, V, R calculus for K_* induced by the one for $\operatorname{Im}_V(W_*\Omega^1_{U/k})$ and $W_*\Omega^1_{\tilde{Y}/k}(\log E)$. We set $Q_* := W_*\Omega^1_{\tilde{Y}/k}(\log E)/K_*$.

Lemma 8.14. *Suppose that $p \neq 2$ or n is even. Then, for all $l \geq 1$, the following map is surjective*:

$$R : K_l \to K_1.$$

Proof. The first case is $p \neq 2$. We need to compute K_1. We may assume that Y has only one singularity as in the proof of Proposition 8.8. Recall that $\tilde{Y}$ is constructed as a sequence of blow-ups $\cdots \to Y_i \to Y_{i-1} \to \cdots \to Y$. We denote by $r_i : Y_i \to Y$ the evident composition; we let D_i be the exceptional divisor of r_i, and E_i denotes the exceptional divisor of $Y_i \to Y_{i-1}$. We would like to understand

$$\jmath_{Y_i \setminus D_i,*}(\mathrm{image}(r_i^*\pi^*\Omega^1_X|_{Y_i \setminus D_i} \to \Omega^1_{Y_i\setminus D_i})) \cap \Omega^1_{Y_{i,\mathrm{sm}}}(\log D_i|_{Y_{i,\mathrm{sm}}}), \tag{8-19}$$

in a neighborhood of $E_i \cap Y_{i,\mathrm{sm}}$, where $Y_{i,\mathrm{sm}}$ is the smooth locus of Y_i, and $\jmath_{Y_i\setminus D_i} : Y_i \setminus D_i \to Y_{i,\mathrm{sm}}$ is the open immersion.

As in the proof of Proposition 8.8, we have coordinates $y, x'_1, \dots, x'_n$ around the singular point of Y_{i-1}, where $x'_j = x_j/y^{i-1}$. We can cover E_i by $n+1$ open sets $V_0, V_1, \dots, V_n$, where V_0 is a hypersurface in the affine space with coordinates $y, x'_1/y, \dots, x'_n/y$, and V_j is a hypersurface in the affine space with coordinates $y/x'_j, x'_1/x'_j, \dots, x'_j, \dots, x'_n/x'_j$, for $j = 1, \dots, n$. On V_0 we have $E_i \cap V_0 = D_i \cap V_0 = V(y)$. Note that if $i = (p^m - 1)/2 + 1$, which is the last blow-up, then $E_i \cap V_0$ is empty.

On V_j we have $E_i \cap V_j = V(x'_j)$ and $D_i \cap V_j = V(y)$ if $j = 1, \dots, n$ and $i \notin \{1, (p^m-1)/2+1\}$, that is, except for the first and the last blow-ups. For the first blow-up ($i = 1$), we have $E_i \cap V_j = D_i \cap V_j = V(x'_j)$. For the last blow-up ($i = (p^m-1)/2+1$), we have $E_i \cap V_j = V(x'_j)$ and $D_i \cap V_j = V(y/x'_j)$.

We claim that the restriction of (8-19) to V_0 is generated by $dx_1/y^i, \dots, dx_n/y^i$, and the restriction of (8-19) to V_j is generated by $dx_1/x_j, \dots, dx_j/x_j, \dots, dx_n/x_j$. It is obvious that all differential forms are contained in the left-hand side of (8-19), and we need to show that they are contained in $\Omega^1_{Y_{i,\mathrm{sm}}}(\log D_i|_{Y_{i,\mathrm{sm}}})$. Indeed, $dx_j/y^i = d(x'_j/y \cdot y^i)/y^i = d(x'_j/y) + i \cdot (x'_j/y) \cdot (dy/y)$, and

$$\frac{dx_k}{x_j} = d\left(\frac{x_k}{x_j}\right) + \frac{x_k}{x_j} \cdot \frac{dx_j}{x_j} = d\left(\frac{x'_k}{x'_j}\right) + \frac{x'_k}{x'_j} \cdot \frac{dx_j}{x_j} = d\left(\frac{x'_k}{x'_j}\right) + \frac{x'_k}{x'_j} \cdot \left(\frac{dx'_j}{x'_j} + (i-1) \cdot \frac{dy}{y}\right).$$

In order to show that the given differential forms are generators, we note that the quotient of $\Omega^1_{Y_{i,\mathrm{sm}}}(\log D_i|_{Y_{i,\mathrm{sm}}}) \cap V_j$ by the module generated by these forms is a quotient of a free rank $= 1$ module. Since the quotient of $\Omega^1_{Y_{\mathrm{sm}}}$ by the image of $\pi^*(\Omega^1_U)$ is free of rank 1, the claim follows.

The case $p = 2$ and n even can be proved in the same way.

In order to prove that $K_l \to K_1$ is surjective, we may argue by induction on i and only consider a neighborhood of $E_i \cap Y_{i,\mathrm{sm}}$ in $Y_{i,\mathrm{sm}}$. We note that dx_j/y^i can be lifted by $d[x_j]/[y^i] \in K_l(V_0)$, and dx_k/x_j can be lifted by $d[x_k]/[x_j] \in K_l(V_j)$. □

Remark 8.15. We do not know whether Lemma 8.14 holds if $p = 2$ and n is odd. We can still describe K_1, but the coordinate changes $x_1^{[1]}, x_1^{[2]}, \dots$ used in the resolution process are incompatible with the multiplicative Teichmüller map and evident liftings do not exist.

8D. Let us assume that $p \neq 2$ or n is even. In view of the lemma, the map

$$\ker\big(W_*\Omega^1_{\tilde{Y}/k}(\log E) \xrightarrow{R} W_1\Omega^1_{\tilde{Y}/k}(\log E)\big) \to \ker(Q_* \xrightarrow{R} Q_1) \tag{8-20}$$

is surjective.

As a consequence of Proposition 8.13 we obtain the following corollary.

Corollary 8.16. *For all $l \leq m$, the composition*

$$\begin{aligned}\mathbb{O}_{\tilde{Y}}/\mathbb{O}_{\tilde{Y}}^{p^{l-1}} &\xrightarrow{dV^{l-2},\cong} \ker(V : W_{l-1}\Omega^1_{\tilde{Y}/k} \to W_l\Omega^1_{\tilde{Y}/k}) \\ &\to \ker\big(V : W_{l-1}\Omega^1_{\tilde{Y}/k}(\log E) \to W_l\Omega^1_{\tilde{Y}/k}(\log E)\big) \to \ker(V : Q_{l-1} \to Q_l)\end{aligned}$$

is surjective on the open set Y_{sm}.

Proof. The first isomorphism follows from [Illusie 1979, Chapitre I, Proposition 3.11]. The second arrow is an isomorphism on Y_{sm}. Set

$$\begin{aligned} A_l &:= \ker(R : \mathrm{Im}_V(W_l\Omega^1_{U/k}) \to \mathrm{Im}_V(W_1\Omega^1_{U/k})), \\ B_l &:= \ker(R : W_l\Omega^1_{Y_{\mathrm{sm}}/k} \to \Omega^1_{Y_{\mathrm{sm}}}), \\ C_l &:= \ker(R : Q_{l|Y_{\mathrm{sm}}} \to Q_{1|Y_{\mathrm{sm}}}). \end{aligned}$$

In view of (8-20) we have a morphism of exact sequences

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathrm{Im}_V(W_{l-1}\Omega^1_{U/k}) & \longrightarrow & W_{l-1}\Omega^1_{Y_{\mathrm{sm}}/k} & \longrightarrow & Q_{l-1|Y_{\mathrm{sm}}} & \longrightarrow & 0 \\ & & \downarrow V & & \downarrow V & & \downarrow V & & \\ 0 & \longrightarrow & A_l & \longrightarrow & B_l & \longrightarrow & C_l & \longrightarrow & 0 \end{array}$$

and the snake lemma and Proposition 8.13 imply the assertion. □

Theorem 8.17. *Let X be a smooth projective variety of dimension n over an algebraically closed field of characteristic p. Suppose that p is odd or n is even. Let L be a line bundle on X, and let $s \in H^0(X, L^{\otimes p^m})$ for $m \geq 1$. Suppose that the p^m cyclic covering $\pi : Y \to X$ corresponding to s has only nondegenerate singularities; let $r : \tilde{Y} \to Y$ be the resolution from Proposition 8.8. Suppose that*

(1) $n \geq 3$,

(2) $H^0(X, L^{\otimes p^m} \otimes K_X) \neq 0$,

(3) *the Frobenius acts bijectively on* $H^{n-1}(V(s), \mathcal{O})$,

(4) $H^n(X, L^{\otimes -j}) = 0$ *for all* $j = 0, \ldots, p^m - 1$, *and*

(5) $H^{n-1}(X, L^{\otimes -j}) = 0$ *for all* $j = 0, \ldots, p^m$.

Then $W_m(k) \subset H^0(\tilde{Y}, W_m\Omega^{n-1})$.

Proof. We have

$$\mathrm{coker}(\pi^*(\Omega^1_U) \to \Omega^1_{Y_{\mathrm{sm}}}) = \pi^*(L^{-1}),$$

and this identity extends to

$$Q_1 = r^*\pi^*(L^{-1})(E')$$

on $\tilde{Y}$, with E' as defined in (8-11). If the singularity of Y is of the form (8-5), (8-6), or (8-7), then Q_1 is generated by dy/y.

In view of Lemmas 8.9 and 8.10, and conditions (2), (4), and (5), we obtain

$$H^{n-1}(\tilde{Y}, Q_1) = 0, \qquad H^n(\tilde{Y}, Q_1) \cong H^n(X, L^{\otimes -p^m}) \neq 0. \tag{8-21}$$

We will work with the short exact sequences

$$0 \to \ker(R : Q_l \to Q_1) \to Q_l \to Q_1 \to 0, \tag{8-22}$$

$$Q_{l-1} \xrightarrow{V} \ker(R : Q_l \to Q_1) \to T_l \to 0, \tag{8-23}$$

where T_l is simply defined to be the cokernel. We claim

$$H^{n-1}(\tilde{Y}, T_l) = 0 = H^n(\tilde{Y}, T_l) \tag{8-24}$$

for all $l \leq m$. The surjectivity of (8-20) yields the surjectivity of the composition

$$\ker\bigl(W_l\Omega^1_{\tilde{Y}/k}(\log E) \xrightarrow{R} \Omega^1_{\tilde{Y}}(\log E)\bigr)/VW_{l-1}\Omega^1_{\tilde{Y}/k}(\log E) \xrightarrow[\cong]{F^{l-1}} B_{l-1}\Omega^1_{\tilde{Y}} \to T_l \tag{8-25}$$

[Illusie 1979, p. 575]. Note that

$$\ker(W_l\Omega^1_{\tilde{Y}/k} \xrightarrow{R} \Omega^1_{\tilde{Y}})/VW_{l-1}\Omega^1_{\tilde{Y}/k} \xrightarrow{\cong} \ker\bigl(W_l\Omega^1_{\tilde{Y}/k}(\log E) \xrightarrow{R} \Omega^1_{\tilde{Y}}(\log E)\bigr)/VW_{l-1}\Omega^1_{\tilde{Y}/k}(\log E)$$

is an isomorphism.

Now we need to find a complex of $W_l(\mathcal{O}_{\tilde{Y}})$-modules

$$R_1 \to R_0 \to \ker(B_{l-1}\Omega^1_{\tilde{Y}} \to T_l),$$

such that the following conditions hold:

- $R_{0|Y_{\mathrm{sm}}} \to \ker(B_{l-1}\Omega^1_{\tilde{Y}} \to T_l)|_{Y_{\mathrm{sm}}}$ is surjective and
- $H^n(\tilde{Y}, R_1) \to H^n(\tilde{Y}, R_0)$ is surjective.

It will follow that $H^n(\tilde{Y}, T_l) = 0 = H^{n-1}(\tilde{Y}, T_l)$. Indeed, we have

$$H^n(\tilde{Y}, B_{l-1}\Omega^1_{\tilde{Y}}) = 0 = H^{n-1}(\tilde{Y}, B_{l-1}\Omega^1_{\tilde{Y}})$$

by induction on l, and using the exact sequence (8-26). The case $l = 2$ follows from assumptions (4) and (5), Lemma 8.9, and the short exact sequence (8-27).

We take

$$R_{0,l} := r^*\pi^* B_{l-1}\Omega^1_X, \qquad R_{1,l} = \ker(R_{0,l} \to B_{l-1}\Omega^1_{\tilde{Y}}).$$

Clearly, the image of $r^*\pi^* B_{l-1}\Omega^1_X$ is contained in $\ker(B_{l-1}\Omega^1_{\tilde{Y}} \to T_l)$. The surjectivity of $R_{0|Y_{\mathrm{sm}}} \to \ker(B_{l-1}\Omega^1_{\tilde{Y}} \to T_l)|_{Y_{\mathrm{sm}}}$ follows from Lemma 8.11 and diagram (8-14).

We claim that $H^n(\tilde{Y}, R_{1,l}) \to H^n(\tilde{Y}, R_{0,l})$ is surjective. We will proceed by induction on l. We have an exact sequence of locally free $\mathcal{O}_X$-modules

$$0 \to \mathrm{Frob}_*^{l-2} B_1\Omega^1_X \to B_{l-1}\Omega^1_X \xrightarrow{C} B_{l-2}\Omega^1_X \to 0, \tag{8-26}$$

where C is the Cartier operator. Therefore,

$$0 \to r^*\pi^*\mathrm{Frob}_*^{l-2} B_1\Omega^1_X \to R_{0,l} \xrightarrow{C} R_{0,l-1} \to 0$$

is exact. Lemma 8.12 shows that $R_{1,l|Y_{\mathrm{sm}}} \xrightarrow{C} R_{1,l-1|Y_{\mathrm{sm}}}$ is surjective; note that under the isomorphism $F^{l-2}d$ from (8-16) the Cartier operator corresponds to the restriction. By induction we need to prove that the image of

$$H^n(\tilde{Y}, r^*\pi^*\mathrm{Frob}_*^{l-2} B_1\Omega_X^1) \to H^n(\tilde{Y}, R_{0,l})$$

is contained in the image of $H^n(\tilde{Y}, R_{1,l})$. Rationality of the resolution r implies

$$\begin{aligned} H^n(\tilde{Y}, r^*\pi^*\mathrm{Frob}_*^{l-2} B_1\Omega_X^1) &= H^n(Y, \pi^*\mathrm{Frob}_*^{l-2} B_1\Omega_X^1) \\ &= H^n(X, \mathrm{Frob}_*^{l-2}(B_1\Omega_X^1) \otimes_{\mathcal{O}_X} \pi_*\mathcal{O}_Y). \end{aligned}$$

In view of the exact sequence

$$0 \to \mathcal{O}_X \xrightarrow{\mathrm{Frob}} \mathrm{Frob}_*\mathcal{O}_X \to B_1\Omega_X^1 \to 0, \tag{8-27}$$

we obtain a surjective map

$$H^n\left(X, \bigoplus_{i=p^{m-l+1}}^{p^m-1} L^{-i\cdot p^{l-1}}\right) \to H^n(\tilde{Y}, r^*\pi^*\mathrm{Frob}_*^{l-2} B_1\Omega_X^1),$$

because

$$\mathrm{Frob}_*^{l-1}\mathcal{O}_X \otimes_{\mathcal{O}_X} \pi_*\mathcal{O}_Y = \bigoplus_{i=0}^{p^m-1} \mathrm{Frob}_*^{l-1}(\mathrm{Frob}^{l-1,*} L^{-i}).$$

For every $p^m > i \geq p^{m-l+1}$, we have two morphisms $\mathrm{Frob}_*^{l-1}(\mathrm{Frob}^{l-1,*}(L^{-i})) \to \mathrm{Frob}_*^{l-1}(\mathrm{Frob}^{l-1,*}(\pi_*\mathcal{O}_Y))$; the first one is induced by $\mathrm{Frob}_*^{l-1}\mathrm{Frob}^{l-1,*}$ applied to $L^{-i} \subset \pi_*\mathcal{O}_Y$. The second one is induced by Frob_*^{l-1} applied to

$$\begin{aligned} \mathrm{Frob}^{l-1,*}(L^{-i}) = L^{-i\cdot p^{l-1}} &\xrightarrow{s^{(i:p^{m+1-l})}} L^{-(i\% p^{m+1-l})\cdot p^{l-1}} \\ &= \mathrm{Frob}^{l-1,*}(L^{-(i\% p^{m+1-l})}) \\ &\to \mathrm{Frob}^{l-1,*}(\pi_*\mathcal{O}_Y), \end{aligned}$$

where the last arrow comes from $L^{-(i\% p^{m+1-l})} \subset \pi_*\mathcal{O}_Y$. Note that after application of $H^n(X, \cdot)$ this map vanishes, because it factors over

$$H^n(X, L^{-(i\% p^{m+1-l})\cdot p^{l-1}}) = 0.$$

Subtracting the two maps yields a morphism

$$r^*\pi^*\mathrm{Frob}_*^{l-1}(\mathrm{Frob}^{l-1,*}(L^{-i})) \to (R_{1,l} \cap r^*\pi^*\mathrm{Frob}_*^{l-2} B_1\Omega_X^1)$$

which shows that the $H^n(X, L^{-i\cdot p^{l-1}})$ piece of $H^n(\tilde{Y}, r^*\pi^*\mathrm{Frob}_*^{l-2} B_1\Omega_X^1)$ is contained in the image of $H^n(\tilde{Y}, R_{1,l})$. This proves claim (8-24).

In view of the short exact sequences (8-22) and (8-23), Corollary 8.16, vanishing of $H^n(\tilde{Y}, \mathcal{O}_{\tilde{Y}}/\mathcal{O}_{\tilde{Y}}^{p^{l-1}})$, and (8-24), we obtain, for all $l \leq m$, a short exact sequence

$$0 \to H^n(\tilde{Y}, Q_{l-1}) \xrightarrow{V} H^n(\tilde{Y}, Q_l) \xrightarrow{R} H^n(\tilde{Y}, Q_1) \to 0. \tag{8-28}$$

This enables us to define

$$\psi_{l-1} : H^n(\tilde{Y}, Q_1) \to H^n(\tilde{Y}, Q_1), \quad a \mapsto F^{l-1}(R^{-1}(a)).$$

It is evident that $\psi_{l-1} = \psi_1^{l-1}$. In view of (8-21) we have

$$H^n(\tilde{Y}, Q_1) \cong H^n(X, L^{-p^m}).$$

Via this identification, the map ψ_1 is given by

$$H^n(X, L^{-p^m}) \to H^n(X, L^{-p^{m+1}}) \xrightarrow{\cdot s^{p-1}} H^n(X, L^{-p^m}),$$

where the first arrow is induced by the p-th power map $L^{-p^m} \to L^{-p^{m+1}}$, $a \mapsto a^p$. Indeed, denoting by $\iota : L^{-p^m} \to \pi_* r_* Q_1$ the evident map, we have a commutative diagram

$$\begin{tikzcd}
L^{-p^m} \arrow[r, "(\cdot)^p"] \arrow[d, "\iota"] & L^{-p^{m+1}} \arrow[r, "s^{p-1}"] & L^{-p^m} \arrow[d, "\iota"] \\
\pi_* r_* Q_1 \arrow[r, "\pi_* r_*(F \circ R^{-1})"] & \pi_* r_* \left(\dfrac{Q_1}{\mathrm{image}(B_1 \Omega^1_{\tilde{Y}})} \right) & \dfrac{\pi_* r_* Q_1}{\mathrm{image}(\pi_* r_* B_1 \Omega^1_{\tilde{Y}})} \arrow[l]
\end{tikzcd}$$

Moreover, ψ_1 equals the composition

$$H^n(\tilde{Y}, Q_1) \xrightarrow{=} H^n(X, \pi_* r_* Q_1) \xrightarrow{\pi_* r_*(F \circ R^{-1})} H^n(X, \pi_* r_*(Q_1/\,\mathrm{image}(B_1 \Omega^1_{\tilde{Y}})))$$
$$\to H^n(\tilde{Y}, Q_1/\,\mathrm{image}(B_1 \Omega^1_{\tilde{Y}})) \xrightarrow{\cong} H^n(\tilde{Y}, Q_1),$$

where the last morphism is the inverse of $H^n(\tilde{Y}, Q_1) \xrightarrow{\eta} H^n(\tilde{Y}, Q_1/\,\mathrm{image}(B_1 \Omega^1_{\tilde{Y}}))$, which is injective because $H^n(\tilde{Y}, Q_1) \xrightarrow{V} H^n(\tilde{Y}, Q_2)$ factors through η.

In the notation of [Chatzistamatiou 2012, Definition 1.3.1], we therefore get

$$H^n_c(X \setminus V(s), \mathbb{O})_s \cong \bigcap_{i \geq 1} \mathrm{image}(\psi_1^i).$$

Since $H^{n-1}(X, \mathbb{O}_X) = 0 = H^n(X, \mathbb{O}_X)$, [Chatzistamatiou 2012, §1.4] implies

$$H^n_c(X \setminus V(s), \mathbb{O})_s \cong H^{n-1}(V(s), \mathbb{O})_s = \bigcap_{i \geq 1} \mathrm{image}(\mathrm{Frob}^i).$$

By using assumption (3), we obtain

$$H^n(\tilde{Y}, Q_l) \cong \bigoplus_{i=1}^{h} W(k)/p^l, \tag{8-29}$$

where $h = \dim_k H^n(X, L^{-p^m})$. Indeed, since the Frobenius acts bijectively on $H^{n-1}(V(s), \mathbb{O}) \cong H^n(X, L^{-p^m})$, ψ_1 is bijective on $H^n(\tilde{Y}, Q_1)$. In view of (8-28), any lifting of a basis of $H^n(\tilde{Y}, Q_1)$ via the map $R : H^n(\tilde{Y}, Q_l) \to H^n(\tilde{Y}, Q_1)$ will be a $W(k)/p^l$-basis of $H^n(\tilde{Y}, Q_l)$.

Finally, let us show that $W_l(k) \subset H^0(\tilde{Y}, W_l\Omega^{n-1}_{\tilde{Y}})$. In view of (8-29), there is a surjective morphism of $W(k)$-modules

$$H^n(\tilde{Y}, W_l\Omega^1_{\tilde{Y}/k}(\log E)) \to W(k)/p^l = W_l(k).$$

From the residue short exact sequence (8-18) we obtain a surjective map

$$H^n(\tilde{Y}, W_l\Omega^1_{\tilde{Y}/k}) \to W_l(k).$$

Ekedahl duality [1984] implies

$$R\Gamma(W_l\Omega^{n-1}_{\tilde{Y}}) \xrightarrow{\cong} R\operatorname{Hom}_{W_l(k)}(R\Gamma(W_l\Omega^1_{\tilde{Y}}), W_l(k)[-n]),$$

hence the claim. □

Remark 8.18. Even for the case $m = 1$ the approach is dual to the one in [Kollár 1995]. With the notation in the proof of Theorem 8.17, we show that the composition

$$H^n(\tilde{Y}, \Omega^1_{\tilde{Y}}) \to H^n(\tilde{Y}, Q_1) \xrightarrow{\cong} H^n(X, L^{\otimes -p^m})$$

is surjective. For the last isomorphism we use $n \geq 3$, because we need to use Lemma 8.10, where vanishing holds for $i > 1$ only. Since we don't use Lemma 8.14 for this part, the argument also works for $p = 2$ and n odd. Taking duals we obtain an inclusion

$$H^0(X, \omega_X \otimes L^{\otimes p^m}) \subset H^0(\tilde{Y}, \Omega^{n-1}_{\tilde{Y}}).$$

This corresponds to a result about extending $(n-1)$-forms from Y_{sm} to $\tilde{Y}$ in [Kollár 1995] (and [Colliot-Thélène and Pirutka 2016a; Okada 2016]).

Acknowledgments

We would like to thank the referees very much for thoroughly reading the paper and for their comments and suggestions, which we think have greatly improved the paper. We are grateful to Kay Rülling, who suggested the statement and proof of Lemma 8.1. This result enabled us to improve an earlier version of our Theorem 8.2 to the statement on higher level torsion orders mentioned above. We are also grateful to Jean-Louis Colliot-Thélène, who very kindly allowed us to include some of the results of his paper [Colliot-Thélène 2016]. This led to a new result (Lemma 1.8) on specialization of decompositions of the diagonal, derived from [Colliot-Thélène 2016, Lemme 2.2], as well as Example 6.9 mentioned above, a version of which appears as [Colliot-Thélène 2016, Théorème 2.4].

References

[Abhyankar 1966] S. S. Abhyankar, *Resolution of singularities of embedded algebraic surfaces*, Pure Appl. Math. **24**, Academic, 1966. MR Zbl

[Bloch and Srinivas 1983] S. Bloch and V. Srinivas, "Remarks on correspondences and algebraic cycles", *Amer. J. Math.* **105**:5 (1983), 1235–1253. MR Zbl

[Chatzistamatiou 2012] A. Chatzistamatiou, "On the Frobenius stable part of Witt vector cohomology", *Math. Ann.* **354**:3 (2012), 1177–1200. MR Zbl

[Chatzistamatiou and Rülling 2011] A. Chatzistamatiou and K. Rülling, "Higher direct images of the structure sheaf in positive characteristic", *Algebra Number Theory* **5**:6 (2011), 693–775. MR Zbl

[Chatzistamatiou and Rülling 2012] A. Chatzistamatiou and K. Rülling, "Hodge–Witt cohomology and Witt-rational singularities", *Doc. Math.* **17** (2012), 663–781. MR Zbl

[Colliot-Thélène 2016] J.-L. Colliot-Thélène, "Non rationalité stable d'hypersurfaces cubiques sur des corps non algébriquement clos", preprint, 2016. To appear in *Proceedings of the International Colloquium on K-Theory* (Mumbai, 2016). arXiv

[Colliot-Thélène and Pirutka 2016a] J.-L. Colliot-Thélène and A. Pirutka, "Cyclic covers that are not stably rational", *Izv. Math.* **80**:4 (2016), 665–677. MR Zbl

[Colliot-Thélène and Pirutka 2016b] J.-L. Colliot-Thélène and A. Pirutka, "Hypersurfaces quartiques de dimension 3: non-rationalité stable", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:2 (2016), 371–397. 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

[Ekedahl 1984] T. Ekedahl, "On the multiplicative properties of the de Rham–Witt complex, I", *Ark. Mat.* **22**:2 (1984), 185–239. MR Zbl

[El Zein 1978] F. El Zein, *Complexe dualisant et applications à la classe fondamentale d'un cycle*, Mém. Soc. Math. France **58**, Société Mathématique de France, 1978. MR Zbl

[Fulton 1984] W. Fulton, *Intersection theory*, Ergebnisse Math. Grenzgebiete (3) **2**, Springer, 1984. MR Zbl

[Gros 1985] M. Gros, *Classes de Chern et classes de cycles en cohomologie de Hodge–Witt logarithmique*, Mém. Soc. Math. France (N.S.) **21**, Société Mathématique de France, 1985. MR Zbl

[Hassett and Tschinkel 2016] B. Hassett and Y. Tschinkel, "On stable rationality of Fano threefolds and del Pezzo fibrations", preprint, 2016. arXiv

[Hassett et al. 2016] B. Hassett, A. Pirutka, and Y. Tschinkel, "Stable rationality of quadric surface bundles over surfaces", preprint, 2016. arXiv

[Illusie 1979] L. Illusie, "Complexe de de Rham–Witt et cohomologie cristalline", *Ann. Sci. École Norm. Sup.* (4) **12**:4 (1979), 501–661. MR Zbl

[Illusie 1990] L. Illusie, "Ordinarité des intersections complètes générales", pp. 376–405 in *The Grothendieck Festschrift*, vol. II, edited by P. Cartier et al., Progr. Math. **87**, Birkhäuser, 1990. MR Zbl

[Kahn 2016] B. Kahn, "Torsion order of smooth projective surfaces", preprint, 2016. arXiv

[Kollár 1995] J. Kollár, "Nonrational hypersurfaces", *J. Amer. Math. Soc.* **8**:1 (1995), 241–249. MR Zbl

[Lam 1980] T. Y. Lam, *The algebraic theory of quadratic forms*, 2nd ed., Math. Lecture Note Series **54**, Benjamin/Cummings, 1980. MR Zbl

[Matsumura 1980] H. Matsumura, *Commutative algebra*, 2nd ed., Math. Lecture Note Series **56**, Benjamin/Cummings, 1980. MR Zbl

[Okada 2016] T. Okada, "Stable rationality of cyclic covers of projective spaces", preprint, 2016. arXiv

[Roitman 1980] A. A. Roitman, "Rational equivalence of 0-cycles", *Math. Notes Acad. Sci. USSR* **28**:1 (1980), 507–510. MR Zbl

[Totaro 2016] B. Totaro, "Hypersurfaces that are not stably rational", *J. Amer. Math. Soc.* **29**:3 (2016), 883–891. MR Zbl

[Voisin 2015] C. Voisin, "Unirational threefolds with no universal codimension 2 cycle", *Invent. Math.* **201**:1 (2015), 207–237. MR Zbl

[Zariski and Samuel 1975] O. Zariski and P. Samuel, *Commutative algebra*, vol. I, 2nd ed., Grad. Texts Math. **28**, Springer, 1975. MR Zbl

Communicated by Jean-Louis Colliot-Thélène
Received 2016-05-23 Revised 2017-06-14 Accepted 2017-07-29

a.chatzistamatiou@uni-due.de *Fakultät Mathematik, Universität Duisburg-Essen, Essen, Germany*

marc.levine@uni-due.de *Fakultät Mathematik, Universität Duisburg-Essen, Essen, Germany*

ALGEBRA AND NUMBER THEORY 11:8 (2017)
dx.doi.org/10.2140/ant.2017.11.1837

Integral canonical models for automorphic vector bundles of abelian type

Tom Lovering

We define and construct integral canonical models for automorphic vector bundles over Shimura varieties of abelian type.

More precisely, we first build on Kisin's work to construct integral canonical models over $\mathbb{O}_E[1/N]$ for Shimura varieties of abelian type with hyperspecial level at all primes not dividing N compatible with Kisin's construction. We then define a notion of an integral canonical model for the standard principal bundles lying over Shimura varieties and proceed to construct them in the abelian type case. With these in hand, one immediately also gets integral models for automorphic vector bundles.

MSC2010: primary 11G18; secondary 14G35.
Keywords: Shimura varieties, automorphic vector bundles, integral models, Shimura, abelian type.

1. Introduction

Since the introduction of the abstract theory of Shimura varieties and their canonical models by Deligne [1971; 1979] following Shimura, and given its promise as a rich generalisation of the classical theory of modular curves, substantial bodies of literature have arisen whose aim is to extend features of the classical theory to this wider context.

One such feature is the existence of smooth integral models at primes not dividing the level, and their subsequent utility for studying the action of Frobenius on the Galois representations arising from Shimura varieties crucial for the Langlands programme. Such results have been available in the PEL type case for a long time, thanks largely to the programme of Kottwitz, but have recently been extended to the much more general case of abelian type Shimura varieties in work culminating with the recent papers of Kisin [2010; 2017].

Another feature is the manifestation of certain automorphic forms as algebraic sections of a vector bundle over a Shimura variety, generalising the classical algebraic description of modular forms. The vector bundles playing the role analogous to the tensor powers of the Hodge bundle in the theory of modular forms are the automorphic vector bundles, and canonical models were defined and shown to exist in some cases by Harris [1985] and more generally by Milne [1988; 1990].

In the present paper we start to draw these two threads of the literature together, working in the case of general[1] abelian type Shimura varieties, first filling a gap in the existing literature and showing that Kisin's good integral models can be spread out to smooth models over $\mathcal{O}_E[1/N]$, then defining a notion of integral canonical models for automorphic vector bundles with a uniqueness property, and finally proving existence in the abelian type case.

More precisely, suppose $(G, \mathfrak{X})$ is a Shimura datum, with reflex field $E = E(G, \mathfrak{X})$, $N > 1$ and suppose G admits a reductive[2] model $\mathcal{G}/\mathbb{Z}[1/N]$. Take an open compact $K = K^N K_N \subset G(\mathbb{A}^\infty)$ where $K^N = \prod_{p\nmid N} \mathcal{G}(\mathbb{Z}_p)$ and $K_N \subset \prod_{p|N} G(\mathbb{Q}_p)$ is open compact. Consider the tower

$$\mathrm{Sh}_{K^N}(G, \mathfrak{X}) := \varprojlim_{K_N} \mathrm{Sh}_{K^N K_N}(G, \mathfrak{X})$$

of quasiprojective E-schemes, which comes equipped with an algebraic $\prod_{p|N} G(\mathbb{Q}_p)$ action (in fact it carries the action of a slightly larger group as in [Deligne 1979]). The main result of Kisin's first integral models paper [2010] tells us that when $(G, \mathfrak{X})$ is of abelian type this tower admits, for every $v \nmid N$, a smooth integral model

[1] We do require a small technical restriction: that $Z(G)^\circ$ is split by a CM field, but feel it should be possible to remove this restriction, and that it ought to be harmless for most applications.

[2] Recall that a (connected) reductive group scheme $G \to S$ is a smooth affine group scheme with connected reductive geometric fibres.

$\mathcal{S}^{\mathrm{Kis},v}_{K^N}/\mathcal{O}_{E,v}$ to which the $\prod_{p|N} G(\mathbb{Q}_p)$-action extends. Note that by a "smooth" model, we mean one which is smooth quasiprojective at any finite level and for which the maps between different finite levels are finite étale.

It is natural to ask whether these all come from a global model over $\mathcal{O}_E[1/N]$. Moreover, Kisin's models also enjoy an "extension property" which characterises them uniquely, so it is natural to ask if furthermore we can find a global model having this extension property. Our first theorem answers this in the affirmative.

Theorem. *With the above setup, suppose* $(G, \mathfrak{X})$ *is of abelian type. Then the tower* $\mathrm{Sh}_{K^N}(G, \mathfrak{X})$ *admits a smooth integral model* $\mathcal{S}_{K^N}(G, \mathfrak{X})/\mathcal{O}_E[1/N]$ *to which the* $\prod_{p|N} G(\mathbb{Q}_p)$*-action extends and having the extension property.*

Moreover, for any $v \nmid N$ *a place of* E, *the model* $\mathcal{S}_{K^N}(G, \mathfrak{X}) \otimes_{\mathcal{O}_E[1/N]} \mathcal{O}_{E,v}$ *is canonically identified*[3] *with the model* $\mathcal{S}^{\mathrm{Kis},v}_{K^N}$ *obtained from Kisin's theory.*

We call these models "integral canonical models" because the extension property guarantees their uniqueness.

We make some remarks about the proof. The obvious direct "patching" argument to obtain the result formally from Kisin's models fails because one cannot automatically get the extension property for the models obtained by spreading out, so instead we give a direct construction. The proof in the Hodge type case very closely follows that of [Kisin 2010]. In the abelian type case we need several new ideas.

Firstly, we have no guarantee that $\prod_{p|N} G(\mathbb{Q}_p)$ acts transitively on the components of $\mathrm{Sh}_{K^N}(G, \mathfrak{X})$ so we replace the group theoretic "Deligne-induction" argument by an argument that uses the extension property to reduce the problem to constructing models for each component individually over the ring of integers of the maximal abelian unramified extension of E. Secondly, to prove the analogue of [Kisin 2010, 3.4.6] without the lemma which follows it, which may not be true in our context, we show that Kisin's "twisting abelian varieties" construction can be carried out using torsors for the centre of G^{der} rather than G, and since such torsors are finite, they are simpler to work with. Finally, we believe in fact that unless $G^{\mathrm{der}} = G^{\mathrm{ad}}$, the group $\Delta(G, G^{\mathrm{ad}})$ [ibid.] is not finite, which is necessary to descend the extension property. Fortunately we are able to prove that the corresponding group at level K^N (our situation) is finite, which perhaps also helps to fill a small gap in the original argument.

We then turn to automorphic vector bundles, and following [Milne 1990, III], the heart of the matter is really to define and construct integral canonical models for the standard principal bundles $P_{K^N}(G, \mathfrak{X}) \to \mathrm{Sh}_{K^N}(G, \mathfrak{X})$. Such bundles are torsors for the quotient

$$G^c = G/Z_{\mathrm{nc}}$$

[3] This identification is as an $\mathcal{O}_{E,v}$-scheme with $\prod_{p|N} G(\mathbb{Q}_p)$-action and equivariant identification ι of its generic fibre with $\mathrm{Sh}_{K^N}(G, \mathfrak{X})$.

by the maximal subtorus $Z_{nc} \subset Z(G)^\circ$ which is $\mathbb{R}$-split but $\mathbb{Q}$-anisotropic. They come equipped with a flat connection ∇, an equivariant $\prod_{p|N} G(\mathbb{Q}_p)$-action and a "filtration" which can be written down by giving a map

$$\gamma : P_{K^N}(G, \mathfrak{X}) \to \mathrm{Gr}_\mu$$

where Gr_μ is the flag variety corresponding to a Hodge cocharacter $\mu : \mathbb{G}_m \to G$ coming from $\mathfrak{X}$. Informally, it is helpful to think of the fibre functors ω attached to these bundles as giving one the sheaf of de Rham cohomology of a family of motives over Sh_{K^N} with its Gauss–Manin connection and Hodge filtration.

Any smooth integral model for $P_{K^N}(G, \mathfrak{X})$ that is a $\mathcal{G}^c$-torsor would give, for each representation $\rho : \mathcal{G}^c \to \mathrm{GL}(V_{\mathbb{Z}[1/N]})$, a $\mathbb{Z}[1/N]$ lattice inside $\omega(V \otimes \mathbb{Q})$. We define an integral canonical model to be one where at "crystalline points" this lattice coincides with a different lattice constructed using Kisin's theory of $\mathfrak{S}$-modules [2006]. Roughly speaking, working on the generic fibre one has a Galois cover $\mathrm{Sh}_{K^{N_p}}(G, \mathfrak{X}) \to \mathrm{Sh}_{K^N}(G, \mathfrak{X})$ with Galois group $\mathcal{G}^c(\mathbb{Z}_p)$ so we obtain attached to ρ a lisse $\mathbb{Z}_p$-sheaf

$$\mathcal{L} := \mathrm{Sh}_{K^{N_p}}(G, \mathfrak{X}) \times V_{\mathbb{Z}_p}/\mathcal{G}^c(\mathbb{Z}_p)$$

on $\mathrm{Sh}_{K^N}(G, \mathfrak{X})$. Restricting this to a crystalline point s, the theory of $\mathfrak{S}$-modules gives a lattice $\mathbb{D}(s^*\mathcal{L}) \subset D_{\mathrm{dR}}(s^*\mathcal{L}[1/p])$. Thus we have defined lattices (at each prime) inside $\omega(V \otimes \mathbb{Q})$, and we say a model

$$(\mathcal{P}_{K^N}(G, \mathfrak{X}), \iota : \mathcal{P}_{K^N}(G, \mathfrak{X}) \otimes_{\mathbb{O}_E[1/N]} E \xrightarrow{\sim} P_{K^N}(G, \mathfrak{X}))$$

is canonical if the lattices it generates agree with these lattices coming from p-adic Hodge theory. We then check that if such a model exists and there are enough crystalline points (which we check in the abelian type case), it is unique up to canonical isomorphism. We also reserve the term "canonical" for models for which the connection, Hecke action and filtration extend, but these seem to be automatic properties of the models satisfying the lattice condition in the abelian type case (and we would assume in general).

Of course, with this definition, our main theorem is the following.

Theorem. *Let $(G, \mathfrak{X})$ be a Shimura datum of abelian type and $\mathcal{G}/\mathbb{Z}[1/N]$ a reductive model for G. Then $P_{K^N}(G, \mathfrak{X})$ has an integral canonical model $(\mathcal{P}_{K^N}, \iota)$.*

We give a quick summary of the proof. For the "special type" case where G is a torus, we use the theory of CM motives to find lattices in the de Rham cohomology of abelian varieties. For the Hodge type case, the universal abelian variety has an integral model so we take as our starting point the sheaf $\mathcal{V}$ of its relative de Rham cohomology, and note that the Hodge tensors of [Kisin 2010, 2.2] $s_{\alpha,\mathrm{dR}} \in \mathcal{V}_E^{\otimes}$ extend to $\mathcal{V}^{\otimes}$, at which point they may be used to define a functor

$$\mathcal{P}_{K^N} := \underline{\mathrm{Isom}}_{s_\alpha}(V, \mathcal{V})$$

which we show is a $\mathcal{G}$-torsor using several ingredients from [Kisin 2010], and is the required integral canonical model.

For passing from the Hodge type to abelian type case, we need a new idea which may be more widely applicable. Suppose $(G_2, \mathfrak{X}_2)$ is an abelian type Shimura datum of interest. If we let $(G, \mathfrak{X})$ be a Hodge type datum such that there is an isogeny $G^{\mathrm{der}} \to G_2^{\mathrm{der}}$ witnessing that $(G_2, \mathfrak{X}_2)$ is of abelian type, we do not have a map relating G_2 to G but only between the derived groups. However, the torsors $P_{K^N}(G, \mathfrak{X})$ cannot be reduced to G^{der} without passing to $\mathbb{C}$, which loses the information we are interested in.

Our solution, inspired by Deligne [1979, 2.5], is to define for each connected Shimura datum $(G^{\mathrm{der}}, \mathfrak{X}^+)$ and field $E \supset E(G^{\mathrm{der}}, \mathfrak{X}^+) := E(G^{\mathrm{ad}}, \mathfrak{X}^{\mathrm{ad}})$ a new Shimura datum $(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})$ with the property that any Shimura datum $(G, \mathfrak{X})$ whose reflex field is contained in E and whose connected Shimura datum is $(G^{\mathrm{der}}, \mathfrak{X}^+)$ admits a canonical map

$$(\mathcal{B}, \mathfrak{X}_{\mathcal{B}}) \to (G, \mathfrak{X}).$$

We then pass from Hodge type to abelian type by first giving a direct construction of a canonical model for $P_{K^N_{\mathcal{B}}}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})$ given one for $P_{K^N}(G, \mathfrak{X})$. With this, we are in business, because if we let $(\mathcal{B}_2, \mathfrak{X}_{\mathcal{B},2})$ be the corresponding pair for $(G_2, \mathfrak{X}_2)$ there is a map

$$(\mathcal{B}, \mathfrak{X}_{\mathcal{B}}) \to (\mathcal{B}_2, \mathfrak{X}_{\mathcal{B},2}) \to (G_2, \mathfrak{X}_2)$$

and we are able to descend the torsor through the map on connected components while pushing out the $\mathcal{B}^c$-action to a G_2^c-action. Finally we translate our results into results for automorphic vector bundles.

While we do not give applications in this paper, we anticipate this construction playing a useful role in several places. It is already used to study integrality of periods in a preprint of Ichino and Prasanna [2016], and we expect it should be useful in much more general contexts along these lines.

Working at a formal completion at a single place the same construction gives families of strongly divisible filtered F-crystals. This theory is developed and used in [Lovering 2017] to establish a new result that the Galois representations formed by taking the cohomology of the Shimura variety with coefficients in the usual lisse sheaves are crystalline, and in appropriate situations Fontaine Laffaille. In particular this gives a new proof of part of local global compatibility at $l = p$ in certain cases, as well as providing a new tool for studying the p-adic geometry of Shimura varieties and p-adic automorphic forms.

2. Integral canonical models for Shimura varieties of abelian type

2.1. *Extension property.*

2.1.1. We first recall the extension property used to characterise Shimura varieties at infinite level. Let R be a domain with field of fractions K. We call S/R a *test scheme* if it is regular and formally smooth over R.[4] Suppose we are given a scheme X/R. We say X *has the extension property* if for any test scheme S/R any map $S_K \to X_K$ extends over R.

The following uniqueness statement is well known.

Lemma 2.1.2. *Suppose Y/K is a scheme. Then if X/R is a model for it over R and is both a test scheme and has the extension property, X is the unique such model up to canonical isomorphism.*

Proof. Let X and X' be two such models. Then we are (as part of the data of a model) given maps

$$X_K \xrightarrow{\sim} Y \xrightarrow{\sim} X'_K.$$

Since X is a test scheme, and since X' has the extension property, this isomorphism extends to an isomorphism $X \xrightarrow{\sim} X'$ of R-schemes. □

We also record the following useful formal properties.

Lemma 2.1.3. *Let R'/R be an étale or indétale (or formally smooth) extension of domains with fraction field extension K'/K. Suppose X/R satisfies the extension property. Then so does $X \otimes_R R'/R'$.*

Conversely, if R'/R is also faithfully flat, then if $X \otimes_R R'/R'$ satisfies the extension property, so does X/R.

Proof. Let S'/R' be a test scheme. Then S' is regular, and it is formally smooth over R since indétale algebras are formally étale. Suppose we are given $S' \otimes_{R'} K' \to X_{R'}$, and first compose it with the map $X_{R'} \to X$. By the extension property for X/R, $S' \otimes_{R'} K' = S' \otimes_R K \to X$ extends to a map $S' \to X$ over R. But since we have a diagram

$$\begin{array}{ccc} S' & \longrightarrow & X \\ \downarrow & & \downarrow \\ \operatorname{Spec} R' & \longrightarrow & \operatorname{Spec} R \end{array}$$

this map must factor through $X \otimes_R R'$, as required, giving the extension property for $X \otimes_R R'/R'$.

For the converse, a test scheme S/R gives rise to a test scheme $S_{R'}/R'$ together with a descent datum θ_S for R'/R. Given a map $S_K \to X_K$ we obtain $S_{K'} \to X_{K'}$

[4] I believe there is still some controversy over the "correct" definition of a test scheme but this should do for our purposes.

compatible with the descent data on both sides. By the extension property this map extends to $S_{R'} \to X_{R'}$, and being a map of descent data this descends to the desired extension $S \to X$. $\square$

Lemma 2.1.4. *Let X/R have the extension property and $Y \to X$ be finite or profinite étale. Then Y/R has the extension property.*

Proof. It clearly suffices to do the finite étale case (the profinite one then following formally). Let $K = \mathrm{Frac}(R)$ and suppose we have a test scheme S and a map $S_K \to Y_K$. Then this map composed with $Y_K \to X_K$ extends to a map $S \to X$. Pulling this back we get $S \times_X Y \to Y$. The map $S_K \to Y_K$ gives a section of $(S \times_X Y)_K \to S_K$. But $S \times_X Y \to S$ is finite étale so such sections extend uniquely and we get the required $S \to Y$. $\square$

Lemma 2.1.5. *Let $Y \to X$ be finite étale, and suppose Y/R has the extension property. Then X/R has the extension property.*

Proof. By the theory of the étale fundamental group and Lemma 2.1.4 we may assume $Y \to X$ is Galois. The argument then follows from that of [Moonen 1998, 3.21.4]. $\square$

We remark that the above result fails for pro(finite étale) extensions. Indeed, Shimura varieties at finite level certainly do not generally have the extension property. For example, for the modular curve at finite level this would imply all elliptic curves have good reduction.

2.2. ***Main theorem.*** Now, let $(G, \mathfrak{X})$ be a Shimura datum of abelian type with reflex field E.

Fix S a finite nonempty set of finite primes containing all those at which G is ramified, set $N = \prod_{p \in S} p$, a reductive integral model $G_{\mathbb{Z}[1/N]}$ of G (which exists by taking an arbitrary integral model, observing that it is reductive at all but finitely many primes and then gluing in models for the remaining primes). We abusively denote this model by G, and let $K^N = \prod_{p \notin S} G(\mathbb{Z}_p)$.

Consider the tower

$$\mathrm{Sh}_{K^N} = \varprojlim_{K_N \subset \prod_{p \in S} G(\mathbb{Q}_p)} \mathrm{Sh}_{K_N K^N}$$

of Shimura varieties over E with infinite level at the primes dividing N but hyperspecial level at all other primes.

Recall that in this context a *smooth integral model* $\mathscr{S}_{K^N}$ for Sh_{K^N} over $\mathbb{O}_E[1/N]$ is an integral model on which $\prod_{p|N} G(\mathbb{Q}_p)$ acts such that whenever $K_N \subset \prod_{p|N} G(\mathbb{Q}_p)$ compact open is sufficiently small that $\mathrm{Sh}_{K_N K^N}$ is a scheme, the model $\mathscr{S}_{K^N}/K_N$ is smooth quasiprojective, and the maps between such finite level schemes induced by Hecke operators are finite étale.

In this chapter we prove the following theorem.

Theorem 2.2.1. *The tower* Sh_{K^N} *has a smooth integral model* $\mathcal{S}_{K^N}/\mathbb{O}_E[1/N]$ *satisfying the extension property. For any* $v \nmid N$ *the localisation of this model at* v *agrees with Kisin's smooth integral model.*

The following together with Lemma 2.1.2 gives us that this model is canonical.

Lemma 2.2.2. *If* $\mathcal{S}_{K^N}$ *is a smooth integral model in the above sense, then it is regular and formally smooth (in the usual sense).*

Proof. That it is formally smooth follows formally because it is a limit of smooth schemes with finite étale transition maps. That it is regular follows from the same argument as [Milne 1992, 2.4]. □

2.2.3. We now set out to prove the theorem, following the outline of Kisin's strategy, first using the modular interpretation of Siegel varieties to get going, then making constructions in the Hodge and abelian cases close enough to his that the smoothness, extension and comparison properties follow by direct comparison or in a very similar way.

2.3. ***Siegel case.*** For this case we recall the following theorem.

Theorem 2.3.1 [Mumford 1965, 7.9]. *If* $n \geq 6^g d\sqrt{g!}$ *then the fine moduli scheme* $\mathcal{A}_{g,d,n}$ *of abelian schemes of dimension* g *together with a polarisation of degree* d *and a level* n *structure exists, and is quasiprojective over* $\mathbb{Z}$.

Moreover, for $N = \mathrm{lcm}(d, n)$ *these moduli schemes are smooth over* $\mathbb{Z}[1/N]$.

2.3.2. We may also (since we may take quotients of quasiprojective schemes by finite free group actions) form moduli $\mathcal{A}_{g,d,K_N}$ with $\mathrm{GSp}_{2g}(\hat{\mathbb{Z}}^N)K_N$-level structures for all $K_N \subset \prod_{p|N} \mathrm{GSp}_{2g}(\mathbb{Q}_p)$ sufficiently small. These are also smooth over $\mathbb{Z}[1/N]$ by étale descent.

2.3.3. Moreover, recall that for $K = K_N\,\mathrm{GSp}_{2g}(\hat{\mathbb{Z}}^N) \subset \mathrm{GSp}_{2g}(\mathbb{A}^\infty)$, the Shimura variety $\mathrm{Sh}_K(\mathrm{GSp}_{2g}, S^\pm)$ is defined over $\mathbb{Q}$ and has a moduli interpretation giving an embedding $\mathrm{Sh}_K(\mathrm{GSp}_{2g}, S^\pm) \hookrightarrow \mathcal{A}_{g,d,K_N}$. We define its integral model

$$\mathcal{S}_K := \overline{\mathrm{Sh}_K(\mathrm{GSp}_{2g}, S^\pm)} \subset \mathcal{A}_{g,d,K_N}.$$

It is well known that such models are smooth and admit an explicit description as moduli schemes. In particular, they carry a universal abelian scheme defined over $\mathbb{Z}[1/N]$. Moreover the transition maps as we vary K_N are finite étale, so we obtain the desired smooth integral models

$$\mathcal{S}_N = \varprojlim_{K_N} \mathcal{S}_K/\mathbb{Z}\big[\tfrac{1}{N}\big].$$

We are required to check the following. Note that in preparation for the Hodge type case we need to observe the following holds in the slightly more general context where we assume that $K = \prod K_p$ as above except for some finitely many $p \nmid N$, K_p is not necessarily hyperspecial but merely maximal compact. Luckily with this remark made the argument goes through unchanged.

Proposition 2.3.4. *The scheme $\mathcal{S}_N$ satisfies the extension property.*

Proof. This follows because the argument of Milne [1992, 2.10] adapts practically unchanged to our situation. For the reader's convenience we sketch the argument. Let S be a test scheme. A map

$$S_{\mathbb{Q}} \to \mathcal{S}_N$$

gives the data of a triple (A, λ, η) where $A/S_{\mathbb{Q}}$ is an abelian scheme, λ a polarisation (defined up to a constant) and η an infinite level structure at the primes $l|N$. Since $N > 1$, the set of such primes is nonempty, and we may let l be one of them.

Now, since S is assumed regular, in particular each of its components is integral. Let S^0 be such a component, and denote by η its generic point. The infinite level structure at l defined over $S_{\mathbb{Q}}$ in particular trivialises the l-adic Tate module of A_η. By the "generalised Neron criterion" [ibid., 2.13] we see that this implies A_η extends over S^0. Since it does so for all components S^0 of S and S is normal (so these components do not meet), we deduce that we have extended A to some $\mathcal{A}/S$.

The polarisation λ also extends by [ibid., 2.14], and the level structures (being at primes away from the characteristic of the base) also obviously extend. This suffices to show the extension property. □

2.4. *Hodge type case.* Suppose we have $(G, \mathfrak{X})$ a Shimura datum of Hodge type, and fix a symplectic embedding

$$i : (G, \mathfrak{X}) \hookrightarrow (\mathrm{GSp}(V_{\mathbb{Q}}, \psi), S^{\pm}).$$

Let N be the product of all primes where G is ramified. Recall that we are fixing an integral model $G/\mathbb{Z}[1/N]$ for G and taking the hyperspecial level $K^N = \prod_{p \nmid N} G(\mathbb{Z}_p)$ away from N. We begin with some group theoretic preliminaries.

Lemma 2.4.1. *Let $V/\mathbb{Q}$ be a finite dimensional vector space. Take $N \geq 1$ and $p \nmid N$, and suppose we have a $\mathbb{Z}[1/pN]$-lattice $\Lambda \subset V$ and a $\mathbb{Z}_p$-lattice $L \subset V \otimes \mathbb{Q}_p$. Then $L \cap \Lambda$ is a $\mathbb{Z}[1/N]$-lattice in V.*

Proof. We first observe that for any $\mathbb{Z}[1/pN]$-basis $e_1, \dots, e_n$ of Λ, for all $i = 1, \dots, n$ and some m sufficiently large $p^m e_i \subset L$. In particular $\Lambda \cap L$ contains $\mathbb{Z}[1/pN]$-bases for Λ. Let us take some such basis $e_1, \dots, e_n$ such that the p-adic volume is maximal (or equivalently such that the index of its $\mathbb{Z}_p$-span in L is minimal). We claim this basis generates $\Lambda' = \Lambda \cap L$ as a free $\mathbb{Z}[1/N]$-module.

Since it's a $\mathbb{Z}[1/pN]$-basis for Λ, its span certainly is a free $\mathbb{Z}[1/N]$-module. Suppose it doesn't generate. Then there is some $y \in \Lambda'$ not in the span of the e_i.

Since the e_i are a $\mathbb{Z}[1/pN]$-basis we may write y uniquely as

$$y = \sum_i \lambda_i e_i,$$

with $\lambda_i \in \mathbb{Z}[1/pN]$. After reordering, let us assume $\lambda_1 = p^{-k}\lambda$ with $\lambda \in \mathbb{Z}[1/N]$ and $k \geq 1$ is the coefficient with the largest p-adic norm amongst the λ_i. Then we can take the new basis $y, e_2, \dots, e_n$ and it visibly has strictly larger p-adic volume, contradicting our original choice of basis and hence the existence of y. $\square$

Proposition 2.4.2. *Let $G/\mathbb{Q}$ be a reductive group unramified away from N, let $N \mid M$ and T be the set of primes dividing M/N. If $2 \in T$, assume further that any factors of G of type B have simply connected derived group.*

Then given choices of reductive models $G^M/\mathbb{Z}[1/M]$ and $G_p/\mathbb{Z}_p$ for $p \in T$ for G, we can find a model $\mathcal{G}/\mathbb{Z}[1/N]$ isomorphic to each of these.

Moreover, if we have a faithful representation $i : G^M \hookrightarrow \mathrm{GL}(V_{\mathbb{Z}[1/M]})$ there is a $\mathbb{Z}[1/N]$-lattice $\Lambda' \subset V_{\mathbb{Z}[1/M]}$ such that i extends to $\tilde{i} : \mathcal{G} \hookrightarrow \mathrm{GL}(\Lambda')$.

Proof. It obviously suffices to consider the case where $M = pN$. Let $i : G^M \hookrightarrow \mathrm{GL}(V_{\mathbb{Z}[1/pN]})$ be a faithful representation. By [Kisin 2010, 2.3.1], and the remark of Madapusi Pera [2012, 4.3, footnote] in the $2 \in T$ case, we can find a $\mathbb{Z}_p$-lattice $\Lambda \subset V \otimes \mathbb{Q}_p$ such that $G_p(\mathbb{Q}_p) = G(\mathbb{Q}_p) = G^M_{\mathbb{Q}_p} \hookrightarrow \mathrm{GL}(V_{\mathbb{Q}_p})$ is induced from a map $G_p \to \mathrm{GL}(\Lambda)$ over $\mathbb{Z}_p$.

By Lemma 2.4.1 we obtain a $\mathbb{Z}[1/N]$-lattice $\Lambda' = \Lambda \cap V_{\mathbb{Z}[1/pN]}$. Of course we can canonically identify $\Lambda' \otimes \mathbb{Z}[1/pN] \cong V_{\mathbb{Z}[1/pN]}$, in the context of which we take $\mathcal{G}/\mathbb{Z}[1/N]$ to be the closure of

$$\mathrm{Im}(G^M \hookrightarrow \mathrm{GL}(V_{\mathbb{Z}[1/pN]}) \hookrightarrow \mathrm{GL}(\Lambda')).$$

We claim this $\mathcal{G}$ does the job. It's evident that $\mathcal{G}|_{\mathbb{Z}[1/pN]} \cong G^M$. Since we have a canonical isomorphism $\Lambda' \otimes \mathbb{Z}_p \cong \Lambda$ and by the other identifications in the construction we also have $\mathcal{G}|_{\mathbb{Z}_p} \cong G_p$. We also need it to be reductive (i.e., smooth affine with connected reductive geometric fibres). Being a closed subgroup of $\mathrm{GL}(\Lambda')$ it's visibly affine, its geometric fibres are reductive by what we already know, and smoothness can be checked fpqc locally, whence it also follows by the identifications we have made.

The second part of the proposition is an immediate consequence of our argument. $\square$

We also echo the remarks in [Madapusi Pera 2012, 4.3], that in the case of G coming from a Hodge type Shimura datum, the condition on factors of type B is always satisfied, by Deligne's classification of symplectic representations. We shall also need the following modification of [Kisin 2010, 2.1.2].

Lemma 2.4.3. *Let $i : (G_1, \mathfrak{X}_1) \hookrightarrow (G_2, \mathfrak{X}_2)$ be an embedding of Shimura data with $K_2^N \subset \prod'_{p \nmid N} G_2(\mathbb{Q}_p) =: G_2(\mathbb{A}^{\infty,N})$ compact open, and $K_1 = K_1^N K_{1,N}$ a compact open of $G_1(\mathbb{A}^\infty)$ such that $K_1^N = K_2^N \cap G_1(\mathbb{A}^{\infty,N})$. Suppose G_1 and G_2 both have centre which is compact modulo its split part.*[5] *Then there exists an open compact subgroup $K_{2,N} \subset \prod_{p|N} G_2(\mathbb{Q}_p)$ such that $K_2 := K_{2,N} K_2^N \supset K_1$ and the induced map of $E(G_1, \mathfrak{X}_1)$-schemes*

$$\mathrm{Sh}_{K_1}(G_1, \mathfrak{X}_1) \to \mathrm{Sh}_{K_2}(G_2, \mathfrak{X}_2)_{E(G_1,\mathfrak{X}_1)}$$

is a closed embedding.

Proof. By the same argument as [Deligne 1971, 1.15] it suffices to check

$$\alpha : \mathrm{Sh}_{K_1^N}(\mathbb{C}) \to \mathrm{Sh}_{K_2^N}(\mathbb{C})$$

is injective.

On the level of complex points (by the assumption on centres, which removes the technicalities involving units) this map is

$$\alpha : G_1(\mathbb{Q})\backslash\mathfrak{X}_1 \times G_1(\mathbb{A}^\infty)/K_1^N \to G_2(\mathbb{Q})\backslash\mathfrak{X}_2 \times G_2(\mathbb{A}^\infty)/K_2^N.$$

We prove this is injective by first noting that

$$G_1(\mathbb{Q})\backslash \prod_{p|N} G_1(\mathbb{Q}_p) \to G_2(\mathbb{Q})\backslash \prod_{p|N} G_2(\mathbb{Q}_p)$$

is injective as in [Deligne 1971, 1.15.3]. Now fix a set of coset representatives of $G_1(\mathbb{Q})$ in $\prod_{p|N} G_1(\mathbb{Q}_p)$ and note that translating by these, the fibres of $G_1(\mathbb{Q})\backslash\mathfrak{X}_1 \times G_1(\mathbb{A}^\infty)/K_1^N \to G_1(\mathbb{Q})\backslash \prod_{p|N} G_1(\mathbb{Q}_p)$ may be identified with $\mathfrak{X}_1 \times G_1(\mathbb{A}^{\infty,N})/K_1^N$. But now since $K_1^N = K_2^N \cap G_1(\mathbb{A}^{\infty,N})$, we have that

$$\mathfrak{X}_1 \times G_1(\mathbb{A}^{\infty,N})/K_1^N \to \mathfrak{X}_2 \times G_2(\mathbb{A}^{\infty,N})/K_2^N,$$

is injective, and the lemma follows. □

2.4.4. We now proceed with the construction in the Hodge type case.

Firstly, by finite-presentedness we note that i is in fact defined over $\mathbb{Z}[1/M]$ for some M divisible by N. By Proposition 2.4.2 we can find a lattice $V_{\mathbb{Z}[1/N]} \subset V$ and $\mathcal{G}/\mathbb{Z}[1/N]$ a reductive model such that (forgetting the symplectic pairing) i is obtained from a map $\mathcal{G} \hookrightarrow \mathrm{GL}(V_{\mathbb{Z}[1/N]})$ and $\mathcal{G}(\mathbb{Z}_p) = G(\mathbb{Z}_p)$ for all $p \nmid N$, in particular giving $K^N = \prod_{p \nmid N} \mathcal{G}(\mathbb{Z}_p)$.

Let K'^N be the stabiliser of $V_{\mathbb{Z}[1/N]}$ in $\prod_{p \nmid N} \mathrm{GSp}(V \otimes \mathbb{Q}_p, \psi)$, noting that K'^N will be maximal compact but need not be hyperspecial at the primes dividing M/N. We also fix a $\mathbb{Z}$-lattice $V_{\mathbb{Z}} \subset V_{\mathbb{Z}[1/N]}$ and note that for all $K_N \subset \prod_{p|N} G(\mathbb{Q}_p)$ sufficiently small $K = K_N K^N$ fixes $V_{\hat{\mathbb{Z}}}$.

[5] We expect the argument can be modified slightly as in Deligne to remove this hypothesis.

