DIOPHANTINE PROBLEMS AND p-ADIC PERIOD MAPPINGS

BRIAN LAWRENCE AND AKSHAY VENKATESH

ABSTRACT. We give an alternative proof of Faltings's theorem (Mordell's conjecture): a curve of genus at least two over a number field has finitely many rational points. Our argument utilizes the set-up of Faltings's original proof, but is in spirit closer to the methods of Chabauty and Kim: we replace the use of abelian varieties by a more detailed analysis of the variation of p-adic Galois representations in a family of algebraic varieties. The key inputs into this analysis are the comparison theorems of p-adic Hodge theory, and explicit topological computations of monodromy.

By the same methods we show that, in sufficiently large dimension and degree, the set of hypersurfaces in projective space, with good reduction away from a fixed set of primes, is contained in a proper Zariski-closed subset of the moduli space of all hypersurfaces. This uses in an essential way the Ax–Schanuel property for period mappings, recently established by Bakker and Tsimerman.

CONTENTS

1. INTRODUCTION

1.1. Let K be a number field. This paper has two main goals.

Firstly, we will give a new proof of the finiteness of K-rational points on a smooth projective K-curve of genus $\geqslant 2$. The proof is closely related to Faltings's proof [15], but is based on a closer study of the variation of p-adic Galois representations in a family; it makes no usage of techniques specific to abelian varieties.

Secondly, we give an application of the same methods to a higher-dimensional situation. Consider the family of degree-d hypersurfaces in $\mathbf{P}^n$ and let $F_{n,d}$ be the complement of the discriminant divisor in this family; we regard $F_{n,d}$ as a smooth $\mathbf{Z}$-scheme. For S a finite set of primes, points of $F_{n,d}(\mathbf{Z}[S^{-1}])$ correspond to proper smooth hypersurfaces of

degree d in $\mathbf{P}^n_{\mathbf{Z}[S^{-1}]}$. It is very reasonable to suppose that $F_{n,d}(\mathbf{Z}[S^{-1}])$ is finite modulo the action of $\mathrm{GL}_{n+1}(\mathbf{Z}[S^{-1}])$ for $d \geqslant 3$ and all n. We shall show at least that, if $n \geqslant n_0$ and $d \geqslant d_0(n)$, then $F_{n,d}(\mathbf{Z}[S^{-1}])$ is contained in a proper Zariski closed subset of $F_{n,d}$ (i.e., there exists a proper $\mathbf{Q}$-subvariety of the generic fiber $(F_{n,d})_{\mathbf{Q}}$ whose rational points contain $F_{n,d}(\mathbf{Z}[S^{-1}])$). To prove this higher-dimensional result, we use a very recent theorem of Bakker and Tsimerman, the Ax–Schanuel theorem for period mappings.

We can obtain a still stronger theorem along a subvariety of $F_{n,d}$ if one has control over monodromy. Namely, if $F^*_{n,d} \subset (F_{n,d})_{\mathbf{Q}}$ is the Zariski closure of integral points, our result actually implies that the Zariski closure of monodromy for the universal family of hypersurfaces must drop over each component of $F^*_{n,d}$. It is possible that this imposes a stronger codimension condition on $F^*_{n,d}$ than simply "proper" but we do not know for sure.

Note that, without the result of Bakker and Tsimerman, one can *still* prove that $F_{n,d}(\mathbf{Z}[S^{-1}])$ lies in a proper $\mathbf{Q}_p$-analytic subvariety of $F_{n,d}(\mathbf{Q}_p)$, but one cannot prove the second statement about $F^*_{n,d}$.

A simple toy case to illustrate the methods is given by the S-unit equation, which we analyze in §4.

1.2. **Outline of the proof.**

Consider a smooth projective family $X \to Y$ over K, where Y is itself a smooth K-variety; we suppose this extends to a family $\pi : \mathcal{X} \to \mathcal{Y}$ over the ring $\mathcal{O}$ of S-integers of K, for some finite set S of places of K (containing all the archimedean places).

For $y \in Y(K)$ call X_y the fiber over y. We want to bound $\mathcal{Y}(\mathcal{O})$, making use of the fact that, if $y \in Y(K)$ extends to $\mathcal{Y}(\mathcal{O})$, then X_y admits a smooth proper model over $\mathcal{O}$. That one can thus reduce Mordell's conjecture to finiteness results for varieties with good reduction was observed by Parshin [31] and then used by Faltings in his proof of the Mordell conjecture [15].

Choosing a rational prime p that is unramified in K and not below any prime of S, write ρ_y for the Galois representation of $G_K = \mathrm{Gal}(\overline{K}/K)$ on the p-adic geometric étale cohomology of X_y, i.e. $H^*_{\mathrm{et}}(X_y \times_K \bar{K}, \mathbf{Q}_p)$. As observed by Faltings, one deduces from Hermite–Minkowski finiteness that, as y varies through $\mathcal{Y}(\mathcal{O})$, there are only finitely many possibilities for the semisimplification of the G_K-representation ρ_y (denoted by ρ_y^{ss}).

We seek to use the fact that, for v a place of K above p, one can understand the restriction $\rho_{y,v}$ of ρ_y to G_{K_v} via p-adic Hodge theory. In the Mordell case, when Y is a projective curve, our argument proceeds by showing that *both* of the following statements hold for suitable choice of X and v:

> (*) The representation ρ_y is semisimple for all but finitely many $y \in Y(K)$, *and* the map

$$(1.1) \qquad y \in Y(K) \longrightarrow \text{isomorphism class of } \rho_{y,v}$$

> has finite fibers.

Faltings proves much stronger statements when X is an abelian scheme over Y, using a remarkable argument with heights: every ρ_y is semisimple *and* ρ_y determines X_y up to isogeny. Our approach gives less, but it gives results in other cases too, such as the hypersurface family discussed above. However, in that setting, the issue of semisimplicity proves harder to control, and what we prove instead is the following hybrid of the two statements in (*): the map

$$(1.2) \qquad y \in Y(K) \longrightarrow \text{restriction of } \rho_y^{\mathrm{ss}} \text{ to } G_{K_v}$$

considered as a mapping from $Y(K)$ to isomorphism classes of G_{K_v}-representations, has fibers that are not Zariski dense. (It is crucial, in the above equation, that we semisimplify ρ_y as a global Galois representation and *then* restrict to G_{K_v}.)

For the remainder of the current §1.2, we will explain (1.1) in more detail.

Our analysis uses p-adic Hodge theory. However we make no use of p-adic Hodge theory in families: we need only the statements over a local field. Under the correspondence of p-adic Hodge theory, the restricted representation $\rho_{y,v}$ corresponds to a filtered ϕ-module, namely the de Rham cohomology of X_y over K_v equipped with its Hodge filtration and a semilinear Frobenius map. The variation of this filtration is described by a period mapping; in this setting, this is a K_v-analytic mapping

(1.3)   residue disk in $Y(K_v) \longrightarrow K_v$-points of a flag variety,

Therefore, the variation of the p-adic representation $\rho_{y,v}$ with y is controlled by (1.3). The basic, and very naive, "hope" of the proof is that injectivity of the period map (1.3) should force (1.1) to be injective.

However, (1.1) does *not* follow directly from injectivity of the period map, that is to say, from Torelli-type theorems.

Different filtrations on the underlying ϕ-module can give filtered ϕ-modules which are abstractly isomorphic, the isomorphism being given by a linear endomorphism commuting with ϕ. Hence, one needs to know not only that the period mapping (1.3) is injective, but that its image has *finite intersection* with an orbit on the period domain of the centralizer $Z(\phi)$ of ϕ. In other words, we must analyze a question of "exceptional intersections" between the image of a period map and an algebraic subvariety.

To illustrate how this is done, let us restrict to the case when Y is a curve. Assuming that we have shown that the $Z(\phi)$-orbit on the ambient flag variety is a proper subvariety, it will then be sufficient to show that the image of (1.3) is in fact Zariski dense. Then the intersection points between the image of (1.3) and a $Z(\phi)$-orbit amount to zeroes of a nonvanishing K_v-analytic function in a residue disc, and are therefore finite.

To check Zariski density, the crucial point is that one can verify the same statement for the *complex* period map:

(1.4)   universal cover of $Y(\mathbf{C}) \longrightarrow \mathbf{C}$-points of a flag variety

To pass between the p-adic and complex period maps, we use the fact that (in suitable coordinates), they satisfy the same differential equation coming from the Gauss Manin connection, and so have the same power series. This is a simple but crucial argument, given in Lemma 3.2. But – over the complex numbers – Zariski density can be verified by topological methods: (1.4) is now equivariant for an action of $\pi_1(Y)$, acting on the right according to the monodromy representation. It is enough to verify that the image of π_1 under the monodromy representation is sufficiently large. In the Mordell case, we show that the monodromy action of $\pi_1(Y)$ extends to a certain mapping class group, and we deduce large monodromy from the same assertion for the mapping class group (where we can use Dehn twists). This monodromy argument is related to computations of Looijenga [28], Grunewald, Larsen, Lubotzky, and Malestein [19], and Salter and Tshishiku [37].

If Y were not a curve, the argument above says only that the intersection of the image of (1.3) and a $Z(\phi)$-orbit is a proper K_v-analytic subvariety of $Y(K_v)$. One wants to get a proper *Zariski-closed subvariety* (for example, this permits one, in principle at least, to make an inductive argument on the dimension, although we do not try to do so here.) We obtain this only by appealing to a remarkable recent result of Bakker and Tsimerman, the

Ax–Schanuel theorem for period mappings: this is a very powerful and general statement about the transcendence of period mappings.

To summarize, we have outlined the strategy of the proof of (1.1). However, we have omitted one crucial ingredient needed in this proof, and also a crucial ingredient needed to get from (1.1) to Mordell:

(a) Showing that the centralizer $Z(\phi)$ of ϕ is not too large, and
(b) Controlling in some *a priori* way the extent to which ρ_y can fail to be semisimple.

We now discuss these issues in turn.

1.3. **Problem (a): controlling the centralizer of** ϕ. As we have explained, we need a method to ensure the centralizer of the crystalline Frobenius ϕ acting on the cohomology of a fiber X_y is not too large. For example, if $K_v = \mathbf{Q}_p$ so that ϕ is simply a $\mathbf{Q}_p$-linear map, we must certainly rule out the possibility that ϕ is a scalar!