2.4.5. Applying the lemma Lemma 2.4.3 in our setting (since G is Hodge type the hypothesis on the centre holds), and letting $E = E(G, \mathfrak{X})$, we obtain a closed embedding of Shimura varieties

$$\mathrm{Sh}_{K^N}(G, \mathfrak{X}) \hookrightarrow \mathrm{Sh}_{K'^N}(\mathrm{GSp}, S^{\pm})_E \subset \mathscr{S}_{N,\mathbb{O}_E[1/N]}.$$

Letting $\overline{\mathrm{Sh}_{K^N}} \subset \mathscr{S}_{N,\mathbb{O}_E[1/N]}$ be the scheme theoretic closure of this map, the extension property for $\mathscr{S}_N$ implies the extension property for $\overline{\mathrm{Sh}_{K^N}}$. Now let $\mathscr{S}_{K^N}$ be the normalisation of $\overline{\mathrm{Sh}_{K^N}}$. Since test schemes are regular and a fortiori normal, the universal property of normalisation implies that $\mathscr{S}_{K^N}$ also has the extension property.

2.4.6. We need to check smoothness at finite level. It suffices to check smoothness in a formal neighbourhood of any closed point. When the point has characteristic zero it is in the generic fibre and smoothness is guaranteed. When it has characteristic p, we observe that at finite level our construction exactly follows that of Kisin [2010, 2.3] and Kim and Madapusi Pera [2016, 3.5] for the case $p = 2$, and so in particular the necessary local rings are smooth. Hence the $\mathscr{S}_{K^N}$ give the required integral canonical models in the Hodge type case, compatible with Kisin's by construction.

2.5. *Abelian type case.* We begin by making some more observations about the Hodge type setting. As before we fix connected reductive $G/\mathbb{Z}[1/N]$ belonging to a Shimura datum $(G, \mathfrak{X})$ of Hodge type with reflex field E, and consider the tower

$$\mathrm{Sh}_{K^N} = \varprojlim_{K_N} \mathrm{Sh}_{K_N K^N}$$

obtained by fixing $K^N = G(\hat{\mathbb{Z}}^N)$ and letting the level at $p|N$ go to infinity.

Lemma 2.5.1. *The connected component* $\mathrm{Sh}^+_{K^N} \subset \mathrm{Sh}_{K^N,\bar{E}}$ *is defined over the maximal abelian extension* E_N/E *unramified away from* N.

Proof. By [Kisin 2010, 2.2.4], and taking a suitable quotient, we see that $\mathrm{Sh}^+_{K^N}$ is defined over the maximal abelian extension E^p/E unramified away from p for all $p \nmid N$. By the identification $\bigcap_{p\nmid N} E^p = E_N$ inside E^{ab}, we deduce that $\mathrm{Sh}^+_{K^N}$ is defined over E_N. □

2.5.2. Let $q \nmid N$ be a prime, and recall the following groups which are used to construct Kisin's integral models, and the following results from [Kisin 2010, 3.3]. We adopt the usual notations where $G^{\mathrm{ad}}(\mathbb{Q})^+$ is the intersection of $G^{\mathrm{ad}}(\mathbb{Q})$ with the connected component $G^{\mathrm{ad}}(\mathbb{R})$, $G(\mathbb{Q})_+$ the inverse image of $G^{\mathrm{ad}}(\mathbb{Q})^+$ in $G(\mathbb{Q})$, $G^{\mathrm{ad}}(\mathbb{Z}_{(q)})^+ = G^{\mathrm{ad}}(\mathbb{Q})^+ \cap G^{\mathrm{ad}}(\mathbb{Z}_{(q)})$ and $G(\mathbb{Z}_{(q)})_+ = G(\mathbb{Q})_+ \cap G(\mathbb{Z}_{(q)})_+$.

We have

$$\mathscr{A}_q(G) := G(\mathbb{A}^{\infty,q})/\overline{Z(\mathbb{Z}_{(q)})} *_{G(\mathbb{Z}_{(q)})_+/Z(\mathbb{Z}_{(q)})} G^{\mathrm{ad}}(\mathbb{Z}_{(q)})^+$$

which acts on $\mathrm{Sh}_{K_q} = \varprojlim_{K^q} \mathrm{Sh}_{G(\mathbb{Z}_q)K^q}$, and the subgroup

$$\mathscr{A}_q^{\circ}(G) := \overline{G(\mathbb{Z}_{(q)})_+}/\overline{Z(\mathbb{Z}_{(q)})} *_{G(\mathbb{Z}_{(q)})_+/Z(\mathbb{Z}_{(q)})} G^{\mathrm{ad}}(\mathbb{Z}_{(q)})^+$$

which acts on a connected component $\mathrm{Sh}^+_{K_q}$. It follows from the argument in [Kisin 2010, 3.3.7] that this is precisely the subgroup sending $\mathrm{Sh}^+_{K_q}$ into itself.

2.5.3. We make the following new remarks. The construction of twisting abelian varieties by a Z-torsor from [Kisin 2010, §3] can be carried out by twisting by a $Z^{\mathrm{der}} = Z(G^{\mathrm{der}})$-torsor instead. Indeed, given $\gamma \in G^{\mathrm{ad}}(\mathbb{Q})^+$ we may take $\mathcal{P}$ to be the fibre of γ along $G^{\mathrm{der}} \to G^{\mathrm{ad}}$. It will suffice to check the following.

Proposition 2.5.4. (1) *With notation as above, if $\mathcal{P}'$ is the fibre of γ along $G \to G^{\mathrm{ad}}$ then $\mathcal{P}' = \mathcal{P} \times^{Z^{\mathrm{der}}} Z$.*

(2) *If V is a $\mathbb{Q}$-vector space with an $\mathcal{O}_Z$-comodule action, $\mathcal{P}$ and $\mathcal{P}'$ as above, there is a natural isomorphism*

$$(V \otimes_{\mathbb{Q}} \mathcal{O}_{\mathcal{P}})^{Z^{\mathrm{der}}} \xrightarrow{\sim} (V \otimes_{\mathbb{Q}} \mathcal{O}_{\mathcal{P}'})^{Z}.$$

Proof. For (1) we can define (using $\mathcal{P}' \subset G$) a map

$$\mathcal{P} \times^{Z^{\mathrm{der}}} Z \ni (p, z) \mapsto p.z \in \mathcal{P}'$$

which is obviously an isomorphism of Z-torsors and proves the claim.

Part (2) can be seen easily by appealing to the general Tannakian framework that Z-torsors over $\operatorname{Spec} \mathbb{Q}$ correspond to fibre functors $\omega : \mathrm{Rep}_{\mathbb{Q}} Z \to \mathrm{Vec}_{\mathbb{Q}}$, and that $\mathcal{P} \times^{Z^{\mathrm{der}}} Z = \mathcal{P}'$ implies that we can factor $\omega_{\mathcal{P}'}$ as

$$\omega_{\mathcal{P}'} : \mathrm{Rep}_{\mathbb{Q}} Z \xrightarrow{\mathrm{Res}} \mathrm{Rep}_{\mathbb{Q}} Z^{\mathrm{der}} \xrightarrow{\omega_{\mathcal{P}}} \mathrm{Vec}_{\mathbb{Q}}.$$

Writing out what this statement actually means algebraically (and writing an arbitrary $\mathcal{O}_Z$-comodule as a filtered colimit of finite dimensional ones), we recover (2). □

This has the technical advantage given by the following lemma. We thank Kestutis Cesnavicius for pointing out to us that the analogous result from [Kisin 2010, 3.4.8] where $G = Z$ is a general group of multiplicative type is false over the base $\mathbb{Z}[1/N]$ in general.

Lemma 2.5.5. *Let R be an integrally closed domain with fraction field K, and G/R a finite group scheme. Then a torsor T/R for G is trivial if and only if T_K is a trivial G_K-torsor.*

Proof. Since G/R is finite, it is proper, which as a property stable under fpqc descent is inherited by T, and so $T(R) = \bigcap_{R_v \subset K} T(R_v) = T(K)$. In particular the latter is nonempty if and only if the former is. □

2.5.6. With these remarks in mind, we can rewrite the action of $\mathscr{A}_q(G)$ on the integral model $\mathscr{S}_{K_q}/\mathbb{O}_{E,v}$ explicitly following [Kisin 2010, 3.4.5]. A point $x \in \mathscr{S}_{K_q}(T)$ gives rise to a triple (A, λ, ϵ^q) where A/T is an abelian scheme up to prime to q-isogeny, λ a weak polarisation of A and ϵ^q a section of $\Gamma(T, \underline{\mathrm{Isom}}(V_{\mathbb{A}^{\infty,q}}, \hat{V}^q(A)))$.

By [Kisin 2010, 3.4.5] and our previous remarks, if we take $(h, \gamma^{-1}) \in \mathscr{A}_q(G)$ and x associated to the triple (A, λ, ϵ^q) then the triple associated to $x.(h, \gamma^{-1})$ is isogenous to

$$(A^{\mathscr{P}}, \lambda^{\mathscr{P}}, \epsilon^{q,\mathscr{P}} \circ \tilde{\gamma} h \tilde{\gamma}^{-1})$$

where $\mathscr{P}$ is the torsor for $Z^{\mathrm{der}} \subset G^{\mathrm{der}}$ given by the fibre of γ, $\tilde{\gamma}$ is an element of $G^{\mathrm{der}}(F)$ for some finite Galois extension $F/\mathbb{Q}$ mapping to γ under $G^{\mathrm{der}}(F) \to G^{\mathrm{ad}}(F)$, and the notations $A^{\mathscr{P}}, \lambda^{\mathscr{P}}, \epsilon^{q,\mathscr{P}}$ are the "twists by $\mathscr{P}$" as defined in [Kisin 2010, 3.1.3].

2.5.7. We introduce the notation $\mathbb{Z}^N := \mathbb{Z}[1/N]$, to provide a slight simplification in situations where the notation quickly becomes messy. Let (where plus notation denotes intersection with $G(\mathbb{R})_+$,$G^{\mathrm{ad}}(\mathbb{R})^+$ as usually defined, and overline notation denotes closures in $\prod_{p|N} G(\mathbb{Q}_p)$)

$$\mathscr{A}^N(G) := \frac{\prod_{p|N} G(\mathbb{Q}_p)}{\overline{Z(\mathbb{Z}^N)}} *_{G(\mathbb{Z}^N)_+/Z(\mathbb{Z}^N)} G^{\mathrm{ad}}(\mathbb{Z}^N)^+,$$

$$\mathscr{A}^{N,\circ}(G) := \frac{\overline{G(\mathbb{Z}^N)_+}}{\overline{Z(\mathbb{Z}^N)}} *_{G(\mathbb{Z}^N)_+/Z(\mathbb{Z}^N)} G^{\mathrm{ad}}(\mathbb{Z}^N)^+,$$

and (where here overline notation means closures in $G(\mathbb{A}^{\infty,q})$)

$$\tilde{\mathscr{A}}_q^N(G) := \frac{\prod_{p|N} G(\mathbb{Q}_p) \times \prod_{p\nmid qN} G(\mathbb{Z}_p)}{\overline{Z(\mathbb{Z}^N)}} *_{G(\mathbb{Z}^N)_+/Z(\mathbb{Z}^N)} G^{\mathrm{ad}}(\mathbb{Z}^N)^+ \subset \mathscr{A}_q(G).$$

Let $K^N = \prod_{p\nmid N} G(\mathbb{Z}_p)$, $K^{Nq} = \prod_{p\nmid Nq} G(\mathbb{Z}_p)$ and

$$\Delta^N = \mathrm{Ker}(\mathscr{A}^{N,\circ}(G) \to \mathscr{A}^{N,\circ}(G^{\mathrm{ad}})).$$

First, a group theoretic lemma following Deligne's Corvallis paper.

Lemma 2.5.8. (1) *The group $\mathscr{A}^{N,\circ}(G)$ is canonically the completion of $G^{\mathrm{ad}}(\mathbb{Z}^N)^+$ with respect to the topology generated by the images of congruence subgroups of G^{der} the form $K_N \times \prod_{p\nmid N} G^{\mathrm{der}}(\mathbb{Z}_p)$ as K_N varies. In particular $\mathscr{A}^{N,\circ}(G^{\mathrm{der}}) := \mathscr{A}^{N,\circ}(G)$ canonically depends only on G^{der}.*

(2) *We can naturally identify*

$$\mathscr{A}^{N,\circ}(G^{\mathrm{der}}) = \overline{G^{\mathrm{der}}(\mathbb{Z}^N)_+} *_{G^{\mathrm{der}}(\mathbb{Z}^N)_+} G^{\mathrm{ad}}(\mathbb{Z}^N)^+$$

(*where the closure is taken in $\prod_{p|N} G^{\mathrm{der}}(\mathbb{Q}_p)$*).

Proof. For (1) we have a natural inclusion $G^{\mathrm{ad}}(\mathbb{Z}^N)^+ \subset \mathscr{A}^{N,\circ}(G)$, and it is easily checked that the image is dense. Moreover, a neighbourhood of the identity is $\overline{G(\mathbb{Z}^N)_+}/\overline{Z(\mathbb{Z}^N)}$ whose topology is generated by that of the congruence subgroups of G with fixed hyperspecial level away from N. But by [Deligne 1979, 2.0.13] this topology is the same as that generated by the congruence subgroups of G^{der} with fixed hyperspecial level away from N. Finally this neighbourhood of the identity is obviously complete, so we are done. From this description (2) follows immediately. □

Our key result is the following, whose proof follows that of [Kisin 2010, 3.4.6].

Proposition 2.5.9. *The group $\mathscr{A}^N(G)$ acts naturally on the integral canonical model $\mathscr{S}_{K^N}$, and Δ^N acts freely.*

Proof. By the extension property, the first part can be checked on the generic fibre, where it follows from the self evident isomorphism

$$\tilde{\mathscr{A}}_q^N(G)/K^{Nq} \xrightarrow{\sim} \mathscr{A}^N(G).$$

This (together with the compatibility with Kisin's construction) gives us the additional information that for $v \mid q$ and $q \nmid N$, the action on $\mathscr{S}_{K^N,v}$ can be described on the level of triples (A, λ, ϵ^q) where ϵ^q is now given modulo K^{Nq}.

Take $(h, \gamma^{-1}) \in \Delta^N \subset \overline{G^{\mathrm{der}}(\mathbb{Z}^N)_+} *_{G^{\mathrm{der}}(\mathbb{Z}^N)_+} G^{\mathrm{ad}}(\mathbb{Z}^N)^+$, and $\tilde{\gamma} \in G^{\mathrm{der}}(F)$ mapping to $\gamma \in G^{\mathrm{ad}}(F)$ with $F/\mathbb{Q}$ finite Galois. We also let $\mathscr{P}$ denote the Z^{der}-torsor of elements of G^{der} mapping to γ. Since $(h, \gamma^{-1}) \in \Delta^N$ we see that $h\tilde{\gamma}^{-1} \in \prod_{p|N} Z^{\mathrm{der}}(\mathbb{Q}_p \otimes F)$.

Suppose $x \in \mathscr{S}_{K^N}(\kappa)$ for some algebraically closed field κ of characteristic q or 0, that $x.(h, \gamma^{-1}) = x$ and associated to x is the triple $(A, \lambda, \epsilon^q/K^{qN})$. We need to show that $(h, \gamma^{-1}) = 1$.

As in the proof of [Kisin 2010, 3.4.6] we can find a unique quasiisogeny $\alpha : A \to A^{\mathscr{P}}$ such that

$$\begin{array}{ccccc}
V \otimes \mathbb{A}^{\infty,q} \otimes F & \xrightarrow{\tilde{\gamma}h\tilde{\gamma}^{-1}} & V \otimes \mathbb{A}^{\infty,q} \otimes F & \xrightarrow{\tilde{\gamma}^{-1}} & V \otimes \mathbb{A}^{\infty,q} \otimes F \\
\Big\downarrow{\scriptstyle \epsilon^q} & & \Big\downarrow{\scriptstyle \epsilon^{q,\mathscr{P}}} & & \Big\downarrow{\scriptstyle \epsilon^q} \\
\hat{V}^q(A) \otimes F & \xrightarrow[\alpha_* \otimes 1]{} & \hat{V}^q(A^{\mathscr{P}}) \otimes F & \xrightarrow[\iota_{\tilde{\gamma}}]{} & \hat{V}^q(A) \otimes F.
\end{array}$$

commutes. This demonstrates that also $h\tilde{\gamma}^{-1} \in \mathrm{Aut}_{\mathbb{Q}}(A) \otimes F$. But the intersection of $\prod_{p|N} Z^{\mathrm{der}}(\mathbb{Q}_p \otimes F)$ and $\mathrm{Aut}_{\mathbb{Q}}(A) \otimes F$ inside $\prod_{p|N} \mathrm{Aut}_{\mathbb{Q}}(A)(\mathbb{Q}_p \otimes F)$ is $Z^{\mathrm{der}}(F)$, so we conclude that

$$h\tilde{\gamma}^{-1} \in Z^{\mathrm{der}}(F).$$

As in [Kisin 2010, 3.4.6] this demonstrates that $\mathscr{P}$ is trivial as a $Z^{\mathrm{der}}_{\mathbb{Q}}$-torsor over $\mathbb{Q}$. But Z^{der} is a finite group, Lemma 2.5.5 implies that $\mathscr{P}_{\mathbb{Z}[1/N]}$ (the torsor given by the inverse image of γ in $G^{\mathrm{der}}_{\mathbb{Z}[1/N]}$) is a trivial $\mathbb{Z}[1/N]$-torsor, so we can assume $\tilde{\gamma} \in G^{\mathrm{der}}(\mathbb{Z}[1/N])_+$ and replacing h by $h\tilde{\gamma}^{-1}$, assume $\gamma = 1$. Moreover, we may take $F = \mathbb{Q}$, and so we deduce $h \in Z^{\mathrm{der}}(\mathbb{Q}) = Z^{\mathrm{der}}(\mathbb{Z}[1/N])$: that is to say, it is trivial as an element of Δ^N, as required. □

We also need the following ingredient.[6]

Lemma 2.5.10. *The group Δ^N is finite.*

Proof. Let $\rho : G^{\mathrm{der}} \to G^{\mathrm{ad}}$ be the usual finite isogeny. By the discussion in [Deligne 1979, 2.0] there is a diagram with exact rows.

$$\begin{array}{ccccc}
\overline{G^{\mathrm{der}}(\mathbb{Z}^N)_+} & \longrightarrow & \mathscr{A}^{N,\circ}(G^{\mathrm{der}}) & \longrightarrow & G^{\mathrm{ad}}(\mathbb{Z}^N)^+/\rho G^{\mathrm{der}}(\mathbb{Z}^N)_+ \\
\downarrow & & \downarrow & & \downarrow \\
\overline{G^{\mathrm{ad}}(\mathbb{Z}^N)^+} & \longrightarrow & \mathscr{A}^{N,\circ}(G^{\mathrm{ad}}) & \longrightarrow & 1
\end{array}$$

Considering the kernels of the vertical maps, this puts Δ^N in an exact sequence between a subgroup of $Z^{\mathrm{der}}(\mathbb{A}_N)$ and the group $G^{\mathrm{ad}}(\mathbb{Z}^N)^+/\rho G^{\mathrm{der}}(\mathbb{Z}^N)_+$. The former as a product of finite groups is visibly finite. The latter group is a subgroup of $G^{\mathrm{ad}}(\mathbb{Z}^N)/\rho G^{\mathrm{der}}(\mathbb{Z}^N)$, which is in turn a subgroup of $H^1_{\mathrm{fppf}}(\mathbb{Z}^N, Z^{\mathrm{der}})$, so it suffices to check that this is finite.

Since Z^{der} is a finite group of multiplicative type, we may find a finite Zariski cover U_i of $\operatorname{Spec}\mathbb{Z}[1/N]$ and finite étale covers $V_i \to U_i$ such that $Z^{\mathrm{der}}|_{V_i}$ is isomorphic to a product of split finite multiplicative groups μ_k. By the Čech to derived functor spectral sequence for this cover, we may reduce our claim to checking that for L a number field and k, M positive integers, the groups $H^1_{\mathrm{fppf}}(\mathbb{O}_L[1/M], \mu_k)$ are finite. But the Kummer exact sequence gives an exact sequence

$$1 \to \mathbb{O}_L\big[\tfrac{1}{M}\big]^*/\mathbb{O}_L\big[\tfrac{1}{M}\big]^{*k} \to H^1_{\mathrm{fppf}}\big(\mathbb{O}_L\big[\tfrac{1}{M}\big], \mu_k\big) \to \operatorname{Pic}\big(\mathbb{O}_L\big[\tfrac{1}{M}\big]\big),$$

and both the outer terms are finite by classical algebraic number theory. □

2.5.11. We now set out to prove the main theorem Theorem 2.2.1 in the abelian type case. Let $(G_2, \mathfrak{X}_2)$ be a Shimura datum of abelian type, with $G_2/\mathbb{Z}[1/N]$ reductive. We first need to relate it to a datum of Hodge type, modifying [Kisin 2010, 3.4.13] slightly.

Lemma 2.5.12. *Let $(H, \mathfrak{Y})$ be a Shimura datum of abelian type with H adjoint. Then there exists a central isogeny $H' \to H$ such that whenever $(G, \mathfrak{X})$ is of Hodge type with $(G^{\mathrm{ad}}, \mathfrak{X}^{\mathrm{ad}}) \cong (H, \mathfrak{Y})$ then G^{der} is a quotient of H'.*

[6] Note that we believe the corresponding result of [Kisin 2010] is not true, so a result like this is perhaps also needed to deduce the extension property in that case.

Assume that H is quasisplit and unramified at all $p \nmid N$. Then there exists a Shimura datum $(G, \mathfrak{X})$ of Hodge type such that $(G^{\mathrm{ad}}, \mathfrak{X}^{\mathrm{ad}}) \cong (H, \mathfrak{Y})$, $G^{\mathrm{der}} = H'$ and G is quasisplit and unramified at all $p \nmid N$.

Proof. Most of this is proved in [Kisin 2010, 3.4.13]; the only thing to check is that we can arrange for G to be quasisplit and unramified at all $p \nmid N$. But since H has this property, clearly H' does, and so it suffices to control the centre, whose ramification is given by the totally imaginary quadratic extension L/F of [Deligne 1979, 2.3.10] which can be chosen arbitrarily. In particular, we may take any prime $q \mid N$, construct L by adjoining a q-th (or 4th if $q = 2$) root of unity and passing to a quadratic subfield. Then note that L/F is unramified at all primes $v \nmid q$, in particular at places over $p \nmid N$. □

Lemma 2.5.13. *Suppose we are given G and G_2, reductive over $\mathbb{Q}$ and unramified away from N, and a central isogeny $f : G^{\mathrm{der}} \to G_2^{\mathrm{der}}$. Suppose we are also given a reductive model $\mathcal{G}_2/\mathbb{Z}[1/N]$ of G_2.*

Then there exists a reductive model $\mathcal{G}/\mathbb{Z}[1/N]$ of G such that f extends to

$$f : \mathcal{G}^{\mathrm{der}} \to \mathcal{G}_2^{\mathrm{der}}.$$

Proof. We do the usual patching argument. Take any integral model $\mathcal{G}/\mathbb{Z}[1/N]$ and note that there will be some M such that $\mathcal{G}[1/M]$ is reductive and $G^{\mathrm{der}} \to G_2^{\mathrm{der}}$ extends to $\mathcal{G}^{\mathrm{der}}[1/M] \to \mathcal{G}_2^{\mathrm{der}}[1/M]$.

By Proposition 2.4.2 we will be done if for every $p \mid M$ and $p \nmid N$ we can find $G_p/\mathbb{Z}_p$ a reductive model such that f extends to $G_p^{\mathrm{der}} \to \mathcal{G}_{2,\mathbb{Z}_p}^{\mathrm{der}}$. But this is the case by the argument of [Kisin 2010, 3.4.14]. □

2.5.14. Now given our $(G_2, \mathfrak{X}_2)$ of abelian type, it gives rise to an adjoint Shimura datum, which by Lemma 2.5.12 is covered by $(G, \mathfrak{X})$ of Hodge type and by Lemma 2.5.13 we may take $G/\mathbb{Z}[1/N]$ reductive and $G^{\mathrm{der}} \to G_2^{\mathrm{der}}$ a central isogeny inducing a morphism $(G^{\mathrm{der}}, \mathfrak{X}^+) \to (G_2^{\mathrm{der}}, \mathfrak{X}_2^+)$ of connected Shimura data. Let $E = E(G, \mathfrak{X}) \subset \overline{\mathbb{Q}}$. Since for every $p \nmid N$, G and G_2 split over an unramified extension of p, we deduce that $E/\mathbb{Q}$ is unramified at all $p \nmid N$, and by Lemma 2.5.1 we see their connected Shimura varieties at levels $K^N := G(\hat{\mathbb{Z}}^N)$ and $K_2^N := G_2(\hat{\mathbb{Z}}^N)$ are defined over E_N/E, the maximal abelian extension of E unramified away from N. Let $\mathbb{O}_N$ be its ring of integers, and note that $\mathbb{O}_N[1/N]/\mathbb{Z}[1/N]$ is indétale.

Now, comparing our description (2.5.8(1)) of $\mathscr{A}^{N,\circ}(G^{\mathrm{der}})$ with Deligne's description [1979, 2.1.6] of the group acting on a connected Shimura variety, it is clear that $\mathscr{A}^{N,\circ}(G)$ acts on $\mathrm{Sh}^+_{K^N}(G, \mathfrak{X})_{E_N}$, and considering the subgroup

$$\Delta^N(G, G_2) := \mathrm{Ker}(\mathscr{A}^{N,\circ}(G^{\mathrm{der}}) \to \mathscr{A}^{N,\circ}(G_2^{\mathrm{der}})) \subset \Delta^N,$$

that the morphism $\mathrm{Sh}^+_{K^N}(G, \mathfrak{X})_{E_N} \to \mathrm{Sh}^+_{K_2^N}(G_2, \mathfrak{X}_2)_{E_N}$ is given by taking the quotient by $\Delta^N(G, G_2)$. Moreover, by the previous section we have an integral model

$\mathcal{S}^+_{K^N}(G,\mathfrak{X})_{\mathbb{O}_N[1/N]}$ for $\mathrm{Sh}^+_{K^N}(G,\mathfrak{X})_{E_N}$ satisfying the extension property, with a free Proposition 2.5.9 action of the finite Lemma 2.5.10 group Δ^N. We may therefore form the quotient by the finite subgroup $\Delta^N(G,G_2)$ and obtain a model $\mathcal{S}^+_{K_2^N}(G_2,\mathfrak{X}_2)_{\mathbb{O}_N[1/N]}$ for $\mathrm{Sh}^+_{K_2^N}(G_2,\mathfrak{X}_2)_{E_N}$ over $\mathbb{O}_N[1/N]$. By Lemma 2.1.5 this model enjoys the extension property, and passing to finite levels we see it is smooth.

2.5.15. Unlike at infinite or K_p-level, we do not know whether $\prod_{p|N} G_2(\mathbb{Q}_p)$ acts transitively on $\pi_0(\mathrm{Sh}_{K^N}(G_2,\mathfrak{X}_2)_{\overline{\mathbb{Q}}})$, so to conclude our proof we need an alternative to the usual "Deligne induction." Noting that our argument up to this point holds for any model $\mathcal{G}_2/\mathbb{Z}[1/N]$ for G_2, the following lemma is enough to conclude our argument.

We will need the Shimura variety

$$\mathrm{Sh}(G,\mathfrak{X}) = \varprojlim_K \mathrm{Sh}_K(G,\mathfrak{X})$$

at full infinite level, together with the usual fixed connected component $\mathrm{Sh}^+_{\overline{\mathbb{Q}}} \subset \mathrm{Sh}(G,\mathfrak{X})_{\overline{\mathbb{Q}}}$ containing the complex point $[x,1]$ for $x \in \mathfrak{X}^+$.

Lemma 2.5.16. *Let $(G,\mathfrak{X})$ be any Shimura datum unramified away from N, $E \supset E(G,\mathfrak{X})$, $K^N = \prod_{p\nmid N} G(\mathbb{Z}_p)$ for some choice of integral model for G, and E_N/E the maximal abelian extension unramified away from N.*

(1) *Let $X^+_{K^N}$ be any component of $\mathrm{Sh}_{K^N}(G,\mathfrak{X})_{\overline{\mathbb{Q}}}$. Then $X^+_{K^N}$ is defined over E_N, and given for every choice of hyperspecial levels of the form $U^N = \prod_{p\nmid N}\mathcal{G}(\mathbb{Z}_p)$ for $\mathcal{G}/\mathbb{Z}[1/N]$ a reductive model for G a smooth integral model $\mathcal{S}^+_{U^N}/\mathbb{O}_N[1/N]$ for $\mathrm{Sh}^+_{U^N,E_N}$ with the extension property, we can construct a smooth integral model for $X^+_{K^N,E_N}$ with the extension property.*

(2) *Given any smooth integral model for every $X^+_{K^N,E_N}$ with the extension property, their disjoint union gives a smooth integral canonical model for Sh_{K^N,E_N} which descends to $\mathbb{O}_E[1/N]$ and to which the $\mathcal{A}^N(G)$ action extends.*

Proof. For (1), let $\pi : \mathrm{Sh}(G,\mathfrak{X})_{\overline{\mathbb{Q}}} \to \mathrm{Sh}_{K^N}(G,\mathfrak{X})_{\overline{\mathbb{Q}}}$ be the canonical projection, and take $a \in G(\mathbb{A}^\infty)$ such that $\pi(\mathrm{Sh}^+.a) = X^+_{K^N}$. Note that this a descends to an identification

$$\mathrm{Sh}_{K^N}(G,\mathfrak{X})_{\overline{\mathbb{Q}}} \xrightarrow{\sim} \mathrm{Sh}_{a^{-1}K^N a}(G,\mathfrak{X})_{\overline{\mathbb{Q}}}$$

defined over E and under which $X^+_{K^N}$ is identified with $\mathrm{Sh}^+_{a^{-1}K^N a,\overline{\mathbb{Q}}}$. Since the latter is defined over E_N by Lemma 2.5.1, the former must be also. Moreover, since $a_p \in G(\mathbb{Z}_p)$ for all but finitely many p, and in all other cases we are taking the conjugate of a hyperspecial subgroup, which are always hyperspecial and so have local reductive models, by Proposition 2.4.2 there exists a reductive model $\mathcal{G}/\mathbb{Z}[1/N]$ for G giving rise to the level $a^{-1}K^N a$. Hence by hypothesis we have a smooth integral model for $\mathrm{Sh}^+_{a^{-1}K^N a,E_N}$ with the extension property. Composing with the isomorphism induced by a gives the required model for $X^+_{K^N,E_N}$.

For (2), we assume we are given for each $X^+_{K^N,E_N}$ some smooth integral model $\mathscr{X}^+_{K^N}$ with the extension property. Letting $\mathscr{S}_{K^N,\mathbb{O}_N[1/N]}$ together with

$$\iota : \mathscr{S}_{K^N,\mathbb{O}_N[1/N]} \otimes_{\mathbb{O}_N[1/N]} E_N \xrightarrow{\sim} \mathrm{Sh}_{K^N,E_N}$$

be their disjoint union, note that it still has the extension property. In particular the $\mathrm{Gal}(E_N/E)$-action on Sh_{K^N,E_N} extends to $\mathscr{S}_{K^N,\mathbb{O}_N[1/N]}$, and since E_N/E is unramified away from N this gives an étale descent datum from $\mathbb{O}_N[1/N]$ to $\mathbb{O}_E[1/N]$. Thus we may descend our model to $\mathscr{S}_{K^N}/\mathbb{O}_E[1/N]$ and by Lemma 2.1.3 this model still has the extension property. The $\mathscr{A}^N(G)$-action also extends to $\mathscr{S}_{K^N,\mathbb{O}_N[1/N]}$ by the extension property, and commutes with the $\mathrm{Gal}(E_N/E)$-action, so it descends, and for any $K_N \subset \prod_{p|N} G(\mathbb{Q}_p)$ we have that $(\mathscr{S}_{K^N}/K_N)\otimes\mathbb{O}_N[1/N]$ is smooth, which is a property stable by fpqc descent and allows us to see that $\mathscr{S}_{K^N}$ is a smooth canonical model. $\square$

3. Automorphic vector bundles and filtered G-bundles

3.1. ***Review of characteristic zero.*** We sketch the main results of [Milne 1990, III], on which our results will build.

3.1.1. Let $(G, \mathfrak{X})$ be a Shimura datum with reflex field E, and $\mu : \mathbb{G}_{m,\bar{E}} \to G_{\bar{E}}$ a Hodge cocharacter of $\mathfrak{X}$. Then we can form the *compact dual* Gr_μ which represents the following functor. Fix a faithful representation $G \hookrightarrow \mathrm{GL}(V)$ and tensors $s_\alpha \in V^\otimes$ such that G is exactly the subgroup fixing these tensors, and note that μ induces a filtration $\mathrm{Fil}^\bullet \subset V \otimes \bar{E}$. For $\pi : S \to \operatorname{Spec} \mathbb{Q}$ we let $\mathscr{V}_S := \pi^* V$ be the constant vector bundle, which also carries tensors $s_{\alpha,S} = \pi^* s_\alpha$. Then as a functor on E-schemes

$$\mathrm{Gr}_\mu(S) = \big\{\text{filtrations } \mathscr{F}^\bullet \text{ of } \mathscr{V}_S \text{ such that } (V_{\bar{s}}, \mathrm{Fil}^\bullet_{\bar{s}}, s_\alpha) \cong (\mathscr{V}_{S,\bar{s}}, \mathscr{F}^\bullet_{\bar{s}}, s_{\alpha,S}) \\ \text{for each geometric point } \bar{s} \in S\big\}.$$

This depends only on the conjugacy class of μ and is representable as an E-scheme. Moreover by construction it carries an algebraic action of G_E, which we write as a right action

$$\mathrm{Gr}_\mu(S) \times G(S) \ni (\mathscr{F}^\bullet, g) \mapsto g^{-1}(\mathscr{F}^\bullet) \in \mathrm{Gr}_\mu(S).$$

3.1.2. Let $Z_{nc} \subset Z(G)$ be the largest subtorus of $Z(G)$ split over $\mathbb{R}$ but with no subtorus split over $\mathbb{Q}$, and $G^c = G/Z_{nc}$. Then there is a G^c-torsor $P = P(G, \mathfrak{X})$ over $\mathrm{Sh}(G, \mathfrak{X})$ with an equivariant $G(\mathbb{A}^\infty)$-action,[7] which can be easily defined analytically over $\mathbb{C}$ and by [Milne 1990, III,4.3] admits a canonical model (in Milne's sense) over E, together with an integrable connection with regular singularities at infinity.

[7]It is natural to ask whether this extends to an equivariant action of Deligne's extension $G(\mathbb{A}^\infty)/\overline{Z(\mathbb{Q})} *_{G(\mathbb{Q})_+/Z(\mathbb{Q})} G^{\mathrm{ad}}(\mathbb{Q})^+$: in fact this group does act but not quite in a way that commutes with the algebraic G^c-action, as we shall later see.

Moreover $P(G, \mathfrak{X})$ has a G-action, via $G \to G^c$, and there is a G-equivariant map [Milne 1990, III,4.6] $\gamma : P(G, \mathfrak{X}) \to \mathrm{Gr}_\mu$ which complex analytically is given by the Hodge filtration coming from $\mathfrak{X}$, but is algebraic and descends to E.

To begin discussing any functoriality of this construction we need a lemma.

Lemma 3.1.3. *Let $f : (G_1, \mathfrak{X}_1) \to (G_2, \mathfrak{X}_2)$ be a morphism of Shimura data. Then there is an induced map*

$$G_1^c \to G_2^c.$$

Proof. This comes down to showing that $f(Z(G_1)_{nc}) \subset Z(G_2)$. Suppose for contradiction we have $\mathbb{R}^* \cong \Psi \subset Z(G_1)(\mathbb{R})$ with $f(\Psi)$ intersecting trivially with $Z(G_2)(\mathbb{R})$. Let $h \in \mathfrak{X}_1$, and recall that $\mathrm{ad}\, f(h(i))$ is a Cartan involution acting on G_2^{ad}, so the real group

$$H = \{g \in G_2^{\mathrm{ad}}(\mathbb{C}) : f(h(i))gf(h(i))^{-1} = \bar{g}\}$$

is compact. On the other hand for any $\psi \in \Psi$ we have

$$h(i)\psi h(i)^{-1} = \psi = \overline{\psi}$$

since ψ is central in G_1, $h(i) \in G_1(\mathbb{R})$, and ψ is real. Hence, we have an embedding of $\Psi \cong \mathbb{R}^*$ into the compact group $H(\mathbb{R})$, which is absurd. □

Note that there is no such functoriality for general group morphisms. For example letting F be a totally real field of degree d acting on itself by multiplication we get a morphism

$$F^\times \hookrightarrow \mathrm{GL}_d$$

failing to have the required property for all $d > 1$.

3.1.4. Automorphic vector bundles are typically parametrised by complex representations of the parabolic subgroup P_μ associated with the Hodge cocharacter μ in the usual fashion (for example one may define $P_\mu \subset G_\mathbb{C}$ as the subgroup preserving the filtrations μ induces on $\mathrm{Rep}\, G_\mathbb{C}$). By a complex analytic interpretation of the Grassmannian it is easy to show these are in correspondence with $G_\mathbb{C}$-equivariant vector bundles on $\mathrm{Gr}_{\mu,\mathbb{C}}$. We therefore take as our input data a G-equivariant vector bundle $\mathcal{J}$ on Gr_μ defined over L/E some number field, and we assume the G-action factors through G^c.[8]

Given this data, we can pull it back along γ to get a $G(\mathbb{A}^\infty) \times G^c$-equivariant vector bundle on $P(G, \mathfrak{X})_L$ and therefore a $G(\mathbb{A}^\infty)$-equivariant vector bundle $\mathcal{V}(\mathcal{J})$ on $\mathrm{Sh}(G, \mathfrak{X})_L$. This is the construction of canonical models for automorphic vector bundles we seek to perform integrally.

[8]This is a reasonable condition to impose since $Z \subset P_\mu$ acts trivially on Gr_μ.

3.1.5. We need a basic functoriality property for which we could not find a direct reference but which follows easily from Milne's definition together with the map on complex points induced by $\mathfrak{X} \times G(\mathbb{A}^\infty) \times G^c(\mathbb{C}) \to \mathfrak{X}' \times G'(\mathbb{A}^\infty) \times G'^c(\mathbb{C})$ which we note relies on Lemma 3.1.3.

Lemma 3.1.6. *Let $f : (G, \mathfrak{X}) \to (G', \mathfrak{X}')$ be a morphism of Shimura data, μ and μ' Hodge cocharacters of $\mathfrak{X}$ and $\mathfrak{X}'$, respectively. Then there is a diagram*

$$\begin{array}{ccccc} \mathrm{Sh}(G, \mathfrak{X}) & \longleftarrow & P(G, \mathfrak{X}) & \longrightarrow & \mathrm{Gr}_\mu \\ \downarrow & & \downarrow & & \downarrow \\ \mathrm{Sh}(G', \mathfrak{X}') & \longleftarrow & P(G', \mathfrak{X}') & \longrightarrow & \mathrm{Gr}_{\mu'} \end{array}$$

defined over $E(G, \mathfrak{X})$ which is G-equivariant and $G(\mathbb{A}^\infty)$-equivariant in the obvious senses.

3.1.7. In the case where $G = T$ a torus we also need the following, the main content of which is due to Blasius. Let p be a prime containing a place v of $E \supset E(T, h)$, ω_{et} the usual fibre functor giving étale local systems on the Shimura variety, and fix $\overline{E}_v$ an algebraic closure of E_v letting $\Gamma_{E_v} := \mathrm{Gal}(\overline{E}_v/E_v)$. Consider the fibre functor coming from p-adic Hodge theory

$$\omega_{v,\mathrm{dR}} : \mathrm{Rep}_{\mathbb{Q}_p}(T^c) \ni V \mapsto (\omega_{\mathrm{et}}(V) \otimes_{\mathbb{Q}_p} B_{\mathrm{dR}})^{\Gamma_{E_v}} \in \mathrm{Vec}_{E_v}.$$

Proposition 3.1.8. *Suppose T is split by a CM field, and let $X = \mathrm{Spec}\, E \subset \mathrm{Sh}_U(T, h)$ be a component of the Shimura variety for $U \subset T(\mathbb{A}^\infty)$ open compact. There is a natural isomorphism between $\omega_{P,X,v} : V \mapsto (V \times^{T^c} P_U(T, h)|_X) \otimes_E E_v$ and $\omega_{v,\mathrm{dR}}$.*

Proof. By Lemma 3.1.6 it suffices to do the case $T = T^c$. We observe that any cocharacter $\mu : \mathbb{G}_{m,E} \to T^c_E$ has weight defined over $\mathbb{Q}$. Indeed, by definition $X_*(T^c)_\mathbb{Q}$ is the summand of $X_*(T)_\mathbb{Q}$ on which either Galois acts trivially or complex conjugation acts via -1, so for any $\mu \in X_*(T^c)$, $\mu + \mu^c$ lands in the summand on which Galois acts trivially.

Combining this with fact that T hence T^c is split by a CM field, the induced Shimura datum (T^c, h^c) is of CM type, and thus admits a characterisation as a moduli space of CM motives M, which we take as those for which the Betti fibre functor $\omega_B = H_B(M(\rho)) : \mathrm{Rep}_\mathbb{Q} T \to \mathrm{Vec}_\mathbb{Q}$ is trivial.

Therefore a point $x \in X(\overline{\mathbb{Q}})$ has attached to it a T-valued CM motive $M : \mathrm{Rep}_\mathbb{Q}(T) \to (CM/\overline{\mathbb{Q}})$, and the fibre functor attached to $x^* P_U(T, h)$ is given by $\omega_P(\rho) := H_{\mathrm{dR}}(M(\rho)/\overline{\mathbb{Q}})$ Noting that for $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/E)$ we have a canonical isomorphism

$$H_{\mathrm{dR}}(M^\sigma(\rho)/\overline{\mathbb{Q}}) \cong H_{\mathrm{dR}}(M(\rho)/\overline{\mathbb{Q}}) \otimes_\sigma \overline{\mathbb{Q}},$$

we see that ω_P carries a canonical descent datum to $X = \operatorname{Spec} E$ which defines $P_U(T,h)|_X$. We may also use the reciprocity law to canonically identify the p-adic étale cohomology fibre functors

$$H_{\mathrm{et}} \circ M = \omega_{\mathrm{et}} : \mathrm{Rep}_{\mathbb{Q}_p}(T_{\mathbb{Q}_p}) \to \mathrm{Vec}_{\mathbb{Q}_p},$$

which in particular have an equivariant Γ_E-action coming from ω_{et}.

Now fix an embedding $\overline{\mathbb{Q}} \hookrightarrow \overline{E}_v$, and a faithful representation $T \hookrightarrow \mathrm{GL}(V)$, and s_α cycles on $V^\otimes$ fixed precisely by T. For any $M \in X(\overline{\mathbb{Q}})$, these give rise to absolute Hodge cycles on $H_B(M(V))^\otimes$, which in turn give rise to de Rham cycles $s_{\alpha,\mathrm{dR}} \in H_{\mathrm{dR}}(M(V))^\otimes$ via the Betti–de Rham comparison, and étale cycles $s_{\alpha,et} \in \omega_{\mathrm{et}}(V)^\otimes$. These $s_{\alpha,et}$ are Γ_E-invariant because the action factors

$$\Gamma_E \to U_p \subset T(\mathbb{Q}_p) \to \mathrm{GL}(\omega_{\mathrm{et}}(V))$$

by construction, and as in [Kisin 2010, 2.2.1] this implies the $s_{\alpha,\mathrm{dR}} \in \omega_{P,X,v}(V)^\otimes$ because an absolute Hodge cycle is determined by either component.

Let us fix Y/L with $\overline{\mathbb{Q}} \supset L$ a finite Galois extension of E such that $M(V)$ is realised in the cohomology of Y, and a place $w|v$ of L determining $L_w \subset \overline{E}_v$. By Blasius' theorem on de Rham cycles [1994], the p-adic Hodge theoretic comparison map

$$H_{\mathrm{et}}(M(V), \mathbb{Q}_p)^\otimes \otimes_{\mathbb{Q}_p} B_{\mathrm{dR}} \xrightarrow{\sim} H_{\mathrm{dR}}(M(V)_{L_w}/L_w)^\otimes \otimes_{L_w} B_{\mathrm{dR}}$$

identifies $s_{\alpha,et}$ with $s_{\alpha,\mathrm{dR}}$. Unravelling the definitions, we see that

$$P_U(T,h)_L = \underline{\mathrm{Isom}}_{s_\alpha}(V_L, H_{\mathrm{dR}}(M(V))_L)$$

and

$$P_{\omega_{v,\mathrm{dR}}} \otimes L_w = \underline{\mathrm{Isom}}_{s_\alpha}(V_{L_w}, (\omega_{\mathrm{et}}(V) \otimes_{L_w} B_{\mathrm{dR}})^{\Gamma_{L_w}}).$$

Thus putting it all together we get a canonical[9] identification

$$\omega_{P,X,v} \otimes L_w \cong \omega_{v,\mathrm{dR}} \otimes L_w.$$

We conclude by checking that this descends to E_v. Indeed, Γ_{E_v} acts canonically on both sides of the p-adic comparison map compatibly, so we have an isomorphism of $\mathrm{Gal}(L_w/E_v)$-modules

$$(\omega_{\mathrm{et}}(V) \otimes_{\mathbb{Q}_p} B_{\mathrm{dR}})^{\Gamma_{L_w}} \xrightarrow{\sim} H_{\mathrm{dR}}(M(V)_{L_w}/L_w) = \omega_{P,X,v}(V) \otimes_{E_v} L_w$$

and taking $\mathrm{Gal}(L_w/E_v)$-invariants therefore $\omega_{v,\mathrm{dR}}(V) = \omega_{P,X,v}(V)$. Finally, we have already observed that the $s_{\alpha,et}$ and $s_{\alpha,\mathrm{dR}}$ are Galois invariant, so these too descend to E_v. □

[9] The choice of s_α does not affect the identification. To see this note that it does not change under adding more tensors so given two sets of choices one can just compare both with the union.

3.2. ***Moduli of μ-filtrations of G.*** In §3.2 and §3.3 we make a brief digression from the theory of Shimura varieties to discuss Grassmannians and filtrations more generally. Let R be a domain with fraction field K of characteristic zero, R'/R an étale cover, and G/R a connected reductive group.

3.2.1. Suppose we are given a cocharacter $\mu : \mathbb{G}_{m,R'} \to G_{R'}$. This cocharacter gives us a parabolic subgroup $P_\mu \subset G_{R'}$, which we can view as a point $x_\mu \in \mathrm{Par}_{G/R}(R')$, where we recall [Conrad 2014, 5.2.9] that $\mathrm{Par}_{G/R}$, the functor assigning to each R'/R the set of parabolic subgroups of $G_{R'}$ defined over R', is a proper smooth scheme over R.

After making a base change $R \subset K \to \overline{K}$ we get a well-known finite decomposition

$$\mathrm{Par}_{G/R} \otimes_R \overline{K} = \mathrm{Par}_{G_{\overline{K}}/\overline{K}} = \coprod_i G_{\overline{K}}/P_i$$

where the P_i are representatives of the finitely many $\overline{K}$ conjugacy classes of parabolic subgroups of $G_{\overline{K}}$.

For $P_\mu \subset G_{R'}$ a parabolic subgroup, we denote its conjugacy class by $[P_\mu]$, in a precise sense we will soon make clear. Let us say that $[P_\mu]$ *is defined over* R if there is a component[10] $Z_\mu \subset \mathrm{Par}_{G/R}$ defined over R such that $x_\mu \in Z_\mu(R')$ and $Z_{\mu,\overline{K}}$ is connected.

This definition in a sense is saying that "being étale locally conjugate to P_μ" is a notion that is defined over R, even if P_μ itself is not. More precisely, we have the following.

Lemma 3.2.2. *Let μ be a cocharacter as above defined over R' such that $[P_\mu]$ is defined over R. For S an R-algebra, and $P \subset G_S$ a parabolic subgroup, P is étale locally conjugate to P_μ if and only if $x_P \in \mathrm{Par}_{G/R}(S)$ factors through Z_μ.*

Proof. The statement may be checked étale locally, so we may work over R', at which point by the construction of $\mathrm{Par}_{G/R}$, $G_{R'}/P_\mu \subset Z_{\mu,R'}$ is a union of components. Working over $R' \otimes_R \overline{K}$ using the fact that the fibres of Z_μ over each generic point are connected, we see that in fact $G_{R'}/P_\mu = Z_{\mu,R'}$, so the desired statement follows immediately from [Conrad 2014, 5.2.8]. □

For us an important context where the above holds will be the following.

Proposition 3.2.3. *Assume R is a Dedekind domain with $K = \mathrm{Frac}\, R$ of characteristic zero, G/R connected reductive and L/K a finite extension. Suppose $\mu : \mathbb{G}_{m,L} \to G_L$ is a cocharacter whose conjugacy class is defined over K.*

[10]Here and in much of what follows we use "component" in the relative sense of a "component" of a morphism $X \to Y$ being a Y-subscheme $X' \subset X$ such that the preimage of any connected open subscheme $U \subset Y$ is connected.

Then there exists an étale cover R'/R and a cocharacter $\mu' : \mathbb{G}_{m,R'} \to G_{R'}$ such that if we write $R' \otimes_R K \cong \prod_i L_i$, there are embeddings $L \hookrightarrow L_i$ such that $\mu' \otimes_R K$ is $G(R' \otimes_R K)$-conjugate to $\mu \otimes_L \prod_i L_i$.

Furthermore, this μ' has the property that $[P_{\mu'}]$ is defined over R and independent of the choice of μ'.

Proof. For existence, take R'/R an étale cover which splits G and whose generic points contain L, and fix embeddings $L \hookrightarrow L_i$. Let $T \subset G_{R'}$ be a split maximal torus and let $T_\mu \subset G_{R' \otimes_R K}$ be a maximal torus containing the image of μ which perhaps after enlarging R' we may assume is also split. Then there exists $g \in G\big(\prod_i L_i\big) = \prod_i G(L_i)$ such that $g T_K g^{-1} = T_\mu$ and so

$$\mu' = g^{-1} \mu g \in \mathrm{Hom}_{R' \otimes_R K}(\mathbb{G}_m, T_K) = \mathrm{Hom}_{R'}(\mathbb{G}_m, T)$$

is a cocharacter with the desired property.

Let us next show that $[P_{\mu'}]$ is defined over R. By definition we see that x_μ and $x_{\mu'}$ lie on the same component of $\mathrm{Par}_{G/R} \otimes (R' \otimes_R K)$, whence certainly on that of $\mathrm{Par}_{G/R} \otimes R'$. But since the conjugacy class of μ is defined over K, x_μ lies on a component $Z_{\mu,K} \subset \mathrm{Par}_{G_K/K}$ which is geometrically connected. Letting Z_μ be its closure in (i.e., the corresponding component of) $\mathrm{Par}_{G/R}$, we obtain our witness to the fact that $x_{\mu'}$ is a defined over R. Also since Z_μ is determined by μ, in particular it does not depend on μ'. □

3.2.4. Let S/R be a scheme. We let Vec_S denote the category of vector bundles (projective finitely generated modules) on S, and $\mathrm{Rep}_R(G)$ the category of algebraic representations $G \to \mathrm{Aut}(V)$ for $V \in \mathrm{Vec}_R = \mathrm{Vec}_{\mathrm{Spec}\, R}$.

A *filtered bundle* over S is a vector bundle M/S together with a decreasing complete exhaustive filtration $F^\bullet \subset M$ by flat submodules such that $\mathrm{gr}^\bullet_F M := \bigoplus_p F^p/F^{p+1}$ is flat over S. These form an exact category Fil_S, allowing us to define a *filtered G-bundle* over S to be a faithful exact tensor functor

$$\mathcal{F} : \mathrm{Rep}_R(G) \to \mathrm{Fil}_S\,.$$

We say that $\mathcal{F}$ is a *filtration of G over S* (or just "filtration of G" if no confusion will arise) if its composite with the forgetful functor $\mathrm{Fil}_S \to \mathrm{Vec}_S$ is naturally isomorphic to the usual forgetful functor $\mathrm{Rep}_R(G) \to \mathrm{Vec}_R \xrightarrow{\otimes \mathbb{O}_S} \mathrm{Vec}_S$.

Let $\mu : \mathbb{G}_{m,R'} \to G_{R'}$ be a cocharacter defined over some étale cover R'/R such that $[P_\mu]$ is defined over R. It induces a grading on each $V \in \mathrm{Rep}_R(G) \otimes R'$, each of which in turn gives such V the structure of a filtered bundle, so we can define a canonical filtration of G associated with μ

$$\mathcal{F}_\mu : \mathrm{Rep}_R(G) \xrightarrow{\mathrm{Fil} \circ \mu_*} \mathrm{Fil}_R(G) \otimes R'.$$

3.2.5. Let us say that a filtration $\mathcal{F}$ of G over S is a *μ-filtration* if étale locally on S we have $\mathcal{F} \cong \mathcal{F}_\mu$. We make the necessary remark that of course given $f : S' \to S$, whenever $\mathcal{F}$ is a μ-filtration of G over S, we may take an étale cover $\{U_\alpha \to S\}$ witnessing that $\mathcal{F}$ is a μ filtration and pulling it back along f it will give an étale cover of S' witnessing that $f^* \circ \mathcal{F}$ is a μ-filtration of G over S'.

Thus we have a natural functor on R'-schemes

$$\mathrm{Gr}_{\mu,R'}(S) = \{\mu\text{-filtrations of } G \text{ over } S\}.$$

Proposition 3.2.6. *This has the following properties*:

(1) *The functor* $\mathrm{Gr}_{\mu,R'}$ *is representable by a smooth proper R'-scheme which in fact canonically descends to* Gr_μ /R.

(2) *The scheme* Gr_μ /R *comes equipped with a natural G-action.*

(3) *If R is a field*, *this agrees with the construction of Section 3.1.1.*

Proof. In the usual fashion μ determines a parabolic $P_\mu \subset G_{R'}$, and we claim that $\mathrm{Gr}_{\mu,R'} \cong G/P_\mu$ which in turn is smooth and projective by [Conrad 2014, 5.2.8], and canonically descends by the definition of $[P_\mu]$ being defined over R.

In [ibid.] it is also shown that G/P_μ represents the functor of subgroups of G which are étale locally conjugate to P_μ, so it suffices to check this coincides with our functor. Given a μ-filtration $\mathcal{F}$ over S we can define the subgroup $P_{\mathcal{F}} \subset G$ of elements which preserve $\mathcal{F}$ (if you like, acting on all representations). Since $\mathcal{F}$ is a μ-filtration, étale locally there is an identification of $\mathcal{F}$ with $\mathcal{F}_\mu$, which conjugates $P_{\mathcal{F}}$ onto P_μ. Thus $\mathcal{F} \mapsto P_{\mathcal{F}}$ gives a map $\mathrm{Gr}_\mu(S) \to G/P_\mu(S)$.

Let us construct an inverse. Given P/S étale locally conjugate to $P_{\mu,S}$, after passing to an étale cover $S' \to S$ there exists $g \in G(S')$ such that $P_{S'} = gP_{\mu,S'}g^{-1}$. In particular $P_{S'}$ is a parabolic subgroup of $G_{S'}$, and letting $\mu_P := g \circ \mu$ we see that $P_{S'}$ preserves the filtration $\mathcal{F}_P$ defined by μ_P of G over S'. We must now check this filtration is independent of the choice of g and so in particular is canonical and descends to S.

Suppose we take $h \in G(S')$ such that $hP_{\mu,S'}h^{-1} = gP_{\mu,S'}g^{-1} = P_{S'}$. Then

$$hg^{-1}P_{S'}gh^{-1} = hP_{\mu,S'}h^{-1} = P_{S'}$$

so $hg^{-1} \in N_G(P)(S') = P(S')$, where the final equality is again by [Conrad 2014, 5.2.8]. It follows that $g\mu$ and $h\mu$ induce the same filtration.

Part (2) is now obvious, since G acts by conjugation on G/P_μ, and part (3) is immediate from [Conrad 2014, 5.2.7 (1)] and an easy verification shows that the G-actions agree. □

3.2.7. In the context of Proposition 3.2.3 where we are given a conjugacy class of cocharacters $\mu : \mathbb{G}_{m,L} \to G_L$ defined over $K = \mathrm{Frac}(R)$ and deduce the existence of a μ' defined over an étale cover of R inducing a conjugacy class of parabolics defined over R and hence a $\mathrm{Gr}_{\mu'}/R$ we use the notation $\mathcal{GR}_\mu := \mathrm{Gr}_{\mu'}$ noting the canonical identification

$$\mathcal{GR}_\mu \otimes_R K \cong \mathrm{Gr}_\mu\,.$$

Lemma 3.2.8. *Suppose $G_1 \to G_2$ is a map of connected reductive groups over R and $\mu_1 : \mathbb{G}_m \to G_{1,R'}$ giving a conjugacy class defined over R and inducing $\mu_2 : \mathbb{G}_m \to G_{1,R'} \to G_{2,R'}$. There is a natural map*

$$\mathrm{Gr}_{\mu_1} \to \mathrm{Gr}_{\mu_2}\,.$$

Proof. Given S/R and $\mathcal{F} \in \mathrm{Gr}_{\mu_1}(S)$, recall that $\mathcal{F}$ is specified by a fibre functor

$$\mathrm{Rep}_R(G_1) \to \mathrm{Fil}_S\,.$$

Composing with the restriction map $\mathrm{Rep}_R(G_2) \to \mathrm{Rep}_R(G_1)$, we get a new fibre functor from $\mathrm{Rep}_R(G_2)$ which it is easy to check gives an element of $\mathrm{Gr}_{\mu_2}(S)$, defining the map required. □

3.3. *Filtered G-bundles.* Fix G, μ as above and suppose we have X/R a scheme and $P \to X$ a G-bundle on X. A *μ-filtration* of P is a G-equivariant map of R-schemes

$$\gamma : P \to \mathrm{Gr}_\mu\,.$$

Lemma 3.3.1. *To give a μ-filtration γ on P is to give a fibre functor*

$$\omega_P^\gamma : \mathrm{Rep}_R(G) \to \mathrm{Fil}_X$$

which étale locally is isomorphic to $\mathcal{F}_\mu$ and such that the composite with the forgetful functor

$$\mathrm{Rep}_R(G) \xrightarrow{\omega_P^\gamma} \mathrm{Fil}_X \to \mathrm{Vec}_X$$

is equal to the fibre functor ω_P defined by P.

Proof. Suppose we are given a μ-filtration $\gamma : P \to \mathrm{Gr}_\mu$ of P. Pulling back the universal μ-filtration of G, we obtain a G-equivariant μ-filtration of G over P, which descends to a μ-filtration of G over X, i.e., we obtain a fibre functor

$$\omega_P^\gamma : \mathrm{Rep}_R(G) \to \mathrm{Fil}_X\,.$$

Since it comes from a G-equivariant μ-filtration of G over P, whose forgetful functor to Vec_P by definition is the canonical one from $\mathrm{Rep}_R(G)$ which descends to ω_P, we see that it satisfies the condition in the lemma.

Conversely, if we are given ω_P^γ satisfying the condition, we can pull it back along $P \to X$ to obtain a G-equivariant μ-filtration on P, which is the same as a G-equivariant map $\gamma : P \to \mathrm{Gr}_\mu$. □

We also need the following criterion for extending μ-filtrations over K to μ-filtrations over R.

Lemma 3.3.2. *Suppose we are given $\mu : \mathbb{G}_{m,L} \to G_L$ for $L/K = \mathrm{Frac}(R)$ a finite extension whose conjugacy class is defined over K, X/R a scheme and $P \to X$ a G-bundle. A μ-filtration*

$$\gamma : P_K \to \mathrm{Gr}_\mu$$

extends to a G-equivariant

$$P \to \mathcal{GR}_\mu$$

if for $\rho : G \hookrightarrow \mathrm{GL}(V)$ a faithful representation on V/R finite projective, the bundle $\mathcal{V} := V \times^G P/X$ carries a filtration $\mathcal{V}^\bullet$ (making $\mathcal{V}$ into a filtered bundle in the above sense) extending that of $\mathcal{V}_K = \omega^\gamma_{P_K}(V)$.

Proof. Let R'/R be an étale cover and $\mu : \mathbb{G}_{m,R'} \to G_{R'}$ a cocharacter conjugate to μ as in Proposition 3.2.3. We let $\mathrm{Gr}^G_{\mu'}$ and $\mathrm{Gr}^{\mathrm{GL}(V)}_{\mu'}$ be the two Grassmannians formed from considering $\mu' : \mathbb{G}_m \to G$ and $\mu' : \mathbb{G}_m \to G \xrightarrow{\rho} \mathrm{GL}(V)$ respectively, both defined over R.

By Lemma 3.2.8 we obtain a map

$$\rho_* : \mathrm{Gr}^G_{\mu'} \to \mathrm{Gr}^{\mathrm{GL}(V)}_{\mu'} .$$

We claim this map is a closed immersion, and this suffices to prove the lemma because we are assuming that $\rho_*(\gamma) \in \mathrm{Gr}^{\mathrm{GL}(V)}_{\mu'} \otimes_R K$ extends to a map $\tilde{\gamma} : P \to \mathrm{Gr}^{\mathrm{GL}(V)}_{\mu'}$, and since $\mathrm{Gr}^{\mathrm{GL}(V)}_{\mu'}$ is flat that the ideal sheaf of $\mathrm{Gr}^G_{\mu'}$ is killed by $\tilde{\gamma}^*$ may be checked on the generic fibre.

Since it is clearly proper (as both the source and target are proper), it suffices to show it is a monomorphism, for which it suffices to check the functor of points is injective. But this is obvious: given two μ'-filtrations of G_T for a test scheme T, if they induce the same filtration on V_T then they are equal, since any other representation of G_T can be embedded in a tensor construction on V_T and will have to receive the induced filtration. □

4. Integral Models for the standard principal bundle

4.1. ***Breuil–Kisin modules and lattices in de Rham cohomology.*** We consider the following adaptation of the results of [Kisin 2006], as packaged in [Kisin 2010, 1.2]. Fix $F = \prod_{i=1}^s F_i$ some finite étale algebra over $\mathbb{Q}_p$, and let κ_i be the residue field of F_i, and $E_i(u)$ the monic minimal polynomial of a uniformiser ϖ_i for F_i over $W(\kappa_i)$. Consider the ring $\mathfrak{S} = \mathfrak{S}_F := \prod_i W(\kappa_i)[\![u]\!]$ and $E(u) = (E_1(u), \dots, E_s(u))$. We equip $\mathfrak{S}$ with the Frobenius φ raising $u \mapsto u^p$ and the canonical Frobenius on each $W(\kappa_i)$.

Define the category $\mathrm{Lisse}^{\mathrm{crys}}_{\mathbb{Z}_p}(F)$ to be that of crystalline constant rank lisse $\mathbb{Z}_p$-sheaves on $\operatorname{Spec} F$, i.e., Galois-stable $\mathbb{Z}_p$-lattices in tuples of crystalline p-adic representations $(\sigma_i : \mathrm{Gal}(\overline{F_i}/F_i) \to \mathrm{GL}_n(\mathbb{Q}_p))$. Define the category $\mathrm{Mod}^{\varphi}_{\mathfrak{S}}$ of $\mathfrak{S}$-*modules* to consist of finite free $\mathfrak{S}$-modules $\mathfrak{M}$ together with a φ-semilinear isomorphism

$$1 \otimes \varphi : \varphi^*(\mathfrak{M})[1/E(u)] \xrightarrow{\sim} \mathfrak{M}[1/E(u)].$$

Theorem 4.1.1. *There is a fully faithful tensor functor*

$$\mathfrak{M} : \mathrm{Lisse}^{\mathrm{crys}}_{\mathbb{Z}_p}(F) \to \mathrm{Mod}^{\varphi}_{\mathfrak{S}_F}$$

compatible with the formation of symmetric and exterior powers, unramified base change $F \to F'$ of finite étale algebras over $\mathbb{Q}_p$, and with the property that

$$\mathbb{D}(L) := \varphi^*\mathfrak{M}(L) \otimes_{\mathfrak{S}_F} \mathcal{O}_F \subset D_{\mathrm{dR}}(L \otimes \mathbb{Q}_p)$$

obtained by tensoring along the map $u \mapsto (\varpi_i)$ is a natural $\mathcal{O}_F$-lattice in $D_{\mathrm{dR}}(L \otimes \mathbb{Q}_p)$. Moreover, if $\mathcal{A}$ is an abelian variety over $\mathcal{O}_F$, and $L = T_p(\mathcal{A}_F)^*$, then this lattice is identified with integral de Rham cohomology

$$\begin{array}{ccc} H^1_{\mathrm{dR}}(\mathcal{A}/\mathcal{O}_F) & \xrightarrow{\subset} & H^1_{\mathrm{dR}}(\mathcal{A}_F/F) \\ \| & & \downarrow\wr \\ \mathbb{D}(L) & \xrightarrow{\subset} & D_{\mathrm{dR}}(L \otimes \mathbb{Q}_p). \end{array}$$

Proof. The first part is just [Kisin 2010, 1.2.1] and the second follows perhaps most quickly from [Bhatt et al. 2016, 1.8 (ii)] (although since we are in the case of abelian varieties the theorem probably also can be deduced directly from the theory of Breuil and Kisin). □

4.1.2. For our application to integral models of Shimura varieties over $\mathcal{O}_E[1/N]$, we also need the following abstract lemmas. We thank one of the anonymous referees for pointing out the ideas for the argument for Lemma 4.1.3 in [Maulik 2014, 6.15].

Lemma 4.1.3. (1) *Let A be a Dedekind domain with fraction field K, X/A a smooth scheme and suppose we are given two vector bundles $\mathcal{L}_1$ and $\mathcal{L}_2$ over X together with an identification*

$$\theta : \mathcal{L}_1 \otimes_A K \xrightarrow{\sim} \mathcal{L}_2 \otimes_A K.$$

Suppose further that θ extends as an isomorphism to the formal completion of X at all maximal ideals of A. Then θ extends over X.

(2) *Let $W = W(k)$ for k a perfect field with fraction field K. Suppose X/W is a smooth scheme, with p-adic completion $\hat{X}$, and special fibre X_0, and suppose we are given vector bundles $\mathscr{L}_1$ and $\mathscr{L}_2$ over $\hat{X}$ together with an identification*

$$\theta : \mathscr{L}_1\big[\tfrac{1}{p}\big] \xrightarrow{\sim} \mathscr{L}_2\big[\tfrac{1}{p}\big].$$

Suppose further that there is a Zariski dense subset $U_0 \subset X_0$ of the special fibre such that each $x_0 \in U_0$ admits a lift $\tilde{x}_0 \in \hat{X}(W(\kappa(x_0)))$ with the property that $\tilde{x}_0^*(\theta)$ extends to an isomorphism $\tilde{x}_0^*(\mathscr{L}_1) \xrightarrow{\sim} \tilde{x}_0^*(\mathscr{L}_2)$. Then θ extends over $\hat{X}$.

Proof. Note that in both cases the claim may be checked locally on X, so we may assume $\mathscr{L}_1, \mathscr{L}_2$ are both free and that $X = \operatorname{Spec} R$ is affine, and by picking bases and considering the matrices of both θ and θ^{-1} the question can be reduced to a question about the matrix coefficients.

For (1), we wish to show matrix values lying in $R \otimes_A K$ in fact lie in R. The matrix values all lie in $R[1/D]$ for some $D \in A$, since R is of finite type over A, and the $\mathscr{L}_i$ of finite rank. But by the hypothesis for each maximal ideal $\mathfrak{p}$ containing D, we know θ and θ^{-1} extend over $\hat{X}_{/\mathfrak{p}}$, implying that the matrix values also lie in $R \otimes_A A_{\mathfrak{p}}$, hence they lie in R as required.

For (2), letting $\hat{R}$ be the p-adic completion of R, we are required to show that a matrix coefficient $f \in \hat{R}[1/p]$ in fact lies in $\hat{R}$. Suppose for contradiction it does not, and let $r > 0$ be minimal such that $F = p^r f \in \hat{R}$. The hypothesis tells us that for any $x_0 \in U_0$ and some lift $\tilde{x}_0$ we have that $\tilde{x}_0^*(f) \in W(\kappa(x_0))$. Therefore $\tilde{x}_0^*(F) \in p^r W(\kappa(x_0))$ and in particular $x_0^*(F) = 0$, or $F \in \mathfrak{m}_{x_0} \subset R \otimes_W k$. Since the set of such x_0 is dense, and $R \otimes_W k$ a reduced algebra of finite type over a field, we deduce that the $F = 0$ restricted to the special fibre. I.e. $F = pF'$ for some $F' \in \hat{R}$. But then $p^{r-1} f = F' \in R$ giving the required contradiction. □

Lemma 4.1.4. *Let R be a PID with fraction field K, X/R a flat scheme, G/R a flat affine group scheme of finite type, and P/X_K a G_K-torsor. Suppose we have two pairs $(\mathscr{P}_i, \iota_i)$ $i = 1, 2$ where $\mathscr{P}_i$ is a G-torsor over X and $\iota_i : \mathscr{P}_i \otimes_R K \xrightarrow{\sim} P$ an isomorphism of G_K-torsors over X_K.*

Then the map $\iota_2^{-1}\iota_1 : \mathscr{P}_1 \otimes K \xrightarrow{\sim} \mathscr{P}_2 \otimes K$ extends over R if and only if for every representation $G \to \mathrm{GL}(V)$ with V/R finite free the composite isomorphism

$$\omega_{\mathscr{P}_1}(V) \otimes K \xrightarrow{\iota_{1*}} \omega_P(V_K) \xrightarrow{\iota_{2*}^{-1}} \omega_{\mathscr{P}_2}(V) \otimes K$$

identifies the lattices $\omega_{\mathscr{P}_1}(V)$ and $\omega_{\mathscr{P}_2}(V)$.

Proof. Recall the natural equivalence [Broshi 2013, 1.2] under our hypotheses between the groupoid of G_X-torsors and the groupoid of fibre functors $\operatorname{Rep}(G) \to \operatorname{Vec}(X)$, and that it is functorial in X/R. In particular, $\iota_2^{-1} \circ \iota_1$ extends over R if and only if the composite induced map $\omega_{\mathscr{P}_1} \otimes_R K \xrightarrow{\sim} \omega_{\mathscr{P}_2} \otimes_R K$ on fibre functors extends over R, and this is the case if and only if the condition on lattices holds. □

4.2. ***Special type case.***

4.2.1. Let (T, h) be the Shimura datum defined by a torus split by a CM field, $E_T := E(T, h)$. Then $\mathrm{Sh}(T, h)/E_T$ is a product of algebraic field extensions, and we recall that its integral canonical model $\mathcal{S}(T, h)$ is that obtained by taking integral closures. In particular, recall that if T is unramified at all $p \nmid N$, we get a unique integral model (also abusively written) $T/\mathbb{Z}[1/N]$ and letting $K^N = \prod_{p \nmid N} T(\mathbb{Z}_p)$ we get $\mathcal{S}_{K^N}(T, h)$ indétale over $\mathbb{O}_{E_T}[1/N]$ with a natural $\prod_{p|N} T(\mathbb{Q}_p)$ action extending that on the generic fibre.

Recall that we also have $P_{K^N}(T, h) \to \mathrm{Sh}_{K^N}(T, h)$ the standard principal T^c-bundle, defined over E_T. Our aim here is to construct for it a canonical integral model. We assume henceforth that T is unramified at $p \nmid N$.

4.2.2. Suppose $\mathcal{P}/\mathcal{S}_{K^N}(T, h)$ together with $\iota : \mathcal{P}_{E_T} \xrightarrow{\sim} P_{K^N}(T, h)$ is such a model. We can study the associated fibre functor

$$\omega_{\mathcal{P}} : \mathrm{Rep}_{\mathbb{Z}[1/N]}(T^c) \ni W \mapsto W \times^{T^c} \mathcal{P} \in \mathrm{Vec}(\mathcal{S}_{K^N}(T, h)).$$

The identification ι realises the image of such as lattices

$$\omega_{\mathcal{P}}(W) \subset \omega_{P_{K^N}(T,h)}(W \otimes \mathbb{Q})$$

in the "de Rham sheaves" defined by $P_{K^N}(T, h)$.

On the other hand, for each prime $q \nmid N$, and $v|q$ a place of E_T we can let $K^{Nq} = \prod_{p \nmid Nq} T(\mathbb{Z}_p)$ and it is immediate from the setup and class field theory that the pro-étale $T^c(\mathbb{Z}_q)$-cover $\mathrm{Sh}_{K^{Nq}}(T, h) \to \mathrm{Sh}_{K^N}$ gives rise to, for each representation W_q of $T^c(\mathbb{Z}_q)$, a crystalline lisse $\mathbb{Z}_q$-sheaf $\omega_{\mathrm{et}}(W_q)$ on Sh_{K^N}. By Theorem 4.1.1 we have associated to such a datum a canonical lattice $\mathbb{D}(\omega_{\mathrm{et}}(W_q)) \subset D_{\mathrm{dR}}(\omega_{\mathrm{et}}(W_q) \otimes_{\mathbb{Z}_q} \mathbb{Q}_q)$.

Finally recall Proposition 3.1.8 which gives a natural identification

$$\theta : \omega_{P_{K^N}(T,h)}(W \otimes \mathbb{Q}) \otimes_{E_T} E_{T,v} \xrightarrow{\sim} D_{\mathrm{dR}}(\omega_{\mathrm{et}}(W_q) \otimes_{\mathbb{Z}_q} \mathbb{Q}_q).$$

We say that the model $\mathcal{P}$ is *canonical* if for every $q \nmid N$ and $W \in \mathrm{Rep}_{\mathbb{Z}[1/N]}(T)$ the two lattices $\omega_{\mathcal{P}}(W) \otimes_{\mathbb{O}_{E_T}[1/N]} \mathbb{O}_{E_T,v}$ and $\mathbb{D}(\omega_{\mathrm{et}}(W_q))$ constructed above are, under the map θ, identified.

Proposition 4.2.3. *With* (T, h) *as above, there exists a unique integral canonical model for* $P_{K^N}(T, h)$.

Proof. We first remark that uniqueness follows directly from (4.1.3 (1)) and Lemma 4.1.4.

For existence, let us first remark that we have the map of Shimura data $(T, h) \to (T^c, h^c)$ giving rise to a map of integral canonical models for the Shimura varieties $i : \mathcal{S}_{K^N}(T, h) \to \mathcal{S}_{K^{cN}}(T^c, h^c)$. Suppose $\mathcal{P}^c$ is an integral canonical model for $P_{K^{cN}}(T^c, h^c)$. Then $i^*\mathcal{P}^c$ is an integral canonical model for $P_{K^N}(T, h)$ because

$i^*\omega_{\mathrm{et},(T^c,h^c)} = \omega_{\mathrm{et},(T,h)}$. Since any $h : \mathbb{S} \to T^c_{\mathbb{R}}$ has weight defined over $\mathbb{Q}$, we are therefore reduced to the case where (T, h) is a CM pair (i.e., where T is split by a CM field and the weight of h is defined over $\mathbb{Q}$).

Let $T \to \mathrm{GL}(W)$ be a representation of $T/\mathbb{Z}[1/N]$. Since (T, h) is a CM pair, $\mathrm{Sh}_{K^N}(T, h)$, a union of copies of $\operatorname{Spec} E$ for some field E, can be interpreted as a moduli space of CM motives with level structures. Let $P_{K^N}(T, h)^0 \to \mathrm{Sh}_{K^N}(T, h)^0 \xrightarrow{\sim} \operatorname{Spec} E$ be any single component together with the restriction of $P_{K^N}(T, h)$ over it. There exists a finite extension F/E and a CM abelian variety A/F for such that $\omega_{P_{K^N}(T,h)^0}(W \otimes \mathbb{Q}) \otimes_E F \subset H^1_{\mathrm{dR}}(A/F)^{\otimes}$ and $\omega_{\mathrm{et}}(W) = H^1_{\mathrm{et}}(A, \mathbb{Z}_p)^{\otimes} \cap \omega_{\mathrm{et}}(W) \otimes \mathbb{Q}_p$ for all $p \nmid N'$ some N' divisible by N. Let $\mathcal{A}/\mathcal{O}_F$ be the Neron model of A, and notice that it is an abelian variety over $\mathcal{O}_F[1/M]$ for some M which we may assume is divisible by N'. Thus we construct an $\mathcal{O}_E[1/M]$-lattice

$$\Lambda' := \omega_{P_{K^N}(T,h)^0}(W \otimes \mathbb{Q}) \cap H^1_{\mathrm{dR}}\big(\mathcal{A}/\mathcal{O}_F\big[\tfrac{1}{M}\big]\big)^{\otimes} \subset \omega_{P_{K^N}(T,h)^0}(W \otimes \mathbb{Q}).$$

The second part of Theorem 4.1.1 assures us that at all $p \nmid M$ this lattice agrees with that obtained via first taking the dual Tate module and applying Breuil–Kisin theory.

Moreover, using the construction of Theorem 4.1.1 together with Lemma 2.4.1 for the finitely many p which divide M but do not divide N, we are able to extend Λ' to a lattice Λ over $\mathcal{O}_F[1/N]$. This Λ (applying the construction for all components) gives us the lattice in $\omega_{P_{K^N}(T,h)}(W \otimes \mathbb{Q})$ we require to prove existence. Note that since we already have uniqueness, we may observe that the construction does not depend on the choices of F and A/F. $\square$

4.3. ***Connections on G-bundles.*** It will help to collect some basic facts about connections on G-bundles. Let S be a scheme and G/S a flat affine group scheme of finite type.

Suppose X/S a scheme, and let $\Delta^2(1) = \Delta^2_{X/S}(1)$ be the first order neighbourhood of the diagonal in $X \times_S X$, $\delta : X \hookrightarrow \Delta(1)$ and $p_1, p_2 : \Delta(1) \to X$ the two projection maps. We also will need $\Delta^3(1)$, the first order neighbourhood of the diagonal in $X \times_S X \times_S X$ and its three projections $p_{12}, p_{23}, p_{13} : \Delta^3(1) \to \Delta^2(1)$.

4.3.1. Recall that if we have a vector bundle $\mathcal{V}$ on X a *connection* on $\mathcal{V}/X$ (relative to S) is given by an isomorphism

$$\nabla : p_1^*\mathcal{V} \xrightarrow{\sim} p_2^*\mathcal{V}$$

such that $\delta^*\nabla = \mathrm{id}$ (under the canonical identification $(\delta \circ p_i)^*\mathcal{V} = \mathrm{id}^*\mathcal{V} \cong \mathcal{V}$). Such a connection is said to be *flat* if

$$p_{13}^*(\nabla) = p_{23}^*(\nabla) \circ p_{12}^*(\nabla).$$

It is well known that these definitions are equivalent to the more usual definitions from differential geometry.

4.3.2. Let $P \to X$ be a G_X-torsor. A *connection* on P is an isomorphism

$$\nabla : p_1^* P \xrightarrow{\sim} p_2^* P$$

such that $\delta^* \nabla = \mathrm{id}_P$, and ∇ is said to be flat, again if

$$p_{13}^*(\nabla) = p_{23}^*(\nabla) \circ p_{12}^*(\nabla).$$

Lemma 4.3.3. *Let X/S be a scheme, and assume S is Dedekind. Let $P \to X$ be a G-bundle, with associated fibre functor ω_P. To give the following pieces of data are equivalent*:

(1) *A (flat) connection on $P \to X$.*

(2) *For each representation $V \in \mathrm{Rep}_S(G)$ a (flat) connection on $\omega_P(V)$ in such a way that the isomorphisms*

$$\omega_P(V \otimes W) \cong \omega_P(V) \otimes \omega_P(W)$$

and

$$\omega_P(V^\vee) \cong \omega_P(V)^\vee$$

are isomorphisms of bundles with connection.

(3) *Given a faithful representation $G \hookrightarrow \mathrm{GL}_S(V)$ and tensors $s_\alpha \in V^\otimes$ such that $G = \mathrm{GL}_{S, s_\alpha}(V)$, a (flat) connection on $\omega_P(V)$ with the property that each $\omega_P(s_\alpha \mathcal{O}_S) \subset \omega_P(V^\otimes) = \omega_P(V)^\otimes$ receives the trivial connection.*

Proof. The equivalence of (1) and (2) is formal given Broshi's Tannakian formalism over a Dedekind scheme [Broshi 2013, 1.2]. Indeed the conditions in (2) are exactly those needed to say that the connections $\nabla_{\omega_P(V)}$ define an isomorphism of tensor functors $p_1^* \omega_P \xrightarrow{\sim} p_2^* \omega_P$, which is the same as an isomorphism of torsors $p_1^* P \xrightarrow{\sim} p_2^* P$, and one easily verifies that the conditions translate across.

It is obvious how to pass from (1) to (3). For going from (3) to (1), we note that P can be canonically identified with the frame bundle

$$P \cong \underline{\mathrm{Isom}}_{s_\alpha}(V \otimes \mathcal{O}_X, P \times^G V),$$

(via $p \mapsto (v \mapsto (p, v))$). Let $\mathcal{V} = \omega_P(V) = P \times^G V$ and $s_{\alpha,1} = (1, s_\alpha) \in \mathcal{V}^\otimes$. If we are given a connection $\nabla : p_1^* \mathcal{V} \xrightarrow{\sim} p_2^* \mathcal{V}$, which is trivial on each line containing $s_{\alpha,1}$, that is to say $\nabla(p_1^* s_{\alpha,1}) = p_2^* s_{\alpha,1}$, we obtain a well-defined isomorphism

$$\nabla_P : p_1^* P \ni \phi \mapsto \nabla \circ \phi \in p_2^* P,$$

giving the desired connection on P. Again it is easy to check that if ∇ is flat then so is ∇_P. $\square$

4.4. *Definition of canonical models.* Let $(G, \mathfrak{X})$ be a Shimura datum with reflex $E = E(G, \mathfrak{X})$, and for what follows assume it admits an integral canonical model over $\mathcal{O}_E[1/N]$ at any level hyperspecial away from N.

4.4.1. We shall need the mild assumption that $Z(G)^\circ$ is split by a CM field, which we impose for the rest of the paper. This can be removed if the arguments relying on CM motives in proving Proposition 3.1.8 and Proposition 4.2.3 can be replaced by arguments that work with greater generality.

4.4.2. Recall Section 3.1.2 that the Shimura variety $\mathrm{Sh}(G, \mathfrak{X})$ is equipped with a *standard principal bundle* $P(G, \mathfrak{X})$, which is a $G(\mathbb{A}^\infty)$-equivariant G^c-torsor defined over E, with a flat connection ∇ and a G-equivariant map $\gamma : P(G, \mathfrak{X}) \to \mathrm{Gr}_\mu$. Suppose also G is quasisplit and unramified away from N, fix $G = G_{\mathbb{Z}[1/N]}$ a reductive model, $K^N = \prod_{p \nmid N} G(\mathbb{Z}_p)$ and for μ the Hodge cocharacter note that over an integral model one has the canonical inclusion $\mathrm{Gr}_\mu \subset \mathcal{GR}_\mu$.

We also let

$$\omega_{\mathrm{et}} : \mathrm{Rep}_{\mathbb{Z}_p}(G^c) \to \mathrm{Lisse}_{\mathbb{Z}_p}(\mathrm{Sh}_K(G, \mathfrak{X}))$$

denote the standard étale $\mathbb{Z}_p$-sheaves coming from the tower at p, and $\omega_{\mathrm{et},\eta} : \mathrm{Rep}_{\mathbb{Q}_p}(G^c) \to \mathrm{Lisse}_{\mathbb{Q}_p}(\mathrm{Sh}_K(G, \mathfrak{X}))$ the corresponding lisse $\mathbb{Q}_p$-sheaves.

4.4.3. We will say that a $(G, \mathfrak{X})$ *has enough crystalline points* if for every prime $\mathfrak{p}$ of E with $\mathfrak{p} | p \nmid N$, and every $K_N \subset \prod_{q|N} G(\mathbb{Q}_q)$ compact open, $K = K_N K^N$, there is a dense subset U_0 of the special fibre of $\mathcal{S}_{K_N K^N}/\mathfrak{p}$ with the property that for each $x_0 \in U_0$ there is a lift $\tilde{x}_0 \in \mathcal{S}_{K_N K^N}(W(\kappa(x_0))$ such that $\tilde{x}_0[1/p]^*\omega_{\mathrm{et},\eta}$ takes values in crystalline representations of $\Gamma_{W(\kappa(x_0))[1/p]}$.

We say $(G, \mathfrak{X})$ has *all the crystalline points* if for every $W(k)$-valued point $x \in \mathcal{S}_{K_N K^N}$ for k a finite field of characteristic $p \nmid N$, $x[1/p]^*\omega_{\mathrm{et},\eta}$ takes values in crystalline representations of $\Gamma_{W(k)[1/p]}$. Clearly in the present context where the base is smooth and so every k-point admits a $W(k)$-lift, if you have all the crystalline points you have enough crystalline points.

Moreover, we note that when $(G, \mathfrak{X})$ is of Hodge type or special type it has all the crystalline points. In the Hodge type case, this is because the relevant Galois representations live inside $H^1_{\mathrm{et}}(\mathcal{A}_x[1/p], \mathbb{Q}_p)^\otimes$, where $\mathcal{A}_x$ is the fibre of the universal abelian scheme at x, and since $\mathcal{A}_x/W(\kappa(x))$ is an abelian scheme, this representation is crystalline. In the special type case it follows from the explicit reciprocity law computation as in Section 4.2.2. We shall later use the methods of §4.6 to establish that $(G, \mathfrak{X})$ of abelian type also has all the crystalline points.

In what follows, a *crystalline point* is understood to be, for some K_N and some finite field k of characteristic $p \nmid N$, a point of $\mathcal{S}_{K_N K^N}(W(k))$ such that $x[1/p]^*\omega_{\mathrm{et},\eta}$ takes values in crystalline representations, and any $p, K_N, k, K_0 = W(k)[1/p]$ appearing will be understood to be part of this data.

4.4.4. We will also say that to give $(G, \mathfrak{X})$ a *de Rham structure* is to give for every crystalline point $x \in \mathcal{S}_{K_N K^N}(W(k))$, an identification

$$D_{\mathrm{dR}} \circ x\left[\tfrac{1}{p}\right]^* \omega_{\mathrm{et},\eta} \xrightarrow{\sim} x\left[\tfrac{1}{p}\right]^* \omega_{P_{K_N K^N}(G,\mathfrak{X})} : \mathrm{Rep}_{\mathbb{Q}_p}(G^c) \to \mathrm{Vec}_{K_0} .$$

In the case where $(G, \mathfrak{X})$ is of abelian type (and we expect in general), there is a canonical de Rham structure. This follows using functoriality and the morphism $(G, \mathfrak{X}) \to (G^c, \mathfrak{X}^c)$ to reduce to the case where the weight is defined over $\mathbb{Q}$. Here we may use Milne's moduli interpretation [1994] in such a situation to interpret $x[1/p]$ as an abelian motive, and get the required identification by the same argument as Proposition 3.1.8. It is also straightforward to check that these canonical de Rham structures are compatible under morphisms of Shimura varieties induced by maps of Shimura data. In what follows, we will usually assume $(G, \mathfrak{X})$ is of abelian type in which case we are always working with reference to the canonical de Rham structure.

4.4.5. Let $(G, \mathfrak{X})$ be a Shimura datum as above equipped with a de Rham structure. An *integral canonical model* $\mathcal{P} = \mathcal{P}_{K^N}(G, \mathfrak{X})$ for $P_{K^N}(G, \mathfrak{X})$ is a G^c-torsor over the integral canonical model $\mathcal{S}_{K^N}(G, \mathfrak{X})$ for $\mathrm{Sh}_{K^N}(G, \mathfrak{X})$ with an identification

$$\iota : \mathcal{P}_{K^N}(G, \mathfrak{X}) \otimes_{\mathbb{O}_E[1/N]} E \xrightarrow{\sim} P_{K^N}(G, \mathfrak{X})$$

and equivariant $\prod_{p|N} G(\mathbb{Q}_p)$-action such that we have the *lattice property* given below for any crystalline point $x \in \mathcal{S}_{K_N K^N}(W(k))$.

Consider

$$\omega_{\mathrm{dR},x} : D_{\mathrm{dR}} \circ x\left[\tfrac{1}{p}\right]^* \omega_{\mathrm{et},\eta} : \mathrm{Rep}_{\mathbb{Q}_p}(G^c) \to \mathrm{Vec}_{K_0} .$$

This functor comes with two canonical lattices. First, there is the lattice given by $\omega_{\mathcal{P}_{K_N K^N}}$ coming via ι and the de Rham structure. Second, there is the lattice $\mathbb{D} \circ x[1/p]^* \omega_{\mathrm{et}}$ coming from the theory Theorem 4.1.1 of $\mathfrak{S}$-modules. The lattice property requires that these lattices are equal. Note that since everything is Hecke equivariant, it suffices to check the lattice property at infinite level, an observation of which we shall make liberal use.

We will also insist that for $\mathcal{P}_{K^N}$ to qualify as a canonical model the connection ∇ and the μ-filtration γ extend to $\mathcal{P}$, although we shall see that in the abelian type situation these additional properties are automatic given the condition on crystalline points.

4.4.6. We remark that this definition is compatible with Section 4.2.2. We also remark that one can make definitions of integral canonical models defined over any unramified-away-from-N extension F/E, and for any such F containing the abelian extension over which $\mathrm{Sh}_{K^N}(G, \mathfrak{X})$ splits into its geometrically connected components we may also define integral canonical models for $P_{K^N}(G, \mathfrak{X})^+_F$. These definitions are made in exactly the same way and such models enjoy the properties we are about to note for the same reasons.

Lemma 4.4.7. *Suppose $(G, \mathfrak{X})$ has enough crystalline points. An integral canonical model $(\mathcal{P}, \iota)$, if it exists, is unique up to canonical isomorphism.*

This is true without assuming ∇ or γ extend.

Proof. Suppose we have $(\mathcal{P}, \iota)$ and $(\mathcal{P}', \iota')$, two integral canonical models for $P_{K^N}(G, \mathfrak{X})$. We will aim to show that $\iota'^{-1} \circ \iota : \mathcal{P} \otimes E \xrightarrow{\sim} \mathcal{P}' \otimes E$ extends over $\mathcal{O}_E[1/N]$. By Lemma 4.1.4 it suffices to check that for any representation of $G^c_{\mathbb{Z}[1/N]}$ on a finite free module V that the composite $(\iota'^{-1} \circ \iota)_* : \omega_{\mathcal{P}}(V) \otimes E \xrightarrow{\sim} \omega_{\mathcal{P}'}(V) \otimes E$ identifies the $\mathcal{O}_E[1/N]$-lattice $\omega_{\mathcal{P}}(V)$ with $\omega_{\mathcal{P}'}(V)$.

By Hecke equivariance we may reduce to checking this at every finite level $K_N K^N$. Noting that $\mathcal{S}_{K_N K^N}$ is smooth, we may invoke our lattice Lemma 4.1.3 to reduce the statement first (4.1.3(1)) to checking equality of lattices over the formal completion of $\mathcal{S}_{K_N K^N}$ at each (maximal) prime $\mathfrak{p}$ of $\mathcal{O}_E[1/N]$, and then (4.1.3 (2)) to checking equality of lattices on lifts of a dense subset of the special fibre at $\mathfrak{p}$. But since $(G, \mathfrak{X})$ has enough crystalline points, and by the lattice condition we have the necessary equality at these points, we get the desired identification.

Note that the connection and μ-filtration are extended from the generic fibre and that they extend is a property, so they have no impact on the uniqueness statement. □

The following functoriality properties can now be read off.

Proposition 4.4.8. *Let $f : (G_1, \mathfrak{X}_1) \to (G_2, \mathfrak{X}_2)$ be a morphism of Shimura data with enough crystalline points and compatible de Rham structures induced by a map $G_{1,\mathbb{Z}[1/N]} \to G_{2,\mathbb{Z}[1/N]}$ of reductive groups over $\mathbb{Z}[1/N]$, and $K_i^N = \prod_{p \nmid N} G_i(\mathbb{Z}_p)$ and $(\mathcal{P}_i, \iota_i)$ an integral canonical model for $P_{K_i^N}(G_i, \mathfrak{X}_i)$, $i = 1, 2$.*

(1) *We can canonically identify*

$$\mathcal{P}_1 \times^{G_1^c} G_2^c \xrightarrow{\sim} f^*\mathcal{P}_2.$$

(2) *The induced diagram*

$$\begin{array}{ccc} \mathcal{P}_1 & \xrightarrow{\gamma_1} & \mathcal{GR}_{\mu_1} \\ \downarrow & & \downarrow \\ \mathcal{P}_2 & \xrightarrow{\gamma_2} & \mathcal{GR}_{\mu_2} \end{array}$$

commutes.

Proof. Let $E = E(G_1, \mathfrak{X}_1)$. Note that by Lemma 3.1.6 we have a natural identification

$$\theta : P_{K_1^N}(G_1, \mathfrak{X}_1) \times^{G_1^c} G_2^c \xrightarrow{\sim} f^* P_{K_2^N}(G_2, \mathfrak{X}_2)$$

and so as in the previous lemma it will suffice to check that

$$f^*(\iota_2^{-1}) \circ \theta \circ \iota_1 : (\mathcal{P}_1 \times^{G_1^c} G_2^c) \otimes_{\mathcal{O}_E[1/N]} E \xrightarrow{\sim} (f^*\mathcal{P}_2) \otimes_{\mathcal{O}_E[1/N]} E$$

exchanges the natural lattices when taken on the level of fibre functors

$$\omega : \mathrm{Rep}_{\mathbb{Z}[1/N]} G_2^c \to \mathrm{Vec}(\mathcal{S}_{K_1^N}(G_1, \mathfrak{X}_1)).$$

As in the previous lemma, we may reduce to working at finite level and checking equality at crystalline points. But this is then immediate from the lattice condition, the compatibility of the étale fibre functors on both Shimura varieties, together with the (consequent) observation that the image of a crystalline point is always a crystalline point. This completes the proof of (1). Now (2) follows directly from the fact that after taking $\otimes_{\mathcal{O}_E[1/N]} E$ the diagram commutes by Lemma 3.1.6. Note that the right hand vertical map is that given by Lemma 3.2.8. □

Lemma 4.4.9. *Suppose $F/E = E(G, \mathfrak{X})$ a Galois extension unramified away from N, and $(\mathcal{P}'/\mathcal{O}_F[1/N], \iota')$ an integral canonical model for $P_{K^N}(G, \mathfrak{X})_F$.*

Then the descent data

$$\mathrm{Gal}(F/E) \ni \sigma \mapsto \theta_\sigma : \sigma^* P_{K^N}(G, \mathfrak{X})_F \xrightarrow{\sim} P_{K^N}(G, \mathfrak{X})_F$$

extend to $\mathcal{P}'$, and the pair $(\mathcal{P}, \iota)$ obtained by étale descent is an integral canonical model for $P_{K^N}(G, \mathfrak{X})$ over $\mathcal{O}_E[1/N]$.

Proof. Take $\sigma \in \mathrm{Gal}(F/E)$. We claim that $\sigma^*\mathcal{P}'_{K^N}$ together with the composite

$$\sigma^*\mathcal{P}'_{K^N} \otimes F \xrightarrow{\sigma^*\iota} \sigma^* P_{K^N}(G, \mathfrak{X})_F \xrightarrow{\sigma} P_{K^N}(G, \mathfrak{X})_F$$

is an integral canonical model. From this claim it is immediate that the descent data extend and in its turn $\mathcal{P}'_{K^N}$ descends to an integral canonical model over $\mathcal{O}_E[1/N]$, because being a crystalline point and the lattices coming from $\mathfrak{S}$-modules are stable under finite unramified base change.

We may pull back the Hecke action, so it remains to check the lattice condition on crystalline points at finite level. Let $x \in \mathcal{S}_{K_N K^N}(G, \mathfrak{X})(W(\kappa))$ be a crystalline point and v the corresponding place of F, and $\sigma^*(x)$ its pullback along $\sigma : \mathcal{S}_{K_N K^N} \xrightarrow{\sim} \mathcal{S}_{K_N K^N}$. Let $\omega_{\mathrm{et},x}$ and $\omega_{\mathrm{et},\sigma^*(x)}$ be the $\mathbb{Z}_p$-linear étale fibre functors coming from the Shimura variety at infinite level at p and restricting to the fibres at $x[1/p]$ and $\sigma^*(x)[1/p]$ respectively, viewing both as taking values in Γ_{K_0}-representations. Note that by construction (since $\mathrm{Sh}(G, \mathfrak{X})$ is defined over E) there is an isomorphism $\omega_{\mathrm{et},\sigma^*(x)} \xrightarrow{\sim} \omega_{\mathrm{et},x}$ covering $\sigma : \sigma^* x \xrightarrow{\sim} x$. In particular, $\sigma^*(x)$ is also crystalline.

For $V \in \mathrm{Rep}_{\mathbb{Z}_p}(G^c)$ it is necessary to check an equality of two lattices in $D_{\mathrm{dR}}(\omega_{\mathrm{et},x}(V)[1/p])$. The first is the usual $\mathbb{D}(\omega_{\mathrm{et},x}(V))$ coming straight from $\mathfrak{S}$-modules, the second obtained as the composite

$$\mathbb{D}(\omega_{\mathrm{et},\sigma^*(x)}(V)) \subset D_{\mathrm{dR}}\big(\omega_{\mathrm{et},\sigma^*(x)}(V)\big[\tfrac{1}{p}\big]\big) \xrightarrow{\sim} D_{\mathrm{dR}}\big(\omega_{\mathrm{et},x}(V)\big[\tfrac{1}{p}\big]\big).$$

Having thus spelt it out, we see that it follows immediately from functoriality of the $\mathbb{D}$ construction. Moreover, the G^c-action, connection and filtration are all defined over E on the generic fibre and it can be checked they extend to $\mathcal{O}_E[1/N]$ étale locally, so the fact they do over F immediately gives them all. $\square$

4.5. ***Hodge type case.*** We turn our attention to existence.

Theorem 4.5.1. *Let $(G, \mathfrak{X})$ be a Shimura variety of Hodge type with G unramified away from N and $K^N = \prod_{p\nmid N} G(\mathbb{Z}_p)$. Then $P_{K^N}(G, \mathfrak{X})$ admits an integral canonical model.*

We let $E = E(G, \mathfrak{X})$ throughout this paragraph, and unless otherwise specified all Shimura varieties Sh_K, $\mathcal{S}_K$ or bundles P_K or $\mathcal{P}_K$ will be understood to be those attached to the Shimura datum $(G, \mathfrak{X})$. We also note that $(G, \mathfrak{X})$ being of Hodge type implies $G = G^c$.

4.5.2. First some more algebraic preliminaries. Let R be a ring, M a finite free R-module, and $M^\otimes$ the direct sum of all R-modules formed from M by the operations of taking duals, tensor products, symmetric powers and exterior powers. If $s_\alpha \in M^\otimes$ is a collection of tensors, we say that they *define* a subgroup $G \subset \mathrm{GL}(M)$ if it is precisely the group which acts trivially on all the s_α.

We need the following version of [Kisin 2010, 1.3.2], whose proof is basically identical.

Proposition 4.5.3. *Let R be a PID with field of fractions K, M a finite free R-module, and $G \subset \mathrm{GL}(M)$ a closed subgroup, flat over R with reductive generic fibre. Then it is defined by a finite collection of tensors s_α.*

Proof. Exactly as in [Kisin 2010, 1.3.2] we can reduce to showing that it suffices to find tensors defining G in some representation of $\mathrm{GL}(M)$ on a finite projective R-module, and we take $I \subset \mathcal{O}_{\mathrm{GL}(M)}$ the ideal of G in the Hopf algebra of $\mathrm{GL}(M)$, which has scheme-theoretic stabiliser precisely G.

By [Waterhouse 1979, 3.3] there is a finite dimensional $\mathrm{GL}(M)_K$-stable subspace $W_\eta \subset \mathcal{O}_{\mathrm{GL}(M)} \otimes K$ containing a finite set of generators for I as an ideal. Let $W = W_\eta \cap \mathcal{O}_{\mathrm{GL}(M)}$, which is visibly $\mathrm{GL}(M)$-stable, and it is finitely generated because $\mathcal{O}_{\mathrm{GL}(M)}$ is a free R-module and letting $e_1, \dots, e_n$ be a basis for W_η, writing them as elements of $\mathcal{O}_{\mathrm{GL}(M)} \otimes K$ we see they lie in $\tilde{W} \otimes K$ for some $\tilde{W} \subset \mathcal{O}_{\mathrm{GL}(M)}$ a finite free R-submodule, and such that $W \subset \tilde{W}$. But R is a PID, so we deduce W is finite free.

Now we see that G is the stabiliser of $W \cap I \subset W$, and the argument can be concluded exactly as in [Kisin 2010, 1.3.2]. $\square$

4.5.4. Now recall Section 2.4.4 that for $(G, \mathfrak{X})$ our Shimura datum of Hodge type with G unramified away from $\mathbb{Z}[1/N]$ (taking $G_{\mathbb{Z}[1/N]} = \mathcal{G}$ from the discussion in

Section 2.4.4) we may choose a symplectic embedding $i : (G, \mathfrak{X}) \hookrightarrow (\mathrm{GSp}_{2g}, S^{\pm})$ extending to a closed immersion

$$i : G_{\mathbb{Z}[1/N]} \hookrightarrow \mathrm{GL}(V_{\mathbb{Z}[1/N]}).$$

Henceforth view V as a $\mathbb{Z}[1/N]$-module, and G as a reductive group over $\mathbb{Z}[1/N]$. By the above proposition we can find $s_\alpha \in V^{\otimes}$ such that $G = \mathrm{Aut}_{s_\alpha}(V)$.

4.5.5. The map i defines (pulling back the universal abelian variety) an abelian variety $\pi : A_{K^N} \to \mathcal{S}_{K^N}$, and we can study its sheaf of relative de Rham cohomology $\mathcal{V} = \mathcal{H}^1_{\mathrm{dR}}(A_{K^N}/\mathcal{S}_{K^N})$ with Gauss–Manin connection ∇.

By the absolute Hodge cycles argument of [Kisin 2010, 2.2.2] we obtain ∇-horizontal $\prod_{p|N} G(\mathbb{Q}_p)$-invariant tensors $s_{\alpha,\mathrm{dR}} \in \mathcal{V}^{\otimes} \otimes_{\mathbb{Z}[1/N]} \mathbb{Q}$, and by [ibid., 2.3.9] for any $v \nmid N$ a finite place of E these sections extend over $\mathcal{O}_{E,(v)}$, which is enough to conclude they in fact lie in $\mathcal{V}^{\otimes}$.

Let us therefore define the sheaf on $(\mathcal{S}_{K^N})_{\mathrm{fppf}}$

$$\mathcal{P}_{K^N} := \underline{\mathrm{Isom}}_{s_\alpha}(V, \mathcal{V})$$

whose sections over $u : U \to \mathcal{S}_{K^N}$ are those trivialisations $\theta : V \otimes_{\mathbb{Z}[1/N]} \mathcal{O}_U \xrightarrow{\sim} u^*\mathcal{V}$ such that (after applying $(-)^{\otimes}$ to both sides) each s_α is identified with $s_{\alpha,\mathrm{dR}}$. Note that since $\prod_{p|N} G(\mathbb{Q}_p)$ acts equivariantly on $\mathcal{V}$ and fixes $s_{\alpha,\mathrm{dR}}$ we may freely pass between infinite and finite level.

Proposition 4.5.6. *The sheaf $\mathcal{P}_{K^N}$ is a G-torsor over $\mathcal{S}_{K^N}$.*

Proof. It obviously suffices to check at finite level $K = K_N K^N$ for some sufficiently small K_N. We claim it also suffices to check in the formal neighbourhood of any closed point $x \in \mathcal{S}_K$.

First we may directly verify that $\mathcal{P}_K$ is representable because passing to a Zariski cover making $\mathcal{V}$ free, and then choosing arbitrary identifications $V \cong \mathcal{V}$, $\mathcal{P}_K$ can be constructed as a closed (hence affine) subvariety of $\mathrm{GL}(V)$. Now suppose we know it is a G-torsor in a formal neighbourhood of every closed point. Then since G is faithfully flat, $\mathcal{P}_K$ is faithfully flat over the formal neighbourhood of every closed point, which is enough to prove that $\mathcal{P}_K/\mathcal{S}_K$ is faithfully flat. It remains to show that the natural map

$$\mathcal{P}_K \times_{\mathcal{S}_K} G_{\mathcal{S}_K} \ni (p, g) \mapsto (p, pg) \in \mathcal{P}_K \times_{\mathcal{S}_K} \mathcal{P}_K$$

is an isomorphism, but this also is the case if and only if it is the case over a formal neighbourhood of each closed point, where again it follows from our assumption.

Thus we are reduced to the study of $\hat{P}_x := \mathcal{P}_K|_{\hat{\mathcal{S}}_{K,x}}$ for closed points $x \in \mathcal{S}_K$. If $\mathrm{char}(x) = 0$, the Betti–de Rham comparison theorem gives an identification $V_{\hat{\mathcal{S}}_{K,x}} \otimes \mathbb{C} \xrightarrow{\sim} \hat{\mathcal{V}}_x \otimes \mathbb{C}$, giving a fpqc-local section. This suffices because it shows

$\hat{P}_x$ is smooth by fpqc descent, and so $\hat{P}_x$ admits a section étale locally, which is enough to see it is a G-torsor.

If $\mathrm{char}(x) = p$, we may follow the proof of [Kisin 2010, 2.3.5] to study $N_x := \mathcal{O}_{\hat{\mathcal{S}}_{K,x}}$ together with the p-divisible group $\mathcal{G} = A_K|_{\hat{\mathcal{S}}_{K,x}}[p^\infty]$ over the formally smooth ring N_x.

First note that x determines a place v of E above p and since E is unramified at p we may identify $\mathcal{O}_{E,v} = W(\kappa(x)) =: W$, over which N_x is canonically an algebra. The crystalline-de Rham comparison gives an identification $\mathcal{V}_{N_x} = \mathbb{D}(\mathcal{G})(N_x)$. We let $\mathcal{G}_0$ denote the restriction of $\mathcal{G}$ to x.

Taking $\tilde{x} : N_x \to W$ a lift, we may follow the argument of [ibid., 2.3.5] to obtain from the Tate module of $\tilde{x}^*\mathcal{G}$ tensors $s_{\alpha,0} \in \mathbb{D}(\mathcal{G}_0)(W)^{\otimes}$ defining a reductive subgroup $G(\tilde{x}) \subset \mathrm{GL}(\mathbb{D}(\mathcal{G}_0)(W))$, which gives rise to an explicit versal deformation ring $R_{G(\tilde{x})}$ of $G(\tilde{x})$-adapted deformations of $\mathcal{G}_0$. This then has the property that we may identify $R_{G(\tilde{x})} \xrightarrow{\sim} N_x$ in such a way that induces an identification of $\mathcal{G}$ with the explicit versal deformation (let's also call it $\mathcal{G}$) one constructs over $R_{G(\tilde{x})}$, so in particular we have

$$\mathcal{V}_{N_x} = \mathbb{D}(\mathcal{G})(R_{G(\tilde{x})}).$$

As in [ibid., 2.3.9] one has by an explicit construction lifts $\tilde{s}_{\alpha,0}$ to $\mathbb{D}(\mathcal{G})(R_{G(\tilde{x})})^{\otimes}$ of $s_{\alpha,0}$ which are identified with $s_{\alpha,\mathrm{dR}} \in \mathcal{V}_{N_x}^{\otimes}$. This explicit construction comes about in [ibid., §1.5] by defining $\mathbb{D}(\mathcal{G})(R_{G(\tilde{x})}) = \mathbb{D}(\mathcal{G}_0)(W) \otimes_W R_{G(\tilde{x})}$ as a module and taking the lift $\tilde{s}_{\alpha,0} = s_{\alpha,0} \otimes 1$. Finally let us note that by [ibid., 1.4.3] (since G is connected) there exists a W-linear isomorphism

$$V_W \xrightarrow{\sim} \mathbb{D}(\mathcal{G}_0)(W)$$

identifying s_α with $s_{\alpha,0}$. In particular this identifies $G_W \xrightarrow{\sim} G(\tilde{x})$ uniquely up to inner automorphisms and the composite

$$V_W \otimes_W R_{G(\tilde{x})} \xrightarrow{\sim} \mathbb{D}(\mathcal{G}_0)(W) \otimes_W R_{G(\tilde{x})} \xrightarrow{\sim} \mathbb{D}(\mathcal{G})(R_{G(\tilde{x})}) = \mathcal{V}_{N_x}$$

gives a section of $\hat{P}_x$ as required. □

4.5.7. We would now like to show it is the desired canonical model for $P_{K^N}(G, \mathfrak{X})$. Firstly we must check that the additional structures extend.

We have already noted that the $\prod_{p|N} G(\mathbb{Q}_p)$-action extends. By Lemma 4.3.3, the connection ∇ on P_{K^N} extends to $\mathcal{P}_{K^N}$ because the Gauss–Manin connection extends to (i.e., can be constructed directly on) $\mathcal{V} = \mathcal{H}^1_{\mathrm{dR}}(\mathcal{A}_{K^N}/\mathcal{S}_{K^N})$ and by the construction of $s_{\alpha,\mathrm{dR}}$ in [Kisin 2010, 2.2] it is clear they are parallel sections. It is also immediate the μ-filtration extends, by Lemma 3.3.2 and that the Hodge filtration is naturally defined on the integral de Rham cohomology $\mathcal{V}$.

Finally, we need to verify that it is indeed a canonical model, and as usual we check the lattice condition at a crystalline point x. But in our situation this is very straightforward: we just use the facts that, taking $\mathscr{A}_x$ to be the pullback of the universal abelian scheme to x, that Theorem 4.1.1 $\mathbb{D}(H^1_{\text{et}}(\mathscr{A}_x, \mathbb{Z}_p)) = H^1_{\text{dR}}(\mathscr{A}_x/W(k))$ and that by compatibility of $\mathbb{D}$ with the p-adic comparison theorems, $\mathbb{D}(x^* s_{\alpha,et}) = x^* s_{\alpha,\text{dR}}$.

4.5.8. For making the transition from Hodge to abelian type, we will need to extend the equivariant $G^c \times \prod_{p|N} G(\mathbb{Q}_p)$-action to an equivariant action of the $\mathbb{Z}^N = \mathbb{Z}[1/N]$-group scheme

$$\mathscr{A}^N_P(G) = \left(G^c \times \frac{\prod_{p|N} G(\mathbb{Q}_p)}{\overline{Z(\mathbb{Z}^N)}} \right) *_{G(\mathbb{Z}^N)_+/Z(\mathbb{Z}^N)} G^{\text{ad}}(\mathbb{Z}^N)^+$$

which sits in an exact sequence

$$1 \to G^c \to \mathscr{A}^N_P(G) \to \mathscr{A}^N(G) \to 1.$$

To do this it suffices to give the $G^{\text{ad}}(\mathbb{Z}^N)^+$-action and then check it is compatible, for which we imitate the approach taken in [Kisin 2010, 3.2] and [Kissin and Pappas 2015, 4.5], with our modifications from Section 2.5.6 and further slight modifications.

To be precise, recall that a point $x \in \mathscr{P}_{K^N}(T)$ gives the data of an quadruple $(A, \lambda, \epsilon_{\text{et}}, \epsilon_{\text{dR}})$ where A/T is an abelian scheme up to an isogeny whose degree is supported on N, λ a weak polarisation of A, ϵ_{dR} a section of $\underline{\text{Isom}}_{s_\alpha}(V_T, \mathscr{V}_T)$, and $\epsilon_{\text{et}} \in \Gamma(T, \underline{\text{Isom}}_{s_\alpha}(V_{\prod_{p|N} \mathbb{Q}_p}, \prod_{p|N} \hat{V}_p(A)))$.

Given $\gamma \in G^{\text{ad}}(\mathbb{Z}^N)^+$, we may form the Z^{der}-torsor $\mathscr{P} = \{g \in G^{\text{der}} | \pi(g) = \gamma\}$. We may take $F/\mathbb{Q}$ Galois such that $\mathscr{P}(\mathbb{O}_F[1/N])$ is nonempty, and take $\tilde{\gamma} \in \mathscr{P}(\mathbb{O}_F[1/N])$. This can be used to define an action just as in [Kisin 2010, 3.2]. Indeed we can define

$$A^{\mathscr{P}}(T) = (A(T) \otimes_{\mathbb{Z}[1/N]} \mathbb{O}_{\mathscr{P}})^{Z^{\text{der}}}$$

and specialise the map

$$A^{\mathscr{P}} \otimes_{\mathbb{Z}[1/N]} \mathbb{O}_{\mathscr{P}} \xrightarrow{\sim} A \otimes_{\mathbb{Z}[1/N]} \mathbb{O}_{\mathscr{P}}$$

at $\tilde{\gamma}$ to obtain

$$\iota_{\tilde{\gamma}} : A^{\mathscr{P}} \otimes_{\mathbb{Z}[1/N]} \mathbb{O}_F\left[\tfrac{1}{N}\right] \xrightarrow{\sim} A \otimes_{\mathbb{Z}[1/N]} \mathbb{O}_F\left[\tfrac{1}{N}\right].$$

These allow one to give the action of γ as taking

$$(A, \lambda, \epsilon_{\text{et}}, \epsilon_{\text{dR}}) \mapsto (A^{\mathscr{P}}, \lambda^{\mathscr{P}}, \epsilon^{\mathscr{P}}_{\text{et}}, \epsilon^{\mathscr{P}}_{\text{dR}}),$$

where all but the last of these are as defined by Kisin. For the action on the de Rham component, we need the following modification of [Kisin 2010, 3.2.5].

Lemma 4.5.9. *With notation as above, the composite*

$$V_T \otimes_{\mathbb{Z}^N} \mathcal{O}_F^N \xrightarrow{\tilde{\gamma}^{-1}} V_T \otimes_{\mathbb{Z}^N} \mathcal{O}_F^N \xrightarrow{\epsilon_{\mathrm{dR}}} H^1_{\mathrm{dR}}(A/T) \otimes_{\mathbb{Z}^N} \mathcal{O}_F^N \xrightarrow{\iota_{\tilde{\gamma}}^{-1}} H^1_{\mathrm{dR}}(A^{\mathcal{P}}/T) \otimes_{\mathbb{Z}^N} \mathcal{O}_F^N$$

is $\mathrm{Gal}(F/\mathbb{Q})$*-invariant, hence induces a section*

$$\epsilon^{\mathcal{P}}_{\mathrm{dR}} : V_T \xrightarrow{\sim} H^1_{\mathrm{dR}}(A^{\mathcal{P}}/T)$$

which takes s_α to $s_{\alpha,\mathrm{dR}}$.

Proof. The cocycle computation argument of [ibid., 3.2.5] goes through unchanged. For the claim about tensors, the only nonobvious point is that $\iota_{\tilde{\gamma}}^{-1}$ preserves the tensors, which one checks directly by working over $\mathbb{C}$, using the comparison with Betti cohomology as in [ibid., 3.2.6]. □

It is also clear that this action is compatible with the $G^c \times \prod_{p|N} G(\mathbb{Q}_p)/\overline{Z(\mathbb{Z}^N)}$-action in the sense that if we are given $g \in G^c(\mathbb{Z}^N)_+$ we may take $F = \mathbb{Q}$ and $\tilde{\gamma} = \gamma = g^{-1}$ and postcompose the whole quadruple by the quasiisogeny ι_γ to identify

$$(A, \lambda, \epsilon_{\mathrm{et}} \circ g, \epsilon_{\mathrm{dR}} \circ g) = (A^{\mathcal{P}}, \lambda^{\mathcal{P}}, \epsilon^{\mathcal{P}}_{\mathrm{et}}, \epsilon^{\mathcal{P}}_{\mathrm{dR}}).$$

4.6. ***Some distinguished Shimura data.*** In this section we digress a little. To get from the Hodge type to the abelian type case we would like to follow Deligne [1979] and pass via connected components, but unfortunately a straightforward reduction of $\mathcal{P}$ to the derived group is not possible to carry out over $\mathcal{O}_E[1/N]$ (roughly speaking because carrying out such a construction requires "trivialising the G^{ab} part of the motivic structure", which can only be done canonically over $\mathbb{C}$ using Betti–de Rham comparison). We address this by constructing some distinguished Shimura data $\mathcal{B}(G^{\mathrm{der}}, \mathfrak{X}^+, E)$ which are in some sense initial among Shimura data whose connected Shimura datum is $(G^{\mathrm{der}}, \mathfrak{X}^+)$ and whose reflex field is contained in E. Pulling back our torsors from the Hodge type case we may then reduce along $\mathcal{B} \to G$ while leaving all motivic structures intact.

4.6.1. Let $(G, \mathfrak{X})$ be a Shimura datum, and E its reflex field. Our goal is to construct a new Shimura datum $(\mathcal{B}, \mathfrak{X}_{\mathcal{B}}) \to (G, \mathfrak{X})$ with $\mathcal{B}^{\mathrm{der}} = G^{\mathrm{der}}$ and reflex field E and show it depends only on these data and $\mathfrak{X}^+$ and not on G.

Since E is the reflex field, we obtain a well-defined

$$\mu_E : \mathbb{G}_{m,E} \to G^{\mathrm{ab}}_E$$

taking the "determinant" (composite with $G \to G^{\mathrm{ab}}$) of any Hodge cocharacter. Taking Weil restriction along $E/\mathbb{Q}$ and the norm map, we get a composite

$$r : E^* := \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{m,E} \to \mathrm{Res}_{E/\mathbb{Q}} G^{\mathrm{ab}}_E \to G^{\mathrm{ab}}.$$

Define $\mathcal{B} = G \times_{G^{\mathrm{ab}}} E^*$, where the first map is the natural $\delta : G \to G^{\mathrm{ab}}$. It is straightforward to check the following.

Lemma 4.6.2. *The scheme $\mathcal{B}$ is a reductive group scheme over $\mathbb{Q}$, sitting in a short exact sequence of $\mathbb{Q}$-group schemes*

$$0 \to G^{\mathrm{der}} \to \mathcal{B} \to E^* \to 0.$$

Suppose now that $(G, \mathfrak{X})$ induces a connected Shimura datum $(G^{\mathrm{der}}, \mathfrak{X}^+)$ with reflex field $E(G^{\mathrm{der}}, \mathfrak{X}^+) := E(G^{\mathrm{ad}}, \mathfrak{X}^{\mathrm{ad}}) \subset E$. We could make the above construction varying over all Shimura data with connected datum $(G^{\mathrm{der}}, \mathfrak{X}^+)$ and reflex field contained in E. The following is an adaptation of the argument in [Deligne 1979, 2.5].

Proposition 4.6.3. *The extension*

$$0 \to G^{\mathrm{der}} \to \mathcal{B} \to E^* \to 0$$

depends only on the connected Shimura datum $(G^{\mathrm{der}}, \mathfrak{X}^+)$ and the field E.

Proof. Let $(G, \mathfrak{X})$ and $(G', \mathfrak{X}')$ be two Shimura data whose reflex fields are both contained in E and which give rise to the same connected Shimura datum. By [Deligne 1979, 2.5.6] we can find a third such Shimura datum $(G'', \mathfrak{X}'')$ which admits maps

$$(G, \mathfrak{X}) \leftarrow (G'', \mathfrak{X}'') \to (G', \mathfrak{X}').$$

It will therefore suffice to show for any morphism $\alpha : (G, \mathfrak{X}, E) \to (G_2, \mathfrak{X}_2, E_2)$ of Shimura data with a field of definition (where $E \to E_2$ represents an inclusion $E_2 \subset E$) that we have an induced natural morphism of extensions

$$\begin{tikzcd}
0 \arrow[r] & G^{\mathrm{der}} \arrow[r] \arrow[d] & \mathcal{B} \arrow[r] \arrow[d] & E^* \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & G_2^{\mathrm{der}} \arrow[r] & \mathcal{B}_2 \arrow[r] & E_2^* \arrow[r] & 0
\end{tikzcd}$$

where the outer maps are the natural ones induced by $\alpha : G \to G_2$ and $N_{E/E_2} : E^* \to E_2^*$. Since in the above special case both these maps are isomorphisms, we will get a canonical identification

$$\mathcal{B} \xleftarrow{\sim} \mathcal{B}'' \xrightarrow{\sim} \mathcal{B}'$$

of extensions, proving the proposition.

Let us verify the claim. Recall that to give a map $\mathcal{B} \to \mathcal{B}_2$ is to give maps $f : \mathcal{B} \to G_2$ and $k : \mathcal{B} \to E_2^*$ that agree when extended to G_2^{ab}. Write $N\mu : E^* \to G^{\mathrm{ab}}$ and $\delta : G \to G^{\mathrm{ab}}$ and $N\mu_2$ and δ_2 for the analogues for G_2. Let us take f to be the natural composite

$$f : \mathcal{B} \xrightarrow{\mathrm{pr}_G} G \to G_2$$

and k to be

$$k : \mathcal{B} \xrightarrow{\mathrm{pr}_{E^*}} E^* \xrightarrow{N_{E/E_2}} E_2^*.$$

We must check that whenever $\delta(g) = N\mu(e)$, we have

$$\delta_2(\alpha(g)) = N\mu_2(N_{E/E_2}(e)).$$

But by functoriality of $(-)^{\mathrm{ab}}$

$$\delta_2(\alpha(g)) = \alpha^{\mathrm{ab}}(\delta(g)) = \alpha^{\mathrm{ab}}(N\mu(e)) = N\mu_2(N_{E/E_2}(e)),$$

where the final equality holds because since α is a morphism of Shimura data we have in particular that $\alpha^{\mathrm{ab}}\mu = \mu_2 : \mathbb{G}_{m,\mathbb{C}} \to G^{\mathrm{ab}}_{2,\mathbb{C}}$, from which it follows that $\alpha^{\mathrm{ab}} \circ N\mu = N\mu_2 \circ N_{E/E_2}$ by functoriality of the norm map between tori. Thus we have a map $\mathcal{B} \to \mathcal{B}_2$, and it is clear from its definition that it fits into the diagram described. □

4.6.4. Our next task is to define a Shimura datum. Take any $h_G \in \mathfrak{X}^+$, and we define a canonical map

$$h_E : \mathbb{C}^* \to E^*(\mathbb{R})$$

as follows. Let $\tau : E \hookrightarrow \mathbb{C}$ be the canonical inclusion of the reflex field E into $\mathbb{C}$, and write $E^*(\mathbb{R}) = E^*_\tau \times \prod_{\tau' \neq \tau} E^*_{\tau'}$. If τ is real, set

$$h_E(z) = (z\bar{z}; 1, 1, \dots, 1).$$

If τ is complex,

$$h_E(z) = (z; 1, \dots, 1)$$

(i.e., the entries away from the place τ are all trivial in both cases).

Proposition 4.6.5. *These give a map $h_G \times h_E : \mathbb{S} \to \mathcal{B}_\mathbb{R}$ which defines a Shimura datum with a natural map $(\mathcal{B}, \mathfrak{X}_\mathcal{B}) \to (G, \mathfrak{X})$. Moreover, it is independent of the choice of h_G, has reflex field E, and only depends on E and the connected Shimura datum $(G^{\mathrm{der}}, \mathfrak{X}^+)$ up to canonical isomorphism.*

Proof. We must check it is defined, which amounts to proving that $r_\mathbb{R}(h_E(z)) = \delta(h_G(z))$. First we note that

$$\delta(h_G(z)) = \mu_E(z)\mu_E(\bar{z}) = \mu_E(z)\overline{\mu_E(z)}$$

where we view μ_E as a character $\mathbb{G}_{m,\mathbb{C}} \to G^{\mathrm{ab}}_\mathbb{C}$ using the canonical embedding τ of E.

On the other hand recall that

$$r_\mathbb{R} = (N_{E/\mathbb{Q}} \circ \mu_E)_\mathbb{R} : E^*_\mathbb{R} \to G^{\mathrm{ab}}_{E\otimes_\mathbb{Q}\mathbb{R}} \to G^{\mathrm{ab}}_\mathbb{R}.$$

If τ is real then we have

$$(z\bar{z}, 1, 1, \dots, 1) \mapsto (\mu_E(z\bar{z}), 1, \dots, 1) \mapsto \mu_E(z\bar{z}) = \mu_E(z)\overline{\mu_E(z)}$$

and if τ is complex then

$$(z,1,1,1,1)\mapsto(\mu_E(z),1,\dots,1)\mapsto\mu_E(z)\overline{\mu_E(z)},$$

where again all occurrences of μ_E are viewed over $\mathbb{C}$ via τ.

In particular we have checked the necessary equality.

It defines a Shimura datum because after projecting along $\mathcal{B}\to G^{\mathrm{ad}}$, $h_G\times h_E$ agrees with h_G. This is independent of the choice of h_G because in general $\mathfrak{X}$ is a union of copies of $\mathfrak{X}^+$ so every other choice for $h_G\in\mathfrak{X}^+$ must be obtained. For the statement about the reflex field, we note that it is immediate from the general fact that $E(\mathcal{B},\mathfrak{X})=E(\mathcal{B}^{\mathrm{ad}},\mathfrak{X}^{\mathrm{ad}}_{\mathcal{B}})E(\mathcal{B}^{\mathrm{ab}},h)=E$.