This issue, that ϕ might have too large a centralizer and thus the map

$$(1.5) \qquad\qquad y \in Y(K_v) \longrightarrow \text{isomorphism class of } \rho_{y,v}$$

might fail to have finite fibers, already occurs in the simplest possible example. When analyzing the S-unit equation, it is natural to take $Y = \mathbf{P}^1 - \{0, 1, \infty\}$ and $X \to Y$ to be the Legendre family, so that X_t is the curve $y^2 = x(x-1)(x-t)$. Unfortunately (1.5) *fails*: for $t \in \mathbf{Z}_p$, if we write ρ_t for the representation of the Galois group $G_{\mathbf{Q}_p}$ on the (rational) Tate module of X_t, then ρ_t belongs to only finitely many isomorphism classes so long as the reduction $\bar{t} \in \mathbf{F}_p$ is not equal to 0 or 1.

Again we proceed in two different ways:

(i) In general, Frobenius is a *semilinear* operator on a vector space over an unramified extension L_w of $\mathbf{Q}_p$; semilinearity alone gives rise to a nontrivial bound (Lemma 2.1) on the size of its centralizer, which, in effect, becomes stronger as $[L_w : \mathbf{Q}_p]$ gets larger.

In the application to Mordell, it turns out that we can always put ourselves in a situation where $[L_w : \mathbf{Q}_p]$ is rather large. This forces the Frobenius centralizer to be small. We explain this at more length below.

(ii) In the case of hypersurfaces, we do not have a way to enlarge the base field as in (i). Our procedure is less satisfactory than in case (i), in that it gives much weaker results:

We are of course able to choose the prime p, and we choose it (via Chebotarev) so that the crystalline Frobenius at p has centralizer that is as small as possible. To do this, we fix an auxiliary prime ℓ, and first use the fact (from counting points over extensions of $\mathbf{F}_p$) that crystalline Frobenius at p has the same eigenvalues as Frobenius on p acting on ℓ-adic cohomology; thus it is enough to choose p such that the latter operator has small centralizer. One can do this via Chebotarev, given a lower bound on the image of the *global* Galois representation, and for this we again use some p-adic Hodge theory (cf. [39]). Another approach, by point-counting, is outlined in Lemma 12.1.

Let us explain point (i) above by example. In our analysis of the S-unit equation in §4, we replace the Legendre family instead by the family with fiber

$$X_t = \coprod_{z^{2^k} = t} \{y^2 = x(x-1)(x-z)\},$$

for a suitable large integer k. In our situation, the corresponding map $t \mapsto [\rho_t]$ will now only have finite fibers, at least on residue disks where $\bar{t}$ is not a square – an example of the importance of enlarging K_v.

Said differently, we have replaced the Legendre family $X \xrightarrow{\ell} \mathbf{P}^1 - \{0, 1, \infty\}$ with a family with the following composite structure:

$$X' \xrightarrow{\ell'} \mathbf{P}^1 - \{0, \mu_{2^k}, \infty\} \to \mathbf{P}^1 - \{0, 1, \infty\}$$

where the second map is given by $u \mapsto u^{2^k}$, and ℓ' is simply the restriction of the Legendre family over $\mathbf{P}^1 - \{0, \mu_{2^k}, \infty\}$. The composite defines a family over $\mathbf{P}^1 - \{0, 1, \infty\}$ with geometrically disconnected fibres, and this disconnectedness is, as we have just explained, to our advantage.

It turns out that the families introduced by Parshin (see [31, Proposition 9]), in his reduction of Mordell's conjecture to Shafarevich's conjecture, automatically have a similar structure. That is to say, if Y is a smooth projective curve, Parshin's families factorize as

$$X \to Y' \to Y,$$

where $Y' \to Y$ is finite étale and $X \to Y'$ is a relative curve.

There is in fact a lot of flexibility in this construction; in Parshin's original construction the covering $Y' \to Y$ is obtained by pulling back multiplication by 2 on the Jacobian, and as such each geometric fiber is a torsor under $H^1(Y_{\bar{K}}, \mu_2)$. We want to ensure that the Galois action on each fiber of $Y' \to Y$ has large image – with reference to the discussion above, this is what allows us to ensure that the auxiliary field L_w is of large degree. We use a variant where each fiber admits a G_K-equivariant map to $H^1(Y_{\bar{K}}, \mathbf{Z}/q\mathbf{Z})$ (for a suitable auxiliary prime q). The Weil pairing alone implies that the Galois action on this is nontrivial, and this (although very weak) is enough to run our argument.

1.4. **Problem (b): how to handle the failure of semisimplicity.** Let $y \in Y(K)$. The

local Galois representation $\rho_y|_{G_{K_v}}$ can certainly be very far from semisimple, and thus we cannot hope to use p-adic Hodge theory alone to constrain semisimplicity.

However, the Hodge weights of a global representation are highly constrained by purity (Lemma 2.9). This means, for example, that any global subrepresentation W of ρ_y corresponds, under p-adic Hodge theory, to a Frobenius-stable subspace $W_{\mathrm{dR}} \subset H^*_{\mathrm{dR}}(X_y \otimes_K K_v)$ whose Hodge filtration is numerically constrained. Now (assuming we have arranged that the Frobenius has small centralizer) there are not too many choices for a Frobenius-stable subspace; on the other hand, the Hodge filtration varies as y varies p-adically. Thus one can at least hope to show that such a "bad" W_{dR} exists only for finitely many $y \in Y(K_v)$. In this way we can hope to show that ρ_y is simple for all but finitely many y.

The purity argument is also reminiscent of an argument at the torsion level in Faltings's proof (the use of Raynaud's results on [15, p. 364]).

We use this argument both for Mordell's conjecture and for hypersurfaces (although for hypersurfaces we prove a much weaker result, just bounding from above the failure of semisimplicity). The linear algebra involved is fairly straightforward for curves (see *Claim 1* and its proof in Section 6) but becomes very unwieldy in the higher-dimensional case. To handle it in a reasonably compact way we use some combinatorics related to reductive groups (§11). However this argument is not very efficient and presumably gives results that are far from optimal.

1.5. **Effectivity; comparison with Chabauty–Kim and Faltings.** It is of interest to compare our method with that of Chabauty, and the nonabelian generalizations thereof due to Kim [24].

Let Y be a projective smooth curve over K with Jacobian J. Fix a finite place v. The classical method of Chabauty proceeds by considering $Y(K)$ as the intersection of global points $J(K)$ on the Jacobian and local points $Y(K_v)$ on the curve, inside $J(K_v)$. If the rank of $J(K)$ is less than the K_v-dimension of J (i.e. the genus of the curve) it is easy to see this intersection is finite.

We can reinterpret this cohomologically. Let T_p be the p-adic Tate module of J, where p is a prime below v. There is a Kummer map $J(K) \otimes \mathbf{Q}_p \to H^1(G_K, T_p)$ and we obtain a mapping

$$Y(K) \longrightarrow H^1(G_K, T_p) = \mathrm{Ext}^1(\mathrm{trivial}, T_p),$$

which, explicitly speaking, sends $y \in Y(K)$ to the extension between the trivial representation and T_p realized by cohomology of the punctured curve $H^1_{\mathrm{et}}(Y - \{y, y_0\})$ for a suitable basepoint y_0. By this discussion, and its local analogue, we get a diagram

(1.6)

$$
\begin{array}{ccccc}
Y(K) & \longrightarrow & J(K) & \xrightarrow{\ S\ } & \text{space of global Galois representations} \\
\downarrow & & \downarrow & & \downarrow \\
Y(K_v) & \longrightarrow & J(K_v) & \xrightarrow{\ S_v\ } & \text{space of local Galois representations.}
\end{array}
$$

(Here the global and local Galois representations are extensions of T_p by the trivial representation.) Kim generalizes this picture, replacing T_p by deeper quotients of $\pi_1(Y)$. The idea of p-adic period mappings also plays a key role in his work, see [24, p. 360], [25, p. 93], [26, Proposition 1.4]. The key difficulty to be overcome is to obtain control over the size of the space of global Galois representations (e.g. the rank of $J(K)$).

Our picture is very much the same: we have a map $y \to \rho_y$ from $Y(K)$ to global Galois representations. In the story just described ρ_y arises from the cohomology of an *open* variety – the curve Y punctured at y and an auxiliary point. In the situation of our paper, ρ_y will arise from the cohomology of a *smooth projective* variety – a covering of Y branched only at y.

What does this gain? Our global Galois representations are now pure and (presumably) semisimple. Therefore our space of global Galois representations should be extremely small. On the other hand, what we lose is that the map S_v is now no longer obviously injective.

Kim has remarked to one of us (A.V.) that it would be of interest to consider combining these methods in some way, in particular that one might replace the role of the pro-unipotent completion of $\pi_1(Y)$ in Kim's analysis by a *relative* completion.

We expect that our method of proof can be made algorithmic in the same sense as the method of Chabauty. For example, given a curve C as above, one would be able to "compute" a finite subset $S \subset C(K_v)$ which contains $C(K)$; "compute" means that there is an algorithm that will compute all the elements of S to a specified p-adic precision in a finite time. However, the resulting method is completely impractical, as we now explain.

Firstly, our argument relies on Faltings's finiteness lemma for Galois representations (Lemma 2.3) to give a finite list of possibilities for ρ_y^{ss}. We expect that Faltings's proof can easily be made algorithmic; but there may be very, very many such representations.

Secondly, we would need to explicitly compute the comparisons furnished by p-adic Hodge theory. For a given local Galois representation ρ_y^{ss}, we need to calculate to some

finite precision the filtered ϕ-module associated to it by the crystalline comparison isomorphism of p-adic Hodge theory. We expect that this should be possible, but we are not aware of any known algorithm to achieve this.

To conclude let us compare our method to Faltings's original proof. That proof gives much more than ours does: it gives the full Shafarevich and Tate conjectures for abelian varieties, as well as semisimplicity of the associated Galois representation. Our proof gives none of these; it gives nothing about the Tate conjecture, and (at least without further effort) it does not give the Shafarevich conjecture but only its restriction to a one-dimensional subfamily of moduli of abelian varieties. Moreover, our proof is also in some sense more elaborate, since it requires the use of tricks and delicate computations to avoid the various complications that we have described. Its only real advantage in the Mordell case seems to be that it is in principle algorithmic in the sense described above. In our view, the real gain of the method is the ability to apply it to families of higher-dimensional varieties. Our results about hypersurfaces are quite modest, but we regard them as a proof of concept for this idea.