To show it only depends on $(G^{\mathrm{der}},\mathfrak{X}^+)$ and E we just return to the argument above for $\mathcal{B}$ being independent. Given $(G,\mathfrak{X})$ and $(G',\mathfrak{X}')$ Shimura data with reflex fields contained in E and the same connected Shimura datum, we can again apply [Deligne 1979, 2.5.6] and find $(G'',\mathfrak{X}'')$ mapping to both. Since the construction of $(\mathcal{B},\mathfrak{X}_{\mathcal{B}})$ is functorial in $(G,\mathfrak{X})$, choosing $h\in\mathfrak{X}^+$ immediately gives a commuting diagram

$$\begin{tikzcd}
\mathbb{S} \arrow[r, equal] \arrow[d] & \mathbb{S} \arrow[r, equal] \arrow[d] & \mathbb{S} \arrow[d] \\
\mathcal{B}_{\mathbb{R}} & \mathcal{B}''_{\mathbb{R}} \arrow[l, "\sim"'] \arrow[r, "\sim"] & \mathcal{B}'_{\mathbb{R}}.
\end{tikzcd}$$

From this it is clear that $(\mathcal{B},\mathfrak{X}_{\mathcal{B}})$ depends only on $(G^{\mathrm{der}},\mathfrak{X})$ and E. □

We would like to check the obvious functorialities for the pair $(\mathcal{B},\mathfrak{X}_{\mathcal{B}})$.

Lemma 4.6.6. *The above construction is functorial in the following senses*:

(1) *Let $u:(G_1,\mathfrak{X}_1^+)\to(G_2,\mathfrak{X}_2^+)$ be a map of connected Shimura data*: *that is*, *a central isogeny $G_1\to G_2$ of semisimple groups such that $\mathfrak{X}_1^{+,ad}=\mathfrak{X}_2^{+,ad}$*, *and $E\supset E(G_1,\mathfrak{X}_1^+)$. Let $(\mathcal{B}_1,\mathfrak{X}_1)$ and $(\mathcal{B}_2,\mathfrak{X}_2)$ be the Shimura data associated to the triples $(G_1,\mathfrak{X}_1^+,E)$ and $(G_2,\mathfrak{X}_2^+,E)$ by the above procedure. Then there is a canonical morphism*

$$u_*:(\mathcal{B}_1,\mathfrak{X}_1)\to(\mathcal{B}_2,\mathfrak{X}_2)$$

of Shimura data also induced by a natural central isogeny $\mathcal{B}_1\to\mathcal{B}_2$.

(2) *Let $(G^{\mathrm{der}},\mathfrak{X}^+)$ be a connected Shimura datum, $E(G^{\mathrm{der}},\mathfrak{X}^+)\subset F\subset E$ two fields*, *and $(\mathcal{B}_F,\mathfrak{X}_F)$ and $(\mathcal{B}_E,\mathfrak{X}_E)$ the associated data. Then the norm map induces a canonical morphism of Shimura data*

$$N_{E/F}:(\mathcal{B}_E,\mathfrak{X}_E)\to(\mathcal{B}_F,\mathfrak{X}_F).$$

Proof. For (1), let $\Delta_1\subset G_1\subset\mathcal{B}_1$ be the kernel of $G_1\to G_2$ viewed as a subgroup of the centre of $\mathcal{B}_1$. Then $(\mathcal{B}_1/\Delta_1,\mathfrak{X}_1/\Delta_1)$ is a Shimura datum with connected Shimura datum $(G_2,\mathfrak{X}_2^+)$ and reflex field contained in E. It therefore receives a natural map $(\mathcal{B}_2, \mathfrak{X}_2) \to (\mathcal{B}_1/\Delta_1, \mathfrak{X}_1/\Delta_1)$ which we claim is an isomorphism. It will suffice to check $\mathcal{B}_2 \to \mathcal{B}_1/\Delta_1$ is an isomorphism, which is immediate because it is a pullback of the identity map on E^*. Thus we obtain the canonical map

$$(\mathcal{B}_1, \mathfrak{X}_1) \to (\mathcal{B}_1/\Delta_1, \mathfrak{X}_1/\Delta_1) \xleftarrow{\simeq} (\mathcal{B}_2, \mathfrak{X}_2)$$

required.

For (2), for any Shimura datum $(G, \mathfrak{X})$ with connected Shimura datum $(G^{\mathrm{der}}, \mathfrak{X}^+)$ and reflex field contained in F, one obviously has a canonical map $\mathcal{B}_E = G \times_{G^{\mathrm{ab}}} E^* \to G \times_{G^{\mathrm{ab}}} F^* = \mathcal{B}_F$ induced by the norm morphism[11] $N_{E/F} : E^* \to F^*$. Visibly $h_F = N_{E/F} \circ h_E$, so this induces the map of Shimura data claimed. □

4.7. *Abelian type case.* We now return to integral canonical models of standard principal bundles. Fix $(G_2, \mathfrak{X}_2)$ a Shimura datum of abelian type unramified away from N, $G_2/\mathbb{Z}[1/N]$ a reductive integral model and $K_2^N = \prod_{p\nmid N} G_2(\mathbb{Z}_p)$. The goal of this section is to prove the following.

Theorem 4.7.1. *There exists an integral canonical model for $P_{K_2^N}(G_2, \mathfrak{X}_2)$.*

4.7.2. As in Section 2.5.14 we may find $(G, \mathfrak{X})$ of Hodge type with G unramified away from N with $j : G^{\mathrm{der}} \to G_2^{\mathrm{der}}$ a central isogeny inducing a morphism of connected Shimura data $(G^{\mathrm{der}}, \mathfrak{X}^+) \to (G_2^{\mathrm{der}}, \mathfrak{X}_2^+)$. By Lemma 2.5.13 we may take $G/\mathbb{Z}[1/N]$ reductive such that j is defined over $\mathbb{Z}[1/N]$.

By the construction of the previous section, we may obtain a diagram of Shimura data

$$\begin{array}{ccc} (\mathcal{B}, \mathfrak{X}_{\mathcal{B}}) & \longrightarrow & (G, \mathfrak{X}) \\ \downarrow & & \\ (\mathcal{B}_2, \mathfrak{X}_{\mathcal{B},2}) & \longrightarrow & (G_2, \mathfrak{X}_2). \end{array}$$

The plan has three stages. We first combine our constructions in the special and Hodge type cases to construct a canonical integral model of $P(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})$. We next pass to connected components and take a quotient to transfer this bundle from a $\mathcal{B}^c$-bundle on $\mathrm{Sh}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})^+$ to a G_2^c-bundle on $\mathrm{Sh}(G_2, \mathfrak{X}_2)^+$ defined over the field E_N over which the connected components split off. We finally then assemble over the whole Shimura variety and descend these models making liberal use of the uniqueness statement for canonical models in place of the extension property from §2.5. We believe this framework could be useful in other contexts where one wants to extend a "G-valued" construction over Hodge type Shimura varieties to abelian type Shimura varieties.

[11] Recall that the norm map $N_{E/F}$ may be defined for example as the composite of $\mathbb{Q}$-group maps $E^* \hookrightarrow \mathrm{Res}_{F/\mathbb{Q}} \mathrm{GL}_F(E) \xrightarrow{\det} F^*$.

4.7.3. We start out by checking this picture respects the integral structures on G and G_2. Since $E := E(G, \mathfrak{X})$ is absolutely unramified at $p \nmid N$ and $E_2 \subset E$ (hence E_2^* and E^* are smooth tori over $\mathbb{Z}[1/N]$) we have a diagram defined over $\mathbb{Z}[1/N]$

$$G \leftarrow \mathcal{B} = G \times_{G^{\mathrm{ab}}} E^* \to \mathcal{B}_2 = G_2 \times_{G_2^{\mathrm{ab}}} E_2^* \to G_2.$$

We remark that the middle map is a composite of Lemma 4.6.6(1) and (2), the first of which involves a change of derived group (and by construction $G^{\mathrm{der}} \to G_2^{\mathrm{der}}$ is defined over $\mathbb{Z}[1/N]$), and the second of which is a norm map $N_{E/E_2} : E^* \to E_2^*$ which clearly respects integral structures. We let K^N, K_2^N and $K_{\mathcal{B}}^N$ denote the obvious corresponding hyperspecial prime to N level structures and note there is a corresponding diagram of integral models of Shimura varieties

$$\mathcal{S}_{K^N}(G, \mathfrak{X}) \leftarrow \mathcal{S}_{K_{\mathcal{B}}^N}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}}) \to \mathcal{S}_{K_2^N}(G_2, \mathfrak{X}_2).$$

Let us, now this diagram is in place, check the following condition Section 4.4.3, which we recall is important in guaranteeing the uniqueness of canonical models.

Proposition 4.7.4. *With notation as above,* $\mathcal{S}_{K_{\mathcal{B}}^N}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})$ *and* $\mathcal{S}_{K_2^N}(G_2, \mathfrak{X}_2)$ *both have all the crystalline points.*

Proof. Let $x \in \mathcal{S}_{K_{\mathcal{B}}^N}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})(W(k))$ and $K_0 = W(k)[1/p]$. Also let $W(k)^{\mathrm{ur}}$ be the ring of integers of the maximal unramified (algebraic) extension K_0^{ur}/K_0. Note that its images in $\mathcal{S}_{K^N}(G, \mathfrak{X})$ and $\mathcal{S}_{\hat{E^{*,N}}}(E^*, h_E)$ are both crystalline (since these Shimura data have all the crystalline points). To check x is crystalline, it will suffice to check $\omega_{\mathrm{et},x}(V)$ is crystalline for V a faithful representation of $\mathcal{B}$. Let us take $G \hookrightarrow \mathrm{GL}(V_1)$ and $E^* \hookrightarrow \mathrm{GL}(V_2)$ and consider the representation

$$\mathcal{B} \subset G \times E^* \hookrightarrow \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \subset \mathrm{GL}(V_1 \oplus V_2).$$

Considering the diagram of Shimura varieties and towers at infinite level at p, it is clear that $\Gamma_{K_0} \to \mathrm{GL}(\omega_{\mathrm{et}}(V_1 \oplus V_2))$ acts via $\Gamma_{K_0} \to \mathrm{GL}(\omega_{\mathrm{et}}(V_1)) \times \mathrm{GL}(\omega_{\mathrm{et}}(V_2))$, each projection of which is crystalline, and this is enough to give the first part of our proposition.

We turn our attention to $\mathcal{S}_{K_2^N}(G_2, \mathfrak{X}_2)$. First, since it receives a map from the Shimura variety $\mathcal{S}_{K_{\mathcal{B}}^N}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})$ and we observe that the image of a crystalline point is crystalline, and moreover at finite level these maps are finite étale, we deduce that in particular all $W(k)^{\mathrm{ur}}$ points on the geometric connected component $\mathcal{S}_{K_2^N}{}^+(G_2, \mathfrak{X}_2)_{W(k)^{\mathrm{ur}}}$ are crystalline.

To show that any point $x \in \mathcal{S}_{K_2^N K_{2,N}}(W(k))$ is crystalline, recall that Kisin's integral model at level $K_{2,p} := G_2(\mathbb{Z}_p)$ over $W(k)$, $\mathcal{S}_{K_{2,p}}(G_2, \mathfrak{X}_2)_{W(k)}$ has a $G(\mathbb{A}^{\infty,p})$-action that acts transitively on geometric connected components and acts equivariantly with regard to the tower at p on the generic fibre. Therefore, taking a lift $\tilde{x}$ of x to level $K_{2,p}$, which we may assume is a $W(k)^{\mathrm{ur}}$ point since $\mathcal{S}_{K_{2,p},W(k)} \to$

$\mathcal{S}_{K_{2,p}K^p,W(k)}$ is formally étale, and taking a translate $\tilde{x}.a$ by a Hecke operator $a \in G(\mathbb{A}^{\infty,p})$ such that $\tilde{x}.a$ lies in the connected component, we may deduce that x is crystalline from the fact that all points on the connected component are, and that the crystalline property can be checked after restricting a Γ_{K_0} representation to $\Gamma_{K_0^{\mathrm{ur}}}$. $\square$

We now proceed with constructing the canonical models. The content of the next step can be extracted as a lemma.

Lemma 4.7.5. *Let S be a scheme, and $G \twoheadrightarrow \Delta \leftarrow H$ a diagram of S-groups, $\mathcal{B} = G \times_\Delta H$ the fibre product group. Let X be an S-scheme and suppose we are given P_G a G-torsor and P_H an H-torsor on X with the property that there is an isomorphism of Δ-torsors $\theta : P_H \times^H \Delta \xrightarrow{\sim} P_G \times^G \Delta$.*

Then there is a unique (up to canonical isomorphism) $\mathcal{B}$-torsor $P_{\mathcal{B}}$ together with torsor isomorphisms $\theta_G : P_{\mathcal{B}} \times^{\mathcal{B}} G \xrightarrow{\sim} P_G$ and $\theta_H : P_{\mathcal{B}} \times^{\mathcal{B}} H \xrightarrow{\sim} P_H$ such that the induced isomorphism

$$\theta_H \circ \theta_G^{-1} : P_G \times^G \Delta \xrightarrow{\sim} P_{\mathcal{B}} \times^{\mathcal{B}} \Delta \xrightarrow{\sim} P_H \times^H \Delta$$

is equal to θ, and it is given by

$$P_{\mathcal{B}} = P_G \times_\theta P_H,$$

with θ_G and θ_H the natural projections.

We use the notation $X \times_\theta Y$ in a situation where we are given maps $X \to Z_X$ and $Y \to Z_Y$ and an isomorphism $\theta : Z_X \xrightarrow{\sim} Z_Y$ between two schemes to mean the limit of the diagram

$$\begin{tikzcd} & & Y \arrow[d] \\ X \arrow[r] & Z_X \arrow[r, "\theta", "\sim"'] & Z_Y. \end{tikzcd}$$

Proof. Let us first show that $P_{\mathcal{B}} = P_G \times_\theta P_H$ is indeed a $\mathcal{B}$-torsor. We define the $\mathcal{B} = G \times_\Delta H$-action

$$(p_G, p_H).(g, h) =: (p_G.g, p_H.h)$$

which is visibly an action. Passing to an étale cover $X' \to X$ over which P_G and P_H admit sections p_G and p_H. Since G surjects onto Δ these may be chosen (perhaps at the cost of passing to a finer étale cover) such that if $\pi_G : P_G \to P_G \times^G \Delta$ and $\pi_H : P_H \to P_H \times^H \Delta$ are the projection maps, we have $\theta(\pi_G(p_G)) = \pi_H(p_H)$, giving a section $(p_G, p_H) \in P_{\mathcal{B}}(X')$, from which we may see immediately that it is a $\mathcal{B}$-torsor. It is also obvious that it has the required property.

Uniqueness is essentially formal, but we give the argument. Suppose we are given $P'_{\mathcal{B}}$ together with $\theta'_G : P'_{\mathcal{B}} \times^{\mathcal{B}} G \xrightarrow{\sim} P_G$ and $\theta'_H : P'_{\mathcal{B}} \times^{\mathcal{B}} H \xrightarrow{\sim} P_H$ satisfying the compatibility given. Then by the universal property of $P_{\mathcal{B}} = P_G \times_\theta P_H$ there exists a unique map $\alpha : P'_{\mathcal{B}} \to P_{\mathcal{B}}$ making everything commute. It is easy to check α is a map of $\mathcal{B}$-torsors, whence it is automatically an isomorphism. □

4.7.6. Let us apply this in the context of our standard principal bundles. If T is a torus unramified away from N we introduce the notation $\hat{T}^N = \prod_{p\nmid N} T(\mathbb{Z}_p)$.

Consider the diagram (whose arrows we have named for convenience)

$$\begin{array}{ccc} \mathcal{S}_{K^N_{\mathcal{B}}}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}}) & \xrightarrow[\delta']{} & \mathcal{S}_{\hat{E}^{*,N}}(E^*, h_E) \\ {\scriptstyle \pi}\big\downarrow & & \big\downarrow{\scriptstyle N\mu} \\ \mathcal{S}_{K^N}(G, \mathfrak{X}) & \xrightarrow[\delta]{} & \mathcal{S}_{\hat{G}^{ab,N}}(G^{\mathrm{ab}}, h^{\mathrm{ab}}). \end{array}$$

By Theorem 4.5.1 we have $\mathcal{P}_{K^N}$ over $\mathcal{S}_{K^N}(G, \mathfrak{X})$ an integral canonical model for $P_{K^N}(G, \mathfrak{X})$, and by Proposition 4.2.3 we have $\mathcal{P}_{\hat{E}^{*,N}}$ and $\mathcal{P}_{\hat{G}^{ab,N}}$ canonical models over the right hand side of the diagram also. By Proposition 4.4.8 there are natural isomorphisms of G^{ab} torsors

$$N\mu^*\mathcal{P}_{\hat{G}^{ab,N}} \cong \mathcal{P}_{\hat{E}^{*,N}} \times^{E^{*c}} G^{\mathrm{ab}} \quad \text{and} \quad \delta^*\mathcal{P}_{\hat{G}^{ab,N}} \cong \mathcal{P}_{K^N} \times^G G^{\mathrm{ab}}.$$

Pulling these back further and using that the diagram commutes and so we have canonical identifications $\pi^*\delta^* \cong (\delta\pi)^* = (N\mu\delta')^* \cong \delta'^*(N\mu)^*$, we obtain a natural isomorphism

$$\theta : \pi^*(\mathcal{P}_{K^N} \times^G G^{\mathrm{ab}}) \xrightarrow{\sim} \delta'^*(\mathcal{P}_{\hat{E}^{*,N}} \times^{E^{*c}} G^{\mathrm{ab}}).$$

Note further that any integral canonical model $\mathcal{P}_{K^N_{\mathcal{B}}}$ for $P_{K^N_{\mathcal{B}}}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})$ will by Proposition 4.4.8 be required to satisfy the conditions of Lemma 4.7.5 with respect to this isomorphism θ. Therefore by Lemma 4.7.5 if it exists it is given by

$$\mathcal{P}_{K^N_{\mathcal{B}}} := \pi^*\mathcal{P}_{K^N} \times_\theta \delta'^*\mathcal{P}_{\hat{E}^{*,N}},$$

together with the identification

$$\begin{aligned} \iota_{\mathcal{B}} : \mathcal{P}_{K^N_{\mathcal{B}}} \otimes_{\mathbb{O}_E[1/N]} E &= (\pi^*\mathcal{P}_{K^N} \times_\theta \delta'^*\mathcal{P}_{\hat{E}^{*,N}}) \otimes_{\mathbb{O}_E[1/N]} E \\ &\xrightarrow{(\iota_G, \iota_{E^*})} \pi^* P_{K^N}(G, \mathfrak{X}) \times_{\theta\otimes E} \delta'^* P_{\hat{E}^{*,N}}(E^*, h_E) = P_{K^N_{\mathcal{B}}}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}}). \end{aligned}$$

Proposition 4.7.7. *As defined above, $(\mathcal{P}_{K^N_{\mathcal{B}}}, \iota_{\mathcal{B}})$ is an integral canonical model for $P_{K^N_{\mathcal{B}}}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})$ with equivariant $\mathcal{A}^N_P(\mathcal{B})$-action.*[12]

Proof. Our task is to show the connection, filtration and $\mathcal{A}^N_P(\mathcal{B})$-actions extend, and check the lattice property.

[12]Recall we defined this group in Section 4.5.8.

We first study the connection. Let $p_i, i = 1, 2$ be the projections from a first order neighbourhood of the diagonal, abusively retaining the same notation for morphisms of schemes and their first order thickenings. Then[13]

$$p_i^*\mathcal{P}_{K^N_{\mathcal{B}}} = p_i^*(\pi^*\mathcal{P}_{K^N} \times_\theta \delta'^*\mathcal{P}_{\hat{E}^*,N}) \cong \pi^* p_i^*\mathcal{P}_{K^N} \times_{p_i^*(\theta)} \delta'^* p_i^*\mathcal{P}_{\hat{E}^*,N},$$

and in these coordinates, $\nabla_{\mathcal{B}} = (\pi^*\nabla_G, \delta'^* \mathrm{id})$ is the connection required so we see directly that it extends.

For the filtrations, it is easy to check using [Conrad 2014, 5.3.4] that in fact we may naturally identify

$$\mathcal{GR}_\mu = G/P_\mu = G^{\mathrm{der}}/(P_\mu \cap G^{\mathrm{der}}) = \mathcal{B}/\mathcal{P}_{\mu_{\mathcal{B}}} = \mathcal{GR}_{\mu_{\mathcal{B}}}$$

with G and $\mathcal{B}$-actions factoring through G^{ad}, so actually the composite

$$\gamma_{\mathcal{B}} : \mathcal{P}_{K^N_{\mathcal{B}}} \to \mathcal{P}_{K^N} \xrightarrow{\gamma} \mathcal{GR}_\mu = \mathcal{GR}_{\mu_{\mathcal{B}}}$$

does the job.

We also note that we already have a natural extension of the Hecke action via

$$\prod_{p|N} \mathcal{B}(\mathbb{Q}_p) \subset \left(\prod_{p|N} G(\mathbb{Q}_p)\right) \times \left(\prod_{p|N} E^*(\mathbb{Q}_p)\right)$$

which acts via its two projections on $\pi^*\mathcal{P}_{K^N} \times_\theta \delta'^*\mathcal{P}_{\hat{E}^*,N}$.

For the lattice property, we borrow the trick from the first part of the proof of Proposition 4.7.4, taking a faithful representation $V = V_1 \oplus V_2$ of $\mathcal{B}$, formed as a sum of faithful representations of G and E^*. It is easy to check from the construction that

$$\omega_{\mathcal{P}_{K^N_{\mathcal{B}}}}(V) = \omega_{\pi^*\mathcal{P}_{K^N}}(V_1) \oplus \omega_{\delta'^*\mathcal{P}_{\hat{E}^*,N}}(V_2)$$

whence the lattice property follows immediately from that already known for the two terms on the right hand side together with the fact that the image of a crystalline point is always crystalline. Now we know the lattice property, we get a uniqueness statement, and the $\mathcal{A}_P^N(\mathcal{B})$-action extends formally. □

4.7.8. We now pass to a connected component

$$\mathrm{Sh}^+_{K^N_{\mathcal{B}}}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}}) = \mathrm{Sh}^+_{K^N}(G, \mathfrak{X}),$$

recalling Lemma 2.5.1 that it is defined over an extension E_N/E unramified away from N and letting $\mathbb{O} := \mathbb{O}_{E_N}[1/N]$. We may therefore extend the above identification to

$$\mathcal{S}^+_{K^N_{\mathcal{B}}}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})_{\mathbb{O}} = \mathcal{S}^+_{K^N}(G, \mathfrak{X})_{\mathbb{O}},$$

[13] We also abuse notation here and confuse π and δ' with their induced maps on first order neighbourhoods of the diagonal.

and restrict $\mathcal{P}_{K^N_{\mathcal{B}}}$ to get the $\mathcal{B}^c$-torsor

$$\mathcal{P}^+_{\mathcal{B}} := \mathcal{P}_{K^N_{\mathcal{B}}} \times_{\mathcal{S}_{K^N_{\mathcal{B}}}(\mathcal{B},\mathfrak{X}_{\mathcal{B}})_{\mathbb{O}}} \mathcal{S}^+_{K^N}(G, \mathfrak{X})_{\mathbb{O}} \to \mathcal{S}^+_{K^N}(G, \mathfrak{X})_{\mathbb{O}}.$$

Recall the group

$$\Delta^N(G, G_2) = \mathrm{Ker}(\mathcal{A}^{N,\circ}(G^{\mathrm{der}}) \to \mathcal{A}^{N,\circ}(G_2^{\mathrm{der}}))$$

with the property that

$$\mathcal{S}^+_{K^N}(G, \mathfrak{X})_{\mathbb{O}}/\Delta^N(G, G_2) = \mathcal{S}^+_{K_2^N}(G_2, \mathfrak{X}_2)_{\mathbb{O}}.$$

Let us also recall the extension of group schemes over $\mathbb{Z}[1/N]$

$$1 \to \mathcal{B}^c \to \mathcal{A}^N_P(\mathcal{B}) \to \mathcal{A}^N(\mathcal{B}) \to 1.$$

Forming the pullback along $\mathcal{A}^{N,\circ}(G^{\mathrm{der}}) = \mathcal{A}^{N,\circ}(\mathcal{B}) \subset \mathcal{A}^N(\mathcal{B})$ we get an extension

$$1 \to \mathcal{B}^c \to \mathcal{A}^{N,\circ}_P(\mathcal{B}) \to \mathcal{A}^{N,\circ}(G^{\mathrm{der}}) \to 1$$

which acts on $\mathcal{P}^+_{\mathcal{B}}$ equivariantly in the obvious fashion.

Lemma 4.7.9. *These group schemes have the following additional properties*:

(1) *The group* $\Xi := \mathrm{Ker}(\mathcal{B}^c \to G_2^c) \subset \mathcal{B}^c$ *is a normal subgroup of* $\mathcal{A}^{N,\circ}_P(\mathcal{B})$.

(2) *The kernel* $\Delta^N_P(\mathcal{B}, G_2) := \mathrm{Ker}(\mathcal{A}^{N,\circ}_P(\mathcal{B}) \to \mathcal{A}^{N,\circ}_P(G_2))$ *is canonically an extension*

$$1 \to \Xi \to \Delta^N_P(\mathcal{B}, G_2) \to \Delta^N(G, G_2) \to 1$$

which acts freely on $\mathcal{P}^+_{\mathcal{B}}$.

Proof. We first remark that (1) is obvious because $\Xi \subset Z(\mathcal{B}^c)$ which commutes with the whole of $\mathcal{A}^{N,\circ}_P(\mathcal{B})$.

For (2), that Ξ is a subgroup is clear and normality follows by (1). Existence of the maps and exactness in the middle follow in the usual way. That the final map is a surjection follows because for any $(g, \gamma^{-1}) \in \Delta^N(G, G_2)$, it is hit by $(1, g, \gamma^{-1}) \in \Delta^N_P(\mathcal{B}, G)$. Finally the action on $\mathcal{P}^+_{\mathcal{B}} \to \mathcal{S}^+_{K^N}(G, \mathfrak{X})_{\mathbb{O}}$ is free because Ξ acts freely on each fibre, and $\Delta^N(G, G_2)$ acts freely on $\mathcal{S}^+_{K^N}(G, \mathfrak{X})_{\mathbb{O}}$ by Proposition 2.5.9. $\square$

4.7.10. Equipped with this lemma we may construct a G_2^c-bundle

$$\mathcal{P}^+_2 := (\mathcal{P}^+_{\mathcal{B}}/\Delta^N_P(\mathcal{B}, G_2)) \times^{\mathcal{B}^c/\Xi} G_2^c \to \mathcal{S}^+_{K_2^N}(G_2, \mathfrak{X}_2)_{\mathbb{O}}$$

with equivariant $\mathcal{A}^{N,\circ}_P(G_2)$-action. Moreover we have

$$\begin{aligned} \iota_2 : \mathcal{P}^+_2 \otimes_{\mathbb{O}} E_N &= (\mathcal{P}^+_{\mathcal{B}} \otimes_{\mathbb{O}} E_N/\Delta^N_P(\mathcal{B}, G_2)) \times^{\mathcal{B}^c/\Xi} G_2^c \\ &\xrightarrow{\iota_{\mathcal{B}}} (P^+_{K^N_{\mathcal{B}}, E_N}/\Delta^N_P(\mathcal{B}, G_2)) \times^{\mathcal{B}^c/\Xi} G_2^c = P^+_{K_2^N, E_N}, \end{aligned}$$

where the final equality follows by working over $\mathbb{C}$ and the argument of [Milne 1988, 7.2].

Proposition 4.7.11. *The pair* $(\mathcal{P}_2^+, \iota_2)$ *is an integral canonical model for the bundle* $P^+_{K_2^N, E_N}(G_2, \mathfrak{X}_2)$.

Proof. The Hecke action extends by construction. Let us check that the connection extends. We know (from direct observation over $\mathbb{C}$) that the $\Delta_P^N(\mathcal{B}, G_2)$ action is horizontal, so the flat connection $\nabla : p_1^*\mathcal{P}_{\mathcal{B}}^+ \xrightarrow{\sim} p_2^*\mathcal{P}_{\mathcal{B}}^+$ descends (and then can obviously be pushed out along $\mathcal{B}^c/\Xi \to G_2^c$) to a flat connection

$$\nabla : p_1^*\mathcal{P}_2^+ \xrightarrow{\sim} p_2^*\mathcal{P}_2^+.$$

Since a similar relationship also relates the connections on the generic fibre we have shown that this ∇ extends over $\mathbb{O}$ the connection on $P_{K_2^N, E_N}$.

To check the lattice condition at a crystalline point x, note the commutative diagram of E_N-schemes, letting $K_{\mathcal{B}}^{Np}$ and K_2^{Np} be the obvious full level structures at all primes except N and p

$$\begin{array}{ccc} \mathrm{Sh}_{K_{\mathcal{B}}^{Np}}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})|_{\mathrm{Sh}^+_{K_{\mathcal{B}}^N}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})} & \longrightarrow & \mathrm{Sh}_{K_2^{Np}}(G_2, \mathfrak{X}_2)|_{\mathrm{Sh}^+_{K_2^N}(G_2, \mathfrak{X}_2)} \\ \big\downarrow & & \big\downarrow \\ \mathrm{Sh}^+_{K_{\mathcal{B}}^N}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}}) & \longrightarrow & \mathrm{Sh}^+_{K_2^N}(G_2, \mathfrak{X}_2). \end{array}$$

Taking a lift $\tilde{x}$ of x to $\mathcal{S}^+_{K_{\mathcal{B}}^N}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})_{\mathbb{O}}$, which is crystalline since the map

$$\mathcal{S}^+_{K_{\mathcal{B}}^N}(\mathcal{B}, \mathfrak{X}_{\mathcal{B}})_{\mathbb{O}} \to \mathcal{S}^+_{K_2^N}(G_2, \mathfrak{X}_2)_{\mathbb{O}}$$

is finite étale and its source has all the crystalline points, the diagram gives an identification

$$\omega_{\mathrm{et}, \tilde{x}} \circ \mathrm{Res}^{G_2^c(\mathbb{Z}_p)}_{\mathcal{B}^c(\mathbb{Z}_p)} = \omega_{\mathrm{et}, x} : \mathrm{Rep}_{\mathbb{Z}_p}(G_2^c) \to \mathrm{Rep}_{\mathbb{Z}_p}(\Gamma_{\mathbb{Q}_p^{\mathrm{ur}}}).$$

But now the lattice condition is immediate, since equality of lattices can be checked after passing to a finite étale cover, and $\mathcal{P}^+_{2,x}$ pulls back to $\mathcal{P}^+_{\mathcal{B},\tilde{x}} \times^{\mathcal{B}^c} G_2^c$ which as a canonical model already has the required property that $\omega_{\mathcal{P}^+_{\mathcal{B},\tilde{x}}} = \mathbb{D} \circ \omega_{\mathrm{et},\tilde{x}}$ by its own lattice condition.

It remains to check the filtration $\gamma_B : \mathcal{P}_{\mathcal{B}}^+ \to \mathcal{GR}_{\mu_{\mathcal{B}}}$ descends to $\gamma_2 : \mathcal{P}_2^+ \to \mathcal{GR}_{\mu_2}$, which amounts to showing that the composite

$$\mathcal{P}_{\mathcal{B}}^+ \to \mathcal{GR}_{\mu_{\mathcal{B}}} \to \mathcal{GR}_{\mu_2}$$

is $\Delta_P^N(\mathcal{B}, G_2)$-invariant. But this is clear; since $\Xi = \mathrm{Ker}(\mathcal{B}^c \to G_2^c) \subset Z(\mathcal{B}^c)$ it acts trivially on $\mathcal{GR}_{\mu_2}$ and $\Delta^N(G, G_2)$ acts trivially because Hecke operators always act trivially on any Gr_μ. $\square$

4.7.12. With this in hand, we are finally able to construct an integral canonical model $\mathcal{P}_2$ for $P_{K_2^N}(G_2, \mathfrak{X}_2)$ in the spirit of Lemma 2.5.16, but with the uniqueness of canonical models used in place of the extension property, because it may be used to canonically extend any isomorphism.

Indeed, let us decompose $\mathcal{S}_{K_2^N,\mathbb{O}} = \coprod_{c\in\pi_0} \mathcal{S}_c$ into components each of which is geometrically integral. By the argument of Lemma 2.5.16(1) we can find for each $c \in \pi_0$ some $\mathcal{G}_c/\mathbb{Z}[1/N]$ a model for G_2 such that we may identify $\mathcal{S}_c \cong \mathcal{S}_{\prod_{p\nmid N}\mathcal{G}_c(\mathbb{Z}_p),\mathbb{O}}{}^+$ via a Hecke operator. For each c we may make such a choice and invoking Proposition 4.7.11 for $\mathcal{G} = \mathcal{G}_c$ and using the Hecke equivariance of P_{K^N} we obtain a canonical model $\mathcal{P}_c^+$ for $P_{K_2^N,E_N}(G_2, \mathfrak{X}_2)|_{\mathcal{S}_c\otimes_{\mathbb{O}} E_N}$. Taking the disjoint union of these we obtain

$$\mathcal{P}_{2,\mathbb{O}} := \coprod_{c\in\pi_0} \mathcal{P}_c^+,$$

an integral canonical model minus a full Hecke action. But by the uniqueness of canonical models (Lemma 4.4.7) it is formal to extend the Hecke operators acting between components. Finally by Lemma 4.4.9 it descends to $\mathcal{P}_2/\mathbb{O}_{E_2}[1/N]$, an integral canonical model for $P_{K_2^N}(G_2, \mathfrak{X}_2)/E_2$. Thus our main theorem has been proved.

4.8. *Automorphic vector bundles.* With our integral canonical models $\mathcal{P}_{K^N}(G, \mathfrak{X})$ constructed in the abelian type case, we should discuss the construction of automorphic vector bundles (recall the discussion Section 3.1.4) in this setting, although it requires no new ideas.

Theorem 4.8.1. *Let $(G, \mathfrak{X})$ be a Shimura datum of abelian type, $\mathcal{G}/\mathbb{Z}[1/N]$ a reductive model for G, $K^N = \prod_{p\nmid N} \mathcal{G}(\mathbb{Z}_p)$ and for μ a Hodge cocharacter of $(G, \mathfrak{X})$, and $L/E = E(G, \mathfrak{X})$ a finite extension.*

Then we have a canonical functor $\mathcal{J} \mapsto \mathcal{V}(\mathcal{J})$ from $\mathcal{G}_{\mathbb{O}_L[1/N]}$-equivariant vector bundles on $\mathcal{GR}_{\mu,\mathbb{O}_L[1/N]}$ to vector bundles on $\mathcal{S}_{K^N}$ which on the generic fibre is identified naturally with that of [Milne 1990, III 5.1].

Proof. Recall that we have the picture

$$\mathcal{S}_{K^N} \leftarrow \mathcal{P}_{K^N} \xrightarrow{\gamma} \mathcal{GR}_\mu.$$

The functor $\mathcal{J} \mapsto \mathcal{V}(\mathcal{J})$ is given by pulling back $\mathcal{J} \mapsto \gamma^*\mathcal{J}$ and then using the usual equivalence between $\mathcal{G}$-equivariant vector bundles on $\mathcal{P}_{K^N}$ and vector bundles on $\mathcal{S}_{K^N}$. This is obviously a functor, and the compatibility with the usual construction [Milne 1990, III 5.1] follows because $\mathcal{P}_{K^N}$ comes with a canonical identification

$$\iota : \mathcal{P}_{K^N} \otimes E \xrightarrow{\sim} P_{K^N}(G, \mathfrak{X})$$

under which γ is an extension of the G-filtration on the RHS. $\square$

It is also easy to deduce from our construction the following additional functoriality property.

Proposition 4.8.2. *We are given a morphism $f : (G_1, \mathfrak{X}_1, \mathcal{G}_1) \to (G_2, \mathfrak{X}_2, \mathcal{G}_2)$ of Shimura data of abelian type together with reductive models for G_i over $\mathbb{Z}[1/N]$, $L/E(G_1, \mathfrak{X}_1)$ finite, and μ_1 and μ_2 Hodge cocharacters for $\mathfrak{X}_1$ and $\mathfrak{X}_2$ respectively. Suppose we are also given $\mathcal{J}_2$ a $\mathcal{G}_2$-equivariant vector bundle on $\mathcal{GR}_{\mu_2,\mathcal{O}_L[1/N]}$. Pulling back $\mathcal{J}_2$ along*

$$\mathcal{GR}_{\mu_1,\mathcal{O}_L[1/N]} \to \mathcal{GR}_{\mu_2,\mathcal{O}_L[1/N]}$$

and restricting the $\mathcal{G}_2$ action to $\mathcal{G}_1$ we obtain a $\mathcal{G}_1$-equivariant vector bundle $\mathcal{J}_1$.

There is a canonical identification of vector bundles over $\mathcal{S}_{\prod_{p\nmid N}\mathcal{G}_1(\mathbb{Z}_p)}(G_1, \mathfrak{X}_1)$

$$\mathcal{V}(\mathcal{J}_1) \cong f^*\mathcal{V}(\mathcal{J}_2).$$

Proof. This is immediate from the construction of $\mathcal{V}(-)$ and Proposition 4.4.8. $\square$

Acknowledgements

The author would like to extend the greatest of thanks to his PhD supervisor Mark Kisin, under whose guidance this project was completed. This project owes much to his sharp insights, but also his optimism, energy and patience.

We would also like to thank Jack Thorne, after a conversation with him in 2013 provided the initial motivation for this project. For other useful conversations we thank Chris Blake, George Boxer, Lukas Brantner, Justin Campbell, Kestutis Cesnavicius, Erick Knight, Ananth Shankar, Jack Shotton, Rong Zhou, and Yihang Zhu.

This paper is part of the author's PhD thesis completed at and funded by Harvard University, and begun while the author was on a Kennedy Scholarship.

References

[Bhatt et al. 2016] B. Bhatt, M. Morrow, and P. Scholze, "Integral p-adic Hodge theory", preprint, 2016. arXiv 1602.03148

[Blasius 1994] D. Blasius, "A p-adic property of Hodge classes on abelian varieties", pp. 293–308 in *Motives* (Seattle, WA, 1991), edited by U. Jannsen et al., Proc. Sympos. Pure Math. **55**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

[Broshi 2013] M. Broshi, "G-torsors over a Dedekind scheme", *J. Pure Appl. Algebra* **217**:1 (2013), 11–19. MR Zbl

[Conrad 2014] B. Conrad, "Reductive group schemes", lecture notes, Stanford University, 2014. Zbl

[Deligne 1971] P. Deligne, "Travaux de Shimura", exposeé 389, pp. 123–165 in *Séminaire Bourbaki,* 1970/1971, Lecture Notes in Math. **244**, Springer, 1971. MR Zbl

[Deligne 1979] P. Deligne, "Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques", pp. 247–289 in *Automorphic forms, representations and L-functions* (Oregon State Univ., Corvallis, Ore., 1977), Proc. Sympos. Pure Math., XXXIII **2**, Amer. Math. Soc., Providence, R.I., 1979. MR Zbl

[Harris 1985] M. Harris, "Arithmetic vector bundles and automorphic forms on Shimura varieties. I", *Invent. Math.* **82**:1 (1985), 151–189. MR Zbl

[Ichino and Prasanna 2016] A. Ichino and K. Prasanna, "Periods of quaternionic Shimura varieties. I", preprint, 2016. arXiv

[Kim and Madapusi Pera 2016] W. Kim and K. Madapusi Pera, "2-adic integral canonical models", *Forum Math. Sigma* **4** (2016), e28, 34. MR Zbl

[Kisin 2006] M. Kisin, "Crystalline representations and F-crystals", pp. 459–496 in *Algebraic geometry and number theory*, edited by V. Ginzburg, Progr. Math. **253**, Birkhäuser, Boston, 2006. MR Zbl

[Kisin 2010] M. Kisin, "Integral models for Shimura varieties of abelian type", *J. Amer. Math. Soc.* **23**:4 (2010), 967–1012. MR Zbl

[Kisin 2017] M. Kisin, "mod p points on Shimura varieties of abelian type", *J. Amer. Math. Soc.* **30**:3 (2017), 819–914. MR Zbl

[Kissin and Pappas 2015] M. Kissin and G. Pappas, "Integral models of Shimura varieties with parahoric level structure", preprint, 2015. arXiv

[Lovering 2017] T. Lovering, "Filtered F-crystals on Shimura varieties of abelian type", preprint, 2017. arXiv

[Madapusi Pera 2012] K. Madapusi Pera, "Toroidal compactifications of integral models of Shimura varieties of Hodge type", preprint, 2012. arXiv

[Maulik 2014] D. Maulik, "Supersingular K3 surfaces for large primes", *Duke Math. J.* **163**:13 (2014), 2357–2425. MR Zbl

[Milne 1988] J. S. Milne, "Automorphic vector bundles on connected Shimura varieties", *Invent. Math.* **92**:1 (1988), 91–128. MR Zbl

[Milne 1990] J. S. Milne, "Canonical models of (mixed) Shimura varieties and automorphic vector bundles", pp. 283–414 in *Automorphic forms, Shimura varieties, and L-functions, Vol. I* (Ann Arbor, MI, 1988)), edited by L. Clozel and J. S. Milne, Perspect. Math. **10**, Academic Press, Boston, 1990. MR Zbl

[Milne 1992] J. S. Milne, "The points on a Shimura variety modulo a prime of good reduction", pp. 151–253 in *The zeta functions of Picard modular surfaces*, edited by R. P. Langlands and D. Ramakrishnan, Univ. Montréal, Montreal, QC, 1992. MR Zbl

[Milne 1994] J. S. Milne, "Shimura varieties and motives", pp. 447–523 in *Motives* (Seattle, WA, 1991), edited by U. Jannsen et al., Proc. Sympos. Pure Math. **55**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

[Moonen 1998] B. Moonen, "Models of Shimura varieties in mixed characteristics", pp. 267–350 in *Galois representations in arithmetic algebraic geometry* (Durham, 1996), edited by A. J. Scholl and R. L. Taylor, London Math. Soc. Lecture Note Ser. **254**, Cambridge Univ. Press, 1998. MR Zbl

[Mumford 1965] D. Mumford, *Geometric invariant theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete (2) **34**, Springer, 1965. MR Zbl

[Waterhouse 1979] W. C. Waterhouse, *Introduction to affine group schemes*, Graduate Texts in Mathematics **66**, Springer, 1979. MR Zbl

Communicated by Brian Conrad
Received 2016-08-29 Revised 2017-06-08 Accepted 2017-07-21

tlovering1729@gmail.com *Department of Mathematics, Harvard University, 1 Oxford St., Cambridge, MA 02138, United States*

ALGEBRA AND NUMBER THEORY 11:8 (2017)
dx.doi.org/10.2140/ant.2017.11.1891

Quasi-Galois theory in symmetric monoidal categories

Bregje Pauwels

Given a ring object A in a symmetric monoidal category, we investigate what it means for the extension $\mathbb{1} \to A$ to be (quasi-)Galois. In particular, we define splitting ring extensions and examine how they occur. Specializing to tensor-triangulated categories, we study how extension-of-scalars along a quasi-Galois ring object affects the Balmer spectrum. We define what it means for a separable ring to have constant degree, which is a necessary and sufficient condition for the existence of a quasi-Galois closure. Finally, we illustrate the above for separable rings occurring in modular representation theory.

Introduction

Classical Galois theory is the study of field extensions l/k through the group of automorphisms of l that fix k. If f is a polynomial over k, the *splitting field of f over k* is the smallest extension over which f decomposes into linear factors. If $f \in k[x]$ is moreover separable, its splitting field is the smallest extension l such that

MSC2010: primary 18BXX; secondary 16GXX, 18GXX.
Keywords: tensor triangulated category, separable, etale, Galois, ring-object, stable category.

$l \otimes_k k[x]/(f) \cong l^{\times \deg(f)}$. The field extension l/k is often called *quasi-Galois*[1] if l is the splitting field of some polynomial in $k[x]$. Then, an algebraic field extension is called *Galois* whenever it is quasi-Galois and separable.

The generalization of Galois extensions from fields to rings originated with Auslander and Goldman [1960, Appendix]; see also Remark 5.6. Grothendieck [SGA 1 1971] took on an axiomatic viewpoint to Galois theory and revealed its relation with the fundamental group. Janelidze [2001] adopted a purely categorical approach which covered the above examples. More recently, Rognes [2008] introduced a Galois theory up-to-homotopy. For more generalizations in various directions, see [Chase and Sweedler 1969; Hess 2009; Kreimer 1967].

In this paper, we adapt some of these ideas to the context of ring objects in an additive symmetric monoidal category $(\mathcal{K}, \otimes, \mathbb{1})$, with special emphasis on tensor-triangulated categories. That is, our analogue of a field extension will be a monoid $\eta : \mathbb{1} \to A$ in $\mathcal{K}$ with associative commutative multiplication $\mu : A \otimes A \to A$. We call A a *ring in* $\mathcal{K}$, and moreover assume that A is *separable*, which means μ has an (A, A)-bilinear right inverse $A \to A \otimes A$.

Separable ring objects play an important (though at times invisible) role in various areas of mathematics. In algebraic geometry, for instance, they appear as étale extensions of quasicompact and quasiseparated schemes; see [Balmer 2016a; Neeman 2015]. More precisely, given a separated étale morphism $f : V \to X$, the object $A := Rf_*(\mathcal{O}_V)$ in $D^{\mathrm{qcoh}}(X)$ is a separable ring, and we can understand $D^{\mathrm{qcoh}}(V)$ as the category of A-modules in $D^{\mathrm{qcoh}}(X)$. In representation theory, we can let $\mathcal{K}(G)$ be the (derived or stable) module category of a group G over a field $\Bbbk$, and consider a subgroup $H < G$ of finite index. Balmer [2015] showed there is a separable ring A^G_H in $\mathcal{K}(G)$ such that the category of A^G_H-modules in $\mathcal{K}(G)$ coincides with $\mathcal{K}(H)$, and such that the restriction functor

$$\mathrm{Res}^G_H : \mathcal{K}(G) \to \mathcal{K}(H)$$

is just extension-of-scalars along A^G_H. In the same vein, extension-of-scalars along a separable ring recovers restriction to a subgroup in equivariant stable homotopy theory, in equivariant *KK*-theory and in equivariant derived categories; see [Balmer et al. 2015]. For more examples of separable rings in stable homotopy categories, we refer to [Baker and Richter 2008; Rognes 2008].

Thus motivated, we study how much Galois theory carries over. Recall that a ring A in $\mathcal{K}$ is *indecomposable* if it does not decompose as a product of nonzero rings. Separable ring objects have a well-behaved notion of degree [Balmer 2014] and our first Galois-flavored result (Theorem 4.5) shows that the number of ring endomorphisms of a separable indecomposable ring in $\mathcal{K}$ is bounded by its degree.

[1] see [Bourbaki 1981, V.9.3]. In the literature, a quasi-Galois extension is sometimes called normal or Galois, probably because these notions coincide when l/k is separable and finite.

Definition. Let A and B be separable rings of finite degree in $\mathcal{K}$. We say B *splits* A if $B \otimes A \cong B^{\times \deg(A)}$ as (left) B-algebras in $\mathcal{K}$. We call an indecomposable ring B a *splitting ring of* A if B splits A and any ring morphism $C \to B$, where C is an indecomposable ring splitting A, is an isomorphism.

Definition. If A is a ring in $\mathcal{K}$ and Γ is a group of ring automorphisms of A, we call A *quasi-Galois in* $\mathcal{K}$ *with group* Γ if the A-algebra homomorphism

$$\lambda_\Gamma : A \otimes A \to \prod_{\gamma \in \Gamma} A$$

defined by $\mathrm{pr}_\gamma \lambda_\Gamma = \mu(1 \otimes \gamma)$ is an isomorphism.

Under mild conditions on $\mathcal{K}$, Corollary 6.10 shows an indecomposable ring B is quasi-Galois in $\mathcal{K}$ for some group Γ if and only if B is a splitting ring of some separable ring A in $\mathcal{K}$. By Theorem 5.9, this happens exactly when B has $\deg(B)$ distinct ring endomorphisms in $\mathcal{K}$. Moreover, Proposition 6.9 shows that every separable ring in $\mathcal{K}$ has (possibly multiple) splitting rings. In particular, l is a splitting field of a separable polynomial f over k if and only if l is a splitting ring of $k[x]/(f)$ in the category k-mod; our terminology matches classical field theory.

If, in addition, we assume that $\mathcal{K}$ is tensor-triangulated, we can say more about the way splitting rings arise. Balmer [2005] introduced a topological space $\mathrm{Spc}(\mathcal{K})$ associated to $\mathcal{K}$, in which every object $x \in \mathcal{K}$ has a support $\mathrm{supp}(x) \subset \mathrm{Spc}(\mathcal{K})$. The *Balmer spectrum* $\mathrm{Spc}(\mathcal{K})$ provides an algebro-geometric approach to the study of triangulated categories, and a complete description of the spectrum is equivalent to a classification of the thick $\otimes$-ideals in the category.

For the remainder of the introduction, we assume $\mathcal{K}$ is tensor-triangulated and nice (say, $\mathrm{Spc}(\mathcal{K})$ is noetherian or $\mathcal{K}$ satisfies Krull–Schmidt). If A is a separable ring in $\mathcal{K}$, the *Eilenberg–Moore category* $A\text{-}\mathrm{Mod}_{\mathcal{K}}$ of A-modules in $\mathcal{K}$ admits a triangulation such that extension-of-scalars $\mathcal{K} \to A\text{-}\mathrm{Mod}_{\mathcal{K}}$ is exact; see [Balmer 2011, Corollary 4.3]. We can thus extend scalars along a separable ring without leaving the tensor-triangulated world or descending to a model category. If A is quasi-Galois with group Γ in $\mathcal{K}$, then Γ acts on $A\text{-}\mathrm{Mod}_{\mathcal{K}}$ and on the spectrum $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$. By Theorem 9.1, the Γ-orbits of $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ are given by $\mathrm{supp}(A) \subset \mathrm{Spc}(\mathcal{K})$. In particular, we recover $\mathrm{Spc}(\mathcal{K})$ from $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ if $\mathrm{supp}(A) = \mathrm{Spc}(\mathcal{K})$, which happens exactly when $A \otimes f = 0$ implies f is $\otimes$-nilpotent for every morphism f in $\mathcal{K}$.

Recall that for a quasi-Galois field extension l/k, any irreducible polynomial $f \in k[x]$ with a root in l splits in l; see [Bourbaki 1981, V.9.3]. Proposition 9.6 provides us with a tensor triangular analogue:

Proposition. Let A be a separable ring in $\mathcal{K}$ such that the spectrum $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ is connected, and suppose B is an A-algebra with $\mathrm{supp}(A) = \mathrm{supp}(B)$. If B is quasi-Galois in $\mathcal{K}$, then B splits A.

Finally, Theorem 9.7 reveals which separable rings have a *quasi-Galois closure* in $\mathcal{K}$. Given $\mathcal{P} \in \mathrm{Spc}(\mathcal{K})$, we consider *the local category* $\mathcal{K}_{\mathcal{P}}$ *at* $\mathcal{P}$, the idempotent completion of the Verdier quotient $\mathcal{K}/\mathcal{P}$. We say a ring A has *constant degree in* $\mathcal{K}$ if the degree of A as a ring in $\mathcal{K}_{\mathcal{P}}$ is the same for every prime $\mathcal{P} \in \mathrm{supp}(A)$.

Theorem. If A has constant degree in $\mathcal{K}$ and the spectrum $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ is connected, then A has a unique splitting ring A^*. Furthermore, $\mathrm{supp}(A) = \mathrm{supp}(A^*)$ and A^* is the quasi-Galois closure of A in $\mathcal{K}$. That is, for any A-algebra B that is quasi-Galois in $\mathcal{K}$ with $\mathrm{supp}(A) = \mathrm{supp}(B)$, there exists a ring morphism $A^* \to$B.

We conclude this paper by computing degrees and splitting rings for the separable rings $A^G_H := \Bbbk(G/H)$ mentioned above. Here, $H < G$ are finite groups and $\Bbbk$ is a field with characteristic p dividing $|G|$. The degree of A^G_H in $\mathrm{D}^b(\Bbbk G\text{-mod})$ is simply $[G : H]$ and A^G_H is quasi-Galois if and only if H is normal in G. Accordingly, the quasi-Galois closure of A^G_H in $\mathrm{D}^b(\Bbbk G\text{-mod})$ is the ring A^G_N, where N is the normal core of H in G (see Corollary 10.11). On the other hand, Proposition 10.13 shows the degree of A^G_H in $\Bbbk G$-stab is the greatest $0 \le n \le [G : H]$ such that there exist distinct $[g_1], \dots, [g_n]$ in $H\backslash G$ with p dividing $|H^{g_1} \cap \cdots \cap H^{g_n}|$. In that case, the splitting rings of A^G_H are exactly the $A^G_{H^{g_1} \cap \cdots \cap H^{g_n}}$ with $g_1, \dots, g_n$ as above.

1. The Eilenberg–Moore category

Definition 1.1. Let $\mathcal{K}$ be an additive category. We say $\mathcal{K}$ is *idempotent-complete* if for all $x \in \mathcal{K}$, any morphism $e : x \to x$ with $e^2 = e$ yields a decomposition $x \cong x_1 \oplus x_2$ under which e becomes $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Every additive category $\mathcal{K}$ can be embedded in an idempotent-complete category $\mathcal{K}^\natural$ in such a way that $\mathcal{K} \hookrightarrow \mathcal{K}^\natural$ is fully faithful and every object in $\mathcal{K}^\natural$ is a direct summand of some object in $\mathcal{K}$. We call $\mathcal{K}^\natural$ the *idempotent-completion* of $\mathcal{K}$, and [Balmer and Schlichting 2001] shows that $\mathcal{K}^\natural$ stays triangulated if $\mathcal{K}$ was.

Notation 1.2. Throughout, $(\mathcal{K}, \otimes, \mathbb{1})$ denotes an idempotent-complete symmetric monoidal category. For objects $x_1, \dots, x_n$ in $\mathcal{K}$ and a permutation $\tau \in S_n$, we also write $\tau : x_1 \otimes \ldots \otimes x_n \to x_{\tau(1)} \otimes \ldots \otimes x_{\tau(n)}$ to denote the isomorphism that permutes the tensor factors.

Definition 1.3. A *ring object* $A \in \mathcal{K}$ is a monoid $(A,\ \mu : A \otimes A \to A,\ \eta : \mathbb{1} \to A)$ with associative multiplication μ and two-sided unit η. We call A *commutative* if $\mu(12) = \mu$. All ring objects in this paper will be commutative and we often simply call A a *ring in* $\mathcal{K}$. For rings A and B in $\mathcal{K}$, a *ring morphism* $f : A \to B$ is a morphism in $\mathcal{K}$ that is compatible with the ring structure.

A *(left) A-module* is a pair $(x \in \mathcal{K},\ \varrho : A \otimes x \to x)$, where the action ϱ is compatible with the ring structure in the usual way. *Right A-modules* as well as *(A, A)-bimodules* are defined analogously.

The *Eilenberg–Moore category* $A\text{-Mod}_{\mathcal{K}}$ has left A-modules as objects and A-linear morphisms, which are defined in the usual way. See [Eilenberg and Moore 1965] or [Mac Lane 1998, Chapter VI] for more details. Every object $x \in \mathcal{K}$ gives rise to a *free A-module* $F_A(x) = A \otimes x$ with action given by

$$\varrho : A \otimes A \otimes x \xrightarrow{\mu \otimes 1} A \otimes x.$$

We call the functor $F_A : \mathcal{K} \to A\text{-Mod}_{\mathcal{K}}$ *the extension-of-scalars*, and write U_A for its forgetful right adjoint:

$$\begin{tikzcd} \mathcal{K} \arrow[d, shift right=2ex, "F_A"'] \arrow[d, phantom, "\dashv"] \\ A\text{-Mod}_{\mathcal{K}} \arrow[u, shift right=2ex, "U_A"'] \end{tikzcd}$$

A ring A in $\mathcal{K}$ is *separable* if the multiplication map μ has an (A, A)-bilinear section $\sigma : A \to A \otimes A$. That is, $\mu\sigma = 1_A$ and the diagram

$$\begin{tikzcd} & A \otimes A \arrow[dl, "\sigma\otimes 1"'] \arrow[d, "\mu"] \arrow[dr, "1\otimes\sigma"] & \\ A \otimes A \otimes A \arrow[dr, "1\otimes\mu"'] & A \arrow[d, "\sigma"] & A \otimes A \otimes A \arrow[dl, "\mu\otimes 1"] \\ & A \otimes A & \end{tikzcd}$$

commutes.

Remark 1.4. The module category $A\text{-Mod}_{\mathcal{K}}$ is idempotent-complete whenever $\mathcal{K}$ is idempotent-complete.

Example 1.5. Let R be a commutative ring and consider the category R-mod of finitely generated R-modules. Let A be a commutative projective R-algebra and suppose A is *separable over R*, that is A is projective as an $A \otimes_R A$-module. Then A is finitely generated as an R-module by [DeMeyer and Ingraham 1971, Proposition 2.2.1], so A defines a separable ring object in R-mod. On the other hand, we can think of $A = A[0]$ as a separable ring object in $\mathrm{D}^{\mathrm{perf}}(R)$, the homotopy category of bounded complexes of finitely generated projective R-modules. Note that the category of A-modules in $\mathrm{D}^{\mathrm{perf}}(R)$ is equivalent to $\mathrm{D}^{\mathrm{perf}}(A)$ by [Balmer 2011, Theorem 6.5].

Notation 1.6. Let A and B be rings in $\mathcal{K}$. The ring structure on $A \otimes B$ is given by $(\mu_A \otimes \mu_B)(23) : (A \otimes B)^{\otimes 2} \to (A \otimes B)$. We write A^e for the enveloping ring $A \otimes A^{\mathrm{op}}$, so that left A^e-modules are just (A, A)-bimodules. We write $A \times B$ for the ring $A \oplus B$ with componentwise multiplication.

Remark 1.7. If A and B are separable rings in $\mathcal{K}$, then so are A^e, $A \otimes B$ and $A \times B$. Conversely, A and B are separable whenever $A \times B$ is separable.

Remark 1.8. Let A be a ring in $\mathcal{K}$. Note that every (left) A-linear endomorphism $A \to A$ is in fact A^e-linear, by commutativity of A. What is more, any two A-linear endomorphisms $A \to A$ commute.

Definition 1.9. We call a nonzero ring A in $\mathcal{K}$ *indecomposable* if the only idempotent A-linear endomorphisms $A \to A$ in $\mathcal{K}$ are the identity 1_A and 0. In other words, A is indecomposable if it does not decompose as a direct sum of nonzero A^e-modules. By the following lemma, this is equivalent to saying A does not decompose as a product of nonzero rings.

Lemma 1.10 [Balmer 2014, Lemma 2.2]. *Let A be a ring in $\mathcal{K}$. Suppose there is an A^e-linear isomorphism $h : A \xrightarrow{\sim} B \oplus C$ for some A^e-modules B, C in $\mathcal{K}$. Then B and C admit unique ring structures under which h becomes a ring isomorphism $h : A \xrightarrow{\sim} B \times C$.*

Let (A, μ, η) be a separable ring in $\mathcal{K}$ with separability morphism σ. In what follows, we define a tensor structure $\otimes_A$ on $A\text{-Mod}_{\mathcal{K}}$ under which extension-of-scalars becomes monoidal. The following results all appear in [Balmer 2014, §1]. For detailed proofs, see [Pauwels 2015, §1.1]. Let (x, ϱ_1) and (y, ϱ_2) be A-modules. Here, we can write ϱ_1 to indicate both a left and right action of A on x, as A is commutative. Seeing how the endomorphism $\upsilon : x \otimes y \to x \otimes y$ given by

$$x \otimes y \xrightarrow{1\otimes\eta\otimes 1} x \otimes A \otimes y \xrightarrow{1\otimes\sigma\otimes 1} x \otimes A \otimes A \otimes y \xrightarrow{\varrho_1\otimes\varrho_2} x \otimes y$$

is idempotent and $\mathcal{K}$ is idempotent-complete, we can define $x \otimes_A y$ as the direct summand $\operatorname{im}(\upsilon)$ of $x \otimes y$. Note that $x \otimes_A y$ is independent, up to canonical isomorphism, of the choice of separability section σ. We get a split coequalizer in $\mathcal{K}$,

$$\begin{tikzcd} x \otimes A \otimes y \arrow[r, shift left, "\varrho_1\otimes 1"] \arrow[r, shift right, "1\otimes\varrho_2"'] & x \otimes y \arrow[r, two heads] & x \otimes_A y, \end{tikzcd}$$

and A acts on $x \otimes_A y$ by

$$A \otimes x \otimes_A y \hookrightarrow A \otimes x \otimes y \xrightarrow{\varrho_1\otimes 1} x \otimes y \twoheadrightarrow x \otimes_A y.$$

Proposition 1.11. *The tensor product $\otimes_A$ yields a symmetric monoidal structure on $A\text{-Mod}_{\mathcal{K}}$ under which F_A becomes monoidal. We will write $\mathbb{1}_A = A$ for the unit object in $A\text{-Mod}_{\mathcal{K}}$.*

Notation 1.12. If A and B are rings in $\mathcal{K}$ and $h : A \to B$ is a ring morphism, we say that B is an *A-algebra*. As usual, we equip B with the A-module structure given by

$$A \otimes B \xrightarrow{h\otimes 1} B \otimes B \xrightarrow{\mu_B} B,$$

and we write $\bar{B}$ for the corresponding object in $A\text{-Mod}_{\mathcal{K}}$, so that $B = U_A(\bar{B})$.

Remark 1.13. Let A be a separable ring in $\mathcal{K}$. There is a one-to-one correspondence between A-algebras B in $\mathcal{K}$ and rings $\bar{B}$ in $A\text{-Mod}_{\mathcal{K}}$. More precisely, if (B, μ, η) is a ring in $\mathcal{K}$ and $h : A \to B$ is a ring morphism, then $(\bar{B},\ \bar{\mu},\ \bar{\eta} := h)$ defines a ring in $A\text{-Mod}_{\mathcal{K}}$, with $\mu : B \otimes B \twoheadrightarrow B \otimes_A B \xrightarrow{\bar{\mu}} B$. Moreover, B is separable in $\mathcal{K}$ if and only if $\bar{B}$ is separable in $A\text{-Mod}_{\mathcal{K}}$.

Remark 1.14. Let A be a separable ring in $\mathcal{K}$ and suppose B is an A-algebra via $h : A \to B$. For every A-module x, we let B act on the left factor of $F_h(x) := \bar{B} \otimes_A x$ as usual. This defines a functor $F_h : A\text{-Mod}_{\mathcal{K}} \to B\text{-Mod}_{\mathcal{K}}$ and the following diagram commutes up to isomorphism:

$$\begin{tikzcd} & \mathcal{K} \arrow[dl, "F_A"'] \arrow[dr, "F_B"] & \\ A\text{-Mod}_{\mathcal{K}} \arrow[rr, "F_h"] & & B\text{-Mod}_{\mathcal{K}} \end{tikzcd}$$

Note also that $F_{gh} \cong F_g F_h$ for any ring morphism $g : B \to C$.