1.6. **Structure of the paper.** §2 contains notation and preliminaries.

We suggest the reader start with §3 and §4 to get a sense of the argument.

§3 sets up the general formalism and the structure of the argument. We relate Galois representations to a p-adic period map using crystalline cohomology, and we connect the p-adic period map to a complex period map and monodromy. The section ends with Proposition 3.4, a preliminary form of our main result.

§4 gives a first application: a proof of the S-unit theorem, using a variant of the Legendre family. This is much simpler than the proof of Mordell and can be considered a "warm-up."

§§5 – 8 give the proof of the Mordell conjecture. §5 describes the strategy of the proof: we apply a certain refined version of Proposition 3.4, formulated as Proposition 5.3, to a specific family of varieties that we call the Kodaira–Parshin family. §6 is the proof of Proposition 5.3. In particular this is where we take advantage of "geometrically disconnected fibers"; the argument also deals with a technical issue relating to semisimplification. In §7 we introduce the Kodaira-Parshin family and §8 is purely topological: it computes the monodromy of the Kodaira–Parshin family.

§§9 – 12 study families of varieties of higher dimension. §9 introduces a recent transcendence result of Bakker and Tsimerman which is needed to study families over a higher-dimensional base. §10 proves the main result, Proposition 10.1, which shows that fibers of good reduction lie in a Zariski-closed subset of the base. The argument however invokes a "general position" result in linear algebra, Proposition 10.6, whose proof takes up §11. In §12 we suggest an alternative argument, not used in the rest of the paper, to bound the size of the Frobenius centralizer.

1.7. **Acknowledgements.** This paper owes, of course, a tremendous debt to the work of Faltings – indeed, all the main tools come from his work. Some of the ideas originated in a learning seminar run at Stanford University on Faltings's proof [15].

The 2017 Stanford PhD thesis [27] of B.L. contained an earlier version of the arguments of this paper. In particular, that thesis presented a proof of the Mordell conjecture conditional on an assumption about monodromy, and verified that assumption for a certain Kodaira–Parshin family in genus 2.

We thank Brian Conrad for many helpful conversations and suggestions. A.V. would like to thank Benjamin Bakker, Andrew Snowden and Jacob Tsimerman for interesting

2. NOTATION AND PREPARATORY RESULTS

We gather here some notation and some miscellaneous lemmas that we will use in the text. We suggest that the reader refer to this section only as necessary when reading the main text.

The following notation will be fixed throughout the paper.

- K a number field
- $\overline{K}$ a fixed algebraic closure of K
- $G_K = \mathrm{Gal}(\overline{K}/K)$ the absolute Galois group
- S a finite set of finite places of K containing all the archimedean places
- $\mathcal{O}_S$ the ring of S-integers
- $\mathcal{O} = \mathcal{O}_S$ when S is understood
- p a (rational) prime number such that no place of S lies above p
- K_w the completion of K at a prime w of $\mathcal{O}$
- $\overline{K}_w$ a fixed algebraic closure of K_w
- $\mathbf{F}_w$ the residue field at w
- q_w the cardinality of $\mathbf{F}_w$
- $\overline{\mathbf{F}}_w$ the residue field of $\overline{K}_w$, which is an algebraic closure of $\mathbf{F}_w$
- $\mathcal{O}_{(w)}$ the localization of $\mathcal{O}$ at w

By a G_K-set we mean a (discretely topologized) set with a continuous action of G_K.

For a variety X over a field E of characteristic zero, we denote by $H^*_{\mathrm{dR}}(X/E)$ the de Rham cohomology of $X \to \mathrm{Spec}(E)$. If $E' \supset E$ is a field extension, we denote by $H^*_{\mathrm{dR}}(X/E')$ the de Rham cohomology of the base-change $X_{E'}$, which is identified with $H_{\mathrm{dR}}(X/E) \otimes_E E'$.

For any scheme S, a *family over S* is an (arbitrary) S-scheme $\pi : Y \to S$. A *curve over S* is a family over S for which π is smooth and proper of relative dimension 1 and each geometric fiber is connected. (Note that we will also make use of "open" curves, for example in §4, but we will avoid using the word "curve" in that context.)

Let $E/\mathbf{Q}_p$ be a finite unramified extension of $\mathbf{Q}_p$, and σ the unique automorphism of E inducing the p-th power map on the residue field. By *ϕ-module (over E)* we will mean a pair (V, ϕ), with V a finite-dimensional E-vector space and $\phi : V \to V$ a map semilinear over σ. A *filtered ϕ-module* will be a triple $(V, \phi, F^i V)$ such that (V, ϕ) is a ϕ-module and $(F^i V)_i$ is a descending filtration on V. We demand that each $F^i V$ be an E-linear subspace

of V but require no compatibility with ϕ. Note that the filtered ϕ-modules arising from Galois representations via p-adic Hodge theory satisfy a further condition, admissibility, but we will make no use of it in this paper (see [18, Exposé III, §4.4] and [18, Exposé III, §5.3.3]).

2.1. **Linear algebra.**

Lemma 2.1. *Suppose that $\sigma : E \to E$ is a field automorphism of finite order e, with fixed field F. Let V be an E-vector space of dimension d, and $\phi : V \to V$ a σ-semilinear automorphism. Define the centralizer $\mathrm{Z}(\phi)$ of ϕ in the ring of E-linear endomorphisms of V via*

$$\mathrm{Z}(\phi) = \{f : V \to V \text{ an } E\text{-linear map}, \ f\phi = \phi f\};$$

it is an F-vector space. Then

$$\dim_F \mathrm{Z}(\phi) = \dim_E \mathrm{Z}(\phi^e),$$

where $\phi^e : V \to V$ is now E-linear. In particular, $\dim_F \mathrm{Z}(\phi) \leqslant (\dim_E V)^2$.

Proof. Let $\bar{F}$ be an algebraic closure of F, and let Σ be the set of F-embeddings $E \hookrightarrow \bar{F}$. Then $\bar{V} = V \otimes_F \bar{F}$ is a $E \otimes_F \bar{F} \simeq \bar{F}^\Sigma$-module, and splitting by idempotents of $E \otimes_F \bar{F}$ we get a decomposition

$$\bar{V} = \bigoplus_{\tau \in \Sigma} \bar{V}^\tau,$$

where $\bar{V}^\tau$ consists of $\bar{v} \in \bar{V}$ such that $e\bar{v} = \tau(e)\bar{v}$ for all $e \in E$. (Here the multiplication $e\bar{v}$ is for the E-module structure, and $\tau(e)\bar{v}$ for the $\bar{F}$-module structure, on $\bar{V}$.) Moreover, ϕ extends to an $\bar{F}$-linear endomorphism $\bar{\phi}$ of $\bar{V}$; this endomorphism carries $\bar{V}^\tau$ to $\bar{V}^{\tau\sigma^{-1}}$.

Fix $\tau_0 \in \Sigma$; then projection to the τ_0 factor induces an isomorphism

$$\mathrm{Z}(\bar{\phi}) \simeq \text{centralizer of } \bar{\phi}^e \text{ on } \bar{V}^{\tau_0}.$$

Now $(\bar{V}^{\tau_0}, \bar{\phi}^e)$ is obtained by base extension $\tau_0 : E \to \bar{F}$ from the E-linear map $\phi^e : V \to V$; in particular, the dimension of the centralizer on the right is the same as $\mathrm{Z}(\phi^e)$, whence the result. $\square$

2.2. **Semisimplicity.**

Lemma 2.2. *Let $H \leqslant G$ be a finite-index inclusion of groups, and let $\rho : H \to \mathrm{GL}_n(F)$ be a semisimple representation of the group H over the characteristic-zero field F. Then the induction $\rho^G = \mathrm{Ind}_H^G \rho$ is also semisimple.*

Proof. This follows readily from the fact that a representation ρ of G is semisimple if and only if its restriction to a finite-index normal subgroup $G_1 \leqslant G$ is semisimple: take G_1 to be the intersection of conjugates of H.

For "if" one can promote a splitting from G_1 to G by averaging; for "only if" we take an irreducible G-representation V, an irreducible G_1-subrepresentation $W \subset V$, and note that G-translates of W must span V, exhibiting $V|_{G_1}$ as a quotient of a semisimple module. $\square$

2.3. **Global Galois representations.**

Lemma 2.3. *(Faltings) Fix integers $w, d \geqslant 0$, and fix K and S as above. There are, up to conjugation, only finitely many semisimple Galois representations $\rho : G_K \to \mathrm{GL}_d(\mathbf{Q}_p)$ such that*

- *(a) ρ is unramified outside S, and*
- *(b) ρ is pure of weight w, i.e. for every prime $\wp \notin S$ the characteristic polynomial of Frobenius at $\wp$ has all roots algebraic, with complex absolute value $q_\wp^{w/2}$.*
- *(c) For $\wp$ as above the characteristic polynomial of Frobenius at $\wp$ has integer coefficients.*

Proof. This is a consequence of Hermite–Minkowski finiteness; see the proof of [15, Satz 5], or [47, V, Proposition 2.7]. □

We want to explain how to adapt this proof to a reductive target group. First we recall the notion of "semisimple" with general reductive target, and some allied notions.

Let K be a field of characteristic zero. First of all, recall that if $\mathbf{G}$ is a reductive algebraic group over K and $\rho : \Gamma \to \mathbf{G}(K)$ is a representation of the group Γ, there are natural notions of "irreducible" and "semisimple" adapted to $\mathbf{G}$, as described by Serre [43, 3.2]:

> the representation ρ is G-ir, or irreducible relative to G, if the image $\rho(\Gamma)$ is not contained in a proper parabolic subgroup $P \leqslant G$ defined over K.

For example, if $\mathbf{G}$ is an orthogonal or symplectic group, this assertion amounts to saying that there is no *isotropic* Γ-invariant subspace. Next

> the representation ρ is G-c.r., or completely reducible relative to G, if for any parabolic subgroup $P \leqslant G$ defined over K containing the image $\rho(\Gamma)$, there exists a Levi factor $L \leqslant P$, defined over K, which also contains this image.

We will also refer to G-c.r. as "semisimple" when the target group is clear. Let $\rho : \Gamma \to \mathbf{G}(K)$ be an arbitrary representation. Let P be a K-parabolic subgroup that contains the image of ρ and which is minimal for this property. Then the projection of ρ to a Levi factor $M \subset P$ is independent, up to G-conjugacy, of the choice of M; see [43, Proposition 3.3]. This resulting representation is called the semisimplification of ρ, relative to the ambient group $\mathbf{G}$, and will be denoted by ρ^{ss}. The Zariski closure of this semisimplification is a reductive group, at least for K in characteristic zero: see [43, Proposition 4.2].[1]

Later on we will use the following observation:

Lemma 2.4. *For any $\gamma \in \Gamma$, $\rho^{\mathrm{ss}}(\gamma)$ and $\rho(\gamma)$ have the same semisimple part up to conjugacy.*

Proof. Indeed, let P be as above, and factorize $P = MU$ into a Levi factor M and U the unipotent radical of P. We must prove that for $p = mu \in P(K)$, with $m \in M(K)$ and $u \in U(K)$, the semisimple parts of p and m are conjugate within P. To prove this take a commuting factorization $p = p^{ss}p^u$, and similarly for m. By functoriality, m^{ss} is the image of p^{ss}. We are reduced to the case of m and p semisimple:

(2.1)   a semisimple element $p = mu$ in $P(K)$ is $P(K)$-conjugate to m,

and clearly it is enough to be able to conjugate p into M.

---

[1] In §4 of [43] the assumption is stated that K is algebraically closed, but this is not used in the proof of Proposition 4.2. Alternately [3, Theorem 5.8] can be used to pass from K to $\bar{K}$.

The element p is contained in some maximal torus T ([7, 10.6,11.10]) which is contained in a Levi subgroup of P. However all Levi subgroups are conjugate under $U(K)$ [7, Proposition 20.5] we may therefore conjugate p into M as desired. $\qquad \square$

In passing we also record:

Lemma 2.5. *Suppose $P = MU$ is a parabolic subgroup of the reductive K-group G.*

Let $S \leqslant P$ be a K-torus, then S is conjugate under $U(K)$ to its projection to M.

In particular, $\chi : \mathbf{G}_m \to P$ be a character; then χ is conjugate, under $P(K)$, to its projection to M.

Proof. We may assume that S is a maximal torus, and then the claim follows from the argument above. $\qquad \square$

Faltings' finiteness theorem continues to apply in this context:

Lemma 2.6. *Let $\mathbf{G} \subset \mathrm{GL}_n$ be a reductive group, K a number field, S a finite set of places. Consider all representations*

$$\rho : G_K \longrightarrow \mathbf{G}(\mathbf{Q}_p)$$

which, when considered as representations into $\mathrm{GL}_n(\mathbf{Q}_p)$, satisfy conditions (a), (b), (c) of Lemma 2.3 (i.e. S-unramified, pure of weight w, integral).

Then there are only finitely many possibilities for the $\mathbf{G}(\mathbf{Q}_p)$-conjugacy class of ρ^{ss}.

Indeed, there are only finitely many possibilities up to $\mathbf{G}(\mathbf{Q}_p)$-conjugacy for pairs $(\mathbf{Q}, \rho : G_{\mathbf{Q}} \to \mathbf{L}_Q(\mathbf{Q}_p))$ where $\mathbf{Q}$ is a $\mathbf{Q}_p$-parabolic subgroup with Levi quotient $\mathbf{L}_Q$, the image of ρ is irreducible in $\mathbf{L}_Q$, and ρ again satisfies the conditions of Lemma 2.3.

Proof. Note first that for such ρ, the $\mathbf{G}$-semisimplification ρ^{ss} is also semisimple considered as a representation with target GL_n (since its Zariski closure is reductive, as noted above).

By Lemma 2.3 is enough to check that, for any fixed such ρ_0, there are only finitely many $\mathbf{G}(\mathbf{Q}_p)$-orbits on the set of $\mathrm{GL}_n(\mathbf{Q}_p)$-conjugates of ρ_0^{ss} with image in $\mathbf{G}$. Let $\mathbf{L}$ be the Zariski closure of the image of ρ_0^{ss}. It is a reductive $\mathbf{Q}_p$-subgroup of $\mathbf{G}$. Then for $g \in \mathrm{GL}_n(\mathbf{Q}_p)$ the image of $\mathrm{Ad}(g)\rho_0$ belongs to $\mathbf{G}$ if, and only if, $\mathrm{Ad}(g)\mathbf{L} \subset \mathbf{G}$. In other words, it is enough to verify that the set

$$\{g \in \mathbf{GL}_n(\mathbf{Q}_p) : \mathrm{Ad}(g)\mathbf{L} \subset \mathbf{G}\}$$

consists of finitely many double cosets under $(\mathbf{G}(\mathbf{Q}_p), \mathbf{L}(\mathbf{Q}_p))$, or equivalently finitely many $\mathbf{G}(\mathbf{Q}_p)$-orbits.

We may replace $\mathbf{L}$ by its connected component, and then it is enough to verify this assertion at the level of Lie algebras, i.e. to prove the same assertion for the set

$$\{g \in \mathrm{GL}_n(\mathbf{Q}_p) : \mathrm{Ad}(g)\mathfrak{l} \subset \mathfrak{g}\}$$

According to Richardson's theorem [35, Theorem 7.1] this forms finitely many $\mathbf{G}$ orbits over the algebraic closure $\overline{\mathbf{Q}_p}$. The result then follows from finiteness of the Galois cohomology $H^1(\mathbf{Q}_p, \mathbf{S})$ for any linear algebraic group $\mathbf{S}$ ([42, III §4, Theorem 4]).

To see the validity of the refinement, note that there are finitely many conjugacy classes of parabolic subgroups $\mathbf{P}$ defined over $\mathbf{Q}_p$, and for each such $\mathbf{P}$ there are – by what we just proved, applied to a Levi factor – only finitely many $\mathbf{P}(\mathbf{Q}_p)$-conjugacy classes of (pure of weight w, unramified outside S) irreducible representations $G_K \to \mathbf{L}_P(\mathbf{Q}_p)$. $\qquad \square$

2.4. **Friendly places.** For our later applications it is convenient to have available a class of "friendly" places of a number field K at which the local behavior of homomorphisms $G_K \to \mathbf{Q}_p^*$ is particularly simple. (Actually, in our applications, it would be enough to do

this for $K = \mathbf{Q}$, for which everything is quite straightforward, and to always use Lemma 2.10 with $K = \mathbf{Q}$. However, it makes our arguments a little easier to write to have friendly places available for a general number field K).

First we recall some structural theory [41, II.3.3]. Let $\mathcal{C} \subset G_{\mathbf{Q}} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ be the conjugacy class of complex conjugation, and let $H^+ = \langle \mathcal{C} \rangle$, the normal subgroup generated by $\mathcal{C}$; there is a unique nontrivial homomorphism $H^+ \to \{\pm 1\}$ and we let H be its kernel. A subfield $K \subset \overline{\mathbf{Q}}$ is totally real if and only if it is fixed by H^+. It is CM if and only if it is fixed by H but not H^+.

For an arbitrary number field $K \subset \overline{\mathbf{Q}}$ let E and E^+ be, respectively, the subfields of K defined by fixed fields of $G_K \cdot H$ and $G_K \cdot H^+$, respectively (where G_K is the Galois group of $\overline{\mathbf{Q}}$ over K). Then E^+ is the largest totally real subfield of K, and either $E^+ = E$ is totally real, *or* E is CM and is the largest CM subfield of K.

Definition 2.7. *(Friendly places). Let K be a number field.*

- *If K has a CM subfield, then let E be its maximal CM subfield and E^+ the maximal totally real subfield of E. In this case, we say that a place v of K is* friendly *if it is unramified over $\mathbf{Q}$, and it lies above a place of E^+ that is inert in E.*
- *If K has no CM subfield, any place v of K which is unramified over $\mathbf{Q}$ will be understood to be friendly.*

Clearly, infinitely many friendly places exist; however, if K has a CM subfield, they have Dirichlet density 0.

Consider, now, a continuous character $\eta : \mathrm{Gal}(\overline{K}/K) \longrightarrow \mathbf{Q}_p^*$, ramified at only finitely many places; by class field theory it corresponds to a homomorphism $\mathbf{A}_K^*/K^* \to \mathbf{Q}_p^*$. In particular, its restriction to places above p gives rise to a homomorphism $\eta_p : (K \otimes \mathbf{Q}_p)^* \longrightarrow \mathbf{Q}_p^*$. As usual, we say this is locally algebraic if it agrees, in a neighbourhood of the identity, with the $\mathbf{Q}_p$-points of an algebraic homomorphism $\mathrm{Res}_{(K \otimes \mathbf{Q}_p)/\mathbf{Q}_p} \mathbf{G}_m \longrightarrow \mathbf{G}_m$ of $\mathbf{Q}_p$-algebraic groups, cf. [41, Chapter III]. This condition is implied by being Hodge–Tate at primes above p, by a theorem of Tate [41, Chapter III, Appendix]. Moreover, since η is finitely ramified, it follows that η_p is trivial on a finite-index subgroup of the units $\mathcal{O}_K^*$, embedded into $(K \otimes \mathbf{Q}_p)^*$.

For such η, we say that η is pure of weight w when it satisfies the condition explained in Lemma 2.3.

Lemma 2.8. *Let v be any friendly place of K, lying above the prime p of $\mathbf{Q}$. For any continuous character $\eta : \mathrm{Gal}(\overline{K}/K) \longrightarrow \mathbf{Q}_p^*$, ramified at only finitely many places, pure of weight w, and locally algebraic at each prime above p, one has*

$$\eta^2|_{K_v^*} = \chi \cdot \mathrm{Norm}_{K_v/\mathbf{Q}_p}^w,$$

where χ has finite order. In particular, w is even and the Hodge–Tate weight of η at the place v equals $w/2$.