Proposition 1.15. *Let A be a separable ring in $\mathcal{K}$ and suppose B is a separable A-algebra, say $\bar{B} \in \mathcal{L} := A\text{-Mod}_{\mathcal{K}}$. There is an equivalence $B\text{-Mod}_{\mathcal{K}} \simeq \bar{B}\text{-Mod}_{\mathcal{L}}$ of symmetric monoidal categories such that*

$$\begin{tikzcd} \mathcal{K} \arrow[r, "F_A"] \arrow[d, "F_B"] & \mathcal{L} \arrow[d, "F_{\bar{B}}"] \\ B\text{-Mod}_{\mathcal{K}} \arrow[r, "\simeq"] & \bar{B}\text{-Mod}_{\mathcal{L}}, \end{tikzcd}$$

commutes up to isomorphism.

2. Separable rings

Proposition 2.1. *Let A be a separable ring in $\mathcal{K}$. If $A \cong B \times C$ for rings B, C in $\mathcal{K}$, then any indecomposable ring factor of A is a ring factor of B or C. In particular, if A can be written as a product of indecomposable A-algebras $A \cong A_1 \times \cdots \times A_n$, this decomposition is unique up to isomorphism.*

Proof. Suppose $A_1 \in \mathcal{K}$ is an indecomposable ring factor of A, say $A \cong A_1 \times A_2$ for some ring A_2 in $\mathcal{K}$. The category $A\text{-Mod}_{\mathcal{K}}$ decomposes as

$$A\text{-Mod}_{\mathcal{K}} \cong A_1\text{-Mod}_{\mathcal{K}} \times A_2\text{-Mod}_{\mathcal{K}},$$

with $\mathbb{1}_A$ corresponding to $(\mathbb{1}_{A_1}, \mathbb{1}_{A_2})$. Accordingly, the A-algebras $\bar{B}$ and $\bar{C}$ correspond to (B_1, B_2) and (C_1, C_2) respectively, with B_i, C_i in $A_i\text{-Mod}_{\mathcal{K}}$ for $i = 1, 2$, such that $\bar{B} \cong B_1 \times B_2$ and $\bar{C} \cong C_1 \times C_2$ in $A\text{-Mod}_{\mathcal{K}}$. Given that $\mathbb{1}_A \cong \bar{B} \times \bar{C}$, we see $\mathbb{1}_{A_1} \cong B_1 \times C_1$, hence $A_1 \cong B_1$ or $A_1 \cong C_1$. ☐

Lemma 2.2. *Let A be a separable ring in $\mathcal{K}$.*

(a) *For every ring morphism $\alpha : A \to \mathbb{1}$, there exists a unique idempotent A-linear morphism $e : A \to A$ such that $\alpha e = \alpha$ and $e\eta\alpha = e$.*

(b) *Suppose $\mathbb{1}$ is indecomposable. If $\alpha_i : A \to \mathbb{1}$ are distinct ring morphisms for $1 \leq i \leq n$, with corresponding idempotent morphisms $e_i : A \to A$ as above, then $e_i e_j = \delta_{i,j} e_i$ and $\alpha_i e_j = \delta_{i,j}\alpha_i$.*

Proof. Let σ be a separability morphism for A. To show (a), consider the A-linear map $e := (\alpha \otimes 1)\sigma : A \to A$. We immediately see that $\alpha e = \alpha(\alpha \otimes 1)\sigma = \alpha\mu\sigma = \alpha$. Idempotence of e follows from the diagram

$$\begin{tikzcd}
A \arrow[r, "\sigma"] \arrow[dd, equal] & A \otimes A \arrow[r, "\alpha\otimes 1"] \arrow[d, "1\otimes\sigma"] & A \arrow[d, "\sigma"] \\
& A \otimes A \otimes A \arrow[r, "\alpha\otimes 1\otimes 1"] \arrow[d, "\mu\otimes 1"] & A \otimes A \arrow[d, "\alpha\otimes 1"] \\
A \arrow[r, "\sigma"] & A \otimes A \arrow[r, "\alpha\otimes 1"] & A
\end{tikzcd}$$

in which the left square commutes by bilinearity of σ. Seeing how

$$\begin{tikzcd}
A \arrow[rr, "\alpha"] \arrow[d, "1\otimes\eta"] & & \mathbb{1} \arrow[rr, "\eta"] & & A \arrow[d, "\sigma"] \\
A \otimes A \arrow[rr, "1\otimes\sigma"] \arrow[d, "\mu"] & & A \otimes A \otimes A \arrow[rr, "\alpha\otimes 1\otimes 1"] \arrow[d, "\mu\otimes 1"] & & A \otimes A \arrow[d, "\alpha\otimes 1"] \\
A \arrow[rr, "\sigma"] & & A \otimes A \arrow[rr, "\alpha\otimes 1"] & & \mathbb{1} \otimes A
\end{tikzcd}$$

commutes, we moreover get $e\eta\alpha = e$. Suppose e' is also an A-linear morphism with $\alpha e' = \alpha$ and $e'\eta\alpha = e'$. Then, $e = e\eta\alpha = e\eta\alpha e' = ee' = e'e = e'\eta\alpha e = e'\eta\alpha = e'$ by Remark 1.8. For (b), let $1 \leq i, j \leq n$. From the commuting diagram

$$\begin{tikzcd}
A \arrow[rr, "\alpha_i"] \arrow[d, "1\otimes\eta"] & & \mathbb{1} \arrow[rr, "\eta"] & & A \arrow[d, "e_j"] \\
A \otimes A \arrow[rr, "1\otimes e_j"] \arrow[d, "\mu"] & & A \otimes A \arrow[rr, "\alpha_i\otimes 1"] \arrow[d, "\mu"] & & A \arrow[d, "\alpha_i"] \\
A \arrow[rr, "e_j"] & & A \arrow[rr, "\alpha_i"] & & \mathbb{1},
\end{tikzcd}$$

we see that $\alpha_i e_j \eta\alpha_i = \alpha_i e_j$. Hence, $(\alpha_i e_j \eta)(\alpha_i e_j \eta) = \alpha_i e_j e_j \eta = \alpha_i e_j \eta$, so the morphism $\alpha_i e_j \eta : \mathbb{1} \to \mathbb{1}$ is idempotent and equals 0 or $1_{\mathbb{1}}$. In the first case, $\alpha_i e_j = \alpha_i e_j \eta \alpha_j = 0$ and $e_i e_j = e_i \eta \alpha_i e_j = 0$, in particular $i \neq j$. On the other hand, if $\alpha_i e_j \eta = 1_{\mathbb{1}}$ we get $\alpha_i e_j = \alpha_i e_j \eta\alpha_i = \alpha_i$ and $\alpha_i e_j = \alpha_i e_j \eta \alpha_j = \alpha_j$, so $i = j$. $\square$

Lemma 2.3. *Let (A, μ_A, η_A) and (B, μ_B, η_B) be separable rings in $\mathcal{K}$.*

(a) *Suppose $f : A \to B$ and $g : B \to A$ are ring morphisms such that $gf = 1_A$. We equip A with the structure of B^e-module via the morphism g. There exists a B^e-linear morphism $\tilde{f} : A \to B$ such that $g\tilde{f} = 1_A$. In particular, A is a direct summand of B as a B^e-module.*

(b) *Suppose A is indecomposable. Let $g_i : B \to A$ be distinct ring morphisms for $1 \le i \le n$ and suppose $f : A \to B$ is a ring morphism with $g_i f = 1_A$. Then $A^{\oplus n}$ is a direct summand of B as a B^e-module, with projections $g_i : B \to A$ for $1 \le i \le n$.*

Proof. Considering the A-module structure on B given by f, we note that $g : B \to A$ is A-linear:

$$\begin{tikzcd}
A \otimes B \arrow[r, "f\otimes 1"] \arrow[d, "1\otimes g"] & B \otimes B \arrow[r, "\mu_B"] \arrow[d, "g\otimes g"] & B \arrow[d, "g"] \\
A \otimes A \arrow[r, equal] & A \otimes A \arrow[r, "\mu_A"] & A
\end{tikzcd}$$

We can thus apply Lemma 2.2 to the ring morphism $\bar{g} : \bar{B} \to \mathbb{1}_A$ in A-$\mathrm{Mod}_{\mathcal{K}}$ and find an idempotent $\bar{B}^e$-linear morphism $\bar{e} : \bar{B} \to \bar{B}$ such that $\bar{g}\bar{e} = \bar{g}$ and $\bar{e}\eta_{\bar{B}}\bar{g} = \bar{e}$. Forgetting the A-action, $U_A(\bar{e}) := e : B \to B$ is idempotent and B^e-linear, with $ge = g$ and $efg = e$. Let $\tilde{f} := ef$. We need to show that $\tilde{f}$ is B^e-linear, where B^e acts on A via g. Left B-linearity of $\tilde{f}$ follows from the commuting diagram

$$\begin{tikzcd}
B \otimes A \arrow[r, "g\otimes 1"] \arrow[dd, "1\otimes f"] & A \otimes A \arrow[r, "\mu_A"] \arrow[d, "f\otimes f"] & A \arrow[d, "f"] \\
 & B \otimes B \arrow[r, "\mu_B"] \arrow[d, "e\otimes 1"] & B \arrow[d, "e"] \\
B \otimes B \arrow[r, "e\otimes 1"] \arrow[d, "1\otimes e"] & B \otimes B \arrow[r, "\mu_B"] & B \arrow[d, equal] \\
B \otimes B \arrow[rr, "\mu_B"] & & B
\end{tikzcd}$$

and right B-linearity follows similarly. Finally, $g\tilde{f} = gef = gf = 1_A$.

For (b), let $g_i : B \to A$ be distinct ring morphisms with $g_i f = 1_A$ for $1 \le i \le n$. As in part (a), we find idempotent B^e-linear morphisms $e_i : B \to B$ and B^e-linear morphisms $\tilde{f}_i := e_i f$ with $g_i \tilde{f}_i = 1_A$ and $e_i = \tilde{f}_i g_i$. In fact, Lemma 2.2(b) shows the e_i are orthogonal. Seeing how $A = \mathrm{im}(e_i)$, we conclude $A^{\oplus n}$ is a direct summand of B as a B^e-module, with projections $g_i : B \to A$ for $1 \le i \le n$. □

Corollary 2.4. *Let A and B be separable rings in $\mathcal{K}$ and suppose B is an A-algebra. The corresponding ring $\bar{B}$ in A-$\mathrm{Mod}_{\mathcal{K}}$ is a ring factor of $F_A(B)$.*

Proof. Applying Lemma 2.3 to the ring morphisms $f : B \xrightarrow{\eta_A \otimes 1_B} A \otimes B$ and g given by the action of A on B, we see that B is a direct summand of $A \otimes B$ as $(A \otimes B)^e$-modules in $\mathcal{K}$. In particular, $\bar{B}$ is a direct summand of $F_A(B)$ as $F_A(B)^e$-modules in A-$\mathrm{Mod}_{\mathcal{K}}$. By Lemma 1.10, $\bar{B}$ admits a ring structure under which $\bar{B}$ becomes a ring factor of $F_A(B)$. This new ring structure on $\bar{B}$ is the original one, seeing how the projection $g : F_A(B) \to \bar{B}$ is a ring morphism for both structures. $\square$

3. Degree of a separable ring

We recall Balmer's definition [2014] of the degree of a separable ring in a tensor-triangulated category, and show the definition works for any idempotent-complete symmetric monoidal category $\mathcal{K}$.

Theorem 3.1. *Let A and B be separable rings in $\mathcal{K}$. Suppose $f : A \to B$ and $g : B \to A$ are ring morphisms such that $gf = 1_A$. There exists a separable ring C in $\mathcal{K}$ and a ring isomorphism $h : B \xrightarrow{\sim} A \times C$ such that $\mathrm{pr}_1\, h = g$. If we equip C with the A-algebra structure coming from $\mathrm{pr}_2\, hf$, it is unique up to isomorphism of A-algebras.*

Proof. This proposition is proved in [Balmer 2014, Theorem 2.4] when $\mathcal{K}$ is a tensor-triangulated category. In our case, Lemma 2.3 yields an isomorphism $h : B \xrightarrow{\sim} A \oplus C$ of B^e-modules with $\mathrm{pr}_1\, h = g$. By Lemma 1.10, A and C admit ring structures under which h becomes a ring isomorphism. This new ring structure on A is the original one, seeing how

$$1_A : A \xrightarrow{f} B \xrightarrow{\mathrm{pr}_1 h} A$$

is a ring morphism. The rest of the proof is identical to the proof in [loc. cit.]. $\square$

Definition 3.2 [Balmer 2014, Definition 3.1]. Let (A, μ, η) be a separable ring in $\mathcal{K}$. Applying Theorem 3.1 to the ring morphisms $f = 1_A \otimes \eta : A \to A \otimes A$ and $g = \mu : A \otimes A \to A$, we find a separable A-algebra A', unique up to isomorphism, and a ring isomorphism $h : A \otimes A \xrightarrow{\sim} A \times A'$ such that $\mathrm{pr}_1\, h = \mu$.

The *splitting tower*

$$\mathbb{1} = A^{[0]} \xrightarrow{\eta} A = A^{[1]} \to A^{[2]} \to \cdots \to A^{[n]} \to A^{[n+1]} \to \cdots$$

is defined inductively by $A^{[n+1]} = (A^{[n]})'$, where we consider $A^{[n]}$ as a ring in $A^{[n-1]}$-$\mathrm{Mod}_{\mathcal{K}}$. We say the *degree* of A is d, writing $\deg_{\mathcal{K}}(A) = d$, if $A^{[d]} \neq 0$ and $A^{[d+1]} = 0$. We say A has *infinite degree* if $A^{[d]} \neq 0$ for all $d \geq 0$.

Remark 3.3. By construction, we have $(A^{[n]})^{[m+1]} \cong A^{[n+m]}$ as $A^{[n+m-1]}$-algebras for all $m \geq 0$ and $n \geq 1$, where we regard $A^{[n]}$ as a ring in $A^{[n-1]}$-$\mathrm{Mod}_{\mathcal{K}}$. In other words, $\deg_{A^{[n-1]}\text{-}\mathrm{Mod}_{\mathcal{K}}}(A^{[n]}) = \deg_{\mathcal{K}}(A) - n + 1$ for $1 \leq n \leq \deg_{\mathcal{K}}(A) + 1$.

Example 3.4. Let R be a commutative ring and suppose A is a commutative projective separable R-algebra. If $\operatorname{Spec} R$ is connected, then the degree of A as a ring in $\mathrm{D}^{\mathrm{perf}}(R)$ (see Example 1.5) recovers its rank as an R-module. This will follow from Proposition 7.9.

Proposition 3.5. *Let A and B be separable rings in $\mathcal{K}$.*

(a) *We have $F_{A^{[n]}}(A) \cong \mathbb{1}_{A^{[n]}}^{\times n} \times A^{[n+1]}$ as $A^{[n]}$-algebras.*

(b) *Let $F : \mathcal{K} \to L$ be an additive monoidal functor. For every $n \geq 0$, the rings $F(A^{[n]})$ and $F(A)^{[n]}$ are isomorphic. In particular, $\deg_{\mathcal{L}}(F(A)) \leq \deg_{\mathcal{K}}(A)$.*

(c) *Suppose A is a B-algebra. Then $\deg_{B\text{-}\mathrm{Mod}_{\mathcal{K}}}(F_B(A)) = \deg_{\mathcal{K}}(A)$.*

Proof. The proofs for (a) and (b) in [op. cit., Theorems 3.7 and 3.9] still hold in our (not necessarily triangulated) setting. To prove (c), note that $A^{[n]}$ is a B-algebra and hence a direct summand of $F_B(A^{[n]}) \cong F_B(A)^{[n]}$. This means $F_B(A)^{[n]} \neq 0$ when $A^{[n]} \neq 0$ so that $\deg_{B\text{-}\mathrm{Mod}_{\mathcal{K}}}(F_B(A)) \geq \deg_{\mathcal{K}}(A)$. □

Lemma 3.6 [Balmer 2014, Lemma 3.11]. *Let $n \geq 1$ and $A := \mathbb{1}^{\times n} \in \mathcal{K}$. There is an isomorphism $A^{[2]} \cong A^{\times(n-1)}$ of A-algebras.*

Proof. We prove there is an A-algebra isomorphism $\lambda : A \otimes A \xrightarrow{\sim} A \times A^{\times(n-1)}$ with $\mathrm{pr}_1 \lambda = \mu_A$. We write $A = \prod_{i=0}^{n-1} \mathbb{1}_i$, $A \otimes A = \prod_{0 \leq i,j \leq n-1} \mathbb{1}_i \otimes \mathbb{1}_j$ and $A^{\times n} = \prod_{k=0}^{n-1} \prod_{i=0}^{n-1} \mathbb{1}_{ik}$ with $\mathbb{1} = \mathbb{1}_i = \mathbb{1}_{ik}$ for all i, k. Define $\lambda : A \otimes A \to A^{\times n}$ by mapping the factor $\mathbb{1}_i \otimes \mathbb{1}_j$ identically to $\mathbb{1}_{i(i-j)}$, with indices in $\mathbb{Z}_n$. Then, λ is an A-algebra isomorphism and $\mathrm{pr}_{k=0} \lambda = \mu_A$. □

Corollary 3.7. *Let $n \geq 1$. Then $\deg_{\mathcal{K}}(\mathbb{1}^{\times n}) = n$ and $(\mathbb{1}^{\times n})^{[n]} \cong \mathbb{1}^{\times n!}$ in $\mathcal{K}$.*

Proof. Let $A := \mathbb{1}^{\times n}$. The result is clear when $n = 1$, and we proceed by induction on n. By Lemma 3.6, we know $A^{[2]} \cong \mathbb{1}_A^{\times(n-1)}$ in $A\text{-}\mathrm{Mod}_{\mathcal{K}}$. Assuming the induction hypothesis, $\deg_{A\text{-}\mathrm{Mod}_{\mathcal{K}}}(A^{[2]}) = n - 1$ and

$$A^{[n]} \cong (A^{[2]})^{[n-1]} \cong \mathbb{1}_A^{\times(n-1)!} \cong (\mathbb{1}^{\times n})^{\times(n-1)!} \cong \mathbb{1}^{\times n!}. \qquad \square$$

Lemma 3.8. *Let A and B be separable rings of finite degree in $\mathcal{K}$. Then,*

(a) $\deg(A \times B) \leq \deg(A) + \deg(B)$

(b) $\deg(A \times \mathbb{1}^{\times n}) = \deg(A) + n$

(c) $\deg(A^{\times t}) = \deg(A) \cdot t$.

Proof. To prove (a), let $n := \deg(A \times B)$ and $C := (A \times B)^{[n]}$. Writing $A' := F_C(A)$ and $B' := F_C(B)$, we know from Proposition 3.5(a) that

$$A' \times B' = F_C(A \times B) \cong \mathbb{1}_C^{\times n}.$$

If we let $D := (A')^{[\deg(A')]}$ and apply F_D to the isomorphism, we get

$$\mathbb{1}_D^{\times \deg(A')} \times F_D(B') \cong \mathbb{1}_D^{\times n}.$$

Similarly, putting $E := (F_D(B'))^{[\deg(F_D(B'))]}$ and applying F_E gives

$$\mathbb{1}_E^{\times \deg(A')} \times \mathbb{1}_E^{\times \deg(F_D(B'))} \cong \mathbb{1}_E^{\times n}.$$

This shows $n = \deg(A') + \deg(F_D(B')) \leq \deg(A) + \deg(B)$ by Proposition 3.5(b). For (b), let $B := A^{[\deg(A)]}$. Then, $F_B(A \times \mathbb{1}^{\times n}) \cong \mathbb{1}_B^{\times \deg(A)} \times \mathbb{1}_B^{\times n}$ and we find

$$\deg(A \times \mathbb{1}^{\times n}) \geq \deg(F_B(A \times \mathbb{1}^{\times n})) = \deg(A) + n.$$

To prove (c), we write $B := A^{[\deg(A)]}$ again and note that $F_B(A^{\times t}) \cong (\mathbb{1}_B^{\times \deg(A)})^{\times t}$. Hence, $\deg(A^{\times t}) \geq \deg(F_B(A^{\times t})) = \deg(A) \cdot t$. □

4. Counting ring morphisms

Lemma 4.1. *Let A be a separable ring in $\mathcal{K}$ and suppose $\mathbb{1}$ is indecomposable. If there are n distinct ring morphisms $A \to \mathbb{1}$, then A has $\mathbb{1}^{\times n}$ as a ring factor. In particular, there are at most* $\deg A$ *distinct ring morphisms $A \to \mathbb{1}$.*

Proof. Let $\alpha_i : A \to \mathbb{1}$ be distinct ring morphisms for $1 \leq i \leq n$. By Lemma 2.3(b), we know that $\mathbb{1}^{\oplus n}$ is a direct summand of A as an A^e-module, with projections $\alpha_i : A \to \mathbb{1}$ for $1 \leq i \leq n$. Moreover, Lemma 1.10 shows that every such summand $\mathbb{1}$ admits a ring structure, under which $\mathbb{1}^{\times n}$ becomes a ring factor of A and the projections α_i are ring morphisms. In fact, these new ring structures on $\mathbb{1}$ are the original one, seeing how $\alpha_i \eta_A = 1_{\mathbb{1}}$ is a ring morphism for every $1 \leq i \leq n$. Finally, Lemma 3.8(b) shows that $\deg(A) \geq n$. □

Proposition 4.2. *Let A and B be separable rings in $\mathcal{K}$ and suppose B is indecomposable. Let $n \geq 1$. The following are equivalent*:

(i) *There are* (*at least*) *n distinct ring morphisms $A \to B$ in $\mathcal{K}$.*

(ii) *The ring $\mathbb{1}_B^{\times n}$ is a ring factor of $F_B(A)$ in B-*$\mathrm{Mod}_{\mathcal{K}}$.

(iii) *There is a ring morphism $A^{[n]} \to B$ in $\mathcal{K}$.*

Proof. Firstly, we claim there is a one-to-one correspondence between ring morphisms $\alpha : A \to B$ in $\mathcal{K}$ and ring morphisms $\beta : F_B(A) \to \mathbb{1}_B$ in B-$\mathrm{Mod}_{\mathcal{K}}$. Indeed, this correspondence sends $\alpha : A \to B$ in $\mathcal{K}$ to the B-algebra morphism

$$B \otimes A \xrightarrow{1_B \otimes \alpha} B \otimes B \xrightarrow{\mu} B,$$

and conversely, $\beta : F_B(A) \to \mathbb{1}_B$ gets mapped to $A \xrightarrow{\eta_B \otimes 1_A} B \otimes A \xrightarrow{\beta} B$ in $\mathcal{K}$.

To show (i) $\Rightarrow$ (ii), note that n distinct ring morphisms $A \to B$ in $\mathcal{K}$ give n distinct ring morphisms $F_B(A) \to \mathbb{1}_B$ in B-$\mathrm{Mod}_{\mathcal{K}}$. By Lemma 4.1, $\mathbb{1}_B^{\times n}$ is a ring factor of $F_B(A)$. For (ii) $\Rightarrow$ (i), suppose $\mathbb{1}_B^{\times n}$ is a ring factor of $F_B(A)$ in B-$\mathrm{Mod}_{\mathcal{K}}$ and consider the projections $\mathrm{pr}_i : F_B(A) \to \mathbb{1}_B$ with $1 \leq i \leq n$. By the claim, there are at least n distinct ring morphisms $A \to B$ in $\mathcal{K}$.

We show (ii) $\Rightarrow$ (iii) by induction on n. The case $n = 1$ has already been proven. Let $n \geq 1$ and suppose $\mathbb{1}_B^{\times(n+1)}$ is a ring factor of $F_B(A)$. By the induction hypothesis, there exists a ring morphism $A^{[n]} \to B$. As usual, we write $\bar{B}$ for the separable ring in $A^{[n]}$-$\mathrm{Mod}_{\mathcal{K}}$ corresponding to the $A^{[n]}$-algebra B in $\mathcal{K}$. The diagram

$$\begin{tikzcd} \mathcal{K} \arrow[r, "F_{A^{[n]}}"] \arrow[d, "F_B"] & A^{[n]}\text{-}\mathrm{Mod}_{\mathcal{K}} \arrow[d, "F_{\bar{B}}"] \\ B\text{-}\mathrm{Mod}_{\mathcal{K}} \arrow[r, "\simeq"] & \bar{B}\text{-}\mathrm{Mod}_{A^{[n]}\text{-}\mathrm{Mod}_{\mathcal{K}}} \end{tikzcd} \tag{4.3}$$

from Proposition 1.15 shows that $F_B(A)$ is mapped to $F_{\bar{B}}(F_{A^{[n]}}(A))$ under the equivalence B-$\mathrm{Mod}_{\mathcal{K}} \simeq \bar{B}$-$\mathrm{Mod}_{A^{[n]}\text{-}\mathrm{Mod}_{\mathcal{K}}}$. It follows that $\mathbb{1}_{\bar{B}}^{\times(n+1)}$ is a ring factor of $F_{\bar{B}}(F_{A^{[n]}}(A))$. On the other hand, by Proposition 3.5(a) we know that

$$F_{\bar{B}}(F_{A^{[n]}}(A)) \cong F_{\bar{B}}(\mathbb{1}_{A^{[n]}}^{\times n} \times A^{[n+1]}) \cong \mathbb{1}_{\bar{B}}^{\times n} \times F_{\bar{B}}(A^{[n+1]}). \tag{4.4}$$

Hence, $\mathbb{1}_{\bar{B}}$ is a ring factor of $F_{\bar{B}}(A^{[n+1]})$ by Proposition 2.1 and we conclude there exists a ring morphism $A^{[n+1]} \to \bar{B}$ in $A^{[n]}$-$\mathrm{Mod}_{\mathcal{K}}$.

To show (iii) $\Rightarrow$ (ii), suppose B is an $A^{[n]}$-algebra and write $\bar{B}$ for the corresponding separable ring in $A^{[n]}$-$\mathrm{Mod}_{\mathcal{K}}$. Using diagram (4.3) again, it is enough to show that $\mathbb{1}_{\bar{B}}^{\times n}$ is a ring factor of $F_{\bar{B}}(F_{A^{[n]}}(A))$. This follows from (4.4). □

Theorem 4.5. *Let A and B be separable rings in $\mathcal{K}$, where A has finite degree and B is indecomposable. There are at most* $\deg(A)$ *distinct ring morphisms from A to B.*

Proof. If there are n distinct ring morphisms from A to B, we know $\mathbb{1}_B^{\times n}$ is a ring factor of $F_B(A)$ by Proposition 4.2. So, $n \leq \deg_{B\text{-}\mathrm{Mod}_{\mathcal{K}}}(F_B(A)) \leq \deg_{\mathcal{K}}(A)$ by Proposition 3.5(b) and Lemma 3.8(b). □

Remark 4.6. The assumption B is indecomposable is necessary in Theorem 4.5. Indeed, $\deg(\mathbb{1}^{\times n}) = n$ but $\mathbb{1}^{\times n}$ has at least $n!$ ring endomorphisms.

5. Quasi-Galois theory

Suppose (A, μ, η) is a nonzero ring in $\mathcal{K}$ and Γ is a finite set of ring endomorphisms of A with $1_A \in \Gamma$. Consider the ring $\prod_{\gamma\in\Gamma} A_\gamma$, where we write $A_\gamma = A$ for all $\gamma \in \Gamma$ to keep track of the different copies of A. We define ring morphisms $\varphi_1 : A \to \prod_{\gamma\in\Gamma} A_\gamma$ by $\mathrm{pr}_\gamma\, \varphi_1 = 1_A$ and $\varphi_2 : A \to \prod_{\gamma\in\Gamma} A_\gamma$ by $\mathrm{pr}_\gamma\, \varphi_2 = \gamma$ for all $\gamma \in \Gamma$. Thus, φ_1 renders the (standard) left A-module structure on $\prod_{\gamma\in\Gamma} A_\gamma$ and we introduce a right A-module structure on $\prod_{\gamma\in\Gamma} A_\gamma$ via φ_2.

Definition 5.1. We will consider the following ring morphism:

$$\lambda_\Gamma = \lambda : A \otimes A \longrightarrow \prod_{\gamma\in\Gamma} A_\gamma \quad \text{with } \mathrm{pr}_\gamma\, \lambda = \mu(1 \otimes \gamma).$$

Note that $\lambda(1\otimes\eta)=\varphi_1$ and $\lambda(\eta\otimes 1)=\varphi_2$,

$$\begin{tikzcd} & A \arrow[dl, shift right, "1\otimes\eta"'] \arrow[dl, shift left, "\eta\otimes 1"] \arrow[dr, shift left, "\varphi_2"] \arrow[dr, shift right, "\varphi_1"'] & \\ A\otimes A \arrow[rr, "\lambda"] & & \prod_{\gamma\in\Gamma} A_\gamma \end{tikzcd} \tag{5.2}$$

so that λ is an A^e-algebra morphism.

Lemma 5.3. *Suppose $\lambda_\Gamma : A\otimes A\to\prod_{\gamma\in\Gamma} A_\gamma$ is an isomorphism.*

(a) *There is an A^e-linear morphism $\sigma : A\to A\otimes A$ such that $\mu(1\otimes\gamma)\sigma=\delta_{1,\gamma}$ for every $\gamma\in\Gamma$. In particular, A is separable.*

(b) *Let $\gamma\in\Gamma$. If there exists a nonzero ring B in $\mathcal{K}$ and ring morphism $\alpha : A\to B$ with $\alpha\gamma=\alpha$, then $\gamma=1$.*

(c) *The separable ring A has degree $|\Gamma|$ in $\mathcal{K}$.*

Proof. To prove (a), consider the A^e-linear morphism $\sigma:=\lambda^{-1}\,\mathrm{incl}_1 : A\to A\otimes A$. The following diagram shows that $\mu(1\otimes\gamma)\sigma=\delta_{1,\gamma}$:

$$\begin{tikzcd} A \arrow[rr, "\sigma"] \arrow[dr, hook, "\mathrm{incl}_1"'] & & A\otimes A \arrow[r, "1\otimes\gamma"] \arrow[dr, "\lambda"'] & A\otimes A \arrow[r, "\mu"] & A \\ & \prod_{\gamma\in\Gamma} A_\gamma \arrow[ur, "\lambda^{-1}"'] & & \prod_{\gamma\in\Gamma} A_\gamma \arrow[ur, two heads, "\mathrm{pr}_\gamma"'] & \end{tikzcd}$$

For (b), suppose $\alpha\gamma=\alpha$ and $\sigma : A\to A\otimes A$ as in (a). We get

$$\alpha=\alpha\mu\sigma=\mu(\alpha\otimes\alpha)\sigma=\mu(\alpha\otimes\alpha)(1\otimes\gamma)\sigma=\alpha\mu(1\otimes\gamma)\sigma=\alpha\delta_{\gamma,1}.$$

Hence, either $\alpha=0$ or $\gamma=1_A$. Finally, given that $F_A(A)\cong\mathbb{1}_A^{\times|\Gamma|}$ in A-$\mathrm{Mod}_{\mathcal{K}}$, Proposition 3.5(c) shows that $\deg(A)=|\Gamma|$. $\square$

Definition 5.4. Suppose that A is a nonzero ring in $\mathcal{K}$ and Γ is a finite group of ring automorphisms of A. We say that A is *quasi-Galois in $\mathcal{K}$ with group Γ* if $\lambda_\Gamma : A\otimes A\to\prod_{\gamma\in\Gamma} A_\gamma$ is an isomorphism. By the above lemma, it follows that A is separable of degree $|\Gamma|$ in $\mathcal{K}$. We also call $F_A:\mathcal{K}\longrightarrow A$-$\mathrm{Mod}_{\mathcal{K}}$ a *quasi-Galois extension with group* Γ.

Example 5.5. Let $A:=\mathbb{1}^{\times n}$ and consider the ring morphism $\gamma:=(1\,2\cdots n)$ which permutes the factors. Then A is quasi-Galois with group $\Gamma=\{\gamma^i\mid 0\le i\le n-1\}\cong\mathbb{Z}_n$. Indeed, the isomorphism $\lambda : A\otimes A\to A^{\times n}$ constructed in the proof of Lemma 3.6 is exactly λ_Γ. In particular, Γ does not always contain all ring automorphisms of A.

Remark 5.6. The Galois theory of commutative rings was introduced by Auslander and Goldman [1960, Appendix], and was further developed by Chase, Harrison and Rosenberg [Chase et al. 1965] and many others. They considered commutative rings $R \subset A$ such that A is separable and projective as an R-algebra. If Γ is a finite group of ring automorphisms of A fixing R, then A is *Galois over R with group Γ* if the maps $R \hookrightarrow A^\Gamma$ and

$$A \otimes_R A \to \prod_{\gamma\in\Gamma} A, \qquad x \otimes y \mapsto (x \cdot \gamma(y))_{\gamma\in\Gamma}$$

are isomorphisms. In particular, A defines a ring object in the categories $R-\mathrm{mod}$ and $\mathrm{D}^{\mathrm{perf}}(R)$ (see Example 1.5), which is quasi-Galois with group Γ.

Lemma 5.7. *Let A be quasi-Galois of degree d in $\mathcal{K}$ with group Γ and suppose $F : \mathcal{K} \to \mathcal{L}$ is an additive monoidal functor. If $F(A) \neq 0$, then $F(A)$ is quasi-Galois of degree d in $\mathcal{L}$ with group $F(\Gamma) = \{F(\gamma) \mid \gamma \in \Gamma\}$. In particular, being quasi-Galois is stable under extension-of-scalars.*

Proof. We immediately see that

$$F(\lambda_\Gamma) : F(A) \otimes F(A) \cong F(A \otimes A) \to \prod_{\gamma\in\Gamma} F(A)$$

is an isomorphism in $\mathcal{L}$, so it suffices to show $\Gamma \cong F(\Gamma)$ and $F(\lambda_\Gamma) = \lambda_{F(\Gamma)}$. Now, λ_Γ is defined by $\mathrm{pr}_\gamma\, \lambda_\Gamma = \mu_A(1_A \otimes \gamma)$, hence $\mathrm{pr}_\gamma\, F(\lambda_\Gamma) = \mu_{F(A)}(1_{F(A)} \otimes F(\gamma))$ for every $\gamma \in \Gamma$. In particular, the morphisms $\mu_{F(A)}(1_{F(A)} \otimes F(\gamma))$ with $\gamma \in \Gamma$ are distinct. This shows the morphisms $F(\gamma)$ with $\gamma \in \Gamma$ are distinct, so that $\Gamma \cong F(\Gamma)$ and $F(\lambda_\Gamma) = \lambda_{F(\Gamma)}$. □

Proposition 5.8. *Suppose A is quasi-Galois in $\mathcal{K}$ with group Γ.*

(a) *If B is a separable indecomposable A-algebra, then Γ acts freely and transitively on the set of ring morphisms from A to B. In particular, there are exactly* $\deg(A)$ *distinct ring morphisms from A to B in $\mathcal{K}$.*

(b) *If A is indecomposable then Γ contains all ring endomorphisms of A.*

Proof. Note that the set S of ring morphisms from A to B is nonempty and Γ acts on S by precomposition. The action is free by Lemma 5.3(b) and transitive because $|S| \leq \deg A = |\Gamma|$ by Theorem 4.5. In particular, if A is indecomposable, then A has exactly $\deg A = |\Gamma|$ ring endomorphisms in $\mathcal{K}$. □

By the above proposition, we can simply say an indecomposable ring A in $\mathcal{K}$ is quasi-Galois, with the understanding that the Galois group Γ contains all ring endomorphisms of A.

Theorem 5.9. *Let A be a separable indecomposable ring of finite degree in $\mathcal{K}$ and write Γ for the set of ring endomorphisms of A. The following are equivalent:*

(i) $|\Gamma| = \deg(A)$.

(ii) $F_A(A) \cong \mathbb{1}_A^{\times t}$ *in A-$\mathrm{Mod}_{\mathcal{K}}$ for some $t > 0$.*

(iii) $\lambda_\Gamma : A \otimes A \to \prod_{\gamma\in\Gamma} A_\gamma$ *is an isomorphism.*

(iv) Γ *is a group and A is quasi-Galois in $\mathcal{K}$ with group Γ.*

Proof. First note that $d := \deg(A) = \deg(F_A(A))$ by Proposition 3.5(c). To show (i) $\Rightarrow$ (ii), recall that $\mathbb{1}_A^{\times d}$ is a ring factor of $F_A(A)$ if $|\Gamma| = d$ by Proposition 4.2. By Lemma 3.8(b), we know $F_A(A) \cong \mathbb{1}_A^{\times d}$. For (ii) $\Rightarrow$ (iii), we note that $t = d$ and consider an A-algebra isomorphism $l : A\otimes A \xrightarrow{\simeq} A^{\times d}$. We define ring endomorphisms

$$\alpha_i : A \xrightarrow{\eta\otimes 1_A} A \otimes A \xrightarrow{l} A^{\times d} \xrightarrow{\mathrm{pr}_i} A, \qquad i = 1, \ldots, d,$$

such that $\mu(1_A \otimes \alpha_i) = \mathrm{pr}_i\, l(\mu \otimes 1_A)(1_A \otimes \eta \otimes 1_A) = \mathrm{pr}_i\, l$ for every i. This shows the α_i are all distinct, so that $\Gamma = \{\alpha_i \mid 1 \le i \le d\}$ by Theorem 4.5 and $l = \lambda_\Gamma$. For (iii) $\Rightarrow$ (iv), we show that every $\gamma \in \Gamma$ is an automorphism. By Lemma 5.3(a), we can find an A^e-linear morphism $\sigma : A \to A\otimes A$ such that $\mu(1\otimes\gamma)\sigma = \delta_{1,\gamma}$ for every $\gamma \in \Gamma$. Let $\gamma \in \Gamma$ and note that $\gamma = \mu(\gamma \otimes 1)(1 \otimes \gamma)\sigma$ so that $(1 \otimes \gamma)\sigma : A \to A \otimes A$ is nonzero. Thus there exists $\gamma' \in \Gamma$ such that

$$\mathrm{pr}_{\gamma'}\, \lambda_\Gamma(1 \otimes \gamma)\sigma = \mu(1 \otimes \gamma')(1 \otimes \gamma)\sigma = \delta_{1,\gamma'\gamma}$$

is nonzero. This means $1 = \gamma'\gamma$ and $\gamma'(\gamma\gamma') = \gamma'$ so $\gamma\gamma' = 1$ by Lemma 5.3(b). Finally, (iv) $\Rightarrow$ (i) is the last part of Lemma 5.3. □

Corollary 5.10. *Let A, B and C be separable rings in $\mathcal{K}$ with $A \cong B \times C$, and suppose B is indecomposable. If $F_A(A) \cong \mathbb{1}_A^{\times d}$, then B is quasi-Galois. In particular, being quasi-Galois is stable under passing to indecomposable ring factors.*

Proof. Consider the decomposition $A\text{-Mod}_{\mathcal{K}} \cong B\text{-Mod}_{\mathcal{K}} \times C\text{-Mod}_{\mathcal{K}}$, under which $F_A(A)$ corresponds to $(F_B(B \times C), F_C(B \times C))$ and $\mathbb{1}_A^{\times d}$ corresponds to $(\mathbb{1}_B^{\times d}, \mathbb{1}_C^{\times d})$. Given that $\mathbb{1}_B$ is indecomposable and $F_B(B)$ is a ring factor of $\mathbb{1}_B^{\times d}$ in $B\text{-Mod}_{\mathcal{K}}$, we know $F_B(B) \cong \mathbb{1}_B^{\times t}$ for some $1 \le t \le d$. The result now follows from Theorem 5.9. □

6. Splitting rings

Definition 6.1. Let A and B be separable rings of finite degree in $\mathcal{K}$. We say B *splits* A if $F_B(A) \cong \mathbb{1}_B^{\times \deg(A)}$ in $B\text{-Mod}_{\mathcal{K}}$. We call an indecomposable ring B a *splitting ring* of A if B splits A and any ring morphism $C \to B$, where C is an indecomposable ring splitting A, is an isomorphism.

Remark 6.2. Let A be a separable ring in $\mathcal{K}$ with $\deg(A) = d$. The ring $A^{[d]}$ in $\mathcal{K}$ splits A by Proposition 3.5(a). Moreover, if B is a separable indecomposable ring in $\mathcal{K}$, then B splits A if and only if B is an $A^{[d]}$-algebra. This follows immediately from Proposition 4.2.

Remark 6.3. Let A be a separable ring in $\mathcal{K}$ with $\deg(A) = d$. The ring $A^{[d]}$ in $\mathcal{K}$ splits itself by Proposition 3.5(a), (b) and Corollary 3.7:

$$F_{A^{[d]}}(A^{[d]}) \cong (F_{A^{[d]}}(A))^{[d]} \cong (\mathbb{1}_{A^{[d]}}^{\times d})^{[d]} \cong \mathbb{1}_{A^{[d]}}^{\times d!}.$$

Lemma 6.4. *Let A be a separable ring in $\mathcal{K}$ that splits itself. If A_1 and A_2 are indecomposable ring factors of A, then any ring morphism $A_1 \to A_2$ is an isomorphism.*

Proof. Let A_1 and A_2 be indecomposable ring factors of A and suppose there is a ring morphism $f : A_1 \to A_2$. We know $F_{A_1}(A) \cong \mathbb{1}_{A_1}^{\times \deg(A)}$ because A splits itself. Meanwhile, $F_{A_1}(A_2)$ is a ring factor of $F_{A_1}(A)$, so that $F_{A_1}(A_2) \cong \mathbb{1}_{A_1}^{\times d}$ for some $d \geq 0$. In fact, $d = \deg(A_2) \geq 1$ by Proposition 3.5(c). Proposition 4.2 shows there exists a ring morphism $g : A_2 \to A_1$. Note that A_1 and A_2 are quasi-Galois by Corollary 5.10, so that the ring morphisms $gf : A_1 \to A_1$ and $fg : A_2 \to A_2$ are isomorphisms by Proposition 5.8(b). □

Definition 6.5. We say $\mathcal{K}$ is *nice* if for every separable ring A of finite degree in $\mathcal{K}$, there are indecomposable rings $A_1, \ldots, A_n$ in $\mathcal{K}$ such that $A \cong A_1 \times \cdots \times A_n$.

Example 6.6. Let G be a group and $\Bbbk$ a field. The categories $\Bbbk G$-mod, $\mathrm{D}^b(\Bbbk G\text{-mod})$ and $\Bbbk G$-stab (see Section 10) are nice categories. More generally, $\mathcal{K}$ is nice if it satisfies Krull–Schmidt.

Example 6.7. Let X be a noetherian scheme and let $\mathrm{D}^{\mathrm{perf}}(X)$ be the derived category of perfect complexes over X with left derived tensor product. By Example 7.4 and Proposition 7.12, $\mathrm{D}^{\mathrm{perf}}(X)$ is nice.

Lemma 6.8. *Suppose $\mathcal{K}$ is nice and let A, B be separable rings of finite degree in $\mathcal{K}$. If B is indecomposable and there exists a ring morphism $A \to B$ in $\mathcal{K}$, then there exists a ring morphism $C \to B$ for some indecomposable ring factor C of A.*

Proof. Since $\mathcal{K}$ is nice, we can write $A \cong A_1 \times \cdots \times A_n$ with A_i indecomposable for $1 \leq i \leq n$. If there exists a ring morphism $A \to B$ in $\mathcal{K}$, Proposition 4.2 shows that $\mathbb{1}_B$ is a ring factor of $F_B(A) \cong F_B(A_1) \times \cdots \times F_B(A_n)$. Since $\mathbb{1}_B$ is indecomposable, it is a ring factor of some $F_B(A_i)$ with $1 \leq i \leq n$ by Proposition 2.1. □

Proposition 6.9. *Suppose $\mathcal{K}$ is nice and let A be a separable ring of finite degree in $\mathcal{K}$. An indecomposable ring B in $\mathcal{K}$ is a splitting ring of A if and only if B is a ring factor of $A^{[\deg(A)]}$. In particular, any separable ring in $\mathcal{K}$ has a splitting ring and at most finitely many.*

Proof. Let $d := \deg(A)$ and suppose B is a splitting ring of A. By Remark 6.2, B is an $A^{[d]}$-algebra. Hence, there exists a ring morphism $C \to B$ for some indecomposable ring factor C of $A^{[d]}$ by Lemma 6.8. Now, $A^{[d]}$ splits A, so C splits A and the ring morphism $C \to B$ is an isomorphism. Conversely, suppose B is an indecomposable ring factor of $A^{[d]}$, so B splits A. Let C be an indecomposable separable ring splitting A and suppose there is a ring morphism $C \to B$. As before, C is an $A^{[d]}$-algebra and there exists a ring morphism $B' \to C$ for some indecomposable ring factor B' of $A^{[d]}$. The composition $B' \to C \to B$ is an isomorphism by Remark 6.3 and Lemma 6.4. In other words, B is a ring factor of the indecomposable ring C, so that $C \cong B$. □

Corollary 6.10. *Suppose $\mathcal{K}$ is nice and B is a separable indecomposable ring of finite degree in $\mathcal{K}$. Then B is quasi-Galois in $\mathcal{K}$ if and only if there exists a nonzero separable ring A of finite degree in $\mathcal{K}$ such that B is a splitting ring of A.*

Proof. Suppose B is indecomposable and quasi-Galois of degree t, so $B^{[2]} \cong \mathbb{1}_B^{\times(t-1)}$ as B-algebras. Then, B is a splitting ring for B because B is a ring factor of $B^{[t]}$:

$$B^{[t]} \cong (B^{[2]})^{[t-1]} \cong (\mathbb{1}_B^{\times(t-1)})^{[t-1]} \cong B^{\times(t-1)!}.$$

Now suppose B is a splitting ring for some A in $\mathcal{K}$, say with $\deg(A) = d > 0$. Seeing how $F_B(B)$ is a ring factor of

$$F_B(A^{[d]}) \cong F_B(A)^{[d]} \cong (\mathbb{1}_B^{\times d})^{[d]} = \mathbb{1}_B^{\times d!},$$

we know $F_B(B) \cong \mathbb{1}_B^{\times t}$ for some $t > 0$. By Theorem 5.9, B is quasi-Galois. □

7. Tensor triangular geometry

Definition 7.1. A *tt-category* $\mathcal{K}$ is an essentially small, idempotent-complete tensor-triangulated category. In particular, $\mathcal{K}$ comes equipped with a symmetric monoidal structure $(\otimes, \mathbb{1})$ such that $x \otimes - : \mathcal{K} \to \mathcal{K}$ is exact for all objects x in $\mathcal{K}$. A *tt-functor* $\mathcal{K} \to \mathcal{L}$ is an exact symmetric monoidal functor.

Throughout the rest of this paper, $(\mathcal{K}, \otimes, \mathbb{1})$ will denote a tt-category.

Remark 7.2. Balmer [2011] proved in that extension along a separable ring object A preserves the triangulation: $(A\text{-Mod}_{\mathcal{K}}, \otimes_A, \mathbb{1}_A)$ is a tt-category, extension-of-scalars F_A becomes a tt-functor and U_A is exact.

Definition 7.3. We briefly recall some tt-geometry and refer the reader to [Balmer 2005] for precise statements and motivation. The *spectrum* $\operatorname{Spc}(\mathcal{K})$ of a tt-category $\mathcal{K}$ is the set of all prime thick $\otimes$-ideals $\mathcal{P} \subsetneq \mathcal{K}$. The *support* of an object x in $\mathcal{K}$ is $\operatorname{supp}(x) = \{\mathcal{P} \in \operatorname{Spc}(\mathcal{K}) \mid x \notin \mathcal{P}\} \subset \operatorname{Spc}(\mathcal{K})$. The complements of these supports $\mathcal{U}(x) := \operatorname{Spc}(\mathcal{K}) - \operatorname{supp}(x)$ form an open basis for the *Zariski topology* on $\operatorname{Spc}(\mathcal{K})$.

Example 7.4. Let X be a noetherian scheme. Then $(\mathrm{D}^{\mathrm{perf}}(X), \otimes^{\mathrm{L}}_{\mathcal{O}_X})$ is a tt-category with spectrum $\operatorname{Spc}(\mathrm{D}^{\mathrm{perf}}(X))$ homeomorphic to X; see [op. cit., Theorem 6.3].

Remark 7.5. The spectrum is functorial. In particular, every tt-functor $F : \mathcal{K} \to \mathcal{L}$ induces a continuous map

$$\mathrm{Spc}(F) : \mathrm{Spc}(\mathcal{L}) \to \mathrm{Spc}(\mathcal{K}).$$

Moreover, for all $x \in \mathcal{K}$, we have

$$(\mathrm{Spc}\, F)^{-1}(\mathrm{supp}_{\mathcal{K}}(x)) = \mathrm{supp}_{\mathcal{L}}(F(x)) \subset \mathrm{Spc}\, \mathcal{L}.$$

Let A be a separable ring in $\mathcal{K}$. We will consider the continuous map

$$f_A := \mathrm{Spc}(F_A) : \mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}}) \to \mathrm{Spc}(\mathcal{K})$$

induced by the extension-of-scalars $F_A : \mathcal{K} \to A\text{-}\mathrm{Mod}_{\mathcal{K}}$.

Theorem 7.6 [Balmer 2016b, Theorem 3.14]. *Let A be a separable ring of finite degree in $\mathcal{K}$. Then*

$$\mathrm{Spc}((A \otimes A)\text{-}\mathrm{Mod}_{\mathcal{K}}) \overset{f_1}{\underset{f_2}{\rightrightarrows}} \mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}}) \xrightarrow{f_A} \mathrm{supp}_{\mathcal{K}}(A) \tag{7.7}$$

is a coequalizer, where f_1, f_2 are the maps induced by extension-of-scalars along the morphisms $1 \otimes \eta$ and $\eta \otimes 1 : A \to A \otimes A$ respectively. In particular, the image of f_A is $\mathrm{supp}_{\mathcal{K}}(A) \subset \mathrm{Spc}(\mathcal{K})$.

Definition 7.8. We call a tt-category $\mathcal{K}$ *local* if $x \otimes y = 0$ implies that x or y is $\otimes$-nilpotent for all $x, y \in \mathcal{K}$. The *local category $\mathcal{K}_{\mathcal{P}}$ at the prime $\mathcal{P} \in \mathrm{Spc}(\mathcal{K})$* is the idempotent completion of the Verdier quotient $\mathcal{K}/\mathcal{P}$. We write $q_{\mathcal{P}}$ for the canonical tt-functor $\mathcal{K} \twoheadrightarrow \mathcal{K}/\mathcal{P} \hookrightarrow \mathcal{K}_{\mathcal{P}}$.

Proposition 7.9 [Balmer 2014, Theorem 3.8]. *Suppose A is a separable ring in $\mathcal{K}$. If the ring $q_{\mathcal{P}}(A)$ has finite degree in $\mathcal{K}_{\mathcal{P}}$ for every $\mathcal{P} \in \mathrm{Spc}(\mathcal{K})$, then A has finite degree and*

$$\deg_{\mathcal{K}}(A) = \max_{\mathcal{P} \in \mathrm{Spc}(\mathcal{K})} \deg_{\mathcal{K}_{\mathcal{P}}}(q_{\mathcal{P}}(A)).$$

Proposition 7.10 [Balmer 2014, Corollary 3.12]. *Let $\mathcal{K}$ be a local tt-category and suppose A, B are separable rings of finite degree in $\mathcal{K}$. Then* $\deg(A \times B) = \deg(A) + \deg(B)$.

Lemma 7.11 [Balmer 2014, Theorem 3.7]. *Let A and B be separable rings in $\mathcal{K}$ and suppose* $\mathrm{supp}(A) \subseteq \mathrm{supp}(B)$. *Then* $\deg_{B\text{-}\mathrm{Mod}_{\mathcal{K}}}(F_B(A)) = \deg_{\mathcal{K}}(A)$.

Proposition 7.12. *Suppose the spectrum* $\mathrm{Spc}(\mathcal{K})$ *of $\mathcal{K}$ is noetherian. Then $\mathcal{K}$ is nice. That is, any separable ring A of finite degree in $\mathcal{K}$ has a decomposition $A \cong A_1 \times \ldots \times A_n$ where $A_1, \ldots, A_n$ are indecomposable rings in $\mathcal{K}$.*

Proof. Let A be a separable ring of finite degree in $\mathcal{K}$. We prove that any ring decomposition of A in $\mathcal{K}$ has at most finitely many nonzero ring factors. Suppose there is a sequence of nontrivial decompositions $A = A_1 \times B_1$, $B_1 = A_2 \times B_2, \ldots,$ with $B_n = A_{n+1} \times B_{n+1}$ for $n \geq 1$. By Proposition 7.10, we know

$$\deg(q_{\mathcal{P}}(B_n)) \geq \deg(q_{\mathcal{P}}(B_{n+1}))$$

for every $\mathcal{P} \in \mathrm{Spc}(\mathcal{K})$. We note that $\deg(q_{\mathcal{P}}(B_n)) \geq i$ if and only if $\mathcal{P} \in \mathrm{supp}(B_n^{[i]})$, so we get $\mathrm{supp}(B_n^{[i]}) \supseteq \mathrm{supp}(B_{n+1}^{[i]})$ for every $i \geq 0$. Since $\mathrm{Spc}(\mathcal{K})$ is noetherian, we can find $k \geq 1$ with $\mathrm{supp}(B_n^{[i]}) = \mathrm{supp}(B_{n+1}^{[i]})$ for every $i \geq 0$ and $n \geq k$. In particular, $\deg(q_{\mathcal{P}}(B_k)) = \deg(q_{\mathcal{P}}(B_{k+1}))$ for every $\mathcal{P} \in \mathrm{Spc}(\mathcal{K})$, so $q_{\mathcal{P}}(A_{k+1}) = 0$ for all $\mathcal{P} \in \mathrm{Spc}(\mathcal{K})$. By Proposition 7.9, we conclude $A_{k+1} = 0$, a contradiction. □

8. Rings of constant degree

Definition 8.1. We say a separable ring A in $\mathcal{K}$ has *constant degree* $d \in \mathbb{N}$ if the degree $\deg_{\mathcal{K}_{\mathcal{P}}} q_{\mathcal{P}}(A)$ equals d for every $\mathcal{P} \in \mathrm{supp}(A) \subset \mathrm{Spc}(\mathcal{K})$.

Lemma 8.2. *Let A be a separable ring of degree d in $\mathcal{K}$. Then A has constant degree if and only if* $\mathrm{supp}(A^{[d]}) = \mathrm{supp}(A)$.

Proof. Note that $\mathrm{supp}(A^{[2]}) \subseteq \mathrm{supp}(A)$ because $A \otimes A \cong A \times A^{[2]}$ in $\mathcal{K}$. Hence $\mathrm{supp}(A^{[d]}) \subseteq \mathrm{supp}(A)$. Now, let $\mathcal{P} \in \mathrm{supp}(A)$. Then $q_{\mathcal{P}}(A)$ has degree d if and only if $q_{\mathcal{P}}(A^{[d]}) \neq 0$, in other words $\mathcal{P} \in \mathrm{supp}(A^{[d]})$. □

Lemma 8.3. *Let A be a separable ring in $\mathcal{K}$ and suppose $F : \mathcal{K} \to \mathcal{L}$ is a tt-functor with $F(A) \neq 0$. If A has constant degree d, then $F(A)$ has constant degree d. Conversely, if $F(A)$ has constant degree d and* $\mathrm{supp}(A) \subset \mathrm{im}(\mathrm{Spc}(F))$, *then A has constant degree d.*

Proof. We first note that $\deg(F(A)) \leq \deg(A)$ by Proposition 3.5(b). Now, if A has constant degree d, then

$$\begin{aligned}\mathrm{supp}_{\mathcal{L}}(F(A)^{[d]}) = \mathrm{supp}_{\mathcal{L}}(F(A^{[d]})) &= \mathrm{Spc}(F)^{-1}(\mathrm{supp}_{\mathcal{K}}(A^{[d]})) \\ &= \mathrm{Spc}(F)^{-1}(\mathrm{supp}_{\mathcal{K}}(A)) = \mathrm{supp}_{\mathcal{L}}(F(A)) \neq \varnothing,\end{aligned}$$

which shows $F(A)$ has constant degree d. Conversely, suppose $F(A)$ has constant degree d and $\mathrm{supp}(A) \subset \mathrm{im}(\mathrm{Spc}(F))$. In particular, $\mathrm{supp}(A^{[d+1]}) \subset \mathrm{im}(\mathrm{Spc}(F))$, so

$$\varnothing = \mathrm{supp}(F(A^{[d+1]})) = \mathrm{Spc}(F)^{-1}(\mathrm{supp}(A^{[d+1]}))$$

implies $\mathrm{supp}(A^{[d+1]}) = \varnothing$. Thus A has degree d. Moreover, seeing how

$$\begin{aligned}\mathrm{Spc}(F)^{-1}(\mathrm{supp}_{\mathcal{K}}(A^{[d]})) &= \mathrm{supp}_{\mathcal{L}}(F(A)^{[d]}) \\ &= \mathrm{supp}_{\mathcal{L}}(F(A)) = \mathrm{Spc}(F)^{-1}(\mathrm{supp}_{\mathcal{K}}(A)),\end{aligned}$$

we can conclude $\mathrm{supp}_{\mathcal{K}}(A^{[d]}) = \mathrm{supp}_{\mathcal{K}}(A)$. □

Proposition 8.4. *Let A be a separable ring in $\mathcal{K}$. Then A has constant degree d if and only if there exists a separable ring B in $\mathcal{K}$ with* $\mathrm{supp}(A) \subset \mathrm{supp}(B)$ *and such that $F_B(A) \cong \mathbb{1}_B^{\times d}$. In particular, if A is quasi-Galois in $\mathcal{K}$ with group Γ, then A has constant degree $|\Gamma|$ in $\mathcal{K}$.*

Proof. If A has constant degree d, we can let $B := A^{[d]}$ and use Proposition 3.5(a). On the other hand, if A and B are separable rings in $\mathcal{K}$ with $\mathrm{supp}(A) \subset \mathrm{supp}(B)$, then Theorem 7.6 and Lemma 8.3 show that A has constant degree d whenever $F_B(A)$ has constant degree d. □

Proposition 8.5. *Let A be a separable ring of constant degree in $\mathcal{K}$ with connected support* $\mathrm{supp}(A) \subset \mathrm{Spc}(\mathcal{K})$. *If B and C are nonzero rings in $\mathcal{K}$ such that $A = B \times C$, then B and C have constant degree and* $\mathrm{supp}(A) = \mathrm{supp}(B) = \mathrm{supp}(C)$.

Proof. Given that A has constant degree d, we claim that for every $1 \le n \le d$,

$$\mathrm{supp}(A) = \mathrm{supp}(B^{[n]}) \sqcup \mathrm{supp}(C^{[d-n+1]}).$$

Fix $1 \le n \le d$ and suppose $\mathcal{P} \in \mathrm{supp}(B^{[n]}) \cap \mathrm{supp}(C^{[d-n+1]})$, so $\deg(q_{\mathcal{P}}(B)) \ge n$ and $\deg(q_{\mathcal{P}}(C)) \ge d-n+1$. By Proposition 7.10, $\deg(q_{\mathcal{P}}(A)) \ge d+1$, which is a contradiction. So far we've proven $\mathrm{supp}(A) \supset \mathrm{supp}(B^{[n]}) \sqcup \mathrm{supp}(C^{[d-n+1]})$. Now, if $\mathcal{P} \in \mathrm{supp}(A) - \mathrm{supp}(B^{[n]})$, we get $\deg(q_{\mathcal{P}}(A)) = d$ and $\deg(q_{\mathcal{P}}(B)) \le n-1$. It follows that $\deg(q_{\mathcal{P}}(C)) \ge d-n+1$, so $\mathcal{P} \in \mathrm{supp}(C^{[d-n+1]})$ and the claim follows.

Assuming A has connected support, we note that for every $1 \le n \le d$, either $\mathrm{supp}(B^{[n]}) = \mathrm{supp}(A)$ or $\mathrm{supp}(B^{[n]}) = \varnothing$. In particular, taking $n = \deg(B)$ and then $n = 1$ shows that $\mathrm{supp}(A) = \mathrm{supp}(B^{[\deg(B)]}) = \mathrm{supp}(B)$. Similarly, we see $\mathrm{supp}(A) = \mathrm{supp}(C^{[\deg(C)]}) = \mathrm{supp}(C)$ by letting $n = d+1-\deg(C)$ and then $n = 1$. In other words, $\mathrm{supp}(A) = \mathrm{supp}(B) = \mathrm{supp}(C)$ and B, C have constant degree. □

9. Quasi-Galois theory and tensor triangular geometry

Let A be a separable ring in $\mathcal{K}$ and suppose Γ is a finite group of ring automorphisms of A. Then, Γ acts on A-$\mathrm{Mod}_{\mathcal{K}}$ (see Remark 1.14) and therefore on the spectrum $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$.

Theorem 9.1. *Suppose A is quasi-Galois in $\mathcal{K}$ with group Γ. Then,*

$$\mathrm{supp}(A) \cong \mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})/\,\Gamma.$$

Proof. Diagram (5.2) yields a diagram of topological spaces

$$\begin{tikzcd}
 & \mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}}) & \\
\mathrm{Spc}((A \otimes A)\text{-}\mathrm{Mod}_{\mathcal{K}}) \arrow[ur, shift left, "f_1"] \arrow[ur, shift right, "f_2"'] \arrow[rr, "\cong", "l"'] & & \mathrm{Spc}(\prod_{\gamma \in \Gamma} A_\gamma\text{-}\mathrm{Mod}_{\mathcal{K}}), \arrow[ul, shift right, "g_2"'] \arrow[ul, shift left, "g_1"]
\end{tikzcd}$$

where f_1, f_2, g_1, g_2 and l are the maps induced by extension-of-scalars along the morphisms $1\otimes\eta$, $\eta\otimes 1$, φ_1, φ_2 and λ respectively (in the notation of Definition 5.1). That is, $g_1, g_2 : \bigsqcup_{\gamma\in\Gamma} \mathrm{Spc}(A_\gamma\text{-}\mathrm{Mod}_{\mathcal{K}}) \to \mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ are continuous maps such that $g_1\,\mathrm{incl}_\gamma$ is the identity and $g_2\,\mathrm{incl}_\gamma$ is the action of γ on $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$. Now, the coequalizer (7.7) turns into

$$\bigsqcup_{\gamma\in\Gamma} \mathrm{Spc}(A_\gamma\text{-}\mathrm{Mod}_{\mathcal{K}}) \overset{g_1}{\underset{g_2}{\rightrightarrows}} \mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}}) \xrightarrow{f_A} \mathrm{supp}(A),$$

which shows $\mathrm{supp}(A) \cong \mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})/\,\Gamma$. □

Remark 9.2. Let A be a ring in $\mathcal{K}$. We call A *nil-faithful* if $F_A(f) = 0$ implies f is $\otimes$-nilpotent for any morphism f in $\mathcal{K}$. By [Balmer 2016b, Proposition 3.15], A is nil-faithful if and only if $\mathrm{supp}(A) = \mathrm{Spc}(\mathcal{K})$. If A is nil-faithful and quasi-Galois in $\mathcal{K}$ with group Γ, Theorem 9.1 recovers $\mathrm{Spc}(\mathcal{K})$ as the Γ-orbits of $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$.