In other words, the restriction of globally pure characters to friendly places is of a standard form. Note that if the coefficients are enlarged from $\mathbf{Q}_p^*$ to $\mathbf{Q}_{p^2}^*$, the statement above is no longer true; an example is given by the idele class character associated to a CM elliptic curve.

The proof of this result is routine. The key point is due to Artin and Weil: an algebraic Hecke character factors through the norm map to the maximal CM subfield.

Proof. Being locally algebraic, η gives rise to an algebraic character of $\mathrm{Res}_{K/\mathbf{Q}} \mathbf{G}_m$, which is trivial on a finite-index subgroup of $\mathcal{O}^*$. Said differently, we obtain a $\mathbf{Q}_p$-rational character $\mathbf{S} \longrightarrow \mathbf{G}_m$ of the Serre torus $\mathbf{S}$; we will denote this also by η. (Note that η is forced

to be $\mathbf{Q}_p$-rational since it carries $\mathbf{S}(\mathbf{Q}_p)$ into $\mathbf{Q}_p^*$). Here $\mathbf{S}$ is the quotient of $\mathrm{Res}_{K/\mathbf{Q}}\mathbf{G}_m$ by the Zariski closure of (a sufficiently deep finite-index subgroup of) the units. Because of the purity assertion, if $\lambda \in K^*$ is a unit at all ramified primes for η, then $\eta(\lambda)$ is an algebraic number all of whose conjugates have absolute value $\mathrm{N}_{K/\mathbf{Q}}(\lambda)^{w/2}$.

The structure of this torus was in effect computed by Weil [45], and in detail by Serre: If K admits no CM subfield, then the norm map $\mathbf{S} \to \mathbf{G}_m$ is in fact an isogeny. So η is (up to finite order) the norm raised to the power $w/2$. The result follows.

Thus we suppose that K has a CM subfield; now let E be the largest CM subfield of K, and let E^+ be the totally real subfield of E. Then the norm map $\mathbf{S} \to \mathbf{S}_E$ is an isogeny; in other words, a suitable power η^k factors through the norm from K to E. Therefore it is enough to prove the Lemma for $K = E$, replacing v by the place of E below it. In particular, by definition, v lies above an inert prime of E/E^+.

Now there is a norm map $\mathbf{S}_E \to \mathbf{G}_m$. Write $x \mapsto \bar{x}$ for the complex conjugation on E. The map $x \mapsto x/\bar{x}$, from E^* to E^*, is trivial on a finite-index subgroup of the units, and its image consists entirely of elements whose norm (to E^+) equals 1. Indeed for any $\mathbf{Q}$-algebra R the rule $x \mapsto x/\bar{x}$ defines a map $(E \otimes R)^* \to (E \otimes R)^*$, corresponding to a unique map of $\mathbf{Q}$-algebraic groups

$$\theta : \mathbf{S}_E \to (\mathrm{Res}_{E/\mathbf{Q}}\mathbf{G}_m)^1$$

where the superscript 1 denotes the kernel of the norm to E^+. Together with the norm map this gives an isogeny $\mathbf{S}_E \longrightarrow \mathbf{G}_m \times (\mathrm{Res}_{E/\mathbf{Q}}\mathbf{G}_m)^1$. Raising the character η to a suitable power we can suppose that it factors through the right-hand side; twisting it by a power of the cyclotomic character, we can arrange that it is trivial on the $\mathbf{G}_m$ factor.

In other words, we are reduced to checking the case where η factors through θ. Now the weights of $x \mapsto \eta(x)$ and $x \mapsto \eta(\bar{x})$ coincide, but their product is trivial; so the weight of η is zero. Also η is trivial on E_v^*: consider

$$E_v^* \subset (E \otimes \mathbf{Q}_p)^* \to \mathbf{S}(\mathbf{Q}_p) \xrightarrow{\theta} (E \otimes \mathbf{Q}_p)^1 = \left(\prod_{w|p} E_w^*\right)^1.$$

The image of E_v^* is contained inside $\{y \in E_v^* : y\bar{y} = 1\}$; this is contained in a $\mathbf{Q}_p$-anisotropic subtorus of $(\mathrm{Res}_{E/\mathbf{Q}}\mathbf{G}_m)^1$. Therefore, any $\mathbf{Q}_p$-rational character of $(\mathrm{Res}_{E/\mathbf{Q}}\mathbf{G}_m)^1$ is trivial upon pullback to E_v^*. This is exactly what we wanted to prove (since, as we just saw, once η factors through θ its weight is zero). $\qquad\square$

2.5. **Reducibility of global Galois representations.** We now give some lemmas which limit the reducibility of a global pure Galois representation. The mechanism is as follows: purity passes to subrepresentations, and then leads to restrictions on the sub-Hodge structure.

For a decreasing filtration $F^\bullet V$ on a vector space V (with $F^0 V = V$) we define the *weight* of the filtration to be

$$(2.2) \qquad \mathrm{weight}_F(V) = \frac{\sum_{p \geqslant 0} p \dim \mathrm{gr}^p(V)}{\dim V},$$

where $\mathrm{gr}^p(V) = F^p(V)/F^{p+1}(V)$ is the associated graded. [2] For the other p-adic Hodge theory terms that appear in the following result, see [8, §6] or [18, Expose III].

---

[2]Here and in Section 10, the symbol p is used abusively to refer to the indexing on a Hodge filtration. We hope this will not cause confusion.

Lemma 2.9. *Let K be a number field and v a friendly place. Let V be a Galois representation of G_K on a $\mathbf{Q}_p$-vector space which is crystalline at all primes above p, and pure of weight w.*

Let $V_{\mathrm{dR}} = (V \otimes_{\mathbf{Q}_p} B_{\mathrm{cris}})^{G_{K_v}}$ be the filtered[3] K_v-vector space that is associated to $\rho|_{K_v}$ by the p-adic Hodge theory functor $\underline{D}_{\mathrm{cris}}$ of [18, Expose III].

Then the weight of the Hodge filtration on V_{dR} equals $w/2$.

Proof. Apply Lemma 2.8 to $\det(V)$. $\qquad\square$

Lemma 2.10. *Let K be a number field, and $L \supset K$ a finite extension. Let $\rho : G_L \to \mathrm{GL}_n(\mathbf{Q}_p)$ be a representation of G_L that is crystalline at all primes above p, and pure of weight w; let $a_u(\rho)$ be the weight of the associated Hodge filtration at each such prime u. Then, for any friendly prime v of K above p,*

$$\sum_{u|v} [L_u : K_v] a_u(\rho) = [L : K]\frac{w}{2}.$$

Proof. We apply Lemma 2.9 to $\mathrm{Ind}_{G_L}^{G_K} \rho$ and to the place v. Applying the functor of p-adic Hodge theory to its restriction to K_v, we obtain

$$(\mathrm{Ind}_L^K \rho \otimes_{\mathbf{Q}_p} B_{\mathrm{dR}})^{G_{K_v}} \simeq \bigoplus_{u|v} (\rho \otimes_{\mathbf{Q}_p} B_{\mathrm{dR}})^{G_{L_u}}$$

(considered now as a filtered K_v-vector space), and its weight is therefore $\frac{\sum_{u|v}[L_u:K_v]a_u(\rho)}{[L:K]}$. $\qquad\square$

2.6. **The affine group** $\mathrm{Aff}(q)$. Let $q \geqslant 3$ be a prime number and let $\mathrm{Sym}(\mathbf{F}_q)$ be the symmetric group on the q elements of $\mathbf{F}_q$. Let $\mathrm{Aff}(q) \subseteq \mathrm{Sym}(\mathbf{F}_q)$ be the subgroup consisting of permutations of $\mathbf{F}_q$ of the form $x \mapsto ax + b$ where $a \in \mathbf{F}_q^*$ and $b \in \mathbf{F}_q$. Thus[4] $\mathrm{Aff}(q) \cong (\mathbf{F}_q)^+ \rtimes (\mathbf{F}_q)^*$; this group has important applications in the theory of qualifying examinations. We shall make extensive use of it as a Galois group for certain auxiliary coverings of curves.

Lemma 2.11. *For any $s \geqslant 1$ consider the map $f : \mathrm{Aff}(q)^{2s} \longrightarrow \mathbf{F}_q^+$ given by*

$$f : \mathbf{g} = (g_1, g_1', \cdots, g_s, g_s') \mapsto [g_1, g_1'] \cdot [g_2, g_2'] \cdot \cdots \cdot [g_s, g_s']$$

(here $[x, y]$ is the commutator $xyx^{-1}y^{-1}$). The image of the map

$$(2.3) \qquad \{\mathbf{g} \in \mathrm{Aff}(q)^{2s} : f(\mathbf{g}) \neq 0,\ \mathbf{g}\ generates\ \mathrm{Aff}(q)\} \to \left[\mathbf{F}_q^*\right]^{2s}$$

(sending each g_i to its image in the abelian quotient $\mathbf{F}_q^*$) consists precisely of those $(2s)$-tuples in $\mathbf{F}_q^*$ whose entries generate $\mathbf{F}_q^*$. The fiber above any point in the image has the same size.

Proof. Note that, for such a fiber to be nonempty, the element $\mathbf{y} = (y_1, y_1', \ldots, y_s, y_s')$ of the target must have the property that the y_i and y_i' generate $\mathbf{F}_q^*$. In this case, any preimage $\mathbf{g} \in \mathrm{Aff}(q)^{2s}$ with the property that $f(\mathbf{g}) \neq 0$ necessarily generates $\mathrm{Aff}(q)$. The fiber of $\mathrm{Aff}(q)^{2s}$ above $\mathbf{y}$ is (in obvious coordinates) an affine space over $\mathbf{F}_q$, and the map f is a nontrivial affine-linear map; each fiber thus has size $q^{2s-1}(q-1)$. $\qquad\square$

---

[3]Here, and in other contexts, we will write V_{dR} even though we are using the crystalline functor, because in our applications it will be helpful to think of it in terms of de Rham cohomology.

[4]We use $\mathbf{F}_q^+$ to denote the additive group $\mathbf{F}_q$.

2.7. **Symplectic groups.** Let K be a field of characteristic zero. As usual if V is a symplectic space over a field K, with nondegenerate alternating bilinear form $\langle -, - \rangle$, we write $\mathrm{Sp}(V)$ for the algebraic group of automorphisms of V preserving the bilinear form.

The following statement is an algebraic version of Goursat's lemma (cf. [34, Lemma 5.2.1]). One uses the fact that the Lie algebra $\mathfrak{sp}_V$ of $\mathrm{Sp}(V)$ is simple, and that all the automorphisms of $\mathfrak{sp}_V$ are inner.

Lemma 2.12. *Suppose G is an algebraic subgroup of $\mathrm{Sp}(V)^N$, satisfying the following conditions.*

- *For $1 \leqslant i \leqslant N$, the projection $\pi_i : G \to \mathrm{Sp}(V)$ onto the i-th factor is surjective.*
- *For $1 \leqslant i, j \leqslant N$, there exists $g \in G$ such that $\pi_i(g)$ and $\pi_j(g)$ are unipotent with fixed spaces of different dimensions.*

Then G is all of $\mathrm{Sp}(V)^N$.

Any unipotent element of $\mathrm{Sp}(V)$ whose fixed space has codimension 1 is of the form

$$(2.4) \qquad\qquad T_v^r : x \mapsto x + r\langle v, x \rangle v$$

for some $v \in V, r \in K$. We call T_v^r a transvection with center v, and write T_v for T_v^1.

Lemma 2.13. *Let V be a symplectic space over $\mathbb{Q}$. Suppose $v_1, v_2 \in V$ are linearly independent and satisfy*

$$\langle v_1, v_2 \rangle \neq 0.$$

The Zariski closure of the subgroup generated by T_{v_1}, T_{v_2} also contains T_v for every $v \in \mathrm{Span}(v_1, v_2)$.

Proof. The subgroup in question preserves the splitting $V = \langle v_1, v_2 \rangle \oplus \langle v_1, v_2 \rangle^\perp$, and so we reduce to the case that V is 2-dimensional. The statement then amounts to the fact that $\mathrm{SL}(2)$ is generated, as an algebraic group, by upper and lower triangular matrices. $\square$

Lemma 2.14. *Let V be a symplectic space over $\mathbb{Q}$. Let S be a set of vectors $v \in V$. Make a graph whose set of vertices are S, having an edge between v_1 and v_2 if and only if $\langle v_1, v_2 \rangle \neq 0$. If this graph is connected, then the Zariski closure of the group generated by the transvections T_v, for $v \in S$, contains T_w for any w in the span of S.*

Proof. We can assume S is finite, and then use induction on $|S|$, using Lemma 2.13 for the inductive step.

In detail: Suppose $S = S_0 \cup \{v\}$, with the graph on S_0 connected. By inductive hypothesis we obtain all transvections centered at vectors in $W := \mathrm{span}(S_0)$. It is enough to verify that the Zariski closure in question contains the transvection T_x for each vector x of the form $w + v$ ($w \in W$); this is so when $\langle w, v \rangle \neq 0$ by the prior Lemma. The condition $\langle w, v \rangle \neq 0$ defines a Zariski-dense subset of W and so we also get the remaining transvections T_x when $\langle w, v \rangle = 0$ in the Zariski closure of them. $\square$

3. FIBERS WITH GOOD REDUCTION IN A FAMILY

In this section we give a general criterion (Proposition 3.4) which controls, in a given family of smooth proper varieties, the collection of fibers that have good reduction outside a fixed set of primes. The Proposition simply translates (using p-adic Hodge theory) the finiteness statement of Lemma 2.3 into a restriction on the image of the period map.

3.1. **Basic notation.** We use notation $K, \mathcal{O}, \mathcal{O}_{(w)}, S, G_K, \mathbf{F}_w$ as in §2.

Let Y be a smooth K-variety, and $\pi : X \to Y$ a proper smooth morphism.

Suppose that this admits a good model over $\mathcal{O}$, i.e. it extends to a proper smooth morphism $\pi : \mathcal{X} \to \mathcal{Y}$ of smooth $\mathcal{O}$-schemes. Suppose, moreover, that all the cohomology sheaves $\mathbf{R}^q \pi_* \Omega^p_{\mathcal{X}/\mathcal{Y}}$ are sheaves of locally free $\mathcal{O}_Y$-modules, and that the same is true of the relative de Rham cohomology $\mathcal{H}^q = \mathbf{R}^q \pi_* \Omega^\bullet_{\mathcal{X}/\mathcal{Y}}$. There is no harm in these assumptions, because the sheaves in question are coherent $\mathcal{O}_Y$-modules which are free over the generic point of $\mathcal{O}$ [11, Theorem 5.5]; so the assumptions can always be achieved by possibly enlarging the set S of primes.

The generic fiber of $\mathcal{H}^q$ is equipped with the Gauss–Manin connection (by [23, Theorem 1]) and, again by enlarging S if necessary, we may suppose that this extends to a morphism

$$\text{(3.1)} \qquad\qquad \mathcal{H}^q \to \mathcal{H}^q \otimes \Omega^1_{\mathcal{Y}/\mathcal{O}}.$$

For any $y \in Y(K)$, we shall denote by $X_y = \pi^{-1}(y)$ the fiber of π above y; it is a smooth proper variety over K. Our goal in this section is to bound $\mathcal{Y}(\mathcal{O})$. We will do this by studying the p-adic properties of the Galois representation attached to X_y, for $y \in \mathcal{Y}(\mathcal{O}) \hookrightarrow Y(K)$. Fixing a degree $q \geqslant 0$, we denote by ρ_y the representation of the Galois group G_K on the étale cohomology group of $(X_y)_{\bar{K}}$:

$$\text{(3.2)} \qquad\qquad \rho_y : G_K \to \operatorname{Aut} H^q_{\text{et}}(X_y \times_K \bar{K}, \mathbf{Q}_p).$$

Fix an archimedean place $\iota : K \hookrightarrow \mathbf{C}$, and fix a finite place $v : K \hookrightarrow K_v$ satisfying:

- if p is the rational prime below v, then $p > 2$, and
- K_v is unramified over $\mathbf{Q}_p$, and
- no prime above p lies in S.

Fix $y_0 \in \mathcal{Y}(\mathcal{O})$. In what follows, we will analyze the set

$$\text{(3.3)} \qquad\qquad U := \{ y \in \mathcal{Y}(\mathcal{O}) : y \equiv y_0 \text{ modulo } v. \}$$

and give criteria for the finiteness of U in terms of the associated period map. Clearly if U is finite for each choice of y_0, then $\mathcal{Y}(\mathcal{O})$ is finite too.

Finally, we put

$$X_0 = \pi^{-1}(y_0)$$

to be the fiber above y_0.

3.2. **The cohomology at the basepoint** y_0**.** For any K-variety Z, we shall denote by $Z_{\mathbf{C}}$ its base change to $\mathbf{C}$ via ι, and by Z_{K_v} its base change to K_v via v.

Let

$$\text{(3.4)} \qquad\qquad V = H^q_{\text{dR}}(X_0/K).$$

Let $d = \dim_K V$. We will also denote by V_v and $V_{\mathbf{C}}$ the K_v- and $\mathbf{C}$-vector spaces obtained by $\otimes_K K_v$ or $\otimes_{(K,\iota)} \mathbf{C}$. Then $V_{\mathbf{C}}$ is naturally identified with the de Rham cohomology of the variety $X_{0,\mathbf{C}}$, which is also (by the comparison theorem) identified with the singular cohomology of $X_{0,\mathbf{C}}$ with complex coefficients:

$$V_{\mathbf{C}} \simeq H^q_{\text{sing}}(X_{0,\mathbf{C}}, \mathbf{C}).$$

In particular, monodromy defines a representation $\mu : \pi_1(Y_{\mathbf{C}}(\mathbf{C}), y_0) \longrightarrow \operatorname{GL}(V_{\mathbf{C}})$, whose Zariski closure we denote by Γ:

$$\text{(3.5)} \qquad\qquad \Gamma = \text{ Zariski closure of } \operatorname{image}(\mu),$$

an algebraic subgroup of $\mathrm{GL}(V_{\mathbf{C}})$. Note that both $V_{\mathbf{C}}$ and Γ depend on the choice of archimedean place ι, although this dependence is suppressed in our notation.

3.3. **The Gauss–Manin connection.** The connection (3.1) allows us to identify the cohomology of nearby fibers. This is true both for the K_v and $\mathbf{C}$ topologies. However, as we now discuss, both identifications can be described as the evaluation of a single power series with K coefficients, which is convergent both for the K_v and $\mathbf{C}$ topology.

Specifically, if we fix a local basis $\{v_1, \ldots, v_r\}$ for $\mathscr{H}^q$ in a neighborhood of some point of the scheme $\mathcal{Y}$, and write $\nabla v_i = \sum_j A_{ij} v_j$, where A_{ij} are sections of $\Omega^1_{\mathcal{Y}}$, then a local section $\sum f_i v_i$ is flat exactly when it solves the equation

$$(3.6) \qquad d(f_i) = -\sum_j A_{ji} f_j.$$

In particular, if $y_0 \in \mathcal{Y}(\mathcal{O})$ and the place v is as before, let $\overline{y_0} \in \mathcal{Y}(\mathbf{F}_v)$ be the reduction, and choose a system of parameters $p, z_1, \ldots, z_m \in \mathcal{O}_{\mathcal{Y}, \overline{y_0}}$ for the local ring of $\mathcal{Y}$ at $\overline{y_0}$; we may do this so that $(z_1, \ldots, z_m)$ generate the kernel of the morphism $\mathcal{O}_{\mathcal{Y}, \overline{y_0}} \to \mathcal{O}_{(v)}$ corresponding to y_0. The completed local ring $\widehat{\mathcal{O}}_{\mathcal{Y}, \overline{y_0}}$ at $\overline{y_0}$ is therefore identified with $\mathcal{O}_v[[z_1, \ldots, z_m]]$, and the image of $\mathcal{O}_{\mathcal{Y}, \overline{y_0}}$ in it is contained in $\mathcal{O}_{(v)}[[z_1, \ldots, z_m]]$.