The following is a tensor-triangular version of Lemma 6.4.

Lemma 9.3. *Let A be a separable ring in $\mathcal{K}$ that splits itself. If A_1 and A_2 are indecomposable ring factors of A, then* $\mathrm{supp}(A_1)\cap\mathrm{supp}(A_2) = \varnothing$ *or* $A_1 \cong A_2$.

Proof. Let A_1 and A_2 be indecomposable ring factors of A and suppose A splits itself. We know $F_{A_1}(A) \cong \mathbb{1}_{A_1}^{\times \deg(A)}$ and hence $F_{A_1}(A_2) \cong \mathbb{1}_{A_1}^{\times t}$ for some $t \geq 0$. In fact, $t = 0$ only if $\mathrm{supp}(A_1 \otimes A_2) = \mathrm{supp}(A_1)\cap\mathrm{supp}(A_2) = \varnothing$. If $t > 0$, we can find a ring morphism $A_2 \to A_1$ by Proposition 4.2. Now Lemma 6.4 shows this is an isomorphism. □

Proposition 9.4. *Suppose $\mathcal{K}$ is nice. Let A be a separable ring in $\mathcal{K}$ with connected support* $\mathrm{supp}(A)$ *and constant degree. Then the splitting ring A^* of A is unique up to isomorphism and* $\mathrm{supp}(A) = \mathrm{supp}(A^*)$.

Proof. Let $d := \deg(A)$. Recall that by Proposition 6.9, the splitting rings of A are exactly the indecomposable ring factors of $A^{[d]}$. We now prove that any two indecomposable ring factors, say A_1 and A_2, of $A^{[d]}$ are isomorphic. Note that $\mathrm{supp}(A) = \mathrm{supp}(A^{[d]})$ is connected and $A^{[d]}$ has constant degree $d!$ by Remark 6.3, so that $\mathrm{supp}(A) = \mathrm{supp}(A_1) = \mathrm{supp}(A_2)$ by Proposition 8.5. Now, Lemma 9.3 shows A_1 and A_2 are isomorphic. □

Remark 9.5. In what follows, we consider a separable ring A in $\mathcal{K}$ and assume the spectrum $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ is connected, which implies that A is indecomposable. Moreover, if the tt-category $A\text{-}\mathrm{Mod}_{\mathcal{K}}$ is *rigid*, $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ is connected if and only if A is indecomposable, see [Balmer 2007, Theorem 2.11]. We note that many tt-categories are rigid, including all examples given in this paper.

Proposition 9.6. *Suppose $\mathcal{K}$ is nice. Let A be a separable ring in $\mathcal{K}$ and suppose* $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ *is connected. Let B be an A-algebra with* $\mathrm{supp}(A) = \mathrm{supp}(B)$. *If B*

is quasi-Galois in $\mathcal{K}$ with group Γ, then B splits A. In particular, the degree of A in $\mathcal{K}$ is constant.

Proof. If B is quasi-Galois in $\mathcal{K}$ for some group Γ, then all of its indecomposable ring factors are also quasi-Galois by Corollary 5.10. What is more, $\operatorname{supp}(B) = f_A(\operatorname{Spc}(A\text{-Mod}_{\mathcal{K}}))$ is connected, so the indecomposable ring factors of B have support equal to $\operatorname{supp}(B)$ by Proposition 8.5. It thus suffices to prove the proposition when B is indecomposable. Now, $F_A(B)$ is quasi-Galois by Lemma 5.7 and $\operatorname{supp}(F_A(B)) = f_A^{-1}(\operatorname{supp}(B)) = \operatorname{Spc}(A\text{-Mod}_{\mathcal{K}})$ is connected. By Corollary 2.4, $\bar{B}$ is an indecomposable ring factor of $F_A(B)$, and all ring factors of $F_A(B)$ have equal support by Proposition 8.5. In fact, Lemma 9.3 shows that $F_A(B) \cong \bar{B}^{\times t}$ for some $t \geq 1$. Forgetting the A-action, we get $A \otimes B \cong B^{\times t}$ in $\mathcal{K}$ and $F_B(A \otimes B) \cong F_B(B^{\times t}) \cong \mathbb{1}_B^{\times dt}$ in B-$\mathrm{Mod}_{\mathcal{K}}$, where $d := \deg(B)$. On the other hand, $F_B(A \otimes B) \cong F_B(A) \otimes_B \mathbb{1}_B^{\times d} \cong (F_B(A))^{\times d}$. It follows that $F_B(A) \cong \mathbb{1}_B^{\times t}$, with $t = \deg(A)$ by Lemma 7.11. □

Theorem 9.7 (Quasi-Galois closure). *Suppose $\mathcal{K}$ is nice. Let A be a separable ring of constant degree in $\mathcal{K}$ and suppose* $\operatorname{Spc}(A\text{-Mod}_{\mathcal{K}})$ *is connected. The splitting ring A^* is the quasi-Galois closure of A. That is, A^* is quasi-Galois in $\mathcal{K}$,* $\operatorname{supp}(A) = \operatorname{supp}(A^*)$ *and for any A-algebra B that is quasi-Galois in $\mathcal{K}$ with* $\operatorname{supp}(A) = \operatorname{supp}(B)$, *there exists a ring morphism $A^* \to B$.*

Proof. Corollary 6.10 and Proposition 9.4 show that A^* is quasi-Galois in $\mathcal{K}$ and $\operatorname{supp}(A) = \operatorname{supp}(A^*)$. Suppose there is an A-algebra B as above. By Proposition 9.6, B splits A, so there exists a ring morphism $A^{[\deg(A)]} \to B$. The result now follows because $A^{[\deg(A)]} \cong A^* \times \cdots \times A^*$ by Proposition 9.4. □

Remark 9.8. By Proposition 9.6, the assumption that A has constant degree is necessary for the existence of a quasi-Galois closure A^* of A with $\operatorname{supp}(A) = \operatorname{supp}(A^*)$.

10. Some modular representation theory

Let G be a finite group and $\Bbbk$ a field with characteristic $p > 0$ dividing $|G|$. We write $\Bbbk G$-mod for the category of finitely generated left $\Bbbk G$-modules. This category is nice, idempotent-complete and symmetric monoidal: the tensor is $\otimes_{\Bbbk}$ with diagonal G-action, and the unit is the trivial representation $\mathbb{1} = \Bbbk$.

We will also work in the bounded derived category $\mathrm{D}^b(\Bbbk G\text{-mod})$ and stable category $\Bbbk G$-stab, which are nice tt-categories. The spectrum $\operatorname{Spc}(\mathrm{D}^b(\Bbbk G\text{-mod}))$ of the derived category is homeomorphic to the homogeneous spectrum $\operatorname{Spec}^h(H^\bullet(G, \Bbbk))$ of the graded-commutative cohomology ring $H^\bullet(G, \Bbbk)$. Accordingly, the spectrum $\operatorname{Spc}(\Bbbk G\text{-stab})$ of the stable category is homeomorphic to the projective support variety $\mathcal{V}_G(\Bbbk) := \operatorname{Proj}(H^\bullet(G, \Bbbk))$; see [Benson et al. 1997].

Notation 10.1. Let $H \le G$ be a subgroup. The $\Bbbk G$-module $A_H = A_H^G := \Bbbk(G/H)$ is the free $\Bbbk$-module with basis G/H and left G-action given by $g \cdot [x] = [gx]$ for every $[x] \in G/H$. The $\Bbbk G$-linear map $\mu : A_H \otimes_{\Bbbk} A_H \to A_H$ is given by

$$\gamma \otimes \gamma' \mapsto \begin{cases} \gamma & \text{if } \gamma = \gamma', \\ 0 & \text{if } \gamma \neq \gamma', \end{cases} \quad \text{for all } \gamma, \gamma' \in G/H.$$

We define $\eta : \mathbb{1} \to A_H$ by sending $1 \in \Bbbk$ to $\sum_{\gamma \in G/H} \gamma \in \Bbbk(G/H)$.

We will write $\mathcal{K}(G)$ to denote any of $\Bbbk G$-mod, $\mathrm{D}^b(\Bbbk G\text{-mod})$ or $\Bbbk G$-stab and consider the object A_H in each of these categories.

Proposition 10.2 [Balmer 2015, Proposition 3.16 and Theorem 4.4]. *Let $H \le G$ be a subgroup. Then,*

(a) *The triple (A_H, μ, η) is a commutative separable ring object in $\mathcal{K}(G)$.*

(b) *There is an equivalence of categories*

$$\Psi_H^G : \mathcal{K}(H) \xrightarrow{\simeq} A_H\text{-}\mathrm{Mod}_{\mathcal{K}(G)}$$

sending $V \in \mathcal{K}(H)$ to $\Bbbk G \otimes_{\Bbbk H} V \in \mathcal{K}(G)$ with A_H-action

$$\varrho : \Bbbk(G/H) \otimes_{\Bbbk} (\Bbbk G \otimes_{\Bbbk H} V) \to \Bbbk G \otimes_{\Bbbk H} V$$

given for $\gamma \in G/H$, $g \in G$ and $v \in V$ by $\gamma \otimes g \otimes v \to \begin{cases} g \otimes v & \text{if } g \in \gamma, \\ 0 & \text{if } g \notin \gamma. \end{cases}$

(c) *The following diagram commutes up to isomorphism*:

$$\begin{tikzcd} & \mathcal{K}(G) \arrow[dl, "\mathrm{Res}_H^G"'] \arrow[dr, "F_{A_H}"] & \\ \mathcal{K}(H) \arrow[rr, "\Psi_H^G", "\simeq"'] & & A_H\text{-}\mathrm{Mod}_{\mathcal{K}(G)}. \end{tikzcd}$$

So, every subgroup $H \le G$ provides an indecomposable separable ring A_H in $\mathcal{K}(G)$, along which extension-of-scalars becomes restriction to the subgroup.

Proposition 10.3. *The ring A_H has degree $[G : H]$ in* $\Bbbk G$-mod *and* $\mathrm{D}^b(\Bbbk G\text{-mod})$.

Proof. Seeing how the fiber functor $\mathrm{Res}_{\{1\}}^G$ is conservative, we get

$$\deg_{\Bbbk G\text{-mod}}(A_H) = \deg_{\Bbbk\text{-mod}}(\mathrm{Res}_{\{1\}}^G(A_H)) = [G : H].$$

The degree of A_H in $\mathrm{D}^b(\Bbbk G\text{-mod})$ is computed in [Balmer 2014, Corollary 4.5]. $\square$

Lemma 10.4. *Let $\mathcal{K}(G)$ denote* $\mathrm{D}^b(\Bbbk G\text{-mod})$ *or* $\Bbbk G$-stab *and consider subgroups $K \le H \le G$. Then* $\mathrm{supp}(A_H) = \mathrm{supp}(A_K) \subset \mathrm{Spc}(\mathcal{K}(G))$ *if and only if every elementary abelian p-subgroup of H is conjugate in G to a subgroup of K.*

Proof. This follows from [Evens 1991, Theorem 9.1.3], seeing how $\operatorname{supp}(A_H) = (\operatorname{Res}^G_H)^*(\operatorname{Spc}(\mathcal{K}(H)))$ can be written as a union of disjoint pieces coming from conjugacy classes in G of elementary abelian p-subgroups of H. □

Notation 10.5. For any two subgroups $H, K \le G$, we write ${}_H[g]_K$ for the equivalence class of $g \in G$ in $H - G/K$, just $[g]$ if the context is clear. We will write $H^g := g^{-1}Hg$ for the conjugate subgroups of H.

Remark 10.6. Let $H, K \le G$ be subgroups and choose a complete set $T \subset G$ of representatives for $H\backslash G/K$. Consider the Mackey isomorphism of G-sets,

$$\coprod_{g\in T} G/(K\cap H^g) \xrightarrow{\cong} G/K \times G/H,$$

sending $[x] \in G/(K\cap H^g)$ to $([x]_K, [xg^{-1}]_H)$. The corresponding ring isomorphism

$$\tau : A_K \otimes A_H \xrightarrow{\cong} \prod_{g\in T} A_{K\cap H^g}$$

in $\mathcal{K}(G)$ [Balmer 2016b, Construction 4.1] is given for $g \in T$ and $x, y \in G$ by

$$\operatorname{pr}_g \tau([x]_K \otimes [y]_H) = \begin{cases} [xk]_{K\cap H^g} & \text{if } {}_H[g]_K = {}_H[y^{-1}x]_K, \\ 0 & \text{otherwise,} \end{cases}$$

with $k \in K$ such that $y^{-1}xkg^{-1} \in H$. This yields an A_K-algebra structure on $A_{K\cap H^t}$ for every $t \in T$, given by

$$A_K \xrightarrow{1\otimes\eta} A_K \otimes A_H \cong \prod_{g\in T} A_{K\cap H^g} \xrightarrow{\operatorname{pr}_t} A_{K\cap H^t},$$

which sends $[x]_K \in G/K$ to

$$\sum_{[k]\in K/K\cap H^t} [xk]_{K\cap H^t} \in A_{K\cap H^t}.$$

In the notation of Proposition 10.2(b), this just means $A_{K\cap H^t} = \Psi^G_K(A^K_{K\cap H^t})$ in $A_K\text{-Mod}_{\mathcal{K}(G)}$. In other words, τ defines an isomorphism

$$F_{A_K}(A_H) \cong \Psi^G_K\Bigl(\prod_{g\in T} A^K_{K\cap H^g}\Bigr)$$

of rings in $A_K\text{-Mod}_{\mathcal{K}(G)}$.

Lemma 10.7. *Let $H < G$. Suppose $x, g_1, g_2, \dots, g_n \in G$ and $1 \le i \le n$. Then*

$${}_H[x]_{H\cap H^{g_1}\cap\cdots\cap H^{g_n}} = {}_H[g_i]_{H\cap H^{g_1}\cap\cdots\cap H^{g_n}}$$

if and only if ${}_H[x] = {}_H[g_i]$.

Proof. It suffices to prove that for $x, y \in G$, we have ${}_H[x]_{H^y} = {}_H[y]_{H^y}$ if and only if ${}_H[x] = {}_H[y]$. This follows because for $[x] = [y]$ in $H - G/H^y$, there are $h, h' \in H$ with $x = hy(y^{-1}h'y) = hh'y$. $\square$

Notation 10.8. We fix a subgroup $H < G$ and a complete set $S \subset G$ of representatives for $H - G/H$. Likewise, if $g_1, g_2, \ldots, g_n \in G$ we will write $S_{g_1,g_2,\ldots,g_n} \subset G$ to denote some complete set of representatives for $H - G/H \cap H^{g_1} \cap \cdots \cap H^{g_n}$.

Recall that $\mathcal{K}(G)$ can denote $\Bbbk G$-mod, $\mathrm{D}^b(\Bbbk G\text{-mod})$ or $\Bbbk G$-stab.

Lemma 10.9. *Let $1 \leq n < [G : H]$. There is an isomorphism of rings*

$$A_H^{[n+1]} \cong \prod_{g_1,\ldots,g_n} A_{H\cap H^{g_1}\cap\cdots\cap H^{g_n}}$$

in $\mathcal{K}(G)$, where the product runs over all $g_1 \in S$ and $g_i \in S_{g_1,\ldots,g_{i-1}}$ for $2 \leq i \leq n$ with ${}_H[1], {}_H[g_1], \ldots, {}_H[g_n]$ distinct in $H\backslash G$.

Proof. By Remark 10.6, we know that

$$A_H \otimes A_H \cong \prod_{g\in S} A_{H\cap H^g} = A_H \times \prod_{\substack{g\in S\\ {}_H[g]\neq {}_H[1]}} A_{H\cap H^g},$$

so Proposition 2.1 shows

$$A_H^{[2]} \cong \prod_{\substack{g\in S\\ {}_H[g]\neq {}_H[1]}} A_{H\cap H^g} \quad \text{in } \mathcal{K}(G).$$

Now suppose

$$A_H^{[n]} \cong \prod_{g_1,\ldots,g_{n-1}} A_{H\cap H^{g_1}\cap\cdots\cap H^{g_{n-1}}}$$

for some $1 \leq n < [G : H]$, where the product runs over all $g_1 \in S$ and $g_i \in S_{g_1,\ldots,g_{i-1}}$ for $2 \leq i \leq n-1$ with ${}_H[1], {}_H[g_1], \ldots, {}_H[g_{n-1}]$ distinct in $H\backslash G$. Then

$$A_H^{[n]} \otimes A_H \cong \prod_{g_1,\ldots,g_{n-1}} A_{H\cap H^{g_1}\cap\cdots\cap H^{g_{n-1}}} \otimes A_H \cong \prod_{g_1,\ldots,g_{n-1}} \prod_{g_n\in S_{g_1,\ldots,g_{n-1}}} A_{H\cap H^{g_1}\cap\cdots\cap H^{g_n}},$$

again by Remark 10.6. We note that every $g_n \in S_{g_1,\ldots,g_{n-1}}$ with either ${}_H[g_n] = {}_H[1]$ or ${}_H[g_n] = {}_H[g_i]$ for $1 \leq i \leq n-1$ provides a copy of $A_H^{[n]}$. By Lemma 10.7, this happens exactly n times. Hence,

$$A_H^{[n]} \otimes A_H \cong \bigl(A_H^{[n]}\bigr)^{\times n} \times \prod_{g_1,\ldots,g_n} A_{H\cap H^{g_1}\cap\cdots\cap H^{g_n}},$$

where the product runs over all $g_1 \in S$ and $g_i \in S_{g_1,\ldots,g_{i-1}}$ for $2 \leq i \leq n$ with distinct ${}_H[1], {}_H[g_1], \ldots, {}_H[g_n]$ in $H\backslash G$. The lemma follows by Proposition 3.5(a). $\square$

Corollary 10.10. *Let $d := [G : H]$. There is an isomorphism of rings*

$$A_H^{[d]} \cong \bigl(A_{\mathrm{norm}_H^G}\bigr)^{\times k(G,H)}, \quad \textit{where } k(G,H) = \frac{d!}{[G : \mathrm{norm}_H^G]},$$

in $\Bbbk G$-mod *and* $\mathrm{D}^b(\Bbbk G\text{-mod})$. *Here*, $\mathrm{norm}_H^G := \bigcap_{g\in G} g^{-1}Hg$ *is the normal core of* H *in* G.

Proof. From the above lemma, we know

$$A_H^{[d]} \cong \prod_{g_1,\dots,g_{d-1}} A_{H\cap H^{g_1}\cap\cdots\cap H^{g_{d-1}}},$$

where the product runs over some $g_1,\dots g_{d-1}\in G$ with

$$\{{}_H[1], {}_H[g_1], \dots, {}_H[g_{d-1}]\} = H - G.$$

This shows $A_H^{[d]} \cong A_{\mathrm{norm}_H^G}^{\times t}$ for some $t \geq 1$. Now, $\deg(A_{\mathrm{norm}_H^G}) = [G : \mathrm{norm}_H^G]$ and $\deg(A_H^{[d]}) = d!$ by Remark 6.3, so $t = d!/[G : \mathrm{norm}_H^G]$ by Lemma 3.8(c). □

Corollary 10.11. *The ring* A_H *in* $\mathrm{D}^b(\Bbbk G\text{-mod})$ *has constant degree* $[G : H]$ *if and only if* norm_H^G *contains every elementary abelian* p*-subgroup of* H. *In that case, its quasi-Galois closure is* $A_{\mathrm{norm}_H^G}$. *Furthermore,* A_H *is quasi-Galois in* $\mathrm{D}^b(\Bbbk G\text{-mod})$ *if and only if* H *is normal in* G.

Proof. By Lemma 8.2, A_H has constant degree $[G : H]$ in $\mathrm{D}^b(\Bbbk G\text{-mod})$ if and only if $\mathrm{supp}(A^{[d]}) = \mathrm{supp}(A) \subset \mathrm{Spc}(\mathrm{D}^b(\Bbbk G\text{-mod}))$. Hence, the first statement follows immediately from Lemma 10.4 and Corollary 10.10. By Proposition 6.9, the splitting ring of A_H is $A_{\mathrm{norm}_H^G}$, so the second statement is Theorem 9.7. Since A_H is an indecomposable ring, it is quasi-Galois if and only if it is its own splitting ring. Thus A_H is quasi-Galois if and only if $A_{\mathrm{norm}_H^G} \cong A_H$, which yields $\mathrm{norm}_H^G = H$ by comparing degrees. □

Remark 10.12. Let $H \leq G$ be a subgroup. Recall that $A_H \cong 0$ in $\Bbbk G$-stab if and only if p does not divide $|H|$. On the other hand, $A_H \cong \Bbbk$ in $\Bbbk G$-stab if and only if H is *strongly* p*-embedded in* G, that is p divides $|H|$ and p does not divide $|H \cap H^g|$ if $g \in G - H$.

Proposition 10.13. *Let* $H \leq G$ *and consider the ring* A_H *in* $\Bbbk G$-stab. *Then,*

(a) *The degree of* A_H *is the greatest* $0 \leq n \leq [G : H]$ *such that there exist distinct* $[g_1],\dots,[g_n]$ *in* $H\backslash G$ *with* p *dividing* $|H^{g_1}\cap\cdots\cap H^{g_n}|$.

(b) *The ring* A_H *is quasi-Galois if and only if* p *divides* $|H|$ *and* p *does not divide* $|H\cap H^g\cap H^{gh}|$ *whenever* $g \in G - H$ *and* $h \in H - H^g$.

(c) *If* A_H *has degree* n, *the degree is constant if and only if there exist distinct* $[g_1],\dots,[g_n]$ *in* $H\backslash G$ *such that* $H^{g_1}\cap\cdots\cap H^{g_n}$ *contains a* G*-conjugate of every elementary abelian* p*-subgroup of* H.
In that case, A_H *has quasi-Galois closure given by* $A_{H^{g_1}\cap\cdots\cap H^{g_n}}$.

Proof. For (a), recall that $\deg(A_H)$ is the greatest n such that $A_H^{[n]} \neq 0$, thus such that there exist distinct ${}_H[1], {}_H[g_1], \dots, {}_H[g_{n-1}]$ with $|H\cap H^{g_1}\cap\cdots\cap H^{g_{n-1}}|$ divisible by p.

To show (b), recall that $F_{A_H}(A_H) \cong \Psi_H^G\big(\prod_{g\in S} A^H_{H\cap H^g}\big)$ by Remark 10.6. It follows that $F_{A_H}(A_H) \cong \mathbb{1}_{A_H}^{\times \deg(A_H)}$ in A_H-$\mathrm{Mod}_{\Bbbk G\text{-stab}}$ if and only if

$$\prod_{g\in S} A^H_{H\cap H^g} \cong \Bbbk^{\times \deg(A_H)}$$

in $\Bbbk H$-stab. So, A_H is quasi-Galois in $\Bbbk G$-stab if and only if $A_H \neq 0$ and for every $g \in G$, either $A^H_{H\cap H^g} = 0$ or $A^H_{H\cap H^g} \cong \Bbbk$ in $\Bbbk H$-stab. By Remark 10.12, this means either p does not divide $|H \cap H^g|$, or p divides $|H \cap H^g|$ but does not divide $|H \cap H^g \cap H^{gh}|$ when $h \in H - H^g$. Equivalently, p does not divide $|H \cap H^g \cap H^{gh}|$ whenever $g \in G - H$ and $h \in H - H^g$.

For (c), suppose A_H has constant degree n. By Proposition 9.4, any indecomposable ring factor of $A_H^{[n]}$ is isomorphic to the splitting ring A_H^*, so Lemma 10.9 shows that the quasi-Galois closure is given by $A_H^* \cong A_{H^{g_1}\cap\cdots\cap H^{g_n}}$ for any distinct ${}_H[g_1], \dots, {}_H[g_n]$ with $|H^{g_1} \cap \cdots \cap H^{g_n}|$ divisible by p. Then, $\mathrm{supp}(A_H) = \mathrm{supp}(A_H^*) = \mathrm{supp}(A_{H^{g_1}\cap\cdots\cap H^{g_n}})$ so $H^{g_1} \cap \cdots \cap H^{g_n}$ contains a G-conjugate of every elementary abelian p-subgroup of H. On the other hand, if there exist distinct $[g_1], \dots, [g_n]$ in $H\backslash G$ such that $H^{g_1} \cap \cdots \cap H^{g_n}$ contains a G-conjugate of every elementary abelian p-subgroup of H, then $\mathrm{supp}(A_H^{[n]}) = \mathrm{supp}(A_{H^{g_1}\cap\cdots\cap H^{g_n}}) = \mathrm{supp}(A_H)$, so the degree of A_H is constant. $\square$

Example 10.14. Let $p = 2$ and suppose $G = S_3$ is the symmetric group on 3 elements $\{1, 2, 3\}$. Consider the subgroup $H := \{(\,), (1\,2)\} \cong S_2$ of permutations fixing $\{3\}$. Its conjugate subgroups in G are the subgroups of permutations fixing $\{1\}$ and $\{2\}$ respectively, so $\mathrm{norm}_H^G = \{(\,)\}$. Now, A_H is a ring of degree 3 in $\mathrm{D}^b(\Bbbk G\text{-mod})$, and we immediately see that $\mathrm{supp}(A_H) = \mathrm{Spc}(\mathrm{D}^b(\Bbbk G\text{-mod}))$ because p does not divide $[G : H]$. Seeing how $\mathrm{supp}(A_H^{[3]}) \subset \mathrm{Spc}(\mathrm{D}^b(\Bbbk G\text{-mod}))$ contains only one point, the ring A_H does not have constant degree in $\mathrm{D}^b(\Bbbk G\text{-mod})$. On the other hand, the ring A_H considered in $\Bbbk G$-stab is quasi-Galois of degree 1, since H is strongly p-embedded in G.

Example 10.15. Let $p = 2$ and suppose $G = S_4$ is the symmetric group on 4 elements $\{1, 2, 3, 4\}$. If $H \cong S_3$ is the subgroup of permutations fixing $\{4\}$, the ring A_H in $\Bbbk G$-stab has constant degree 2. Indeed, the intersections $H \cap H^g$ with $g \in G - H$ each fix two elements of $\{1, 2, 3, 4\}$ pointwise, so p does not divide $[H : H \cap H^g]$; thus $\mathrm{supp}(A_H^{[2]}) = \mathrm{supp}(A_H)$. Furthermore, the intersections $H \cap H^{g_1} \cap H^{g_2}$ with $[1], [g_1], [g_2]$ distinct in $H\backslash G$ are trivial, so $A_H^{[3]} = 0$ in $\Bbbk G$-stab. The quasi-Galois closure of A_H in $\Bbbk G$-stab is A_{S_2}, with S_2 embedded in H.

Acknowledgement

I am very thankful to my advisor Paul Balmer for valuable ideas and instructive comments. I'd also like to thank the referee for detailed and very helpful suggestions.

References

[Auslander and Goldman 1960] M. Auslander and O. Goldman, "The Brauer group of a commutative ring", *Trans. Amer. Math. Soc.* **97** (1960), 367–409. MR

[Baker and Richter 2008] A. Baker and B. Richter, "Galois extensions of Lubin–Tate spectra", *Homology Homotopy Appl.* **10**:3 (2008), 27–43. MR Zbl

[Balmer 2005] P. Balmer, "The spectrum of prime ideals in tensor triangulated categories", *J. Reine Angew. Math.* **588** (2005), 149–168. MR Zbl

[Balmer 2007] P. Balmer, "Supports and filtrations in algebraic geometry and modular representation theory", *Amer. J. Math.* **129**:5 (2007), 1227–1250. MR Zbl

[Balmer 2011] P. Balmer, "Separability and triangulated categories", *Adv. Math.* **226**:5 (2011), 4352–4372. MR Zbl

[Balmer 2014] P. Balmer, "Splitting tower and degree of tt-rings", *Algebra Number Theory* **8**:3 (2014), 767–779. MR Zbl

[Balmer 2015] P. Balmer, "Stacks of group representations", *J. Eur. Math. Soc.* (*JEMS*) **17**:1 (2015), 189–228. MR Zbl

[Balmer 2016a] P. Balmer, "The derived category of an étale extension and the separable Neeman–Thomason theorem", *J. Inst. Math. Jussieu* **15**:3 (2016), 613–623. MR Zbl

[Balmer 2016b] P. Balmer, "Separable extensions in tensor-triangular geometry and generalized Quillen stratification", *Ann. Sci. Éc. Norm. Supér.* (*4*) **49**:4 (2016), 907–925. MR Zbl

[Balmer and Schlichting 2001] P. Balmer and M. Schlichting, "Idempotent completion of triangulated categories", *J. Algebra* **236**:2 (2001), 819–834. MR Zbl

[Balmer et al. 2015] P. Balmer, I. Dell'Ambrogio, and B. Sanders, "Restriction to finite-index subgroups as étale extensions in topology, KK-theory and geometry", *Algebr. Geom. Topol.* **15**:5 (2015), 3025–3047. MR

[Benson et al. 1997] D. J. Benson, J. F. Carlson, and J. Rickard, "Thick subcategories of the stable module category", *Fund. Math.* **153**:1 (1997), 59–80. MR Zbl

[Borceux and Janelidze 2001] F. Borceux and G. Janelidze, *Galois theories*, Cambridge Studies in Advanced Mathematics **72**, Cambridge University Press, 2001. MR Zbl

[Bourbaki 1981] N. Bourbaki, *Éléments de mathématique*, Lecture Notes in Mathematics **864**, Masson, Paris, 1981. MR Zbl

[Chase and Sweedler 1969] S. U. Chase and M. E. Sweedler, *Hopf algebras and Galois theory*, Lecture Notes in Mathematics **97**, Springer, Berlin, 1969. MR Zbl

[Chase et al. 1965] S. U. Chase, D. K. Harrison, and A. Rosenberg, "Galois theory and Galois cohomology of commutative rings", pp. 15–33 Mem. Amer. Math. Soc. No. **52**, 1965. MR Zbl

[DeMeyer and Ingraham 1971] F. DeMeyer and E. Ingraham, *Separable algebras over commutative rings*, Lecture Notes in Mathematics **181**, Springer, Berlin, 1971. MR Zbl

[Eilenberg and Moore 1965] S. Eilenberg and J. C. Moore, "Foundations of relative homological algebra", pp. 39 Mem. Amer. Math. Soc. No. **55**, 1965. MR Zbl

[Evens 1991] L. Evens, *The cohomology of groups*, Clarendon, New York, 1991. MR Zbl

[Hess 2009] K. Hess, "Homotopic Hopf–Galois extensions: foundations and examples", pp. 79–132 in *New topological contexts for Galois theory and algebraic geometry* (*BIRS 2008*) (Banff, 2008), edited by A. Baker and B. Richter, Geom. Topol. Monogr. **16**, Geom. Topol. Publ., Coventry, 2009. MR Zbl

[Kreimer 1967] H. F. Kreimer, "A Galois theory for noncommutative rings", *Trans. Amer. Math. Soc.* **127** (1967), 29–41. MR Zbl

[Mac Lane 1998] S. Mac Lane, *Categories for the working mathematician*, 2nd ed., Graduate Texts in Mathematics **5**, Springer, New York, 1998. MR Zbl

[Neeman 2015] A. Neeman, "Separable monoids in $\mathrm{D}_{\mathrm{qc}}(X)$", *J. Reine Angew. Math.* (2015).

[Pauwels 2015] B. Pauwels, *Quasi-Galois theory in triangulated categories*, Ph.D. thesis, University of California, Los Angeles, 2015, Available at https://search.proquest.com/docview/1732406955.

[Rognes 2008] J. Rognes, "Galois extensions of structured ring spectra: Stably dualizable groups", pp. viii+137 Mem. Amer. Math. Soc. **192**, 2008. MR Zbl

[SGA 1 1971] A. Grothendieck, *Revêtements étales et groupe fondamental* (Séminaire de Géométrie Algébrique du Bois Marie 1960–1961), vol. 1960/61, Lecture Notes in Math. **224**, Springer, Berlin, 1971. MR

Communicated by Dave Benson
Received 2016-09-01 Revised 2017-05-29 Accepted 2017-07-09

bregje.pauwels@anu.edu.au *Mathematical Sciences Institute, The Australian National University, Acton ACT 2601, Australia*

ALGEBRA AND NUMBER THEORY 11:8 (2017)
dx.doi.org/10.2140/ant.2017.11.1921

$\mathfrak{p}$-rigidity and Iwasawa μ-invariants

Ashay A. Burungale and Haruzo Hida

Let F be a totally real field with ring of integers O and p be an odd prime unramified in F. Let $\mathfrak{p}$ be a prime above p. We prove that a mod p Hilbert modular form associated to F is determined by its restriction to the partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$ ($\mathfrak{p}$-rigidity). Let K/F be an imaginary quadratic CM extension such that each prime of F above p splits in K and λ a Hecke character of K. Partly based on $\mathfrak{p}$-rigidity, we prove that the μ-invariant of the anticyclotomic Katz $\mathfrak{p}$-adic L-function of λ equals the μ-invariant of the full anticyclotomic Katz p-adic L-function of λ. An analogue holds for a class of Rankin–Selberg p-adic L-functions. When λ is self-dual with the root number -1, we prove that the μ-invariant of the cyclotomic derivatives of the Katz $\mathfrak{p}$-adic L-function of λ equals the μ-invariant of the cyclotomic derivatives of the Katz p-adic L-function of λ. Based on previous works of the authors and Hsieh, we consequently obtain a formula for the μ-invariant of these $\mathfrak{p}$-adic L-functions and derivatives. We also prove a $\mathfrak{p}$-version of a conjecture of Gillard, namely the vanishing of the μ-invariant of the Katz $\mathfrak{p}$-adic L-function of λ.

1. Introduction

Zeta values seem to suggest deep phenomena in Mathematics. They seem to mysteriously encode deep arithmetic information. They also seem to suggest surprising modular and Iwasawa-theoretic phenomena.

MSC2010: primary 11G18; secondary 19F27.
Keywords: Hilbert modular Shimura variety, Hecke stable subvariety, Iwasawa μ-invariant, Katz p-adic L-function.

Sometimes, they are a sum of evaluations of modular forms at CM points. Such an expression for critical Hecke L-values and conjectural nontriviality of the corresponding anticyclotomic p-adic L-function, suggested to the second named author a linear independence of mod p Hilbert modular forms. Based on Chai's theory of Hecke-stable subvarieties of a Shimura variety (see [Chai 1995; 2003; 2006]), this guess was proven in [Hida 2010]. Let $\mathfrak{p}$ be a prime above p (see the abstract). Recently, the expression and conjectural nontriviality of the corresponding anticyclotomic $\mathfrak{p}$-adic L-function suggested to us a rather surprising rigidity property of mod p Hilbert modular forms. Partly based on the rigidity, we obtain intriguing equalities of Iwasawa μ-invariants of seemingly independent p-adic L-functions.

Let F be a totally real field of degree d and O the integers ring. Let p be an odd prime unramified in F. Let $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ be the primes above p. Fix two embeddings $\iota_\infty\colon \bar{\boldsymbol{Q}} \to \boldsymbol{C}$ and $\iota_p\colon \bar{\boldsymbol{Q}} \to \boldsymbol{C}_p$. Let v_p be the p-adic valuation of $\boldsymbol{C}_p$ normalised such that $v_p(p) = 1$. Let $\mathbb{F}$ be an algebraic closure of $\mathbb{F}_p$.

Let $\mathrm{Sh}_{/\mathbb{F}}$ be the Kottwitz model of the prime-to-p Hilbert modular Shimura variety associated to F. We refer to Section 2B for the definition. Here we only mention that in the moduli interpretation for $\mathrm{Sh}_{/\boldsymbol{Z}_{(p)}}$, the full prime-to-$p$ level structure appears. Let $x \in \mathrm{Sh}$ be a closed ordinary point. From Serre–Tate deformation theory, a p^∞-level structure on x induces a canonical isomorphism

$$\mathrm{Spf}(\widehat{\mathcal{O}}_{\mathrm{Sh},x}) \simeq \prod_i \widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}_i}. \tag{1-1}$$

Let $\mathfrak{p} = \mathfrak{p}_i$, for some i. Let f be a mod p Hilbert modular form in the sense of Section 2D. In view of the irreducibility of the connected components of Sh, the form f is determined by its restriction to $\mathrm{Spf}(\widehat{\mathcal{O}}_{\mathrm{Sh},x})$. In fact, we have the following rigidity result.

Theorem 1.1 ($\mathfrak{p}$-rigidity). *Let $F/\boldsymbol{Q}$ be a totally real extension with integer ring O, p an odd prime unramified in F and $\mathfrak{p}|p$ a prime in F. Let f be a nonzero mod p Hilbert modular form over F as above.*

Then, f does not vanish identically on the partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$. In particular, a mod p Hilbert modular form is determined by its restriction to the partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$.

We now describe the results regarding the Iwasawa μ-invariants.

Let K be a totally imaginary quadratic extension of F. Let $h_?$ denote the class number of ?, for $? = K, F$. Let c denote the complex conjugation on $\boldsymbol{C}$ which induces the unique nontrivial element of $\mathrm{Gal}(K/F)$ via ι_∞. We assume the following hypothesis:

(*ord*) Every prime of F above p splits in K.

The condition (*ord*) guarantees the existence of a p-adic CM type Σ, i.e., Σ is a CM type of K such that the p-adic places induced by elements in Σ via ι_p are disjoint from the ones induced by Σc. We fix such a CM type. We also identify it with the set of infinite places of F. Let K_∞^- (resp. K_∞^+) be the anticyclotomic $\boldsymbol{Z}_p^d$-extension (resp. cyclotomic $\boldsymbol{Z}_p$-extension) of K and $K_{\mathfrak{p},\infty}^- \subset K_\infty^-$ be the $\mathfrak{p}$-anticyclotomic subextension, i.e., the maximal subextension unramified outside the primes above $\mathfrak{p}$ in K. Let $K_{\mathfrak{p},\infty} = K_{\mathfrak{p},\infty}^- K_\infty^+$. Let $\Gamma^\pm := \mathrm{Gal}(K_\infty^\pm/K)$, $\Gamma_\mathfrak{p}^- := \mathrm{Gal}(K_{\mathfrak{p},\infty}^-/K)$ and $\Gamma_\mathfrak{p} := \mathrm{Gal}(K_{\mathfrak{p},\infty}/K)$.

Let $\mathfrak{C}$ be a prime-to-p integral ideal of K. Let λ be an arithmetic Hecke character over K. Suppose that $\mathfrak{C}$ is the prime-to-p conductor of λ. Associated to this data, a natural $(d+1)$-variable Katz p-adic L-function

$$L_{\Sigma,\lambda} = L_{\Sigma,\lambda}(T_1, \dots, T_d, S) \in \bar{\boldsymbol{Z}}_p[\![\Gamma]\!]$$

is constructed in [Katz 1978; Hida and Tilouine 1993]. Here, the T_i are the anticyclotomic variables and S is the cyclotomic variable. The Katz p-adic L-function interpolates critical Hecke L-values $L(0, \lambda\chi)$ as χ varies over certain Hecke characters over K factoring through the Ray class group with conductor $\mathfrak{C}p^\infty$ (see [Hida and Tilouine 1993, Theorem II]). Let $L_{\Sigma,\lambda}^- \in \bar{\boldsymbol{Z}}_p[\![\Gamma^-]\!]$ (resp. $L_{\Sigma,\lambda,\mathfrak{p}}^- \in \bar{\boldsymbol{Z}}_p[\![\Gamma_\mathfrak{p}^-]\!]$) be the anticyclotomic (resp. $\mathfrak{p}$-anticyclotomic) projection obtained from the projection $\pi^- : \bar{\boldsymbol{Z}}_p[\![\Gamma]\!] \twoheadrightarrow \bar{\boldsymbol{Z}}_p[\![\Gamma^-]\!]$ (resp. $\pi_\mathfrak{p}^- : \bar{\boldsymbol{Z}}_p[\![\Gamma]\!] \twoheadrightarrow \bar{\boldsymbol{Z}}_p[\![\Gamma_\mathfrak{p}^-]\!]$). Let $L_{\Sigma,\lambda,\mathfrak{p}} \in \bar{\boldsymbol{Z}}_p[\![\Gamma_\mathfrak{p}]\!]$ be obtained from the projection $\pi_\mathfrak{p} : \bar{\boldsymbol{Z}}_p[\![\Gamma]\!] \twoheadrightarrow \bar{\boldsymbol{Z}}_p[\![\Gamma_\mathfrak{p}]\!]$. We call $L_{\Sigma,\lambda,\mathfrak{p}}$ the Katz $\mathfrak{p}$-adic L-function to emphasise the consideration of the $\mathfrak{p}$-component. This is a slightly nontraditional terminology as the construction is still under the same embedding ι_p.

For the notion of Iwasawa μ-invariants, we refer to [Hida 2010, §1]. Here we only mention that the μ-invariant measures nontriviality modulo p. We now state our results regarding the Iwasawa μ-invariants of p-adic L-functions and derivatives arising in the context of the Iwasawa theory of an arithmetic Hecke character over a CM field as above.

The μ-invariant of $L_{\Sigma,\lambda,\mathfrak{p}}^-$ is given by the following theorem.

Theorem 1.2. *Let $F/\boldsymbol{Q}$ be a totally real extension, p an odd prime unramified in F and $\mathfrak{p}|p$ a prime in F. Let K/F be a p-ordinary CM quadratic extension. Let λ be an arithmetic Hecke character over K. Let $L_{\Sigma,\lambda}^-$ (resp. $L_{\Sigma,\lambda,\mathfrak{p}}^-$) be the corresponding anticyclotomic Katz p-adic (resp. anticyclotomic Katz $\mathfrak{p}$-adic) L-function as above. Then, we have*

$$\mu(L_{\Sigma,\lambda}^-) = \mu(L_{\Sigma,\lambda,\mathfrak{p}}^-).$$

In most of the cases, $\mu(L_{\Sigma,\lambda}^-)$ has been explicitly determined (see [Hida 2010; Hsieh 2014a]). Thus, we obtain a formula for $\mu(L_{\Sigma,\lambda,\mathfrak{p}}^-)$.

We show a result analogous to Theorem 1.2 for a class of Rankin–Selberg anticyclotomic $\mathfrak{p}$-adic L-functions.

When λ is self-dual with the root number -1, all the Hecke L-values appearing in the interpolation property of $L^-_{\Sigma,\lambda}$ vanish. Accordingly, $L^-_{\Sigma,\lambda}$ and $L^-_{\Sigma,\lambda,\mathfrak{p}}$ identically vanish. The anticyclotomic arithmetic information contained in $L^-_{\Sigma,\lambda}$ and $L^-_{\Sigma,\lambda,\mathfrak{p}}$ may seem to have disappeared. However, we can look at the cyclotomic derivatives

$$L'_{\Sigma,\lambda} = \Big(\frac{\partial}{\partial S} L_{\Sigma,\lambda}(T_1, \dots, T_d, S)\Big)\Big|_{S=0} \tag{1-2}$$

and $L'_{\Sigma,\lambda,\mathfrak{p}}$ (defined analogously).

The μ-invariant of $L'_{\Sigma,\lambda,\mathfrak{p}}$ is given by the following theorem.

Theorem 1.3. *Let $F/\mathbb{Q}$ be a totally real extension, p an odd prime unramified in F and $\mathfrak{p}|p$ a prime in F. Let K/F be a p-ordinary CM quadratic extension. Let λ be an arithmetic self-dual Hecke character over K with root number -1. Let $L'_{\Sigma,\lambda}$ (resp. $L'_{\Sigma,\lambda,\mathfrak{p}}$) be the corresponding cyclotomic derivative of the Katz p-adic (resp. Katz $\mathfrak{p}$-adic) L-function as above.*

Suppose that $p \nmid h^-_K$, where h^-_K is the relative class number given by $h^-_K = h_K/h_F$. Then, we have

$$\mu(L'_{\Sigma,\lambda}) = \mu(L'_{\Sigma,\lambda,\mathfrak{p}}).$$

In most of the cases, $\mu(L'_{\Sigma,\lambda})$ has been explicitly determined (see [Burungale 2015]). Thus, we obtain a formula for $\mu(L'_{\Sigma,\lambda,\mathfrak{p}})$.

Finally, we consider the $\mathfrak{p}$-adic L-function. We have the following $\mathfrak{p}$-version of a conjecture of Gillard [1991, Conjecture (i)] regarding the vanishing of the μ-invariant of Katz p-adic L-function.

Theorem 1.4. *Let $F/\mathbb{Q}$ be a totally real extension, p an odd prime unramified in F and $\mathfrak{p}|p$ a prime in F. Let K/F be a p-ordinary CM quadratic extension. Let λ be an arithmetic Hecke character over K. Let $L_{\Sigma,\lambda,\mathfrak{p}}$ be the corresponding cyclotomic Katz $\mathfrak{p}$-adic L-function as above. Then, we have*

$$\mu(L_{\Sigma,\lambda,\mathfrak{p}}) = 0.$$

We now describe the strategy of the proof of Theorem 1.1. Some of the notation used here is not followed in the rest of the article.

We begin with general remarks on rigidity. Here is a geometric reason for this rigidity, which we refer to as L-rigidity. Let $\widehat{S}$ be a formal torus over a field which is the residue field of a mixed characteristic ring. Suppose that we have a rational structure coming from an algebraic subscheme S over the mixed characteristic ring whose formal completion at a closed point $x \in S$ gives $\widehat{S}$. Here S is not necessarily an algebraic torus. We suppose that there exists a positive dimensional transcendental linear subvariety L of $\widehat{S}$ with strictly smaller dimension, i.e., a nontrivial formal subtorus which does not equal the formal completion along x of an algebraic subscheme of S. Transcendence of L implies the "algebraic" Zariski closure of L in S is the entire S. This may not happen for the formal Zariski closure. Let $A = \mathcal{O}_{S,x}$ and $L = \mathrm{Spf}(B)$. Thus, the transcendence of L is equivalent to the injectivity of the natural morphism $A \to B$ given by $\phi \mapsto \phi|_L$ for $\phi \in A$ and hence the L-rigidity, i.e., ϕ is determined by its restriction to the formal subtorus L. The rigidity remains true for $\phi \circ a$ for any automorphism a of $\widehat{S}$. It turns out that often the rigidity also remains true for $\sum_i \phi_i \circ a_i$ for a well chosen set of automorphisms a_i of $\widehat{S}$ and $\phi_i \in \Gamma(S, \mathcal{O}_S)$, i.e.,

$$\sum_i \phi_i \circ a_i = 0 \Rightarrow \phi_i = 0$$

for all i. As the notion of well chosen may vary from context to context, we only mention that $\{a_i\}$'s typically satisfy a transcendental property which we later specify in this context. We study the case for the Serre–Tate deformation space $\widehat{S}$ with rational structure induced from the Hilbert modular Shimura variety. The formal subtorus we study is the partial $\mathfrak{p}$-deformation subspace of $\widehat{S}$. We consider certain $\{a_i\}_i$ such that the differences $\{a_i a_j^{-1}\}_{i \neq j}$ are "transcendental" automorphisms of the Serre–Tate deformation space (see Section 4A).

Let $G = \mathrm{Res}_{O/\boldsymbol{Z}}(\mathrm{GL}_2)$. The group $G(\boldsymbol{Z}_{(p)})$ acts on the prime-to-p Hilbert modular Shimura variety $\mathrm{Sh}_{/\boldsymbol{Z}_{(p)}}$. We refer to Section 3B for the action. Here we only mention that in terms of the moduli interpretation, the action corresponds to the one on the level structure. Let V be an irreducible component of Sh containing x. Let $H_x(\boldsymbol{Z}_{(p)})$ be the stabiliser of x in $G(\boldsymbol{Z}_{(p)})$. It acts on $\mathrm{Spec}(\mathcal{O}_{V,x})$ and thus on the Serre–Tate deformation space $\mathrm{Spf}(\widehat{\mathcal{O}}_{V,x})$. In view of the description of the action on the Serre–Tate coordinates, we observe that the formal subtorus

$$\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}} \subset \mathrm{Spf}(\widehat{\mathcal{O}}_{V,x})$$

is stable under the action of $H_x(\boldsymbol{Z}_{(p)})$. Chai and the second named author have proven that a positive dimensional closed irreducible subvariety of V containing x and stable under $H_x(\boldsymbol{Z}_{(p)})$ equals V itself. In this sense, the formal subtorus $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$ is transcendental in the Shimura variety. Recall that the Igusa tower is étale over V. As a mod p Hilbert modular form is an algebraic function on the Igusa tower, we prove $\mathfrak{p}$-rigidity (Theorem 1.1) based on the transcendence.

We now describe the strategy of the proof of Theorem 1.2. Some of the notation used here is not followed in the rest of the article.

Let us first recall the second named author's strategy [2010] to determine $\mu(L^{-}_{\Sigma,\lambda})$. Let $O_p = O \otimes \mathbb{Z}_p$. Let

$$G_{\Sigma,\lambda} \in \bar{\boldsymbol{Z}}_p[\![T_1, \ldots, T_d]\!]$$

be the power series expansion of the measure $L^-_{\Sigma,\lambda}$ regarded as a p-adic measure on O_p with support in $1+pO_p$, given by

$$\begin{aligned} G_{\Sigma,\lambda} &= \int_{1+pO_p} t^y \, dL^-_{\Sigma,\lambda}(y) \\ &= \sum_{(k_1,\dots,k_d)\in(\mathbf{Z}_{\geq 0})^d} \left(\int_{1+pO_p} \binom{y}{k_1,\dots,k_d} dL^-_{\Sigma,\lambda}(y) \right) T_1^{k_1} \cdots T_d^{k_d}. \end{aligned} \tag{1-3}$$

The starting point is the observation that there are classical Hilbert modular Eisenstein series $(f_{\lambda,i})_i$ such that

$$G_{\Sigma,\lambda} = \sum_i a_i \circ (f_{\lambda,i}(t)), \tag{1-4}$$

where $f_{\lambda,i}(t)$ is the t-expansion of $f_{\lambda,i}$ around a well chosen CM point y with the CM type (K,Σ) on the Hilbert modular Shimura variety Sh and a_i is an automorphism of the Serre–Tate deformation space $\mathrm{Spf}(\widehat{\mathcal{O}}_{\mathrm{Sh},y})$, i.e., $a_i \in H_y(\mathbf{Z}_p)$ (see Section 3B). Based on Chai's study of Hecke-stable subvarieties of a Shimura variety, the second named author has proven the linear independence of $(a_i \circ f_{\lambda,i})_i$ modulo p. It follows that

$$\mu(L^-_{\Sigma,\lambda}) = \min_i \mu(f_{\lambda,i}(t)).$$

Let $G_{\Sigma,\lambda,\mathfrak{p}} \in \bar{\mathbf{Z}}_p[\![\Gamma^-_{\mathfrak{p}}]\!]$ be the analogous power series expansion of the measure $L^-_{\Sigma,\lambda,\mathfrak{p}}$. Based on the action of p-adic differential operators on the t-expansion of a p-adic Hilbert modular form around an ordinary point in terms of the partial Serre–Tate coordinates, we show that

$$G_{\Sigma,\lambda,\mathfrak{p}} = \sum_i a_{i,\mathfrak{p}} \circ (f_{\lambda,i}(t_{\mathfrak{p}})), \tag{1-5}$$

where $a_{i,\mathfrak{p}}$ is the projection of a_i to $H_y(\mathbf{Z}_p)_{\mathfrak{p}}$ (see Section 3B) and $f_{\lambda,i}(t_{\mathfrak{p}})$ is the $\mathfrak{p}$-adic Serre–Tate expansion of f around y (see Section 6A). Based on $\mathfrak{p}$-rigidity and Chai's theory of Hecke-stable subvarieties of a Shimura variety, we prove $\mathfrak{p}$-independence, i.e., the linear independence of $\big(a_{i,\mathfrak{p}} \circ (f_{\lambda,i}(t_{\mathfrak{p}}))\big)_i$ modulo p. It follows that

$$\mu(L^-_{\Sigma,\lambda,\mathfrak{p}}) = \min_i \mu(f_{\lambda,i}(t_{\mathfrak{p}})).$$

In view of the $\mathfrak{p}$-rigidity, we have

$$\mu(f_{\lambda,i}(t)) = \mu(f_{\lambda,i}(t_{\mathfrak{p}})).$$

This concludes the proof of Theorem 1.2.

The strategy of the proof of Theorem 1.3 is similar to the above strategy. It involves p-adic modular forms $f'_{\lambda,i}$ arising from the p-adic derivative of $f_{\lambda N^s,i}$

with N being the norm Hecke character. In this sense, the proof involves $\mathfrak{p}$-rigidity for nonclassical p-adic modular forms. Finally, Theorem 1.4 is proven based on Theorem 1.2 and the proof of results in [Hida 2011; Burungale and Hsieh 2013]. As in [Hida 2010], we would like to emphasise that Chai's theory plays an underlying role in all of the above results.

In this article, we give an elementary construction of p-adic differential operators on the space of p-adic Hilbert modular forms avoiding the use of Gauss–Manin connection in Katz's construction. This is based on the ideas of the second named author in the early nineties. Based on this construction, we determine the action of p-adic differential operators on the t-expansion of a p-adic Hilbert modular form around an ordinary point in terms of the partial Serre–Tate coordinates.

Along with the p-adic Gross–Zagier formula, Theorem 1.2 and Theorem 1.3 have application towards generic nonvanishing of p-adic heights on CM abelian varieties (see [Burungale and Disegni $\geq$ 2017]). This provides evidence for Schneider's conjecture on the nonvanishing of p-adic heights in the CM case. In view of [Burungale 2017], the underlying ideas also have application towards generic nonvanishing of the p-adic Abel–Jacobi image of generalised Heegner cycles in the non-CM case. We refer to [Burungale 2016b] for a survey.

Let $\mathfrak{p}_i$ and $\mathfrak{p}_j$ be primes above p as above. Theorem 1.2 implies an intriguing equality

$$\mu(L^{-}_{\Sigma,\lambda,\mathfrak{p}_i}) = \mu(L^{-}_{\Sigma,\lambda,\mathfrak{p}_j})$$

of Iwasawa μ-invariants. This is rather surprising as theses μ-invariants could be nonzero and one in general does not expect any relation between the p-adic L-functions $L_{\Sigma,\lambda,\mathfrak{p}_i}$. These p-adic L-functions correspond to independent variables whose number may vary with i. As far as we know, Theorem 1.2 is a first phenomena possibly suggesting a relation. It would be interesting to see whether an analogue holds for Iwasawa λ-invariants. One can perhaps first collect experimental data. In the case of self-duality and the root number being -1, the equality of the μ-invariants persists even after taking the cyclotomic derivative. It seems tempting to suppose that a deeper phenomena mediates the relation.

In view of the anticyclotomic main conjectures, Theorem 1.2 would imply an equality of the corresponding algebraic μ-invariants. Note that the underlying Selmer groups correspond to rather different local conditions. In many cases, the anticyclotomic main conjecture has been proven (see [Hida 2006; 2009b]). It would be interesting to prove the equality of the algebraic μ-invariants directly. This sort of equality does not seem to be conjectured in the literature.

In [Burungale $\geq$ 2017], the first named author proves the analogue of $\mathfrak{p}$-rigidity and $\mathfrak{p}$-independence for quaternionic modular forms over totally real fields. These results concern a quaternion algebra which is not totally definite. In the near future, he hopes to consider an analogue of Theorem 1.2 for a class of quaternionic Rankin–Selberg p-adic L-functions. In [Harris et al. 2006], a construction of p-adic L-functions for unitary Shimura varieties is announced (see [Eischen et al. 2016]). In such a case, the number of variables of the p-adic L-function is typically less than the dimension of the Shimura variety. The variables of the p-adic L-function may correspond to an analogue of the $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$-variables. In the future, we hope to consider this question starting with the case of $U(n, 1)$ Shimura varieties.

Formulating a $\mathfrak{p}$-rigidity type statement for a PEL Shimura variety may have an independent interest. In characteristic zero, we hope to explore rigidity based on the approach in [Burungale and Hida 2016].

The article is organised as follows. In Section 2, we recall basic facts about the Hilbert modular Shimura variety Sh. In Section 3, we prove Theorem 1.1. In Sections 3A–3B, we firstly recall some facts about Serre–Tate deformation theory of an ordinary closed point in Sh. In Section 3C, we prove the theorem. In Section 4, we prove $\mathfrak{p}$-rigidity, i.e., the linear independence of mod p Hilbert modular forms restricted to the partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$. In Section 5A, we give an elementary construction of p-adic differential operators on the space of p-adic Hilbert modular forms. In Section 5B, we use it to compute the action of the p-adic differential operators on the t-expansion of a p-adic Hilbert modular form around an ordinary point in terms of the partial Serre–Tate coordinates. In Section 6, we consider Iwasawa μ-invariants as in Theorems 1.2–1.4. In Section 6A, we determine the μ-invariant of certain anticyclotomic $\mathfrak{p}$-adic L-functions (see Theorem 1.2). In Section 6B, we determine the μ-invariant of the cyclotomic derivative $L'_{\Sigma,\lambda,\mathfrak{p}}$ of the Katz $\mathfrak{p}$-adic L-function, when the branch character λ is self-dual with the root number -1 (see Theorem 1.3). In Section 6C, we prove a $\mathfrak{p}$-version of a conjecture Gillard regarding the vanishing of the μ-invariant of the Katz p-adic L-function (see Theorem 1.4).

Notation. We use the following notation unless otherwise stated.

For a number field L, let $\mathbb{A}_L$ be the adele ring and $\mathbb{A}_L^f$ the finite adeles of L. Let h_L denote the ideal class number. Let G_L be the absolute Galois group of L and G_L^{ab} the maximal abelian quotient. Let $\mathrm{rec}_L : \mathbb{A}_L^\times \to G_L^{\mathrm{ab}}$ be the geometrically normalised reciprocity law.

2. Hilbert modular Shimura variety

In this section, we recall basic facts about Hilbert modular Shimura varieties. We follow [Hida 2004].

2A. ***Setup.*** In this subsection, we recall a basic setup regarding Hilbert modular Shimura varieties.

Let $G = \mathrm{Res}_{F/\boldsymbol{Q}}\, GL_2$ and $h_0 : \mathrm{Res}_{\boldsymbol{C}/\boldsymbol{R}}\, \mathbb{G}_m \to G_{/\boldsymbol{R}}$ be the morphism of real group schemes arising from

$$a + bi \mapsto \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

where $a + bi \in \boldsymbol{C}^\times$. Let X be the set of $G(\boldsymbol{R})$-conjugacy classes of h_0. We have a canonical isomorphism $X \simeq (\boldsymbol{C} - \boldsymbol{R})^I$, where I is the set of real places of F. The pair (G, X) satisfies Deligne's axioms for a Shimura variety. It gives rise to a tower $(\mathrm{Sh}_K = \mathrm{Sh}_K(G, X))_K$ of quasiprojective smooth varieties over $\boldsymbol{Q}$ indexed by open compact subgroups K of $G(\boldsymbol{A}^f)$. The pro-algebraic variety $\mathrm{Sh}_{/\boldsymbol{Q}}$ is the projective limit of these varieties. The complex points of these varieties are given as follows

$$\mathrm{Sh}_K(\boldsymbol{C}) = G(\boldsymbol{Q})\backslash X \times G(\boldsymbol{A}^f)/K, \quad \mathrm{Sh}(\boldsymbol{C}) = G(\boldsymbol{Q})\backslash X \times G(\boldsymbol{A}^f)/\overline{Z(\boldsymbol{Q})}. \tag{2-1}$$

Here, $\overline{Z(\boldsymbol{Q})}$ is the closure of the center $Z(\boldsymbol{Q})$ in $G(\boldsymbol{A}^f)$ under the adélic topology. From (2-1) and the general theory of Shimura varieties, it follows that $\mathrm{Sh}_{/\boldsymbol{Q}}$ is endowed with an action of $G(\boldsymbol{A}^f)$ (see [Hida 2004, §4.2]). This gives rise to the Hecke action.

2B. ***p-integral model.*** In this subsection, we briefly recall a canonical p-integral smooth model $\mathrm{Sh}^{(p)}_{/\boldsymbol{Z}_{(p)}}$ of the Shimura variety $\mathrm{Sh}/G(\boldsymbol{Z}_p)_{/\boldsymbol{Q}}$.

Hilbert modular Shimura variety $\mathrm{Sh}_{/\boldsymbol{Q}}$ represents a functor $\mathcal{F}$ classifying abelian schemes having multiplication by O along with additional structure, where O is the ring of integers of F (see [Hida 2004, §4.2; Shimura 1963]). As in the introduction, let p be an odd prime unramified in F. Under this hypothesis, a p-integral interpretation $\mathcal{F}^{(p)}$ of $\mathcal{F}$ leads to a p-integral smooth model of $\mathrm{Sh}/G(\boldsymbol{Z}_p)_{/\boldsymbol{Q}}$.

The functor $\mathcal{F}^{(p)}$ is given by

$$\begin{aligned} \mathcal{F}^{(p)} : \mathrm{SCH}_{/\boldsymbol{Z}_{(p)}} &\to \mathrm{SETS}, \\ S &\mapsto \{(A, \iota, \bar{\lambda}, \eta^{(p)})_{/S}\}/\sim . \end{aligned} \tag{2-2}$$

Here,

(PM1) A is abelian scheme over S of dimension of d.

(PM2) $\iota : O \hookrightarrow \mathrm{End}_S\, A$ is an algebra embedding.

(PM3) $\bar{\lambda}$ is the polarisation class of a homogeneous polarisation λ up to scalar multiplication by $\iota(O^\times_{(p),+})$, where $O_{(p),+} := \{a \in O_{(p)} \mid \sigma(a) > 0,\ \forall \sigma \in I\}$. Also, the Rosati involution of $\mathrm{End}_S\, A$ takes $\iota(l)$ to $\iota(l^*)$, for $l \in O$.

(PM4) Let $\mathcal{T}^{(p)}(A)$ be the prime-to-p Tate module $\varprojlim_{p\nmid N} A[N]$. Then $\eta^{(p)}$ is a prime-to-p level structure given by an O-linear isomorphism

$$\eta^{(p)} : O^2 \otimes_{\boldsymbol{Z}} \widehat{\boldsymbol{Z}}^{(p)} \simeq \mathcal{T}^{(p)}(A),$$

where $\widehat{\boldsymbol{Z}}^{(p)} = \prod_{l \neq p} \boldsymbol{Z}_l$.

(PM5) Let $\mathrm{Lie}_S(A)$ be the relative Lie algebra of A. There exists an $O\otimes_{\boldsymbol{Z}}\mathcal{O}_S$-module isomorphism

$$\mathrm{Lie}_S(A) \simeq O \otimes_{\boldsymbol{Z}} \mathcal{O}_S,$$

locally under the Zariski topology of S.

The notation $\sim$ denotes up to a prime-to-p isogeny.

Theorem 2.1 (Kottwitz). *Let the notation and assumptions be as above. Then, the functor $\mathcal{F}^{(p)}$ is represented by a pro-algebraic scheme $\mathrm{Sh}^{(p)}(G,X)_{/\boldsymbol{Z}_{(p)}}$. Moreover, there exists an isomorphism given by*

$$\mathrm{Sh}^{(p)} \times \boldsymbol{Q} \simeq \mathrm{Sh}\,/G(\boldsymbol{Z}_p)_{/\boldsymbol{Q}}.$$

(see [Hida 2004, §4.2.1]).

The pro-algebraic scheme $\mathrm{Sh}^{(p)}(G,X)_{/\boldsymbol{Z}_{(p)}}$ is usually referred as the Kottwitz model. In what follows, we let $\mathrm{Sh}^{(p)}_{/\boldsymbol{Z}_{(p)}}$ denote $\mathrm{Sh}^{(p)}(G,X)_{/\boldsymbol{Z}_{(p)}}$ for simplicity of notation.

2C. ***Igusa tower.*** In this subsection, we briefly recall the notion of p-ordinary Igusa tower over the p-integral model (see Section 2B).

Let $\overline{\boldsymbol{Q}}$ be an algebraic closure of $\boldsymbol{Q}$ and $\overline{\boldsymbol{Q}}_p$ be an algebraic closure of $\boldsymbol{Q}_p$. We fix a complex embedding $\iota_\infty : \overline{\boldsymbol{Q}} \hookrightarrow \boldsymbol{C}$ and a p-adic embedding $\iota_p : \overline{\boldsymbol{Q}} \hookrightarrow \overline{\boldsymbol{Q}}_p$.

Let $\mathcal{W}$ be the strict Henselisation inside $\overline{\boldsymbol{Q}}$ of the local ring of $\boldsymbol{Z}_{(p)}$ corresponding to ι_p. Let $\mathbb{F}$ be the residue field of $\mathcal{W}$. Note that $\mathbb{F}$ is an algebraic closure of $\mathbb{F}_p$.

Let $\mathrm{Sh}^{(p)}_{/\mathcal{W}} = \mathrm{Sh}^{(p)} \times_{\boldsymbol{Z}_{(p)}} \mathcal{W}$ and $\mathrm{Sh}^{(p)}_{/\mathbb{F}} = \mathrm{Sh}^{(p)}_{/\mathcal{W}} \times_{\mathcal{W}} \mathbb{F}$.

From now, let Sh denote $\mathrm{Sh}^{(p)}_{/\mathbb{F}}$. Let $\mathcal{A}$ be the universal abelian scheme over Sh.

Let $\mathrm{Sh}^{\mathrm{ord}}$ be the subscheme of Sh on which the Hasse-invariant does not vanish. It is an open dense subscheme. Over $\mathrm{Sh}^{\mathrm{ord}}$, the connected part $\mathcal{A}[p^m]^\circ$ of $\mathcal{A}[p^m]$ is étale-locally isomorphic to $\mu_{p^m} \otimes_{\boldsymbol{Z}_p} O^*$ as an O_p-module, where $O^* = O^{-1}\mathfrak{d}_F^{-1}$, $\mathfrak{d}_F$ is the different of $F/\boldsymbol{Q}$ and $O_p = O \otimes \boldsymbol{Z}_p$.

We now define the Igusa tower. For $m \in \mathbb{N}$, the m-th layer of the Igusa tower over $\mathrm{Sh}^{\mathrm{ord}}$ is defined by

$$\mathrm{Ig}_m = \underline{\mathrm{Isom}}_{O_p}(\mu_{p^m} \otimes_{\boldsymbol{Z}_p} O^*, \mathcal{A}[p^m]^\circ). \tag{2-3}$$

Note that the projection $\pi_m : \mathrm{Ig}_m \to \mathrm{Sh}^{\mathrm{ord}}$ is finite and étale. The full Igusa tower over $\mathrm{Sh}^{\mathrm{ord}}$ is defined by

$$\mathrm{Ig} = \mathrm{Ig}_\infty = \varprojlim \mathrm{Ig}_m = \underline{\mathrm{Isom}}_{O_p}(\mu_{p^\infty} \otimes_{\boldsymbol{Z}_p} O^*, \mathcal{A}[p^\infty]^\circ). \tag{2-4}$$

(Ét) Note that the projection $\pi : \mathrm{Ig} \to \mathrm{Sh}^{\mathrm{ord}}$ is étale.

Let x be a closed ordinary point in Sh. We have the following description of the level p^∞-structure on the corresponding p-divisible group $A_x[p^\infty]$.

(PL) Let η_p° be a level p^∞-structure on $A_x[p^\infty]^\circ$. For the primes $\mathfrak{p}$ in O dividing p, it is a collection of level $\mathfrak{p}^\infty$ structures $\eta_{\mathfrak{p}}^\circ$, given by isomorphisms

$$\eta_{\mathfrak{p}}^\circ : O_{\mathfrak{p}}^* \simeq A_x[\mathfrak{p}^\infty]^\circ,$$

where $O_{\mathfrak{p}}^* = O^* \otimes O_{\mathfrak{p}}$. The Cartier duality and the polarisation $\bar{\lambda}_x$ induces an isomorphism

$$\eta_{\mathfrak{p}}^{\text{ét}} : O_{\mathfrak{p}} \simeq A_x[\mathfrak{p}^\infty]^{\text{ét}}.$$

Thus, we get a level p^∞-structure $\eta_p^{\text{ét}}$ on $A_x[p^\infty]^{\text{ét}}$ from η_p°.

Let V be an irreducible component of Sh and V^{ord} be $V \cap \text{Sh}^{\text{ord}}$. Let I be the inverse image of V^{ord} under π. In [Hida 2004, Chapter 8; 2009a], it has been shown that

(Ir) I is an irreducible component of Ig.

2D. ***Mod p modular forms.*** In this subsection, we briefly recall the notion of mod p modular forms on an irreducible component of the Hilbert modular Shimura variety (see Section 2B).

Let V and I be as in Section 2C. Let B be an $\mathbb{F}$-algebra. The space of mod p modular forms on V over B is defined by

$$M(V, B) = H^0(I_{/B}, \mathcal{O}_{I_{/B}}), \tag{2-5}$$

where $I_{/B} := I \times_{\mathbb{F}} B$. In view of Sections 2B–2C, we have the following geometric interpretation of mod p modular forms.

A mod p modular form is a function f of isomorphism classes of $\tilde{x} = (x, \eta_p^\circ)_{/B'}$, where B' is a B-algebra, $x = (A, \iota, \bar{\lambda}, \eta^{(p)})_{/B'} \in \mathcal{F}^{(p)}(B')$ and

$$\eta_p^\circ : \mu_{p^\infty} \otimes_{\mathbf{Z}_p} O^* \simeq A[p^\infty]^\circ$$

is an O_p-linear isomorphism, such that the following conditions are satisfied.

(G1) $f(\tilde{x}) \in B'$.

(G2) If $\tilde{x} \simeq \tilde{x}'$, then $f(\tilde{x}) = f(\tilde{x}')$, where $\tilde{x} \simeq \tilde{x}'$ means $x \simeq x'$ and the corresponding isomorphism between A and A' induces an isomorphism between η_p° and $\eta_p'^\circ$.

(G3) $f(\tilde{x} \times_{B'} B'') = h(f(\tilde{x}))$ for any B-algebra homomorphism $h : B' \to B''$.

We also have the key notion of q-expansion and q-expansion principle for mod p modular forms (see [Hida 2004, Theorem 4.21]).

2E. ***p-adic modular forms.*** In this subsection, we briefly recall the notion of p-adic modular forms on an irreducible component of the Hilbert modular Shimura variety (see Section 2B).

Let W denote the Witt ring $W(\mathbb{F})$. The construction of the Igusa tower in Section 2C is well defined for the base W. Let $I_{/W}$ be the irreducible component of the Igusa tower $\mathrm{Ig}_{/W}$ over an irreducible component $V_{/W}$ of the Shimura variety $\mathrm{Sh}_{/W}$. Let C be a p-adically complete local W-algebra with maximal ideal $\mathfrak{m}_C$. The space of p-adic modular forms on V over C is defined by

$$M(V, C) = H^0(I_{/C}, \mathcal{O}_{I_{/C}}), \tag{2-6}$$

where $I_{/C} := I_{/W} \times_{/W} C$.

By definition, a p-adic modular form over C modulo $\mathfrak{m}_C$ is a mod p modular form.

We have an analogous moduli interpretation as in Section 2D and also the q-expansion principle, for p-adic modular forms (see [Hida 2010, §4.1]).

3. $\mathfrak{p}$-rigidity

In this section, we prove the rigidity of mod p modular forms (see Theorem 1.1). In Sections 3A–3B, we firstly recall basic facts about Serre–Tate deformation theory of an ordinary closed point on the Hilbert modular variety (see Section 2B). In Section 3C, we prove the rigidity.

3A. ***Serre–Tate deformation theory.*** In this subsection, we briefly recall Serre–Tate deformation theory of an ordinary closed point on the Hilbert modular variety (see Section 2B). We follow [Hida 2004, §8.2; 2010, §2; 2013b, §1].

Let the notation and assumptions be as in Section 2. Let x be a closed point in $\mathrm{Sh}^{\mathrm{ord}}$ carrying $(A_x, \iota_x, \bar{\lambda}_x, \eta_x^{(p)})_{/\mathbb{F}}$. Let V be the irreducible component of Sh containing x.

Let CL_W be the category of complete local W-algebras with residue field $\mathbb{F}$. Let $\mathcal{D}_{/W}$ be the fiber category over CL_W of deformations of $x_{/\mathbb{F}}$ defined as follows. Let $R \in \mathrm{CL}_W$. The objects of $\mathcal{D}_{/W}$ over R consist of $x'^* = (x', \iota_{x'})$, where $x' \in \mathcal{F}^{(p)}(R)$ and

$$\iota_{x'} : x' \times_R \mathbb{F} \simeq x.$$

Let x'^* and x''^* be in $\mathcal{D}_{/W}$ over R. By definition, a morphism ϕ between x'^* and x''^* is a morphism (still denoted by) ϕ between x' and x'' satisfying [Hida 2004, (7.3)] and the following condition.

(M) Let ϕ_0 be the special fiber of ϕ. The automorphism $\iota_{x''} \circ \phi_0 \circ \iota_{x'}^{-1}$ of x equals the identity.

Let $\widehat{\mathcal{F}}_x$ be the deformation functor given by

$$\begin{aligned} \widehat{\mathcal{F}}_x : \mathrm{CL}_{/W} &\to \mathrm{SETS}, \\ R &\mapsto \{x'^{*}_{/R} \in \mathcal{D}\}/\simeq. \end{aligned} \tag{3-1}$$

The notation $\simeq$ denotes up to an isomorphism.

Recall, $R \in \mathrm{CL}_W$. As R is a projective limit of local W-algebras with nilpotent maximal ideal, we can (and do) suppose that R is a local Artinian W-algebra with nilpotent maximal ideal m_R. Let $x'^{*}_{/R} \in \mathcal{D}$ and A denote $A_{x'}$. By Drinfeld's theorem (see [Hida 2004, §8.2.1]), $A[p^\infty]^\circ(R)$ is killed by p^{n_0} for sufficiently large n_0. Let $y \in A(\mathbb{F})$ and $\tilde{y} \in A(R)$ such that $\tilde{y}_0 = y$, where $\tilde{y}_0$ denotes the special fiber of $\tilde{y}$ (as $A_{/R}$ is smooth, such a lift always exists). By definition, $\tilde{y}$ is determined modulo $\ker(A(R) \to A(\mathbb{F})) = A[p^\infty]^\circ(R)$. Thus, for $n \geq n_0$, "p^n"$y_0 := p^n\tilde{y}$ is well defined. From now, we suppose that $n \geq n_0$. If $y \in A[\mathfrak{p}^n](\mathbb{F})$, then "$p^n$"$y \in A[\mathfrak{p}^\infty]^\circ(R)$. Strictly speaking, we apply the idempotent $e_\mathfrak{p}$ corresponding to $\mathfrak{p}$ so that $e_\mathfrak{p}$"p^n"$y \in A[\mathfrak{p}^\infty]^\circ(R)$. We let "$p^n$"$y$ denote $e_\mathfrak{p}$"p^n"y for simplicity of notation.

Thus, we have a homomorphism

$$\text{“}p^n\text{”} : A[\mathfrak{p}^n](\mathbb{F}) \to A[\mathfrak{p}^\infty]^\circ(R). \tag{3-2}$$

We also have the commutative diagram

$$\begin{array}{ccccc} A[\mathfrak{p}^{n+1}]^{\text{ét}}(R) & \xrightarrow{\simeq} & A[\mathfrak{p}^{n+1}](\mathbb{F}) & \xrightarrow{\text{“}p^{n+1}\text{”}} & A[\mathfrak{p}^\infty]^\circ(R) \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle =} \\ A[\mathfrak{p}^{n}]^{\text{ét}}(R) & \xrightarrow{\simeq} & A[\mathfrak{p}^{n}](\mathbb{F}) & \xrightarrow{\text{“}p^{n}\text{”}} & A[\mathfrak{p}^\infty]^\circ(R). \end{array}$$

Passing to the projective limit, this gives rise to a homomorphism

$$\text{“}p^\infty\text{”} : A[\mathfrak{p}^\infty](\mathbb{F}) \to A[\mathfrak{p}^\infty]^\circ(R). \tag{3-3}$$

(CC) For $\varprojlim y_n \in \varprojlim A[\mathfrak{p}^n](\mathbb{F})$, let

$$y = \varprojlim y_n \in A[\mathfrak{p}^\infty](\mathbb{F}) \simeq A_x[\mathfrak{p}^\infty]^{\text{ét}}.$$

The isomorphism is induced by $\iota_{x'}$.

Let $q_{n,\mathfrak{p}}(y_n) =$ "p^n"y_n and $q_\mathfrak{p}(y) = \lim q_{n,\mathfrak{p}}(y_n)$. By definition,

$$q_\mathfrak{p}(y) \in A[\mathfrak{p}^\infty]^\circ(R) \simeq \mathrm{Hom}(A_x^\vee[\mathfrak{p}^\infty]^{\text{ét}}, \widehat{\mathbb{G}}_m(R)),$$

where $A_x^\vee$ is the dual of the abelian variety A_x. Let $q_{A,\mathfrak{p}}$ be the pairing given by

$$\begin{aligned} q_{A,\mathfrak{p}} : A_x[\mathfrak{p}^\infty]^{\text{ét}} \times A_x^\vee[\mathfrak{p}^\infty]^{\text{ét}} &\to \widehat{\mathbb{G}}_m(R), \\ q_{A,\mathfrak{p}}(y, z) &= q_\mathfrak{p}(y)(z). \end{aligned} \tag{3-4}$$

We have the following fundamental result.

Theorem 3.1 (Serre–Tate). (1) *There exists a canonical isomorphism*

$$\widehat{\mathcal{F}}_x(R) \simeq \prod_{\mathfrak{p}} \mathrm{Hom}_{\mathbf{Z}_p}\big(A_x[\mathfrak{p}^\infty]^{\text{ét}} \times A_x^\vee[\mathfrak{p}^\infty]^{\text{ét}}, \widehat{\mathbb{G}}_m(R)\big) \tag{3-5}$$

given by $x'^* \mapsto q_A := \prod_{\mathfrak{p}} q_{A_{x'},\mathfrak{p}}$.

(2) *The deformation functor* $\widehat{\mathcal{F}}_x$ *is represented by the formal scheme* $\widehat{S}_{/W} :=$ Spf$(\widehat{\mathcal{O}}_{V,x})$. *A level* p^∞*-structure as in* (PL), *gives rise to a canonical isomorphism of the deformation space* $\widehat{S}_{/W}$ *with the formal torus* $\prod_{\mathfrak{p}} \widehat{\mathbb{G}}_m \otimes_{\mathbf{Z}_p} O_{\mathfrak{p}}$ (see [Hida 2013b, Proposition 1.2]).

Let x_{ST} be the universal deformation of the closed ordinary point x.

We now recall key underlying notions.

Definition 3.2. Let x be a closed ordinary point on the Hilbert modular variety as above. Recall that a level p^∞-structure on x (see (PL)) gives rise to a canonical isomorphism of the deformation space $\widehat{S}_{/W}$ with the formal torus $\prod_{\mathfrak{p}} \widehat{\mathbb{G}}_m \otimes_{\mathbf{Z}_p} O_{\mathfrak{p}}$ (see part (2) of Theorem 3.1). Under this identification, let

$$t = (t_{\mathfrak{p}})_{\mathfrak{p}}$$

be the coordinates of the deformation space $\widehat{S}_{/W}$, where $t_{\mathfrak{p}}$ is the coordinate of $\widehat{\mathbb{G}}_m \otimes_{\mathbf{Z}_p} O_{\mathfrak{p}}$. We call $t = (t_{\mathfrak{p}})_{\mathfrak{p}}$ the Serre–Tate coordinates of the deformation space $\widehat{S}_{/W}$.

We have $\widehat{S} = \mathrm{Spf}(\widehat{W[O]})$, where $S = \mathbb{G}_m \otimes O^*$, $W[O] = W[X(S)]$ and $\widehat{W[O]}$ is the completion at the augmentation ideal. Here, $X(S)$ is the character group of S. Note that $W[O]$ is the ring consisting of formal finite sums $\sum_{\xi \in O} a(\xi) t^\xi$, where $a(\xi) \in W$ and t is the coordinate of $\mathbb{G}_m$. Here, t^ξ is the character given by

$$t^\xi(t \otimes u) = t^{\mathrm{Tr}_{F/\mathbf{Q}}(\xi u)}$$

for $u \in O^*$.

Definition 3.3. Let f be a mod p modular form over $\mathbb{F}$ (see Section 2D). A level p^∞-structure η_p° of x gives rise to a canonical level p^∞-structure $\eta_{p,\mathrm{ST}}^\circ$ of the universal deformation x_{ST}. We denote by

$$f((x_{\mathrm{ST}}, \eta_{p,\mathrm{ST}}^\circ)) \in \widehat{\mathbb{F}[O]}$$

the t-expansion of f around x.

We have the following t-expansion principle.

(t-expansion principle) The above t-expansion of a mod p modular form f around a closed ordinary point determines f uniquely (see (Ir)).

We have an analogous t-expansion principle for p-adic modular forms (see [Hida 2010, §4.1]).

3B. ***Reciprocity law for the deformation space.*** In this subsection, we recall the action of the local algebraic stabiliser of a closed ordinary point x on the Serre–Tate coordinates of the deformation space of x. This can be considered as an infinitesimal analogue of Shimura's reciprocity law.

Let $g \in G(\mathbf{Z}_{(p)})$ act on Sh through the right multiplication on the prime-to-p level structure, i.e.,

$$(A, \iota, \bar{\lambda}, \eta^{(p)})_{/S} \mapsto (A, \iota, \bar{\lambda}, \eta^{(p)} \circ g)_{/S}$$

(see Section 2B).

Recall that x is a closed ordinary point in Sh with a p^∞-level structure η_p^{ord}. Let (K_x, Σ_x) be the CM-type of x. We suppose that $\iota : O \hookrightarrow \mathrm{End}(A)$ extends to $\iota_x : \mathcal{O} \hookrightarrow \mathrm{End}(A)$, where $\mathcal{O}$ is the ring of integers of K_x. Let $H_x(\mathbf{Z}_{(p)})$ be the stabiliser of x in $G(\mathbf{Z}_{(p)})$. Note that

$$H_x(\mathbf{Z}_{(p)}) = (\mathrm{Res}_{\mathcal{O}_{(p)}/\mathbf{Z}_{(p)}} \mathbb{G}_m)(\mathbf{Z}_{(p)}) = \mathcal{O}_{(p)}^\times, \tag{3-6}$$

where $\mathcal{O}_{(p)} = \mathcal{O} \otimes \mathbf{Z}_{(p)}$ (see [Hida 2010, §3.2; Shimura 1998]). We call $H_x(\mathbf{Z}_{(p)})$ the local algebraic stabiliser of x.

Let c_x be the complex conjugation of K_x.

As x is ordinary, Σ_x is a p-ordinary CM type. When considered as a p-adic CM type, we denote it by $\Sigma_{x,p}$. Let $\boldsymbol{p} = \prod_{v \in \Sigma_{x,p}} \mathfrak{p}_v$, for the primes $\mathfrak{p}_v$ associated to the valuation $v \in \Sigma_{x,p}$. Note that $\mathcal{O}_{\boldsymbol{p}} = O_p = \prod_{\mathfrak{p}} O_{\mathfrak{p}}$ and $\mathcal{O}_p = \mathcal{O}_{\boldsymbol{p}} \times \mathcal{O}_{\boldsymbol{p}^{c_x}}$. Let $H_x(\mathbf{Z}_p)_{\mathfrak{p}}$ be the $\mathfrak{p}$-component $O_{\mathfrak{p}}^\times$ of $H_x(\mathbf{Z}_p) = \mathcal{O}_p^\times$. We have a natural inclusion $\mathcal{O}_{(p)} \subset \mathcal{O}_p$. Thus, we regard $H_x(\mathbf{Z}_{(p)}) \subset H_x(\mathbf{Z}_p)$. Let $H_x(\mathbf{Z}_{(p)})_{\mathfrak{p}}$ be the projection of $H_x(\mathbf{Z}_{(p)})$ to $H_x(\mathbf{Z}_p)_{\mathfrak{p}}$. Note that we have an isomorphism $H_x(\mathbf{Z}_{(p)}) \simeq H_x(\mathbf{Z}_{(p)})_{\mathfrak{p}}$. Let $\alpha \in H_x(\mathbf{Z}_{(p)})$. Let $\alpha_{\mathfrak{p}}$ be the projection of α to the $\mathfrak{p}$-component $O_{\mathfrak{p}}^\times$ of $O_{\boldsymbol{p}}^\times$. As $H_x(\mathbf{Z}_{(p)})$ stabilises x, it follows that α acts on $\mathrm{Spec}(\mathcal{O}_{V,x})$ and thus on $\mathrm{Spf}(\widehat{\mathcal{O}}_{V,x})$. In particular, it acts on the Serre–Tate coordinates (see Definition 3.2). The action is given by the following lemma.

Lemma 3.4 [Hida 2010, Lemma 3.3]. *The endomorphism α acts on the Serre–Tate coordinates $t = (t_{\mathfrak{p}})_{\mathfrak{p}}$ by*

$$t \mapsto t^{\alpha^{1-c_x}}$$

for $t^{\alpha^{1-c_x}} = (t_{\mathfrak{p}}^{\alpha_{\mathfrak{p}}^{1-c_x}})_{\mathfrak{p}}$.

We have the following immediate corollary.

Corollary 3.5. *The partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}} \subset \mathrm{Spf}(\widehat{\mathcal{O}}_{V,x})$ is stable under the action of the local algebraic stabiliser $H_x(\mathbf{Z}_{(p)})$.*

This simple corollary is crucial for $\mathfrak{p}$-rigidity. It may lead to rigidity-style phenomena with different flavour.

3C. $\mathfrak{p}$-*rigidity.* In this subsection, we prove the rigidity of mod p modular forms (see Theorem 1.1).

Recall that x is a closed ordinary point in V with p^∞-level structure η_p° and $V = \varprojlim V_K$, where V_K is the projection of V to $\mathrm{Sh}_K = \mathrm{Sh}/K$ for K small and maximal at p. This gives rise to a closed point $\tilde{x} = (x, \eta_p^\circ)$ in the Igusa tower I over x. In view of (Ét), there exists a canonical isomorphism

$$\widehat{\mathcal{O}}_{V,x} \simeq \widehat{\mathcal{O}}_{I,\tilde{x}}.$$

Thus, a mod p modular form can be considered as a function on $\mathrm{Spec}(\mathcal{O}_{V,x})$.

Let f be a nonzero mod p modular form. We suppose that f is a nonunit in $\mathcal{O}_{V,x}$. Let $\mathfrak{b} \subset \mathcal{O}_{V,x}$ be the zero ideal of f, i.e., $\mathfrak{b} = (f) \cap \mathcal{O}_{V,x}$. As f is an algebraic function on I, it follows that $V(\mathfrak{b})$ is nonempty not only in $\mathrm{Spf}(\widehat{\mathcal{O}}_{V,x})$ but also in $\mathrm{Spec}(\mathcal{O}_{V,x})$. In particular, we have $\mathfrak{b} \neq 0$. Let X be the Zariski closure of $\mathrm{Spec}(\mathcal{O}_{V,x}/\mathfrak{b})$ in V. Note that $X \subset V$ is a closed irreducible pro-subscheme containing x and $X = \varprojlim X_K$, where X_K is the projection of X to V_K.

We start with a preparatory lemma.

Lemma 3.6. *Let Y_i be a family of closed subschemes of an irreducible noetherian scheme Y such that there exists a closed point $y \in Y_i$, for all i, i.e., $y \in \bigcap_i Y_i$. Suppose that the intersection of the* $\mathrm{Spf}(\widehat{\mathcal{O}}_{Y_i,y})$ *(viewed inside* $\mathrm{Spf}(\widehat{\mathcal{O}}_{Y,y})$*) is positive dimensional. Then, the intersection of the Y_i is positive dimensional.*

Proof. It suffices to consider the affine case.

Let A be a noetherian integral domain and $R_i = A/a_i$, for ideals a_i. Take nonunits $t_1, \ldots, t_r$ in A and let $(t) = (t_1, \ldots, t_r)$. Suppose that

$$V(t) = \mathrm{Spec}(R/(t)) \subset V_i = \mathrm{Spec}(R/a_i), \quad \text{for all } i.$$

This implies $a_i \subset (t)$. In particular, $\sum_i a_i \subset (t)$.

If A is local noetherian with maximal ideal m, by faithful flatness of $\widehat{A} = \varprojlim A/m^n$ over A,

$$\widehat{\sum_i a_i} = \sum_i \widehat{a_i}.$$

So, the above argument can be applied to $(t) \in \widehat{m}$, $\widehat{A}$ and $\widehat{(A/a_i)}$.

This finishes the proof as

$$\dim\Bigl(\bigcap_i \mathrm{Spec}(A/a_i)\Bigr) = \dim\Bigl(\mathrm{Spec}\Bigl(A/\sum_i a_i\Bigr)\Bigr) = \dim\Bigl(\mathrm{Spec}\Bigl(\widehat{A}/\sum_i \widehat{a_i}\Bigr)\Bigr)$$
$$= \dim\Bigl(\bigcap_i \mathrm{Spec}(\widehat{A}/\widehat{a_i})\Bigr) \geq \dim(\widehat{A}/(t)). \qquad \square$$

We are now ready to prove the rigidity.

Theorem 3.7 ($\mathfrak{p}$-rigidity). *Let the notation and assumptions be as above. The partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$ is not a formal subscheme of* $\mathrm{Spf}(\widehat{\mathcal{O}}_{X,x})$ (*viewed inside* $\mathrm{Spf}(\widehat{\mathcal{O}}_{V,x})$).

Proof. Suppose that it is a formal subscheme.

We now consider $Z = \bigcap_{\alpha \in H_x(\mathbb{Z}_{(p)})} \alpha(X)$ (viewed inside V). As above, recall that $Z = \varprojlim Z_K$.

Note that $\mathrm{Spf}(\widehat{\mathcal{O}}_{\alpha(X),x}) = \alpha(\mathrm{Spf}(\widehat{\mathcal{O}}_{X,x}))$ (viewed inside $\mathrm{Spf}(\widehat{\mathcal{O}}_{X,x})$). It follows that if $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}} \subset \mathrm{Spf}(\widehat{\mathcal{O}}_{X,x})$, then $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}} \subset \mathrm{Spf}(\widehat{\mathcal{O}}_{\alpha(X),x})$ (see Corollary 3.5). In particular, we have $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}} \subset \mathrm{Spf}(\widehat{\mathcal{O}}_{\alpha(X)_K,x})$. In the last equality, $\alpha(X)_K$ is the projection of $\alpha(X)$ to V_K. By Lemma 3.6, Z_K is positive dimensional. As the projection $\pi_K : V \to V_K$ is étale, we conclude that Z itself is positive dimensional.

Thus, Z is a closed irreducible pro-subscheme of V containing x and stable under $H_x(\mathbb{Z}_{(p)})$. We conclude that $Z = V$ (see [Hida 2010, Proposition 3.8]). It follows that $\mathfrak{b} = 0$. This is a contradiction for f being nonzero as noted in the paragraph before Lemma 3.6. □

In particular, Theorem 1.1 holds. The following consequence is immediate.

Corollary 3.8. *Let g be a p-adic modular form. Then*,

$$\mu(g) = \mu(g|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}}).$$

Proof. Let g be nonzero and defined over a p-adically complete local W-algebra C.

If g is a unit in $\mathcal{O}_{V,x}$, the corollary follows instantly. We thus suppose that g is a nonunit in $\mathcal{O}_{V,x}$.

In view of the t-expansion principle, we have

$$\mu(g) = \mu(g|_{\prod_{\mathfrak{p}} \widehat{\mathbb{G}}_m \otimes_{\mathbb{Z}_p} O_{\mathfrak{p}}}).$$

Let $c \in C$ such that $\mu(g/c) = 0$. By definition, g/c modulo the maximal ideal $\mathfrak{m}_C$ is a nonzero mod p modular form. In view of Theorem 3.7, it thus follows that

$$\mu(g/c|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}}) = 0. \qquad \square$$

4. $\mathfrak{p}$-independence

In this section, we consider $\mathfrak{p}$-independence, i.e., a linear independence of mod p modular forms restricted to the partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$ (see Section 3A). In Section 4A, we first state the formulation and in Section 4B, we prove the independence.

4A. ***Formulation.*** In this subsection, we give a formulation of the linear independence of mod p modular forms restricted to the partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$ (Section 3A).

Let the notation and assumptions be as in Section 3. Recall that x is a closed ordinary point in V with p^∞-level structure η_p°. This gives rise to a closed point $\tilde{x} = (x, \eta_p^\circ)$ in the Igusa tower I over x and a canonical isomorphism

$$\mathrm{Spf}(\widehat{\mathcal{O}}_{V,x}) \simeq \prod_{\mathfrak{p}} \widehat{\mathbb{G}}_m \otimes_{\mathbf{Z}_p} O_{\mathfrak{p}} \tag{4-1}$$

(see Theorem 3.1).

Recall that we have a canonical isomorphism $\widehat{\mathcal{O}}_{V,x} \simeq \widehat{\mathcal{O}}_{I,\tilde{x}}$ (see (Ét)). Thus, the p-completed stabiliser $H_x(\mathbf{Z}_p)$ acts on $\widehat{\mathcal{O}}_{I,\tilde{x}}$.

Let f be a mod p modular form and $a \in H_x(\mathbf{Z}_p)$. Note that

$$(a(f))|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}} = a_{\mathfrak{p}}(f|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}})$$

(see Corollary 3.5). Here $a_{\mathfrak{p}}$ is as in Section 3B.

To provide context for the following independence, note that there exists $a \in H_x(\mathbf{Z}_p)$ such that $a_{\mathfrak{p}_i} \in H_x(\mathbf{Z}_{(p)})_{\mathfrak{p}_i}$ and $a_{\mathfrak{p}_j} \notin H_x(\mathbf{Z}_{(p)})_{\mathfrak{p}_j}$, for $j \neq i$. Indeed,

$$H_x(\mathbf{Z}_p) = O_p^\times = \prod_{\mathfrak{p}} O_{\mathfrak{p}}^\times$$

and we may choose a to be nonidentity precisely at the $\mathfrak{p}_i$-th component.

Let n be a positive integer. For $1 \le i \le n$, let $a_i \in H_x(\mathbf{Z}_p)$ such that $(a_i a_j^{-1})_{\mathfrak{p}} \notin H_x(\mathbf{Z}_{(p)})_{\mathfrak{p}}$ for all $i \neq j$ (see Section 3B). Let $f_1, \dots, f_n$ be n nonconstant mod p modular forms on V (see Section 2D).

Our formulation of the linear independence is the following.

Theorem 4.1 ($\mathfrak{p}$-independence). *Let the notation and assumptions be as above, and suppose that $(a_i a_j^{-1})_{\mathfrak{p}} \notin H_x(\mathbf{Z}_{(p)})_{\mathfrak{p}}$ for all $i \neq j$. Then, the $(a_{i,\mathfrak{p}} \circ (f_i|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}}))_i$ are linearly independent in the partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$.*

Note that $a_{i,\mathfrak{p}} \circ (f_i|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}})$ is not necessarily the restriction of a mod p modular form to $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$.

The above independence is $\mathfrak{p}$-analogue of the linear independence in [Hida 2010, §3.5]. As expounded in Section 4B, the approach in [Hida 2010, §3] is fundamental to the $\mathfrak{p}$-independence.

4B. $\mathfrak{p}$-*independence.* In this subsection, we prove the linear independence of mod p modular forms restricted to the partial Serre–Tate deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$ (see Section 4A).

The approach is based on $\mathfrak{p}$-rigidity and Chai's theory of Hecke-stable subvarieties of a mod p Shimura variety adopted for local algebraic stabilisers in [Hida 2010, §3]. For a detailed treatment of the latter, we refer to [Chai 1995; 2003; 2006; Hida 2010].

Let n be a positive integer. In this subsection, any tensor product is taken n-times.

We consider an $\mathbb{F}$-algebra homomorphism

$$\phi_I : \mathcal{O}_{I,\tilde{x}} \otimes_{\mathbb{F}} \cdots \otimes_{\mathbb{F}} \mathcal{O}_{I,\tilde{x}} \to \widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}} \tag{4-2}$$

given by

$$f_1 \otimes \cdots \otimes f_n \mapsto \prod_{i=1}^{i=n} a_{i,\mathfrak{p}} \circ (f_i|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}}). \tag{4-3}$$

As we are interested in the linear independence of $(a_{i,\mathfrak{p}}(f_i|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}}))_i$, we consider $\mathfrak{b}_I := \ker(\phi_I)$.

Similarly, we consider an $\mathbb{F}$-algebra homomorphism

$$\phi = \phi_V : \mathcal{O}_{V,x} \otimes_{\mathbb{F}} \cdots \otimes_{\mathbb{F}} \mathcal{O}_{V,x} \to \widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}} \tag{4-4}$$

given by

$$h_1 \otimes \cdots \otimes h_n \mapsto \prod_{i=1}^{i=n} a_{i,\mathfrak{p}} \circ (h_i|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}}). \tag{4-5}$$

In view of Theorem 3.7, it follows that ϕ_I and ϕ_V are both nontrivial.

(EQ) We note that ϕ is equivariant with the $H_x(\boldsymbol{Z}_{(p)})$-action.

Let $\mathfrak{b}_I = \ker(\phi_V)$ and $\mathfrak{b} = \ker(\phi_V)$.

Lemma 4.2. $\mathfrak{b}_I = 0$ *if and only if* $\mathfrak{b} = 0$.

Proof. In view of (Ét), we have an étale morphism

$$\pi^m : \mathcal{O}_{V,x} \otimes_{\mathbb{F}} \cdots \otimes_{\mathbb{F}} \mathcal{O}_{V,x} \to \mathcal{O}_{I,\tilde{x}} \otimes_{\mathbb{F}} \cdots \otimes_{\mathbb{F}} \mathcal{O}_{I,\tilde{x}}.$$

Note that $\mathfrak{b}_I$ is the unique prime ideal of $\mathcal{O}_{I,\tilde{x}} \otimes_{\mathbb{F}} \cdots \otimes_{\mathbb{F}} \mathcal{O}_{I,\tilde{x}}$ over $\mathfrak{b}$. $\square$

As ϕ is equivariant with the $H_x(\boldsymbol{Z}_{(p)})$-action (see (EQ)), it follows that $\mathfrak{b}$ is a prime ideal of $\mathcal{O}_{V,x} \otimes_{\mathbb{F}} \cdots \otimes_{\mathbb{F}} \mathcal{O}_{V,x}$ stable under the diagonal action of $H_x(\boldsymbol{Z}_{(p)})$. Let Y be the Zariski closure of $\mathrm{Spec}(\mathcal{O}_{V,x} \otimes_{\mathbb{F}} \cdots \otimes_{\mathbb{F}} \mathcal{O}_{V,x}/\mathfrak{b})$ in V^n. Thus, Y is a closed irreducible subscheme of V^n containing x^n stable under the diagonal action of $H_x(\boldsymbol{Z}_{(p)})$. We also have an analogue of the commutative diagram [Hida 2010, (3.22)] with $\widehat{\mathcal{O}}_S$ replaced by $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}$. For $n \geq 2$, the subscheme Y thus satisfies the hypothesis in [Hida 2010, Corollary 3.19].

Theorem 4.3. *The subscheme Y equals V^n.*

Proof. When $n = 1$, this is nothing but [Hida 2010, Proposition 3.8].

We thus suppose that $n \geq 2$. From [Hida 2010, Corollary 3.19], we have two possibilities, namely $Y = V^{n-2} \times \Delta_{\alpha,\beta}$ (up to a permutation of the factors) for some $\alpha, \beta \in H_x(\boldsymbol{Z}_{(p)})$, or $Y = V^n$. The skewed diagonal $\Delta_{\alpha,\beta}$ is given by

$$\Delta_{\alpha,\beta} = \{(\alpha(v), \beta(v)) \mid v \in V\}.$$

Suppose that $Y = V^{n-2} \times \Delta_{\alpha,\beta}$ (up to a permutation of the factors). Let (t, t') be the Serre–Tate coordinates of the last two factors $\Delta_{\alpha,\beta}$ at (x, x), respectively. It follows that

$$t^{\beta^{1-c_x}} = t'^{\alpha^{1-c_x}}.$$

On the other hand, from the definition of Y it follows that

$$t_{\mathfrak{p}}^{a_{n,\mathfrak{p}}} = t_{\mathfrak{p}}'^{a_{n-1,\mathfrak{p}}}.$$

Thus, $(a_n a_{n-1}^{-1})_{\mathfrak{p}} = (\beta\alpha^{-1})_{\mathfrak{p}}^{1-c_x} \in H_x(\boldsymbol{Z}_{(p)})_{\mathfrak{p}}$. This contradicts the hypothesis on the a_i. We conclude that $Y = V^n$. □

We have the following immediate consequence.

Corollary 4.4. *Theorem 4.1 holds.*

Proof. In view of Theorem 4.3, it follows that $\mathfrak{b} = 0$. Thus, $\mathfrak{b}_I = 0$ (see Lemma 4.2). □

5. *p*-adic differential operators

In this section, we consider p-adic operators on the space of p-adic Hilbert modular forms. This is a p-adic analogue of the Maass–Shimura differential operators on complex Hilbert modular forms. In Section 5A, we give an elementary construction of these operators. In Section 5B, we compute their action on the t-expansion of a p-adic Hilbert modular form around an ordinary point in terms of the partial Serre–Tate coordinates. For a more detailed account in the elliptic modular case, we refer to [Hida 2013a, §1.3.6]. In Section 6, we will use these results to compute the power series expansion of anticyclotomic Katz p-adic L-functions and variants (see Section 1).

5A. ***Elementary construction.*** In this subsection, we give an elementary construction of p-adic differential operators on the space of p-adic Hilbert modular forms.

Let the notation and assumptions be as in Section 2. For the geometric definition of classical and p-adic modular forms on the Hilbert modular Shimura variety, we refer to [Hida 2004, §4.2; 2010, §4.1].

Recall that $\mathbb{F}$ denotes an algebraic closure of $\mathbb{F}_p$, W the Witt ring $W(\mathbb{F})$, $\iota_p : \overline{\boldsymbol{Q}} \hookrightarrow \boldsymbol{C}_p$ a p-adic embedding and $\iota_\infty : \overline{\boldsymbol{Q}} \hookrightarrow \boldsymbol{C}$ a complex embedding. Let $\mathcal{W}$ denote $\iota_p^{-1}(W)$. Possibly enlarging $\mathcal{W}$, we suppose that $\tau(O) \subset \mathcal{W}$, for all $\tau \in \Sigma$.

In this subsection, we suppose the prime-to-p level of classical or p-adic Hilbert modular forms is one. This is only to simplify the notation.

Let $G_\kappa(\Gamma_1(p^m), \mathcal{W})$ be the space of classical Hilbert modular forms of weight κ and level $\Gamma_1(p^m)$ over $\mathcal{W}$, where $\kappa \in \boldsymbol{Z}_{\geq 0}[\Sigma]$ and m is a nonnegative integer. Let $f \in G_\kappa(\Gamma_1(p^m), \mathcal{W})$. Via ι_∞, we regard f as a Hilbert modular form over $\boldsymbol{C}$. Let

$z = (z_\tau)_{\tau\in\Sigma}$ be the complex variables of the Hilbert modular Shimura variety or those of $\mathfrak{H}^\Sigma$, where $\mathfrak{H}$ is the upper half plane (see [Hida 2010, §4.1]). Let $(\mathfrak{a}, \mathfrak{b})$ be a pair giving rise to a cusp of the Hilbert modular Shimura variety (see [Hida 2010, §4.1]). The Fourier expansion of f at the cusp corresponding to $(\mathfrak{a}, \mathfrak{b})$ is given by

$$f(z) = \sum_{\xi\in O} a(\xi) e_F(\xi z)$$

for $e_F(\xi z) = \exp\bigl(2\pi i \sum_{\tau\in\Sigma} \tau(\xi) z_\tau\bigr)$.

Let $\phi : O/p^r O \to \mathcal{W}$ be an arbitrary function with the normalised Fourier transform ϕ^* given by

$$\phi^*(y) = \frac{1}{p^{rd/2}} \sum_{u\in O/p^r O} \phi(u) e_F(yu/p^r), \tag{5-1}$$

where $y \in O/p^r O$ and $e_F(w) = \exp(2\pi i \operatorname{Tr}_{F/\mathbb{Q}}(w))$ for $w \in F$.

Let $f|\phi$ be the classical Hilbert modular form given by

$$f|\phi(z) = \sum_{u'=(\sigma_1(u),\ldots,\sigma_d(u)),u\in O/p^r O} \phi^*(-u) f(z + u'/p^r), \tag{5-2}$$

where $z + u'/p^r = (z_\tau + \tau(u)/p^r)_\tau$.

Note that $f|\phi \in G_\kappa(\Gamma_1(p^s), \mathcal{W})$, where $s = \max(m, 2r)$. In view of the Fourier inversion formula, it follow that the Fourier expansion of $f|\phi$ at the cusp corresponding to $(\mathfrak{a}, \mathfrak{b})$ is given by

$$f|\phi(z) = \sum_{\xi\in O} \phi(\xi) a(\xi) e_F(\xi z). \tag{5-3}$$

Let $(\phi_n^\sigma : O/p^n O \to \mathcal{W})_n$ be a sequence of functions such that

$$\phi_n^\sigma(\xi) \equiv \sigma(\xi) \pmod{p^n \mathcal{W}}.$$

In view of the q-expansion principle, it follow that the sequence $(f|\phi_n^\sigma)_n$ of classical Hilbert modular forms converges p-adically to a p-adic Hilbert modular form $d^\sigma f$ whose formal Fourier expansion at the cusp corresponding to $(\mathfrak{a}, \mathfrak{b})$ is given by

$$d^\sigma f(z) = \sum_{\xi\in O} \sigma(\xi) a(\xi) e_F(\xi z). \tag{5-4}$$

In other words, the operator d^σ equals the Maass–Shimura differential operator

$$\delta_0^\sigma = \frac{1}{2\pi i} \frac{\partial}{\partial z_\sigma}$$

on $G_\kappa(\Gamma_1(p^m), \mathcal{W})$.

This construction extends to the space of p-adic Hilbert modular forms over W as follows. Let $V(W)$ be the space of p-adic Hilbert modular forms of prime-to-p

level one over W. Via ι_p, we can regard $G_\kappa(\Gamma_1(p^m), \mathcal{W})$ as a subspace of $V(W)$ (see [Hida 2004, §8.1]). The space $G_\kappa(\Gamma_1(p^\infty), \mathcal{W}) = \bigcup_m G_\kappa(\Gamma_1(p^m), \mathcal{W})$ is p-adically dense in $V(W)$ (see [Hida 2004, §8.1]). Thus, the differential operator d^σ extends to $V(W)$.

Remark. In [Katz 1978, Chapter II], the operator d^σ is constructed based on the Gauss–Manin connection of the universal abelian scheme over the Shimura variety. The above approach can be generalised for a class of PEL Shimura varieties.

5B. ***Action on the t-expansion.*** In this subsection, we compute the action of the differential operator on the t-expansion of a p-adic Hilbert modular form around a p-ordinary point in terms of the partial Serre–Tate coordinates.

Let $(\zeta_{p^n} = \exp(2\pi i/p^n))_n \in \overline{\boldsymbol{Q}}$ be a compatible system of p-power roots of unity. Via ι_p, we regard it as a compatible system in $\boldsymbol{C}_p$.

Let $\mathfrak{p}$ be the prime corresponding to σ via ι_p and Σ_p be the set of places above p in F. For $\mathfrak{q} \in \Sigma_p$, let $\Sigma_{\mathfrak{q}} \subset \Sigma$ be the subset giving rise to $\mathfrak{q}$ under ι_p.

Let $W_n = W[\mu_{p^n}]$ and m_n be the maximal ideal. We have

$$(\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}^*)(W_n) = (1 + m_n) \otimes O_{\mathfrak{p}}^* \tag{5-5}$$

and

$$\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{q}}^* = \prod_{\tau \in \Sigma_{\mathfrak{q}}} \widehat{\mathbb{G}}_m. \tag{5-6}$$

Let $u \in O$ and $\alpha(u/p^m) \in G(\boldsymbol{A}^f)$ such that

$$\alpha(u/p^m)_{\mathfrak{p}} = \begin{bmatrix} 1 & u/p^m \\ 0 & 1 \end{bmatrix}$$

and $\alpha(u/p^m)_{\mathfrak{l}} = 1$, for $\mathfrak{l} \neq \mathfrak{p}$.

Let us recall some notation. Let $\pi : \mathrm{Ig} \to \mathrm{Sh}$ be the Igusa tower over W. Let $x \in \mathrm{Sh}^{\mathrm{ord}}_{/\mathbb{F}}$ be a closed point and $\tilde{x}$ be a point above it in Sh. Let $\widehat{S}_{/W}$ be the deformation space of $\tilde{x}$. For $q \in \Sigma_p$, recall $t_{\mathfrak{q}} = t \otimes 1_{\mathfrak{q}} \in \widehat{\mathbb{G}}_m \otimes O_{\mathfrak{q}}$ is the Serre–Tate coordinate of the partial deformation space $\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{q}}$ (see Section 3A). We regard $1 \otimes u \in \widehat{\mathbb{G}}_m \otimes O_{\mathfrak{q}}$ via the image $u \in O_{\mathfrak{q}}$.

We start with a preparatory lemma.

Lemma 5.1. *The isogeny action of $\alpha(u/p^m)$ on the Igusa tower $\pi : \mathrm{Ig} \to \mathrm{Sh}_{/W_m}$ preserves the deformation space $\widehat{S}$ and induces $t_{\mathfrak{p}} \mapsto \zeta_{p^m} t \otimes u$ (see* (5-5)*) and $t_{\mathfrak{p}'} \mapsto t_{\mathfrak{p}'}$, for $\mathfrak{p}' \neq \mathfrak{p}$.*

Proof. Let $x_{\mathrm{ST}} = (\mathcal{A}_x, \cdot)$ be the universal deformation and $x_0 = (A_0, \cdot)$ be the origin of the deformation space $\widehat{S}$. In particular, we have

$$A_0[\mathfrak{p}^m] = \mu_{p^m} \otimes O_{\mathfrak{p}}^* \oplus O_{\mathfrak{p}}/\mathfrak{p}^m O_{\mathfrak{p}}. \tag{5-7}$$

By the universal level structure, we have an exact sequence

$$0 \longrightarrow \mu_{p^m} \otimes O_{\mathfrak{p}}^* \longrightarrow \mathcal{A}_x[\mathfrak{p}^m] \xrightarrow{h} O_{\mathfrak{p}}/\mathfrak{p}^m O_{\mathfrak{p}} \longrightarrow 0.$$

A section of the morphism h determines an $O_{\mathfrak{p}}$-cyclic subgroup C_u isomorphic to $O_{\mathfrak{p}}/\mathfrak{p}^m O_{\mathfrak{p}}$ defined over W_m, which specialises at the origin A_0 to a cyclic subgroup generated by $(\zeta_{p^m} \otimes u, \gamma_0)$ in $A_0[\mathfrak{p}^m]$ (see (5-7)). Here γ_0 denotes the image of 1 under the identification $A_0[\mathfrak{p}^m]^{\text{ét}} = O_{\mathfrak{p}}/\mathfrak{p}^m O_{\mathfrak{p}}$.

The isogeny action $\alpha(u/p^m)$ corresponds to the isogeny $\mathcal{A}_x \to \mathcal{A}_{x,u} := \mathcal{A}_x/C_u$. By an argument similar to the proof of [Brakočević 2011, Lemmas 7.1, 7.2], it follows that the $\mathfrak{p}$-Serre–Tate coordinates of $\mathcal{A}_{x,u}$ is given by $\zeta_{p^m} t \otimes u$. For $\mathfrak{p}' \neq \mathfrak{p}$, in view of the construction of the Serre–Tate coordinates (see Section 3A), it follows that the $\mathfrak{p}'$-Serre–Tate coordinate is given by $t_{\mathfrak{p}'}$. □

For $\tau \in \Sigma_{\mathfrak{q}}$, let $t_{\mathfrak{q}}^{\tau}$ be the τ-component of $t_{\mathfrak{q}}$ (see (5-6)).

Proposition 5.2. *The action of the p-adic differential operator d^{σ} on the deformation space $\widehat{S}$ is given by*

$$t_{\mathfrak{p}}^{\sigma} \frac{\partial}{\partial t_{\mathfrak{p}}^{\sigma}}.$$

Proof. As a p-adic Hilbert modular form is a p-adic limit of classical Hilbert modular forms, it suffices to verify the proposition for classical Hilbert modular forms.

Let $f \in G_{\kappa}(\Gamma_1(p^m), \mathcal{W})$ and the t-expansion of f around x be given by

$$f(t) = \sum_{\omega} a(\omega) \prod_{\mathfrak{q} \in \Sigma_p} t_{\mathfrak{q}}^{\omega_{\mathfrak{q}}}. \tag{5-8}$$

Here by $t_{\mathfrak{q}}^{\omega_{\mathfrak{q}}}$, we mean the character $t^{\omega_{\mathfrak{q}}}$ (see Section 3A) and the summation is over $\omega \in O$ (see [Hida 2010, pp. 106–107]). To emphasise the notion of t-expansion around a point, we use the indexing notation ω instead of the traditional notation ξ for q-expansion around a cusp.

We have

$$f|\phi = \sum_{u \in O/p^r O} \phi^*(-u) f|\alpha(u/p^r),$$

as $\alpha(u/p^m)$ acts on $\mathfrak{H}^{\Sigma}$ by $z \mapsto z + u'/p^r$ (see (5-2)).

In view of Lemma 5.1, it follows that

$$\begin{aligned} f|\phi_n^{\sigma}(t) &= \sum_{\omega} a(\omega) \Bigg(\sum_{u \in O/p^n O} (\phi_n^{\sigma})^*(-u) \zeta_{p^n}^{\mathrm{Tr}_{F/\mathbb{Q}}(u\omega_{\mathfrak{p}})} \Bigg) t_{\mathfrak{p}}^{\omega_{\mathfrak{p}}} \prod_{\mathfrak{p}' \neq \mathfrak{p}} t_{\mathfrak{p}'}^{\omega_{\mathfrak{p}'}} \\ &= \sum_{\omega} \phi_n^{\sigma}(\omega_{\mathfrak{p}}) a(\omega) \prod_{\mathfrak{q} \in \Sigma_p} t_{\mathfrak{q}}^{\omega_{\mathfrak{q}}}. \end{aligned} \tag{5-9}$$

The last equality follows from the Fourier inversion formula.

Thus, we have

$$d^\sigma f(t) = \lim_n f|\phi_n^\sigma(t) = \sum_\omega \sigma(\omega_\mathfrak{p})a(\omega) \prod_{\mathfrak{q}\in\Sigma_p} t_\mathfrak{q}^{\omega_\mathfrak{q}} = t_{\mathfrak{p},\sigma}\frac{\partial f}{\partial t_{\mathfrak{p},\sigma}}. \qquad \square$$

We have the following immediate consequence.

Corollary 5.3. *The p-adic differential operator d^σ is $\widehat{S}$-invariant.*

Remark. The above corollary is proven in [Katz 1981, §4.3] based on the computation of the Gauss–Manin connection in terms of the Serre–Tate coordinates. The above approach can be generalised for a class of PEL Shimura varieties.

6. Iwasawa μ-invariants

In this section, we consider Iwasawa μ-invariants as in Theorems 1.2–1.4. In Section 6A, we determine the μ-invariant of certain anticyclotomic $\mathfrak{p}$-adic L-functions (see Theorem 1.2). In Section 6B, we determine the μ-invariant of the cyclotomic derivative $L'_{\Sigma,\lambda,\mathfrak{p}}$ of the Katz $\mathfrak{p}$-adic L-function, when the branch character λ is self-dual with the root number -1 (see Theorem 1.3). In Section 6C, we prove a $\mathfrak{p}$-version of a conjecture Gillard regarding the vanishing of the μ-invariant of Katz p-adic L-function (see Theorem 1.4).

6A. ***μ-invariant of anticyclotomic $\mathfrak{p}$-adic L-functions.*** In this subsection, we first obtain a formula for the μ-invariant anticyclotomic Katz $\mathfrak{p}$-adic L-function (see Theorem 1.2). Towards the end, we comment on a similar formula for a class of Rankin–Selberg anticyclotomic $\mathfrak{p}$-adic L-functions.

Let the notation and assumptions be as in the introduction. Let $\mathfrak{C}$ be a prime-to-p ideal of K. Let $Z(\mathfrak{C})$ be the Ray class group of K modulo $\mathfrak{C}p^\infty$. Let $Z(\mathfrak{C})^-$ be the anticyclotomic quotient. The reciprocity law $\mathrm{rec}_K : (\boldsymbol{A}_K^f)^\times \to Z(\mathfrak{C})^-$ induces the isomorphism

$$\mathrm{rec}_K : \varprojlim_n K^\times(\boldsymbol{A}_F^f)^\times\backslash(\boldsymbol{A}_K^f)^\times/U_K(\mathfrak{C}p^n) \xrightarrow{\sim} Z(\mathfrak{C})^-$$

for $U_K = (\mathcal{O}_K \otimes \widehat{\boldsymbol{Z}})^\times$ and

$$U_K(\mathfrak{C}p^n) = \{u \in U_K \mid u \equiv 1 \pmod{\mathfrak{C}p^n}\}.$$

Let Γ^- be the maximal $\boldsymbol{Z}_p$-free quotient of $Z(\mathfrak{C})^-$ and $\Gamma^-_\mathfrak{p}$ be the $\mathfrak{p}$-part of Γ^- (see Section 1).

Let Γ' and $\Gamma'_\mathfrak{p}$ be the open subgroups of Γ^- generated by the images via rec_K of

$$O_p^\times \times \prod_{v|D_{K/F}} K_v^\times \quad \text{and} \quad O_\mathfrak{p}^\times \times \prod_{v|D_{K/F}} K_v^\times,$$

respectively. Let Σ_p be the places above p in K induced by the p-ordinary CM type Σ. The reciprocity law rec_K at Σ_p induces an injective map

$$\mathrm{rec}_{\Sigma_p} : 1 + pO_p \hookrightarrow O_p^\times = \bigoplus_{w \in \Sigma_p} \mathcal{O}_{K_w}^\times \xrightarrow{\mathrm{rec}_K} Z(\mathfrak{C})^-$$

with finite cokernel as $p \nmid D_F$. It induces isomorphisms $\mathrm{rec}_{\Sigma_p} : 1 + pO_p \xrightarrow{\sim} \Gamma'$ and $\mathrm{rec}_{\Sigma_p} : 1 + \mathfrak{p}O_\mathfrak{p} \xrightarrow{\sim} \Gamma'_\mathfrak{p}$. Via these isomorphisms, we identify Γ' (resp. $\Gamma'_\mathfrak{p}$) with the subgroup $\mathrm{rec}_{\Sigma_p}(1 + pO_p)$ (resp. $\mathrm{rec}_{\Sigma_p}(1 + \mathfrak{p}O_\mathfrak{p})$) of the anticyclotomic quotient $Z(\mathfrak{C})^-$.

In [Katz 1978; Hida and Tilouine 1993], a $\bar{\boldsymbol{Z}}_p$-valued p-adic measure $\mathcal{L}_{\mathfrak{C},\Sigma}$ on $Z(\mathfrak{C})$ is constructed. It interpolates a class of critical Hecke L-values corresponding to Hecke characters over K with prime-to-p conductor $\mathfrak{C}$ (see [Katz 1978; Hida and Tilouine 1993; Burungale 2016a]). Let λ be an arithmetic Hecke character over K with prime-to-p conductor $\mathfrak{C}$. Let $\mathcal{L}^-_{\Sigma,\lambda}$ (resp. $\mathcal{L}^-_{\Sigma,\lambda,\mathfrak{p}}$) be the p-adic measure on Γ^- (resp. $\Gamma^-_\mathfrak{p}$) obtained by the push-forward of $\mathcal{L}_{\mathfrak{C},\Sigma}$ along λ.

Recall that the μ-invariant $\mu(\varphi)$ of a $\bar{\boldsymbol{Z}}_p$-valued p-adic measure φ on a p-adic group H is defined by

$$\mu(\varphi) = \inf_{U \subset H \text{ open}} v_p(\varphi(U)).$$

The μ-invariants of the above measures are related by the following theorem.

Theorem 6.1. *Let the notation and assumptions be as above. Then, we have*

$$\mu(\mathcal{L}^-_{\Sigma,\lambda}) = \mu(\mathcal{L}^-_{\Sigma,\lambda,\mathfrak{p}}).$$

Proof. Let $L^-_{\Sigma,\lambda}$ (resp. $L^-_{\Sigma,\lambda,\mathfrak{p}}$) be the power series of the measure $\mathcal{L}^-_{\Sigma,\lambda}$ (resp. $\mathcal{L}^-_{\Sigma,\lambda,\mathfrak{p}}$) in the sense of (1-3).

We first suppose that $p \nmid h_K^-$, where h_K^- is the relative class number given by $h_K^- = h_K/h_F$ and $h_?$ is the class number of ?, for ? = F, K. We thus have

$$Z(\mathfrak{C})^- \simeq \Gamma^-.$$

We now describe the approach of the second named author to determine $\mu(L^-_{\Sigma,\lambda})$ (see [Hida 2010]).

Under the hypothesis, there exist a finite number of classical Hilbert modular Eisenstein series $(f_{\lambda,i})_i$ such that

$$L^-_{\Sigma,\lambda} = \sum_i a_i \circ (f_{\lambda,i}(t)) \tag{6-1}$$

up to an automorphism of $\bar{\boldsymbol{Z}}_p[[\Gamma^-]]$. Here $f_{\lambda,i}(t)$ is the t-expansion of $f_{\lambda,i}$ around a well chosen CM point x with the CM type (K, Σ) on the Hilbert modular Shimura variety Sh and a_i is an automorphism of the deformation space of x in Sh for each i

(see [Hida 2010, Theorem 5.1; Hsieh 2014a, §5.2]). Moreover, the a_i satisfy the hypothesis in Theorem 4.1.

For $\kappa \in \mathbf{Z}_{\geq 0}[\Sigma]$, let ν_κ be the p-adic character of Γ' such that

$$\nu_\kappa(\mathrm{rec}_{\Sigma_p}(y)) = y^\kappa$$

for $y \in 1 + pO_p$. Let d^κ be the p-adic differential operator corresponding to κ (see Section 5A). From (6-1), we thus conclude

$$\left(d^\kappa \sum_i a_i \circ (f_{\lambda,i}(t))\right)\bigg|_{t=1} = \int_{\Gamma^-} \nu_\kappa \, d\mathcal{L}^-_{\Sigma,\lambda}. \tag{6-2}$$

In view of the linear independence of $(a_i \circ (f_{\lambda,i}))_i$ (see [Hida 2010, Theorem 3.20]), it follows that

$$\mu(\mathcal{L}^-_{\Sigma,\lambda}) = \min_i \mu(f_{\lambda,i}(t)) = \min_i \mu(f_{\lambda,i}). \tag{6-3}$$

We now turn towards the anticyclotomic Katz $\mathfrak{p}$-adic L-function.

For a p-adic Hilbert modular form f, let $f(t_\mathfrak{p})$ be obtained from $f(t)$ by substituting $t_{\mathfrak{p}'} = 1$, for all $\mathfrak{p}' \neq \mathfrak{p}$.

Let

$$f^-_{\Sigma,\lambda,\mathfrak{p}} = \sum_i a_{i,\mathfrak{p}} \circ (f_{\lambda,i}(t_\mathfrak{p})). \tag{6-4}$$

Recall that $\Sigma_\mathfrak{p} \subset \Sigma$ denotes the subset of infinite places of F corresponding to $\mathfrak{p}$, via ι_p. For $\sigma \in \Sigma_\mathfrak{p}$, let $d^{\sigma'}$ be the formal differential operator given by

$$f(t_\mathfrak{p}) \mapsto t_\mathfrak{p}^\sigma \frac{\partial f}{\partial t_\mathfrak{p}^\sigma}.$$

In view of Proposition 5.2, it follows that

$$d^{\sigma'}(f(t_\mathfrak{p})) = (d_\sigma f)(t_\mathfrak{p}). \tag{6-5}$$

From now on, let $\kappa \in \mathbf{Z}_{\geq 0}[\Sigma_\mathfrak{p}]$. Let $d^{\kappa'}$ be the corresponding formal differential operator.

We now have

$$d^{\kappa'}(f^-_{\Sigma,\lambda,\mathfrak{p}})|_{t_\mathfrak{p}=1} = (d^\kappa L^-_{\Sigma,\lambda})|_{t=1} = \int_{\Gamma^-} \nu_\kappa \, d\mathcal{L}^-_{\Sigma,\lambda} = \int_{\Gamma^-_\mathfrak{p}} \nu_\kappa \, d\mathcal{L}^-_{\Sigma,\lambda,\mathfrak{p}}. \tag{6-6}$$

The first two equalities follow from (6-5) and (6-2), respectively. The last equality follows from the fact that for $\kappa \in \mathbf{Z}_{\geq 0}[\Sigma_\mathfrak{p}]$, the character ν_κ factors through $\Gamma^-_\mathfrak{p}$.