Fix a basis $\{\overline{v}_1, \ldots, \overline{v}_r\}$ for $\mathscr{H}^q$ at $\overline{y_0}$, which we assume to be compatible with the Hodge filtration, i.e. each step of the Hodge filtration $F^i \mathscr{H}^q$ at $\overline{y_0}$ is spanned by a subset of $\{\overline{v}_i\}$. Then by lifting we obtain a similar basis $\{v_1, \ldots, v_r\}$ for $\mathscr{H}^q$ over the local ring $\mathcal{O}_{\mathcal{Y}, \overline{y_0}}$ of $\mathcal{Y}$ at $\overline{y_0}$. With respect to such a basis v_i, the coefficients A_{ij} of (3.6) are of the form $A_{ij} = \sum_{k=1}^m a_{ij,k} dz_k$, where $a_{ij,k} \in \mathcal{O}_{\mathcal{Y}, \overline{y_0}}$. In particular, the coefficients of $a_{ij,k}$, considered as formal power series in the z_i, lie in $\mathcal{O}_{(v)}$.

We may write down a formal solution to (3.6), where the f_i are given by formal power series in $K[[z_1, \ldots, z_m]]$. By direct computation we see that these are v-adically absolutely convergent for $|z_i|_v < |p|_v^{1/(p-1)}$ (where p is the residue characteristic of $\mathcal{O}_v$) and ι-adically absolutely convergent for sufficiently small $|z_i|_{\mathbf{C}}$.

By assumption, we have $p > 2$, and v is unramified above p. Thus we obtain an identification

$$(3.7) \qquad \mathrm{GM} : H^q_{\mathrm{dR}}(\mathcal{X}_{y_0}/K_v) \overset{\sim}{\to} H^q_{\mathrm{dR}}(\mathcal{X}_y/K_v)$$

whenever $y \in \mathcal{Y}(\mathcal{O}_v)$ satisfies $y \equiv y_0$ modulo v, and

$$(3.8) \qquad \mathrm{GM} : H^q_{\mathrm{dR}}(X_{y_0, \mathbf{C}}/\mathbf{C}) \overset{\sim}{\to} H^q_{\mathrm{dR}}(X_{y, \mathbf{C}}/\mathbf{C}),$$

when $y \in Y_{\mathbf{C}}(\mathbf{C})$ is sufficiently close to y_0. In the coordinates of the basis v_i fixed above, GM is given by an $r \times r$ matrix with entries

$$A_{ij}(z_1, \ldots, z_m) \in \mathcal{O}_{(v)}[[z_1, \ldots, z_m]],$$

convergent in the regions noted above.

The fiber over the $\mathcal{O}$-point y_0 of $\mathcal{Y}$ gives a smooth proper $\mathcal{O}$-model $\mathcal{X}_0$ for X_0. For $y \in Y(\mathcal{O}_v)$ with $y \equiv y_0$ modulo v, we have a commutative diagram

(3.9)
$$
\begin{array}{ccc}
V_v = H^q_{\mathrm{dR}}(X_{y_0}/K_v) & & \\
& \searrow{\scriptstyle\sim} & \\
\downarrow{\scriptstyle\text{GM}} & & H^q_{\mathrm{cris}}(\overline{\mathcal{X}_0}) \otimes_{\mathcal{O}_v} K_v. \\
& \nearrow{\scriptstyle\sim} & \\
H^q_{\mathrm{dR}}(X_y/K_v) & &
\end{array}
$$

where GM denotes the map induced by the Gauss–Manin connection, H^q_{cris} is the crystalline cohomology of $\overline{\mathcal{X}_0}$ (as a reference for crystalline cohomology, see [5, 6]), the diagonal arrows are the canonical identification [6, Corollary 7.4] of crystalline cohomology with the de Rham cohomology of a lift, and the commutativity of the diagram can be deduced from the results of [5, Chapter V] (see Proposition 3.6.4 and prior discussion).

This crystalline cohomology $V_v = H^q_{\mathrm{dR}}(X_0/K)$ is equipped with a Frobenius operator

$$\phi_v : V_v \longrightarrow V_v,$$

which is semilinear with respect to the Frobenius on the unramified extension $K_v/\mathbf{Q}_p$. By the isomorphisms of (3.9), this ϕ_v acts on $H^q_{\mathrm{dR}}(X_y/K_v)$ and $H^q_{\mathrm{dR}}(X_{y_0}/K_v)$ as well, *in a manner compatible with the map* GM.

3.4. **The period mappings in a neighbourhood of** y. Now $V = H^q_{\mathrm{dR}}(X_0/K)$ is equipped with a Hodge filtration:

(3.10) $$V = F^0 V \supset F^1 V \supset \ldots$$

Let $\mathcal{H}$ be the K-variety parameterizing flags in V with the same dimensional data as (3.10), and let $h_0 \in \mathcal{H}(K)$ be the point corresponding to the Hodge filtration on V.

Base changing by means of v and ι, we get a K_v-variety $\mathcal{H}_v$ and a $\mathbf{C}$-variety $\mathcal{H}_{\mathbf{C}}$. We denote by $h_0^\iota \in \mathcal{H}_{\mathbf{C}}(\mathbf{C})$ the image of h_0.

Let $\Omega_{\mathbf{C}}$ be a contractible analytic neighbourhood of $y_0 \in Y_{\mathbf{C}}^{\mathrm{an}}$. The Gauss-Manin connection defines an isomorphism $H_{\mathrm{dR}}(X_t/\mathbf{C}) \simeq H_{\mathrm{dR}}(X_0/\mathbf{C})$ for each $t \in \Omega_{\mathbf{C}}$. In particular, the Hodge structure on the cohomology of X_t defines a point of $\mathcal{H}_{\mathbf{C}}(\mathbf{C})$; this gives rise to the complex period map

$$\Phi_{\mathbf{C}} : \Omega_{\mathbf{C}} \longrightarrow \mathcal{H}_{\mathbf{C}}(\mathbf{C}).$$

Indeed, $\Phi_{\mathbf{C}}$ extends to a map from the universal cover of $Y_{\mathbf{C}}^{\mathrm{an}}$ to $\mathcal{H}_{\mathbf{C}}(\mathbf{C})$ and this map is equivariant for the monodromy action of $\pi_1(Y_{\mathbf{C}}^{\mathrm{an}}, y_0)$ on $\mathcal{H}_{\mathbf{C}}(\mathbf{C})$. We conclude that the image of the period map can be bounded below by monodromy.

Lemma 3.1. *Suppose given a family* $X \to Y$, *and take notation as above; in particular,* Γ *is the Zariski closure of monodromy, and* $h_0^\iota = \Phi_{\mathbf{C}}(y_0)$. *Then we have the containment*

(3.11) $$\Gamma \cdot h_0^\iota \subset \text{the Zariski closure of } \Phi_{\mathbf{C}}(\Omega_{\mathbf{C}}) \text{ inside } \mathcal{H}_{\mathbf{C}}.$$

Proof. The preimage $\Phi_{\mathbf{C}}^{-1} Z$ of any algebraic subvariety $Z \subset \mathcal{H}_{\mathbf{C}}$, with $Z \supset \Phi_{\mathbf{C}}(\Omega_{\mathbf{C}})$, is a complex-analytic subvariety of $\widetilde{Y_{\mathbf{C}}^{\mathrm{an}}}$ containing Ω and thus all of $\widetilde{Y_{\mathbf{C}}^{\mathrm{an}}}$; therefore

$$\pi_1(Y_{\mathbf{C}}, y_0) \cdot h_0^\iota \subset Z$$

and then Z contains the Zariski closure of the right hand side, which is $\Gamma \cdot h_0^\iota$.   □

We need a v-adic analogue. Again, if $y \in \mathcal{Y}(\mathcal{O}_v)$ satisfies $y \equiv y_0$ modulo v, the Gauss–Manin connection (3.9) allows one to identify the Hodge filtration on $H^q_{\mathrm{dR}}(X_y/K_v)$ with a filtration on V_v, and thus with a point of $\mathcal{H}(K_v)$. This gives rise to a K_v-analytic function

$$\Phi_v : \Omega_v \longrightarrow \mathcal{H}(K_v), \text{ where } \Omega_v = \{y \in \mathcal{Y}(\mathcal{O}_v) : y \equiv y_0 \text{ modulo } v\}.$$

The following simple Lemma plays a crucial role. It allows us to analyze the Zariski closure of the p-adic period map in terms of the Zariski closure of the complex period map; for the latter we can use monodromy.

Lemma 3.2. *Suppose given power series $B_0, \ldots, B_N \in K[[z_1, \ldots, z_m]]$ such that all B_i are absolutely convergent, with no common zero, both in the v-adic and complex disks*

$$U_v = \{\underline{z} : |z_i|_v < \epsilon\} \text{ and } U_{\mathbf{C}} = \{\underline{z} : |z_i|_{\mathbf{C}} < \epsilon\}.$$

Write

$$\underline{B}_v : U_v \to \mathbf{P}^N_{K_v}$$

$$\underline{B}_{\mathbf{C}} : U_{\mathbf{C}} \to \mathbf{P}^N_{\mathbf{C}}$$

for the corresponding maps.

Then there exists a K-subscheme $\mathcal{Z} \subset \mathbf{P}^N$ whose base extension to K_v (respectively $\mathbf{C}$) gives the Zariski closure of $\underline{B}_{\mathbf{C}}(U_{\mathbf{C}}) \subset \mathbf{P}^N_{\mathbf{C}}$ (respectively $\underline{B}_v(U_v) \subset \mathbf{P}^N_{K_v}$). In particular, these Zariski closures have the same dimension.

Proof. We take I the ideal of $\mathcal{Z}$ to be that generated by all homogeneous polynomials $Q \in K[x_0, \ldots, x_N]$ such that $Q(B_0, \ldots, B_N)$ is identically zero.