In other words, $d^{\kappa'}(f^-_{\Sigma,\lambda,\mathfrak{p}})|_{t_\mathfrak{p}=1}$ interpolates the κ-th moment of the measure $\mathcal{L}^-_{\Sigma,\lambda,\mathfrak{p}}$. Thus,

$$f^-_{\Sigma,\lambda,\mathfrak{p}} = L^-_{\Sigma,\lambda,\mathfrak{p}}$$

up to an automorphism of $\bar{\mathbf{Z}}_p[[\Gamma^-_\mathfrak{p}]]$.

In view of the linear independence of $(a_{i,\mathfrak{p}} \circ (f_{\lambda,i}|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}}))_i$ (see Theorem 4.1), it follows that

$$\mu(L^-_{\Sigma,\lambda,\mathfrak{p}}) = \min_i \mu(f_{\lambda,i}(t_{\mathfrak{p}})) = \min_i \mu(f_{\lambda,i}|_{\widehat{\mathbb{G}}_m \otimes O_{\mathfrak{p}}}). \tag{6-7}$$

In view of Corollary 3.8, this finishes the proof of the theorem for the case $p \nmid h^-_K$.

When $p \mid h^-_K$, the power series $L^-_{\Sigma,\lambda}$ restricted to an explicit finite open cover of Γ^- is still of the form (5-1) (see [Hsieh 2014a, §5.2]). Thus, a similar argument proves the Theorem. □

In most of the cases, $\mu(L^-_{\Sigma,\lambda})$ has been explicitly determined [Hida 2010; Hsieh 2014a]. Thus, we obtain a formula for $\mu(L^-_{\Sigma,\lambda,\mathfrak{p}})$.

Remark. A class of Rankin–Selberg anticyclotomic p-adic L-functions for Hilbert modular forms is constructed in [Hsieh 2014b]. It also satisfies a property analogous to (6-1) [Hsieh 2014b, §6.2]. Thus, by an argument similar to the proof of Theorem 6.1, we get an analogue of Theorem 6.1. In the near future, the first named author hopes to consider Rankin–Selberg anticyclotomic p-adic L-functions for quaternionic modular forms. In these situations, the underlying Shimura variety turns out to be a quaternionic Shimura variety arising from a quaternion algebra over a totally real field which is not totally definite.

6B. ***μ-invariant of the cyclotomic derivative of the Katz 𝔭-adic L-function.*** In this subsection, we determine the μ-invariant of the cyclotomic derivative of the Katz 𝔭-adic L-function, when the branch character λ is self-dual with the root number -1 (see Theorem 1.3).

Let K^+_∞ be the cyclotomic $\mathbf{Z}_p$-extension of K and $K_{\mathfrak{p},\infty} = K^-_{\mathfrak{p},\infty} K^+_\infty$. Let $\Gamma = \mathrm{Gal}(K^-_\infty K^+_\infty/K)$ and $\Gamma_{\mathfrak{p}} = \mathrm{Gal}(K_{\mathfrak{p},\infty}/K)$. Let λ be an arithmetic Hecke character over K. Let $\mathcal{L}_{\Sigma,\lambda}$ (resp. $\mathcal{L}_{\Sigma,\lambda,\mathfrak{p}}$) be the p-adic measure on Γ (resp. $\Gamma_{\mathfrak{p}}$) obtained by the pull-back of $\mathcal{L}_{\mathfrak{C},\Sigma}$ (see Section 6A) along λ. We call $\mathcal{L}_{\Sigma,\lambda,\mathfrak{p}}$ the Katz 𝔭-adic L-function with branch character λ. Let $L_{\Sigma,\lambda}(T_1, T_2, \dots, T_d, S) \in \bar{\mathbf{Z}}_p[\![\Gamma]\!]$ (resp. $L_{\Sigma,\lambda,\mathfrak{p}}(\,\cdot\,, S) \in \bar{\mathbf{Z}}_p[\![\Gamma_{\mathfrak{p}}]\!]$) be the power series of $\mathcal{L}_{\Sigma,\lambda}$ (resp. $\mathcal{L}_{\Sigma,\lambda,\mathfrak{p}}$). Here, $T_1, \dots, T_d$ are the anticyclotomic variables and S is the cyclotomic variable.

In this subsection, we now suppose that λ is self-dual, i.e.,

$$\lambda|_{\mathbf{A}_F^\times} = \tau_{K/F} |\cdot|_{\mathbf{A}_F},$$

where $\tau_{K/F}$ is the quadratic character associated to K/F and $|\cdot|_{\mathbf{A}_F}$ is the adelic norm. In particular, the global root number of λ is ± 1. Now, suppose that the global root number is -1. In view of the functional equation of the Hecke L-function, this root number condition forces all the Hecke L-values appearing in the interpolation property of $\mathcal{L}^-_{\Sigma,\lambda}$ to vanish. Accordingly, we have $\mathcal{L}^-_{\Sigma,\lambda} = 0$. This also follows from the functional equation of $\mathcal{L}_{\Sigma,\lambda}$ [Hida and Tilouine 1993, §5]. In particular, we have $\mathcal{L}^-_{\Sigma,\lambda,\mathfrak{p}} = 0$.

We can consider the cyclotomic derivatives

$$L'_{\Sigma,\lambda} = \left(\frac{\partial}{\partial S} L_{\Sigma,\lambda}(T_1, \dots, T_d, S)\right)\Big|_{S=0} \tag{6-8}$$

and $L'_{\Sigma,\lambda,\mathfrak{p}}$ (defined analogously).

The μ-invariants of these derivatives are related by the following theorem.

Theorem 6.2. *Let the notation and assumptions be as above. In addition, suppose that $p \nmid h^-_K$, where h^-_K is the relative class number given by $h^-_K = h_K/h_F$. Then,*

$$\mu(L'_{\Sigma,\lambda}) = \mu(L'_{\Sigma,\lambda,\mathfrak{p}}).$$

Proof. We follow the notation in the proof of Theorem 6.1.

In the proof of [Burungale 2015, Theorem 3.2], it is shown there are p-adic Hilbert modular forms $(f'_{\lambda,i})_i$ such that

$$L'_{\Sigma,\lambda} = \frac{1}{\log_p(1+p)} \sum_i a_i \circ (f'_{\lambda,i}(t)), \tag{6-9}$$

up to an automorphism of $\bar{\mathbf{Z}}_p[[\Gamma^-]]$. More precisely, $f'_{\lambda,i}$ is the p-adic derivative of $f_{\lambda N^s,i}$ at $s = 0$ for N the norm Hecke character over K. Being a p-adic limit of classical Hilbert modular forms, it is a p-adic Hilbert modular form. We refer to the proof of [Burungale 2015, Theorem 3.2] for details.

In view of (6-4), (6-6) and by a similar argument as in the proof of [Burungale 2015, Theorem 3.2], it follows that

$$L'_{\Sigma,\lambda,\mathfrak{p}} = \frac{1}{\log_p(1+p)} \sum_i a_{i,\mathfrak{p}} \circ (f'_{i,\lambda}(t_{\mathfrak{p}})), \tag{6-10}$$

up to an automorphism of $\bar{\mathbf{Z}}_p[[\Gamma^-_{\mathfrak{p}}]]$.

We finish the proof in the same way as in Theorem 6.1. □

In most of the cases, $\mu(L'_{\Sigma,\lambda})$ has been explicitly determined [Burungale 2015, Theorem A]. Thus, we obtain a formula for $\mu(L'_{\Sigma,\lambda,\mathfrak{p}})$.

Remark. When $p \mid h^-_K$, we do not know an expression for $L_{\Sigma,\lambda}$ in terms of the t-expansion of certain Hilbert modular forms. Such an expression seems to be essential in the above approach.

6C. ***$\mathfrak{p}$-version of a conjecture of Gillard.*** In this subsection, we prove a $\mathfrak{p}$-version of a conjecture Gillard regarding the vanishing of the μ-invariant of Katz p-adic L-function (see Theorem 1.4).

Let λ be an arithmetic Hecke character over K. Recall that we have the Katz $\mathfrak{p}$-adic L-function $\mathcal{L}_{\Sigma,\lambda,\mathfrak{p}} \in \bar{\mathbf{Z}}_p[[\Gamma_{\mathfrak{p}}]]$ (see Section 6B). As a consequence of Theorem 6.1

and the results of [Hida 2011; Burungale and Hsieh 2013], we prove a $\mathfrak{p}$-version of a conjecture of Gillard [1991, Conjecture (i)] regarding the vanishing of the μ-invariant of the Katz p-adic L-function. The conjecture was originally formulated for the $(d+1)$-variable Katz p-adic L-function.

Theorem 6.3. *Let the notation and assumptions be as above. Then, we have*

$$\mu(\mathcal{L}_{\Sigma,\lambda,\mathfrak{p}}) = 0.$$

Proof. Let $\mathfrak{X}^+$ be the set consisting of finite order characters $\epsilon : \Gamma^+ \to \mu_{p^\infty}$. For every $\epsilon \in \mathfrak{X}^+$, we regard ϵ as a Hecke character.

In [Hida 2011; Burungale and Hsieh 2013], it has been shown that

$$\liminf_{\epsilon\in\mathfrak{X}^+} \mu(\mathcal{L}^-_{\Sigma,\lambda\epsilon}) = 0. \tag{6-11}$$

In view of Theorem 6.1, this finishes the proof:

$$0 \le \mu(\mathcal{L}_{\Sigma,\lambda,\mathfrak{p}}) \le \liminf_{\epsilon} \mu(\mathcal{L}^-_{\Sigma,\lambda\epsilon}). \qquad \square$$

The theorem evidently implies the main results of [Hida 2011; Burungale and Hsieh 2013].

Acknowledgements

We thank Adebisi Agboola, Ralph Greenberg, Mahesh Kakde, Chandrashekhar Khare, Michael Harris, Ming-Lun Hsieh, Davide Reduzzi, Romyar Sharifi, Richard Taylor, Jacques Tilouine, Kevin Ventullo and Bin Zhao for interesting conversations about the topic. During the preparation of the first draft, the first named author was a graduate student in UCLA and he is grateful to friends back then for the stimulating atmosphere. Finally, we are grateful to the referees for helpful comments and instructive suggestions.

References

[Brakočević 2011] M. Brakočević, “Anticyclotomic p-adic L-function of central critical Rankin–Selberg L-value”, *Int. Math. Res. Not.* **2011**:21 (2011), 4967–5018. MR Zbl

[Burungale 2015] A. A. Burungale, “On the μ-invariant of the cyclotomic derivative of a Katz p-adic L-function”, *J. Inst. Math. Jussieu* **14**:1 (2015), 131–148. MR Zbl

[Burungale 2016a] A. A. Burungale, “An $l \neq p$-interpolation of genuine p-adic L-functions”, *Res. Math. Sci.* **3** (2016), Paper No. 16, 26. MR Zbl

[Burungale 2016b] A. A. Burungale, “Non-triviality of generalised Heegner cycles over anticyclotomic towers: a survey”, pp. 279–306 in *p-adic aspects of modular forms*, edited by B. Balasubramanyam et al., World Sci., Hackensack, NJ, 2016. MR Zbl

[Burungale 2017] A. A. Burungale, “On the non-triviality of the p-adic Abel–Jacobi image of generalised Heegner cycles modulo p, II: Shimura curves”, *J. Inst. Math. Jussieu* **16**:1 (2017), 189–222. MR Zbl

[Burungale ≥ 2017] A. Burungale, "$\mathfrak{p}$-rigidity and $\mathfrak{p}$-independence of quaternionic modular forms modulo p", preprint.

[Burungale and Disegni ≥ 2017] A. Burungale and D. Disegni, "On the non-vanishing of p-adic heights for CM abelian varieties, and the arithmetic of Katz p-adic L-functions", preprint.

[Burungale and Hida 2016] A. A. Burungale and H. Hida, "André–Oort conjecture and nonvanishing of central L-values over Hilbert class fields", *Forum Math. Sigma* **4** (2016), e20, 26. MR Zbl

[Burungale and Hsieh 2013] A. Burungale and M.-L. Hsieh, "The vanishing of μ-invariant of p-adic Hecke L-functions for CM fields", *Int. Math. Res. Not.* **2013**:5 (2013), 1014–1027. MR Zbl

[Chai 1995] C.-L. Chai, "Every ordinary symplectic isogeny class in positive characteristic is dense in the moduli", *Invent. Math.* **121**:3 (1995), 439–479. MR Zbl

[Chai 2003] C.-L. Chai, "Families of ordinary abelian varieties: canonical coordinates, p-adic monodromy,Tate-linear subvarieties and Hecke orbits", preprint, 2003, http://www.math.upenn.edu/~chai/papers_pdf/fam_ord_av.pdf.

[Chai 2006] C.-L. Chai, "Hecke orbits as Shimura varieties in positive characteristic", pp. 295–312 in *International Congress of Mathematicians, II*, Eur. Math. Soc., Zürich, 2006. MR Zbl

[Eischen et al. 2016] E. Eischen, M. Harris, J. Li, and C. Skinner, "p-adic L-functions for unitary groups", preprint, 2016. arXiv

[Gillard 1991] R. Gillard, "Remarques sur l'invariant mu d'Iwasawa dans le cas CM", *J. Théor. Nombres Bordeaux* (2) **3**:1 (1991), 13–26. MR Zbl

[Harris et al. 2006] M. Harris, J.-S. Li, and C. M. Skinner, "p-adic L-functions for unitary Shimura varieties, I: Construction of the Eisenstein measure", *Doc. Math.* Extra Vol. (2006), 393–464. MR Zbl

[Hida 2004] H. Hida, *p-adic automorphic forms on Shimura varieties*, Springer, 2004. MR Zbl

[Hida 2006] H. Hida, "Anticyclotomic main conjectures", *Doc. Math.* Extra Vol. (2006), 465–532. MR Zbl

[Hida 2009a] H. Hida, "Irreducibility of the Igusa tower", *Acta Math. Sin.* (*Engl. Ser.*) **25**:1 (2009), 1–20. MR Zbl

[Hida 2009b] H. Hida, "Quadratic exercises in Iwasawa theory", *Int. Math. Res. Not.* **2009**:5 (2009), 912–952. MR Zbl

[Hida 2010] H. Hida, "The Iwasawa μ-invariant of p-adic Hecke L-functions", *Ann. of Math.* (2) **172**:1 (2010), 41–137. MR Zbl

[Hida 2011] H. Hida, "Vanishing of the μ-invariant of p-adic Hecke L-functions", *Compos. Math.* **147**:4 (2011), 1151–1178. MR Zbl

[Hida 2013a] H. Hida, *Elliptic curves and arithmetic invariants*, Springer, 2013. MR Zbl

[Hida 2013b] H. Hida, "Local indecomposability of Tate modules of non-CM abelian varieties with real multiplication", *J. Amer. Math. Soc.* **26**:3 (2013), 853–877. MR Zbl

[Hida and Tilouine 1993] H. Hida and J. Tilouine, "Anti-cyclotomic Katz p-adic L-functions and congruence modules", *Ann. Sci. École Norm. Sup.* (4) **26**:2 (1993), 189–259. MR Zbl

[Hsieh 2014a] M.-L. Hsieh, "On the μ-invariant of anticyclotomic p-adic L-functions for CM fields", *J. Reine Angew. Math.* **688** (2014), 67–100. MR Zbl

[Hsieh 2014b] M.-L. Hsieh, "Special values of anticyclotomic Rankin–Selberg L-functions", *Doc. Math.* **19** (2014), 709–767. MR Zbl

[Katz 1978] N. M. Katz, "p-adic L-functions for CM fields", *Invent. Math.* **49**:3 (1978), 199–297. MR Zbl

[Katz 1981] N. Katz, "Serre–Tate local moduli", pp. 138–202 in *Algebraic surfaces* (Orsay, France, 1976–78), Lecture Notes in Math. **868**, Springer, 1981. MR Zbl

[Shimura 1963] G. Shimura, "On analytic families of polarized abelian varieties and automorphic functions", *Ann. of Math.* (2) **78** (1963), 149–192. MR Zbl

[Shimura 1998] G. Shimura, *Abelian varieties with complex multiplication and modular functions*, Princeton Mathematical Series **46**, Princeton University Press, 1998. MR Zbl

Communicated by John Henry Coates
Received 2016-09-26 Revised 2016-11-21 Accepted 2017-02-06

ashayburungale@gmail.com *Département de Mathématiques, LAGA, Institute Galilée, Université Paris 13, 93430 Villetaneuse, France*

hhida@ucla.edu *Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, United States*

ALGEBRA AND NUMBER THEORY 11:8 (2017)
dx.doi.org/10.2140/ant.2017.11.1953

A Mordell–Weil theorem for cubic hypersurfaces of high dimension

Stefanos Papanikolopoulos and Samir Siksek

Let $X/\mathbb{Q}$ be a smooth cubic hypersurface of dimension $n \geq 1$. It is well-known that new rational points may be obtained from old ones by secant and tangent constructions. In view of the Mordell–Weil theorem for $n = 1$, Manin (1968) asked if there exists a finite set S from which all other rational points can be thus obtained. We give an affirmative answer for $n \geq 48$, showing in fact that we can take the generating set S to consist of just one point. Our proof makes use of a weak approximation theorem due to Skinner, a theorem of Browning, Dietmann and Heath-Brown on the existence of rational points on the intersection of a quadric and cubic in large dimension, and some elementary ideas from differential geometry, algebraic geometry and numerical analysis.

1. Introduction

Let $X \subset \mathbb{P}^{n+1}$ be a smooth cubic hypersurface over $\mathbb{Q}$ of dimension n. Let $\ell \subset \mathbb{P}^{n+1}$ be a line defined over $\mathbb{Q}$. If ℓ is not contained in X then $\ell \cdot X = P + Q + R$ where P, Q, $R \in X$. If any two of P, Q, R are rational then so is the third. If $S \subseteq X(\mathbb{Q})$, we write $\operatorname{Span}(S)$ for the subset of $X(\mathbb{Q})$ generated from S by successive secant and tangent constructions. More formally, we define a sequence

$$S = S_0 \subseteq S_1 \subseteq S_2 \subseteq \cdots \subseteq X(\mathbb{Q})$$

by letting S_{n+1} be the set of points $R \in X(\mathbb{Q})$ such that either $R \in S_n$, or for some $\mathbb{Q}$-line $\ell \not\subset X$ we have $\ell \cdot X = P + Q + R$ where P, $Q \in S_n$. Then $\operatorname{Span}(S) := \bigcup S_n$. Manin [1974, page 3] (see also [Kanevsky and Manin 2001] and [Manin 1997]) asks if there is some finite subset $S \subset X(\mathbb{Q})$ such that $\operatorname{Span}(S) = X(\mathbb{Q})$.

Theorem 1. *Let X be a smooth cubic hypersurface of dimension $n \geq 48$ defined over $\mathbb{Q}$. Then there exists a point $A \in X(\mathbb{Q})$ such that* $\operatorname{Span}(A) = X(\mathbb{Q})$.

Siksek is supported by the EPSRC *LMF: L-Functions and Modular Forms* Programme Grant EP/K034383/1.

MSC2010: primary 14G05; secondary 11G35.

Keywords: cubic hypersurfaces, rational points, Mordell–Weil problem.

We are grateful to Tim Browning, Simon Rydin Myerson, Michael Stoll and Damiano Testa for valuable discussions, and to the referees for useful remarks. We thank Yuri Manin for drawing our attention to the distinction between weak and strong Mordell–Weil generation (discussed below).

Some remarks. It appears that the Mordell–Weil problem for cubic surfaces was first posed by Segre [1943]. In the sixties Manin posed the same question for cubic hypersurfaces of arbitrary dimension. Prior to Theorem 1, so far as we know, the only positive result is for cubic surfaces endowed with a skew pair of rational lines [Siksek 2012, Theorem 1]. More recently, Manin [2012] made a distinction between "weak Mordell–Weil generation" where both secant and tangent operations are allowed, and "strong Mordell–Weil generation" where only secant operations are allowed. On the basis of computational experiments carried out by Vioreanu [2009], Manin expects that the weak version of the Mordell–Weil property holds for dimension 2, but that the strong version probably fails. In the language of [Manin 2012], our Theorem 1 establishes the weak Mordell–Weil property for cubic hypersurfaces of dimension ≥ 48; tangent operations are crucial to our proofs, and we are unable to adapt them to prove the strong Mordell–Weil property.

Notation. Throughout $X \subset \mathbb{P}^{n+1}$ is a smooth cubic hypersurface of dimension n defined over $\mathbb{Q}$ (for now $n \geq 2$). Thus there is some non-zero homogeneous cubic polynomial $F \in \mathbb{Q}[x_0, \dots, x_{n+1}]$ such that X is given by the equation

$$X : \quad F(x_0, \dots, x_{n+1}) = 0. \tag{1}$$

For $P \in X$ we let $T_P X$ denote the tangent plane to X at P:

$$T_P X : \quad \nabla F(P) \cdot (x_0, \dots, x_{n+1}) = 0.$$

The Gauss map on X sends P to $T_P X \in \mathbb{P}^{n+1*}$. We let $X_P := X \cap T_P X$. Thus

$$X_P : \quad \begin{cases} F(x_0, \dots, x_n) = 0, \\ \nabla F(P) \cdot (x_0, \dots, x_n) = 0. \end{cases}$$

In Section 4 we introduce the second fundamental form $\Pi_P X$, and the Hessian $H_F(P)$. We write $\mathbb{G}(n+1, 1)$ for the Grassmannian parametrizing lines in $\mathbb{P}^{n+1}$. Throughout the terms "open" and "closed" will pertain to the real topology, unless prefixed by "Zariski".

A sketch of the proof of Theorem 1. We show in Section 5 that if $B \in X(\mathbb{Q})$ is not an Eckardt point then $X_B(\mathbb{Q}) \subseteq \mathrm{Span}(B)$ (the definition of Eckardt points is given in Section 4). Fix $B \in X(\mathbb{Q})$ that is non-Eckardt. Given $D \in X(\mathbb{Q})$, we ask if there is $C \in X_B(\mathbb{Q})$ such that $D \in X_C(\mathbb{Q})$? If so, then provided C is non-Eckardt, we have $D \in \mathrm{Span}(C) \subseteq \mathrm{Span}(B)$. The answer to this question is positive provided the variety $Y_{B,D} \subset \mathbb{P}^{n+1}$ given by

$$Y_{B,D} : \quad \begin{cases} F(x_0, \dots, x_{n+1}) = 0, \\ \nabla F(x_0, \dots, x_{n+1}) \cdot D = 0, \\ \nabla F(B) \cdot (x_0, \dots, x_{n+1}) = 0. \end{cases} \tag{2}$$

has a rational point. A theorem of Browning, Dietmann and Heath-Brown allows us to deduce the existence of a rational point under some conditions, the most important being that n is large, and that $Y_{B,D}$ has a smooth real point. By considering the second fundamental form, and using a theorem on weak approximation for cubic hypersurfaces due to Skinner, we shall show the existence of a point $B \in X(\mathbb{Q})$ and a non-empty open $U \subseteq X(\mathbb{R})$, so that $Y_{B,D}$ has a smooth real point for all $D \in U$. It follows (with a little care) that $U \cap X(\mathbb{Q}) \subseteq \mathrm{Span}(B)$. Once the existence of such a set U is established, we use Mordell–Weil operations to enlarge U and quickly complete the proof of Theorem 1.

2. Some results from analytic number theory

Weak approximation.

Theorem 2 [Skinner 1997]. *Suppose $n \geq 15$. Then X satisfies weak approximation.*

This means that $X(\mathbb{Q})$ is dense in $X(A_{\mathbb{Q}})$ where $A_{\mathbb{Q}}$ denotes the adeles. It follows that $X(\mathbb{Q})$ is dense in $X(\mathbb{R})$; a fact we use repeatedly in the proof of Theorem 1.

Corollary 2.1. *Suppose $n \geq 15$. Let $U, V \subseteq X(\mathbb{R})$ be disjoint open sets. Let $A' \in U$, $B' \in V$, and let $\ell' \not\subset X$ be an $\mathbb{R}$-line such that $\ell' \cdot X = 2A' + B'$. Then there are $A \in U \cap X(\mathbb{Q})$, $B \in V \cap X(\mathbb{Q})$ and a $\mathbb{Q}$-line $\ell \not\subset X$ such that $\ell \cdot X = 2A + B$.*

Proof. The projectivized tangent bundle $\mathcal{T}_X$ of X parametrizes pairs (P, ℓ) with $P \in X$ and ℓ a line tangent to X at P. We make use of the fact that $\mathcal{T}_X$ is locally trivial; thus there is a Zariski open $\mathcal{U}$ containing A', and a local isomorphism $\varphi : \mathcal{U} \times \mathbb{P}^{n-1} \to \mathcal{T}_X$ such that $\varphi(P, \alpha) = (P, \ell_{P,\alpha})$ where $\ell_{P,\alpha}$ is a line tangent to X at P. Moreover, as A' is real we take φ to be defined over $\mathbb{R}$. Let $W = \mathcal{U}(\mathbb{R}) \cap U$ which is necessarily an open neighbourhood of A'. Let $\alpha \in \mathbb{P}^{n-1}(\mathbb{R})$ so that $\ell' = \ell_{A',\alpha}$. By Theorem 2 we can find $\{A_i\} \subset W \cap X(\mathbb{Q})$ converging to A'. Write $\ell_i = \ell_{A_i,\alpha}$. Then $\{\ell_i\}$ converges to ℓ' in $\mathbb{G}(n+1, 1)(\mathbb{R})$. In particular, for sufficiently large i, the line ℓ_i meets V. Let $A = A_i \in U \cap X(\mathbb{Q})$ for any such large i. Choosing a line $\ell/\mathbb{Q}$ tangent to X at A that sufficiently approximates ℓ_i completes the proof. $\square$

Intersections of a cubic with a quadric. Let $Q, C \in \mathbb{Q}[x_1, \dots, x_k]$ be a pair of forms of degrees 2 and 3 respectively, such that

$$Y : \quad C(x_1, \dots, x_k) = 0, \quad Q(x_1, \dots, x_k) = 0$$

is a complete intersection $Y \subset \mathbb{P}^{k-1}$. Using the circle method, Browning, Dietmann and Heath-Brown establish various sufficient conditions for Y to have a rational point. We recount one of their theorems [Browning et al. 2015, Theorem 4] which will be essential to our proof of Theorem 1. Define $\mathrm{ord}_Q(C)$ as the least non-negative integer m such that $C = LQ + C'$, with linear form $L \in \mathbb{Q}[x_1, \ldots, x_k]$, and such that there is a matrix $\boldsymbol{T} \in \mathrm{GL}_k(\mathbb{Q})$ with $C'(\boldsymbol{T}(x_1, \ldots, x_k)) \in \mathbb{Q}[x_1, \ldots, x_m]$.

Theorem 3 (Browning, Dietmann and Heath-Brown). *With notation as above, suppose $k \geq 49$ and $\mathrm{ord}_Q(C) \geq 17$. If Y has a smooth real point then $Y(\mathbb{Q}) \neq \varnothing$.*

Corollary 2.2. *Let $f, q, l \in \mathbb{Q}[x_0, \ldots, x_{n+1}]$, be forms of degree* 3, 2, 1. *Write*

$$Z : \quad f(x_0, \ldots, x_{n+1}) = q(x_0, \ldots, x_{n+1}) = l(x_0, \ldots, x_{n+1}) = 0$$

for their common locus of zeros in $\mathbb{P}^{n+1}$. Suppose that

(i) *the cubic hypersurface in $\mathbb{P}^{n+1}$ defined by f is smooth*;

(ii) *Z has a smooth real point*;

(iii) $n \geq 48$.

Then Z has a rational point.

Proof. By a non-singular change of variable, we may suppose that $l = x_0$. Let

$$f'(x_1, \ldots, x_{n+1}) = f(0, x_1, \ldots, x_{n+1}), \quad q'(x_1, \ldots, x_{n+1}) = q(0, x_1, \ldots, x_{n+1}).$$

We may therefore consider Z as being given in $\mathbb{P}^n$ as the common locus of $f' = q' = 0$. Suppose $\mathrm{ord}_{q'}(f') \leq 16$. Then a further non-singular change of variables allows us to write

$$f = x_0 q_0 + l' q' + h(x_1, \ldots, x_{16}).$$

where q_0 is a quadratic form, l' is a linear form, and h is a cubic form. Now as $n \geq 48$, there is a common zero in $\mathbb{P}^{n+1}$ to

$$x_0 = x_1 = \cdots = x_{16} = l' = q_0 = q' = 0.$$

This gives a singular point on the cubic hypersurface $f = 0$ in $\mathbb{P}^{n+1}$ contradicting (i). We may thus suppose that $\mathrm{ord}_{q'}(f') \geq 17$. A similar argument shows that $f' = q' = 0$ defines a complete intersection in $\mathbb{P}^n$. By (ii) this intersection has a smooth real point. Applying Theorem 3 with $k = n + 1$ completes the proof. □

3. A numerical stability criterion

Newton–Raphson. We need a rigorous version of the multivariate Newton–Raphson method. The following result is part of Theorem 5.3.2 of [Stoer and Bulirsch 1980].

Here $\|\cdot\|$ denotes the usual Euclidean norm (both for vectors and for matrices). For differentiable $\boldsymbol{f} = (f_1, \ldots, f_n) : \mathbb{R}^n \to \mathbb{R}^n$, denote the Jacobian matrix by $J_{\boldsymbol{f}}$:

$$J_{\boldsymbol{f}} := \left(\frac{\partial f_i}{\partial x_j}\right)_{i,j=1,\ldots,n}.$$

Theorem 4. *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be open, $\mathcal{C}_0$ be convex such that $\overline{\mathcal{C}_0} \subseteq \mathcal{C}$, and let $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable for all $\boldsymbol{x} \in \mathcal{C}_0$ and continuous for all $\boldsymbol{x} \in \mathcal{C}$.*

For $\boldsymbol{x}_0 \in \mathcal{C}_0$ let r, α, β, γ, h be given positive numbers with the following properties:

$$\mathcal{B}_r(\boldsymbol{x}_0) := \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x}_0\| < r\} \subseteq \mathcal{C}_0, \quad h := \alpha\beta\gamma/2 < 1, \quad r := \alpha/(1-h),$$

and let $\boldsymbol{f}$ satisfy:

(i) $\|J_{\boldsymbol{f}}(\boldsymbol{x}) - J_{\boldsymbol{f}}(\boldsymbol{y})\| \le \gamma \|\boldsymbol{x} - \boldsymbol{y}\|$ *for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}_0$*;

(ii) $J_{\boldsymbol{f}}(\boldsymbol{x})^{-1}$ *exists and satisfies* $\|J_{\boldsymbol{f}}(\boldsymbol{x})^{-1}\| \le \beta$ *for all $\boldsymbol{x} \in \mathcal{C}_0$*;

(iii) $\|\boldsymbol{f}(\boldsymbol{x}_0) \cdot J_{\boldsymbol{f}}(\boldsymbol{x}_0)^{-1}\| \le \alpha$.

Then beginning at $\boldsymbol{x}_0$ each point

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_k) \cdot J_{\boldsymbol{f}}(\boldsymbol{x}_k)^{-1}, \quad k = 0, 1, 2, \ldots$$

is well-defined and belongs to $\mathcal{B}_r(\boldsymbol{x}_0)$. Moreover the limit $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{\xi}$ *exists, belongs to $\overline{\mathcal{B}_r(\boldsymbol{x}_0)}$ and satisfies $\boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{0}$.*

Stability. For $f \in \mathbb{R}[x_1, \ldots, x_n]$ we shall let $\|f\|_c$ denote the maximum of the absolute values of the coefficients of f.

Lemma 3.1. *Let $g_1, \ldots, g_m \in \mathbb{R}[x_1, \ldots, x_n]$ be polynomials with $m \le n$. Let $\boldsymbol{\zeta} \in \mathbb{R}^n$ be a common zero of $g_1, \ldots, g_m$, such that $\nabla g_1(\boldsymbol{\zeta}), \ldots, \nabla g_m(\boldsymbol{\zeta})$ are linearly independent. Let $\varepsilon > 0$ be given. There is $\delta > 0$ such that if $f_1, \ldots, f_m \in \mathbb{R}[x_1, \ldots, x_n]$ satisfy $\|f_i - g_i\|_c < \delta$, then there is $\boldsymbol{\xi} \in \mathbb{R}^n$ such that*

(a) $\boldsymbol{\xi}$ *is a common zero to* $f_1, \ldots, f_m$;

(b) $\nabla f_1(\boldsymbol{\xi}), \ldots, \nabla f_m(\boldsymbol{\xi})$ *are linearly independent*;

(c) $\|\boldsymbol{\xi} - \boldsymbol{\zeta}\| < \varepsilon$.

Proof. Choose $\boldsymbol{v}_{m+1}, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n$ so that $\nabla g_1(\boldsymbol{\zeta}), \ldots, \nabla g_m(\boldsymbol{\zeta}), \boldsymbol{v}_{m+1}, \ldots, \boldsymbol{v}_n$ is a basis. Let

$$g_i(\boldsymbol{x}) = \boldsymbol{v}_i \cdot (\boldsymbol{x} - \boldsymbol{\zeta}), \quad i = m+1, \ldots, n.$$

Then $\boldsymbol{\zeta}$ is a common zero to $g_1, \ldots, g_n$ and $\nabla g_1(\boldsymbol{\zeta}), \ldots, \nabla g_n(\boldsymbol{\zeta})$ are linearly independent. Let $\boldsymbol{g} = (g_1, \ldots, g_n)$. Then $J_{\boldsymbol{g}}(\boldsymbol{\zeta})$ is invertible. We shall fix $f_i = g_i$ for $i = m+1, \ldots, n$, and let $\boldsymbol{f} = (f_1, \ldots, f_n)$. We will apply Theorem 4 with $\boldsymbol{x}_0 = \boldsymbol{\zeta}$. There is some $\delta_0 > 0$ such that if $\|f_i - g_i\|_c < \delta_0$ then $J_{\boldsymbol{f}}(\boldsymbol{x}_0)$ is invertible.

Choose $0 < r_0 \le \varepsilon$ so that condition (ii) of the theorem is satisfied for some $\beta > 0$, with $\mathcal{C}_0 = \overline{\mathcal{B}_{r_0}(\boldsymbol{x}_0)}$. Condition (i) holds for some $\gamma > 0$ by the multivariate Taylor Theorem. Let $\alpha = \|\boldsymbol{f}(\boldsymbol{x}_0) \cdot J_{\boldsymbol{f}}(\boldsymbol{x}_0)^{-1}\|$, which depends on $\boldsymbol{f}$. Now $\boldsymbol{g}(\boldsymbol{x}_0) = \boldsymbol{0}$, so clearly if $\delta \to 0$, then $\alpha \to 0$. Therefore for sufficiently small $\delta < \delta_0$, we have $h := \alpha\beta\gamma/2 < 1$ and $r := \alpha/(1-h) < r_0$. By the theorem, there is $\boldsymbol{\xi} \in \overline{\mathcal{B}_r(\boldsymbol{x}_0)}$ such that $\boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{0}$. By construction $\boldsymbol{\xi}$ satisfies (a), (b), (c). □

Smooth real points on the varieties $Y_{B,D}$.

Lemma 3.2. *Let B, D' be points belonging to $X(\mathbb{R})$ such that the variety $Y_{B,D'} \subset X \subset \mathbb{P}^{n+1}$ given by* (2) *has a smooth real point C'. Let $V \subseteq X(\mathbb{R})$ be an open neighbourhood of C'. Then there is an open neighbourhood $U \subseteq X(\mathbb{R})$ of D', such that for every $D \in U$, the variety $Y_{B,D}$ has a smooth real point $C \in V$.*

Proof. We may suppose that B, C', D' are contained in the affine patch $x_0 = 1$. Let G_1, G_2, G_3 be the three polynomials defining $Y_{B,D'}$ in (2) and let g_1, g_2, $g_3 \in \mathbb{R}[x_1, \dots, x_{n+1}]$ be their dehomogenizations by $x_0 = 1$. Write f_1, f_2, f_3 for the corresponding polynomials in $\mathbb{R}[x_1, \dots, x_{n+1}]$ defining $Y_{B,D} \cap \{x_0 = 1\}$ with $D \in X(\mathbb{R}) \cap \{x_0 = 1\}$. Of course $f_1 = g_1$, $f_3 = g_3$, and moreover

$$\|f_2 - g_2\|_c \le \mu \cdot \|D - D'\|_\infty$$

where $\mu > 0$ is a constant and $\|\cdot\|_\infty$ denotes the infinity norm in the affine patch $x_0 = 1$ (which we identify with $\mathbb{R}^{n+1}$). Now $C' \in \mathbb{R}^{n+1}$ is a common zero for g_1, g_2, g_3 with $\nabla g_1(C')$, $\nabla g_2(C')$, $\nabla g_3(C')$ linearly independent (as C' is now a smooth point on the affine patch $Y_{B,D'} \cap \{x_0 = 1\}$). Let $\varepsilon > 0$ be sufficiently small so that $\mathcal{B}_\varepsilon(C') \cap X(\mathbb{R})$ is contained in V. Applying Lemma 3.1, we know that if $\|D - D'\|_\infty$ is sufficiently small then there is a non-zero vector $C \in \mathcal{B}_\varepsilon(C')$ that is a common zero for f_1, f_2, f_3 with $\nabla f_1(C)$, $\nabla f_2(C)$, $\nabla f_3(C)$ linearly independent. This completes the proof. □

Lemma 3.3. *Suppose $n \ge 48$. Let $B \in X(\mathbb{Q})$. Suppose $D' \in X(\mathbb{R})$ such that $Y_{B,D'}$ has a smooth real point C'. Then there is a non-empty open $U \subseteq X(\mathbb{R})$ such that if $D \in U \cap X(\mathbb{Q})$, then $Y_{B,D}(\mathbb{Q}) \neq \varnothing$.*

Proof. Let U be as in Lemma 3.2. Then $Y_{B,D}$ is defined over $\mathbb{Q}$ and has a smooth real point for every $D \in U \cap X(\mathbb{Q})$. Now the lemma follows from Corollary 2.2. □

4. A little geometry

Lines on X.

Lemma 4.1. *Let ℓ be a line contained in X and $P \in \ell$. Then $\ell \subset T_P X$.*

This is well-known; for a proof, see [Siksek 2012, Lemma 2.1].

The second fundamental form. Let $P \in X$. Associated to P is a quadratic form (well-defined up to multiplication by a non-zero scalar) known as the second fundamental form which we denote by $\Pi_P X$, and which is defined as the differential of the Gauss map (e.g., [Griffiths and Harris 1979; Harris 1992, Chapter 17]. For our purpose the following explicit recipe given in [Griffiths and Harris 1979, pages 369–370] is useful. By carrying out a non-singular change of coordinates we may suppose P is $(1:0:\ldots:0)$, and the tangent plane $T_P X$ to X at P given by $x_{n+1}=0$. Then X has the equation $F=0$ with

$$F = x_0^2 x_{n+1} + x_0 q(x_1, \ldots, x_{n+1}) + c(x_1, \ldots, x_{n+1}) \tag{3}$$

where q and c are homogeneous of degree 2 and 3 respectively. Write $z_1 = x_1/x_0, \ldots, z_{n+1} = x_{n+1}/x_0$. We can take $z_1, \ldots, z_n$ as local coordinates for X at P, and then X is given by the local equation

$$z_{n+1} = q'(z_1, \ldots, z_n) + (\text{higher order terms}).$$

Here $q'(z_1, \ldots, z_n) = -q(z_1, \ldots, z_n, 0)$. The second fundamental form $\Pi_P X$ is the quadratic form $q'(dz_1, \ldots, dz_n)$ (up to scaling). We shall only be concerned with the rank and signature of $\Pi_P X$, which are precisely the rank and signature of $q(x_1, \ldots, x_n, 0)$ and so we will take this as the second fundamental form. We may therefore view it as the restriction of q to $T_P X$. The following follows easily from the above description and the implicit function theorem.

Lemma 4.2. *Suppose $\Pi_P X$ has full rank n.*

(i) *If $\Pi_P X$ is definite then there is an open neighbourhood $U \subseteq X(\mathbb{R})$ such that $U \cap X_P(\mathbb{R}) = \{P\}$.*

(ii) *If $\Pi_P X$ is indefinite then for every open neighbourhood $U \subseteq X(\mathbb{R})$ of P the intersection contains a real manifold of dimension $n-1$.*

Lemma 4.3. *There is a non-empty subset $U_1 \subseteq X(\mathbb{R})$, open in the real topology, such that for $P \in U_1$ the second fundamental form $\Pi_P X$ is indefinite of full rank.*

Proof. A theorem of Landsberg [1994, Theorem 6.1] asserts that at a general point on smooth hypersurface of degree ≥ 2, the second fundamental form has full rank. Thus there is a Zariski open $\mathcal{U} \subset X$ such that $\Pi_P X$ has full rank for $P \in \mathcal{U}$.

A straightforward application of Bertini's Theorem shows the existence of a real 3-dimensional linear subvariety $\Lambda \subset \mathbb{P}^{n+1}$ such that $X' = \Lambda \cap X$ is a smooth real cubic surface. A classical theorem of Schläfli asserts that the number of real lines on smooth real cubic surface is either 3, 7, 15 or 27. Let $\ell \subset \Lambda \cap X$ be a real line. Recall that a point on a smooth real surface is called hyperbolic if the Gaussian curvature at the point is negative. By [Siksek 2012, Lemma 2.2] all but at most two points of $\ell(\mathbb{R})$ are hyperbolic for X'. Let $Q \in \ell(\mathbb{R})$ be a hyperbolic point for X'.

The determinant of the second fundamental form $\Pi_Q X'$ is the Gaussian curvature of X' at Q, which is negative. It follows that the binary quadratic form $\Pi_Q X'$ is indefinite. Now $\Pi_Q X'$ is the restriction of $\Pi_Q X$ to $T_Q X'$ and so $\Pi_Q X$ is indefinite. Thus there is a neighbourhood $V \subseteq X(\mathbb{R})$ of Q, open in the real topology, such that $\Pi_P X$ is indefinite for $P \in V$. Now V is necessarily Zariski-dense in X. Thus $V \cap \mathcal{U}(\mathbb{R})$ is non-empty (as well as being open in the real topology). The proof is complete upon letting $U_1 = V \cap \mathcal{U}(\mathbb{R})$. □

The Hessian. Given $P \in X$, the Hessian of F evaluated at P is given by the $(n+2) \times (n+2)$ matrix

$$H_F(P) = \left(\frac{\partial^2 F}{\partial x_i \partial x_j}(P) \right)_{i,j=0,\dots,n+1}.$$

Of course the Hessian is well-defined up to multiplication by a non-zero scalar.

Lemma 4.4. *Let $P \in X$ and suppose $\Pi_P(X)$ has full rank n. Then $H_F(P)$ has full rank $n+2$.*

Proof. Starting from (3), an easy computation shows that the determinant of the Hessian at P is (up to sign) the determinant of $q(x_1, \dots, x_n, 0)$. □

Eckardt points. We call $P \in X$ an *Eckardt point* if $X_P := X \cap T_P X$ is a cone with vertex at P. Note that if $n = 2$ and P is an Eckardt point then X_P consists of three lines meeting at P; in this case $\Pi_P X$ vanishes identically.

For a proof of the following classical theorem see [Coskun and Starr 2009, Section 2].

Theorem 5. *The set of Eckardt points on X is finite.*

Components of a real cubic hypersurface. We summarize some well-known facts about components of real cubic hypersurfaces. Everything we need is actually contained in [Viro 1998, Section 4.3]. A smooth real cubic hypersurface has either one or two connected components. If it has two connected components then one of these is two-sided, and homeomorphic to S^n, and the other is one-sided and homeomorphic to $\mathbb{R}P^n$. If a line intersects the two-sided component then it intersects it in two points, and intersects the odd-sided component in one point.

Lemma 4.5. *Suppose $X(\mathbb{R})$ has two connected components. Then $\Pi_P X$ is definite of full rank for all P belonging to the two-sided component.*

Proof. Let P be a point on the two-sided component. Suppose $\Pi_P X$ is indefinite or not of full rank. Then there is a real line $\ell \subset T_P X$ (along which $\Pi_P X$ vanishes) that meets X with multiplicity ≥ 3 at P. As this is impossible for points on the two-sided component, we have a contradiction. □

5. Mordell–Weil generation: first steps

Proposition 5.1. *Let $P \in X(\mathbb{Q})$ be a non-Eckardt point. Then the set $X_P(\mathbb{Q})$ (considered as a subset of $X(\mathbb{Q})$) is contained in* Span(P).

The following lemma follows from the definitions.

Lemma 5.2. *Let $P \in X(\mathbb{Q})$ and let $Q \in X_P(\mathbb{Q})$ be distinct from P. Suppose the line ℓ joining P to Q is not contained in X. Then $Q \in$* Span(P).

For the proof of Proposition 5.1 it remains to show that $Q \in \mathrm{Span}(P)$ in the case $\ell \subset X$. For $n = 2$ this is [Siksek 2012, Lemma 3.2], so we suppose for the remainder of this section that $n \geq 3$.

Lemma 5.3. *Any hyperplane section of X is absolutely irreducible.*

Proof. Let $L = 0$ be a hyperplane such that $X \cap \{L = 0\}$ is absolutely reducible. Then we can write $F = LQ + L'Q'$ where L, L' are homogeneous linear, and Q, Q' are homogeneous quadratic. As $n \geq 3$, the variety $L = L' = Q = Q' = 0$ has a point $R \in \mathbb{P}^{n+1}$. It follows that R is a singular on X giving a contradiction. □

Lemma 5.4. *Let $P \in X(\mathbb{Q})$ be a non-Eckardt point. Let $Q \in X_P(\mathbb{Q})$ with $T_Q X \neq T_P X$. Then $Q \in$* Span(P).

Proof. Let $\mathcal{W} \subseteq X_P$ be the subvariety consisting of lines through P contained in X_P. As P is a non-Eckardt point, $\mathcal{W}$ is a proper subvariety. Moreover, by Lemma 5.3, the tangent plane section X_P is irreducible, and so $\dim(\mathcal{W}) < \dim(X_P)$. Let $\mathcal{U} = X_P - \mathcal{W}$ which is Zariski dense in X_P.

Let $\mathcal{V} := X_P \setminus (X_P \cap T_Q)$. As $T_Q \neq T_P$, this is a dense open subset of X_P. Let $\iota : \mathcal{V} \to \mathcal{V}$ be the involution given as follows. If $R \in \mathcal{V}$ we join R to Q by the line $\ell_{R,Q}$ and we let $\iota(R)$ be the third point of intersection of this line with X. We note that $\ell_{R,Q} \not\subset X$, since otherwise it will be contained in $T_Q X$ by Lemma 4.1. Now $(\mathcal{V} \cap \mathcal{U}) \cap \iota(\mathcal{V} \cap \mathcal{U})$ is a Zariski dense subset of the rational variety X_P. This dense subset must contain a rational point R. Then $R, \iota(R) \notin \mathcal{W}$ and so $R, \iota(R) \in \mathrm{Span}(P)$ by Lemma 5.2. Finally the line joining R with $\iota(R)$ passes through Q and is not contained in X. Thus $Q \in \mathrm{Span}(R)$. □

Proof of Proposition 5.1. Let $Q \in X_P(\mathbb{Q})$. We would like to show that $Q \in \mathrm{Span}(P)$. Thanks to Lemmas 5.2 and 5.4, we may suppose there is a $\mathbb{Q}$-line $\ell \subset X$ containing P, Q, and $T_Q X = T_P X$. Now the line ℓ contains at most finitely many Eckardt points by Theorem 5. Moreover, the Gauss map on a smooth hypersurface has finite fibres [Harris 1992, Lecture 15]. Thus there is a non-Eckardt $R \in \ell(\mathbb{Q})$ with $T_R X \neq T_P X$. It follows that $R \in \mathrm{Span}(P)$. Moreover, $Q \in \ell \subset T_R$ by Lemma 4.1 and so $Q \in \mathrm{Span}(R)$ (again by Lemma 5.4). This completes the proof. □

Lemma 5.5. *Suppose $n \geq 48$. Let $B \in X(\mathbb{Q})$ so that X_B does not contain points that are Eckardt for X. Suppose $D' \in X(\mathbb{R})$ such that $Y_{B,D'}$ has a smooth real point C'. Then there is a non-empty open $U \subseteq X(\mathbb{R})$ such that $U \cap X(\mathbb{Q}) \subseteq \mathrm{Span}(B)$.*

Proof. Take U to be as in Lemma 3.3. Let $D \in U \cap X(\mathbb{Q})$. By the conclusion of Lemma 3.3 we see that $Y_{B,D}$ has a rational point C. From the equations defining $Y_{B,D}$ in (2) we have that $C \in X_B(\mathbb{Q})$ and $D \in X_C(\mathbb{Q})$. Moreover, neither B nor C (both contained in X_B) are Eckardt points. Applying Proposition 5.1, we have $C \in \mathrm{Span}(B)$ and $D \in \mathrm{Span}(C)$ completing the proof. □

6. A smoothness criterion

Lemma 6.1. *Let $B \in X$, Let $C' \in X_B$ and $D' \in X_{C'}$. Suppose*

(i) $T_{C'}X \neq T_B X$;

(ii) $H_F(C')$ *has full rank, where H_F is the Hessian matrix*;

(iii) D' *does not belong to the line*

$$\{ (\lambda \nabla F(B) + \mu \nabla F(C')) \cdot H_F(C')^{-1} \quad : \quad (\lambda : \mu) \in \mathbb{P}^1 \}. \tag{4}$$

Then C' is a smooth point on the variety $Y_{B,D'} \subset \mathbb{P}^{n+1}$ given by (2).

Proof. As $C' \in X_B$ and $D' \in X_{C'}$ we see that $C' \in Y_{B,D'}$. We need to show that C' is a smooth point on $Y_{B,D'}$. Write

$$f(x_0, \ldots, x_{n+1}) = \nabla F(x_0, \ldots, x_{n+1}) \cdot D', \quad g = \nabla F(B) \cdot (x_0, \ldots, x_{n+1}).$$

To show that C' is smooth on $Y_{B,D'}$ it is enough to show that $\nabla F(C')$, $\nabla f(C')$ and $\nabla g(C')$ are linearly independent. A straightforward computation shows that

$$\nabla f(C') = D' \cdot H_F(C'), \quad \nabla g(C') = \nabla F(B).$$

Suppose

$$\varepsilon D' \cdot H_F(C') + \lambda \nabla F(B) + \mu \nabla F(C') = 0.$$

By assumptions (ii) and (iii) we see that $\varepsilon = 0$. However, $\nabla F(B)$ and $\nabla F(C')$ are linearly independent by assumption (i), and so $\lambda = \mu = 0$. □

7. Proof of Theorem 1

In this section $n \geq 48$.

Lemma 7.1. *There is $A \in X(\mathbb{Q})$ and a non-empty open $U \subseteq X(\mathbb{R})$ such that*:

(i) $U \cap X(\mathbb{Q}) \subseteq \mathrm{Span}(A)$;

(ii) $\mathrm{Span}(A)$ *contains at least one point in every connected component of* $X(\mathbb{R})$.

Proof. Suppose first that $X(\mathbb{R})$ is connected. Let $U_1 \subseteq X(\mathbb{R})$ be the non-empty open subset whose existence is guaranteed by Lemma 4.3: for every $P \in U_1$, the second fundamental form $\Pi_P X$ is indefinite of full rank. It follows from Theorem 5 that the set of points P with X_P containing an Eckardt point is a proper subset of X that is closed in the Zariski topology. Thus we may replace U_1 by a non-empty open set $U_2 \subseteq U_1$ such that for every $P \in U_2$, the subvariety X_P does not contain points that are Eckardt for X. Fix $B \in U_2 \cap X(\mathbb{Q})$ whose existence is guaranteed by Theorem 2. The hypersurface X is smooth of degree 3, and so the Gauss map $X \to X^*$ has finite fibres [Harris 1992, Lecture 15]. We can therefore take an open neighbourhood $U_3 \subseteq U_2$ of B such that for all $C' \in U_3$ with $C' \neq B$, we have $T_B X \neq T_{C'} X$. By Lemma 4.2, the intersection $U_3 \cap X_B(\mathbb{R})$ contains a real manifold of dimension $n-1$; choose $C' \in U_3 \cap X_B(\mathbb{R})$ with $C' \neq B$. As the second fundamental form has full rank on U_3, we see from Lemma 4.4 that $H_F(C')$ is of full rank $n+2$. Now again by Lemma 4.2, the intersection $U_3 \cap X_{C'}(\mathbb{R})$ contains a manifold of real dimension $n-1$, and so we can find $D' \in U_3 \cap X_{C'}(\mathbb{R})$ that avoids the line (4). The points B, C', D' satisfy the conditions of Lemma 6.1. Thus C' is a smooth point on $Y_{B,D'}$. By Lemma 5.5, there is a non-empty open U such that $U \cap X(\mathbb{Q}) \subseteq \operatorname{Span}(B)$. We simply take $A = B$, and the proof is complete in the case when $X(\mathbb{R})$ is connected.

Now suppose $X(\mathbb{R})$ has two connected components. Let $U_2 \subseteq X(\mathbb{R})$ be as above. From Lemma 4.5 we know that U_2 is contained in the one-sided component. Let $B' \in U_2$. Let ℓ' be a real line passing through B' and tangent to the two-sided component at a point A'. By Corollary 2.1, there is a point $A \in X(\mathbb{Q})$ belonging to the two-sided component and a line ℓ defined over $\mathbb{Q}$ such that $\ell \cdot X = 2A + B$ where $B \in U_2 \cap X(\mathbb{Q})$. Now $B \in \operatorname{Span}(A)$ and $\operatorname{Span}(A)$ contains points belonging to both components of $X(\mathbb{R})$. From the above argument there is a non-empty open $U \subseteq X(\mathbb{R})$ such that $U \cap X(\mathbb{Q}) \subseteq \operatorname{Span}(B) \subseteq \operatorname{Span}(A)$. □

Lemma 7.2. *Let $A \in X(\mathbb{Q})$ be as in Lemma 7.1. Then there is an open $W \subseteq X(\mathbb{R})$ such that $W \cap X(\mathbb{Q}) = \operatorname{Span}(A)$.*

Proof. Let U be as in Lemma 7.1. We may suppose $\operatorname{Span}(A) \not\subset U$, otherwise we simply take $W = U$ and there is nothing to prove. Let $P \in \operatorname{Span}(A)$ that does not belong to U. By Theorem 2, there is some $P' \in U \cap X(\mathbb{Q})$ such that $P' \notin T_P X$. Let ℓ be the line joining P to P'. The line ℓ is not contained in $T_P X$ and so, by Lemma 4.1, not contained in X. Let $P'' \in X(\mathbb{Q})$ be the third point of intersection of ℓ with X. Since $P \in \operatorname{Span}(A)$ and $P' \in U \cap X(\mathbb{Q}) \subseteq \operatorname{Span}(A)$, we have $P'' \in \operatorname{Span}(A)$. Observe that ℓ is not contained in the tangent plane of P'' (for otherwise ℓ would be contained in X). Now there is some non-empty open $U' \subset U$ containing P' that is disjoint from the tangent plane of P''. For a point $R \in U'$, let $\varphi(R)$ denote the third point of intersection of the (real) line joining R to P''. Then the map $\varphi : U' \to X(\mathbb{R})$ is continuous and injective. By the invariance of domain theorem [Bredon 1993, Corollary IV.19.9], the image $\varphi(U')$ is open. We shall let $W_P = \varphi(U')$. Clearly $P \in W_P$ and $W_P \cap X(\mathbb{Q}) \subseteq \mathrm{Span}(A)$. The lemma follows on taking

$$W = U \cup \bigcup_{P \in \mathrm{Span}(A) \setminus U} W_P. \qquad \square$$

Lemma 7.3. *Let W be as in Lemma 7.2, and write $\overline{W}$ for its closure. Then $\overline{W}$ is closed under secant operations: if $P, Q \in \overline{W}$ are distinct, and if the line ℓ joining them is not contained in X, then $R \in \overline{W}$ where $\ell \cdot X = P + Q + R$.*

Proof. By Theorem 2 there exist $\{P_k\}, \{Q_k\} \subset W \cap X(\mathbb{Q})$, with $P_k \neq Q_k$, that converge respectively to P, Q. Write $F \subset \mathbb{G}(n+1, 1)$ for the Fano scheme of lines on X. Then the real points of F are closed in $\mathbb{G}(n+1, 1)(\mathbb{R})$. As $\ell \notin F(\mathbb{R})$, we see for large enough k that the line $\ell_k/\mathbb{Q}$ joining P_k, Q_k is not contained in X. Let $\ell_k \cdot X = P_k + Q_k + R_k$. Then $\{R_k\}$ converges to R. Moreover, P_k, $Q_k \in W \cap X(\mathbb{Q}) \subseteq \mathrm{Span}(A)$. Hence $R_k \in \mathrm{Span}(A) \subset W$ and so $R \in \overline{W}$. $\square$

Lemma 7.4. *Let A, W be as above. Then $\overline{W} = X(\mathbb{R})$.*

Proof. We claim that $\overline{W}$ is open. From that it follows that $\overline{W}$ is a union of connected components of $X(\mathbb{R})$. As $\mathrm{Span}(A) \subset \overline{W}$ contains points from every component, the lemma follows from the claim.

To prove the claim we mimic the argument in the proof of Lemma 7.2. Let $P \in \overline{W}$. Let $P' \in W$ such that $P' \notin T_P X$, and let ℓ be the line joining P to P'. As $\overline{W}$ is closed under secant operations, $P'' \in \overline{W}$ where $\ell \cdot X = P + P' + P''$. Now there is some non-empty open $W' \subset W$ containing P' that is disjoint from the tangent plane of P''. For a point $R \in W'$, let $\varphi(R)$ denote the third point of intersection of the (real) line joining R to P''. Then the map $\varphi : W' \to X(\mathbb{R})$ is continuous and injective, and thus the image $\varphi(W')$ is open. Clearly $\varphi(W')$ contains P and is contained in $\overline{W}$ (as the latter is closed under secant operations). $\square$

Proof of Theorem 1. Let A, W be as above. In particular, $\mathrm{Span}(A) = W \cap X(\mathbb{Q})$ and $\overline{W} = X(\mathbb{R})$. We write $\partial W = X(\mathbb{R}) \setminus W$. We note that ∂W is the complement of an open dense set, and therefore nowhere dense.

We want to show that $X(\mathbb{Q}) = \mathrm{Span}(A)$. Let $P \in X(\mathbb{Q})$. Then there is a Zariski open $\mathcal{U} \subset X$ and an involution $\iota : \mathcal{U} \to \mathcal{U}$ that sends $R \in \mathcal{U}$ to the third point of the line joining R to P. Choose $R \in \mathcal{U}(\mathbb{R}) \cap X(\mathbb{Q})$ such that $R \notin \partial W \cup \iota(\partial W)$. Then $R, \iota(R) \in \mathrm{Span}\, A$ and so $P \in \mathrm{Span}(A)$. $\square$

References

[Bredon 1993] G. E. Bredon, *Topology and geometry*, Graduate Texts in Mathematics **139**, Springer, New York, 1993. MR Zbl

[Browning et al. 2015] T. D. Browning, R. Dietmann, and D. R. Heath-Brown, "Rational points on intersections of cubic and quadric hypersurfaces", *J. Inst. Math. Jussieu* **14**:4 (2015), 703–749. Corrigendum in **14**:4 (2015), 751. MR Zbl

[Coskun and Starr 2009] I. Coskun and J. Starr, "Rational curves on smooth cubic hypersurfaces", *Int. Math. Res. Not.* **2009**:24 (2009), 4626–4641. MR Zbl

[Griffiths and Harris 1979] P. Griffiths and J. Harris, "Algebraic geometry and local differential geometry", *Ann. Sci. École Norm. Sup.* (4) **12**:3 (1979), 355–452. MR Zbl

[Harris 1992] J. Harris, *Algebraic geometry: a first course*, Graduate Texts in Mathematics **133**, Springer, New York, 1992. MR Zbl

[Kanevsky and Manin 2001] D. Kanevsky and Y. Manin, "Composition of points and the Mordell–Weil problem for cubic surfaces", pp. 199–219 in *Rational points on algebraic varieties*, edited by E. Peyre and Y. Tschinkel, Progr. Math. **199**, Birkhäuser, Basel, 2001. MR Zbl

[Landsberg 1994] J. M. Landsberg, "On second fundamental forms of projective varieties", *Invent. Math.* **117**:2 (1994), 303–315. MR Zbl

[Manin 1974] Y. I. Manin, *Cubic forms: algebra, geometry, arithmetic*, North-Holland Mathematical Library **4**, North-Holland, Amsterdam, 1974. Second edition, 1986. MR Zbl

[Manin 1997] Y. I. Manin, "Mordell–Weil problem for cubic surfaces", pp. 313–318 in *Advances in mathematical sciences: CRM's 25 years* (Montreal, 1994), CRM Proc. Lecture Notes **11**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[Manin 2012] Y. I. Manin, "Combinatorial cubic surfaces and reconstruction theorems", pp. 99–118 in *Zeta functions in algebra and geometry*, edited by A. Campillo et al., Contemp. Math. **566**, Amer. Math. Soc., Providence, RI, 2012. MR

[Segre 1943] B. Segre, "A note on arithmetical properties of cubic surfaces", *J. London Math. Soc* **18** (1943), 24–31. MR Zbl

[Siksek 2012] S. Siksek, "On the number of Mordell–Weil generators for cubic surfaces", *J. Number Theory* **132**:11 (2012), 2610–2629. MR Zbl

[Skinner 1997] C. M. Skinner, "Forms over number fields and weak approximation", *Compositio Math.* **106**:1 (1997), 11–29. MR Zbl

[Stoer and Bulirsch 1980] J. Stoer and R. Bulirsch, *Introduction to numerical analysis*, Springer, 1980. MR Zbl

[Vioreanu 2009] B. G. Vioreanu, "Mordell–Weil problem for cubic surfaces, numerical evidence", pp. 223–240 in *Arithmetic geometry*, edited by H. Darmon et al., Clay Math. Proc. **8**, Amer. Math. Soc., Providence, RI, 2009. MR

[Viro 1998] O. Viro, "Mutual position of hypersurfaces in projective space", pp. 161–176 in *Geometry of differential equations*, edited by A. Khovanskii et al., Amer. Math. Soc. Transl. Ser. 2 **186**, Amer. Math. Soc., Providence, RI, 1998. MR Zbl

Communicated by Brian Conrad
Received 2016-10-10 Revised 2017-07-13 Accepted 2017-08-11

spapanik21@gmail.com *Mathematics Institute, University of Warwick, Coventry, United Kingdom*

s.siksek@warwick.ac.uk *Mathematics Institute, University of Warwick, Coventry, United Kingdom*

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 11 No. 8 2017

1937-0652(2017)11:8;1-D