To verify the claim (for K_v; the proof for $\mathbf{C}$ is identical) we just need to verify that if a homogeneous polynomial $Q_v \in K_v[x_0, \ldots, x_N]$ vanishes on $\underline{B}_v(U_v)$ then Q_v lies in the K_v-span of I. But if Q_v vanishes on $\underline{B}_v(U_v)$ then $Q_v(B_0, \ldots, B_N) \equiv 0$ in $K_v[[z_1, \ldots, z_m]]$. The identical vanishing of $Q_v(B_0, \ldots, B_N)$ is an infinite system of linear equations on the coefficients of Q_v, with coefficients in K. Any K_v-solution of such a linear system is, of course, a K_v-linear combination of K-solutions. $\square$

By embedding $\mathcal{H}$ into a projective space $\mathbf{P}^N$, and applying the prior two Lemmas, we deduce:

Lemma 3.3. *The dimension of the Zariski closure (in the K_v-variety $\mathcal{H}_{K_v}$) of $\Phi_v(\Omega_v)$ is at least the (complex) dimension of $\Gamma \cdot h_0^\iota$.*

In particular, if $\mathcal{H}_v^{\mathrm{bad}} \subset \mathcal{H}_v$ is a Zariski-closed subset of dimension less than $\dim_{\mathbf{C}}(\Gamma \cdot h_0^\iota)$, then $\Phi_v^{-1}(\mathcal{H}_v^{\mathrm{bad}})$ is contained in a proper K_v-analytic subset of Ω_v, by which we mean a subset cut out by v-adic power series converging absolutely on Ω_v.

One can do better than this using the results of Bakker and Tsimerman, replacing "proper K_v-analytic" by "Zariski-closed." See §9. We do not need this improvement for the applications to Mordell.

3.5. **Hodge structures.** We use p-adic Hodge theory to relate Galois representations to crystalline cohomology. A good reference is [8] or [13].

For each $y \in U$ the representation ρ_y (see (3.2)) is crystalline upon restriction to K_v, because of the existence of the model $\mathcal{X}_y$ for X_y. By p-adic Hodge theory, there is [8, Proposition 9.1.9] a fully faithful embedding of categories:

(3.12)   crystalline representations of Gal_{K_v} on $\mathbf{Q}_p$ vector spaces $\hookrightarrow \mathcal{FL}$,

where the objects of $\mathcal{FL}$ are triples (W, ϕ, F) of a K_v-vector space W, a Frobenius-semilinear automorphism $\phi : W \to W$, and a descending filtration F of W. The morphisms in the category $\mathcal{FL}$ are morphisms of K_v-vector spaces that respect ϕ and filtrations [18, Expose III, §4.3].

By the crystalline comparison theorem of Faltings [16], the embedding (3.12) carries ρ_y to the triple $(H_{\mathrm{dR}}^q(X_y/K_v), \phi_v, \text{Hodge filtration for } X_y)$. But (3.9) induces an isomorphism in $\mathcal{FL}$:

$$(H_{\mathrm{dR}}^q(X_y/K_v), \phi_v, \text{Hodge filtration for } X_y) \simeq (V_v, \phi_v, \Phi_v(y)),$$

As a sample result of what we can now show, we give the following. We will use the method of proof again and again, so it seems useful to present it in the current simple context.

Proposition 3.4. *Notation as above: in particular $X \to Y$ is a smooth proper family over K, V is the degree q de Rham cohomology of a given fiber X_0 above $y_0 \in Y(K)$, $\mathcal{H}$ a space of flags in V,*

$$\Phi_v : \{y \in \mathcal{Y}(\mathcal{O}_v) : y \equiv y_0\} \longrightarrow \mathcal{H}(K_v)$$

is the v-adic period mapping, $\Gamma \subset \mathrm{GL}(V_{\mathbf{C}})$ is the Zariski closure of the monodromy group, and $h_0 = \Phi(y_0)$ is the image of y_0 under the period mapping.

Suppose that

$$(3.13) \qquad\qquad \dim_{K_v}\left(\mathrm{Z}(\phi_v^{[K_v:\mathbf{Q}_p]})\right) < \dim_{\mathbf{C}} \Gamma \cdot h_0^{\iota}$$

where the left-hand side $\mathrm{Z}(\dots)$ denotes the centralizer, in $\mathrm{Aut}_{K_v}(V_v)$, of the K_v-linear operator $\phi_v^{[K_v:\mathbf{Q}_p]}$.

Then the set

$$(3.14) \qquad\qquad \{y \in Y(\mathcal{O}) : y \equiv y_0 \text{ modulo } v, \rho_y \text{ semisimple}\}$$

is contained in a proper K_v-analytic subvariety of the residue disk of $Y(K_v)$ at y_0.

Proof. For any y as in (3.14) the Galois representation ρ_y belongs to a finite set of isomorphism classes (Lemma 2.3). By our previous discussion the triple $(V_v, \phi_v, \Phi_v(y))$ also belongs to a finite set of isomorphism classes (now in the category $\mathcal{FL}$). Choosing representatives (V_v, ϕ_v, h_i) for these isomorphism classes, we must have

$$\Phi_v(y) \in \bigcup_i \mathrm{Z}(\phi_v) \cdot h_i,$$

where $\mathrm{Z}(\phi_v)$ is the subgroup of elements in $\mathrm{GL}_{K_v}(V_v)$ which commute with ϕ_v.

Now certainly $\mathrm{Z}(\phi_v) \subset \mathrm{Z}(\phi_v^{[K_v:\mathbf{Q}_p]})$, and the right-hand side is now the K_v-points of a K_v-algebraic subgroup of $\mathrm{GL}_{K_v}(V_v)$. Therefore, any y as in (3.14) is contained in the preimage, under Φ_v, of a proper Zariski-closed subset of $\mathcal{H}_v$ with dimension the left hand side of (3.13). This is obviously a K_v-analytic subvariety as asserted. It is proper because of Lemma 3.3. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

In conclusion we note that we really have bounded $\mathcal{Y}(\mathcal{O})$ rather than the set of $y \in Y(K)$ for which the abstract Galois representation ρ_y has good reduction outside S. To bound the latter set, we would have to deal with the possibility that such y would be non-integral at S; this would require a more detailed analysis "at infinity" and we have not attempted it.

4. THE S-UNIT EQUATION

As a first application, and a warm-up to the more complicated case of curves of higher genus, we will show finiteness of the set of solutions to the S-unit equation. This argument is not logically necessary for the later proofs but we hope it will serve as a useful introduction to them.

Theorem 4.1. *The set*
$$U = \{t \in \mathcal{O}_S^* : 1 - t \in \mathcal{O}_S^*\}$$
is finite.

4.1. Reductions. We begin with some elementary reductions.

We may freely enlarge both S and K. Thus, we may suppose that S contains all primes above 2 and that K contains the 8th roots of unity. Let m be the largest power of 2 dividing the order of the group of roots of unity in K. By assumption $m \geqslant 8$.

First of all, it suffices to prove finiteness of the set
$$U_1 = \{t \in \mathcal{O}_S^* : 1 - t \in \mathcal{O}_S^*, t \notin (K^*)^2\},$$

because $U \subset U_1 \cup U_1^2 \cup U_1^4 \cup \cdots \cup U_1^m$. To see this, we take $t \in U$ and try to repeatedly extract its square root; observe that such a square root, if in K, also belongs to U. If we cannot extract an mth root of t, we are done; otherwise, write $t = t_1^m$ and adjust t_1 by a primitive mth root of unity to ensure that t_1 is nonsquare.

Suppose that $t \in U_1$. Since t is a nonsquare and $\mu_m \subset K$ the order of t in the group $(K^*)/(K^*)^m$ is exactly m. Otherwise there is some proper divisor $k > 1$ of m, and an element $a \in K^*$, such that $t^k = a^m$, i.e. $t \in a^{m/k}\mu_k$, contradicting the fact that t is nonsquare.

Fixing $t^{1/m}$ an mth root of t in $\overline{K}$, the field $K(t^{1/m})$ is Galois over K, and Kummer theory guarantees that its Galois group is $\mathbf{Z}/m\mathbf{Z}$. There are (Hermite–Minkowski) only finitely many possibilities for $K(t^{1/m})$. Enumerate them; call them $L_1, \ldots, L_r$, say. Each L_i is a cyclic degree-m extension of K, and it is sufficient to prove finiteness of the set

$$(4.1) \qquad U_{1,L} = \{t \in U_1, K(t^{1/m}) \simeq L\}.$$

for a fixed field $L \in \{L_1, \ldots, L_r\}$; here we understand $K(t^{1/m}) = K[x]/(x^m - t)$.

Fix an L as above. L is cyclic of degree m over K. Choose a prime v of K such that:

   (i) the class of Frobenius at v generates $\mathrm{Gal}(L/K)$;
   (ii) the prime p of $\mathbf{Q}$ below v is unramified in K.
   (iii) no prime of S lies above p.

In particular, v is inert in L/K; thus, if $t \in U_{1,L}$ then t is not a square in K_v, for otherwise $L \otimes_K K_v \simeq K_v[x]/(x^m - t)$ would not be not a field.

In summary, it is enough to prove the following lemma.

Lemma 4.2. *Suppose K contains the 8th roots of unity, and S contains all primes above 2. Fix a cyclic field extension L/K, a place $v \notin S$ as above, and a basepoint $t_0 \in \mathcal{O}_S$. Let $U_{1,L}$ be as above. Then the set*

$$(4.2) \qquad \{t \in U_{1,L} : t \equiv t_0 \ modulo \ v\}$$

is finite.

The proof of this Lemma will occupy the rest of the section. Throughout the proof, p is the prime of $\mathbf{Q}$ below v, and "Tate module" always refers to p-adic Tate module.

4.2. **A variant of the Legendre family.** As discussed in the Introduction, we apply Proposition 3.4 not to the Legendre family, but to a modification of it: Let $\mathcal{Y} = \mathbf{P}^1_{\mathcal{O}} - \{0, 1, \infty\}$ (where $0, 1, \infty$ denote the corresponding sections over $\operatorname{Spec} \mathcal{O}$) and let $\mathcal{Y}' = \mathbf{P}^1_{\mathcal{O}} - \{0, \mu_m, \infty\}$; let $\pi : \mathcal{Y}' \to \mathcal{Y}$ be the map $u \mapsto u^m$.

Let $\mathcal{X} \to \mathcal{Y}'$ be the Legendre family, so that its fiber over t is the curve $y^2 = x(x - 1)(x - t)$; and consider the composite

$$\mathcal{X} \longrightarrow \mathcal{Y}' \xrightarrow{\pi} \mathcal{Y}.$$

We will apply our prior results to the family $\mathcal{X} \to \mathcal{Y}$; also, as before, we denote by X and Y the fibers of $\mathcal{X}$ and $\mathcal{Y}$ over $\operatorname{Spec}(K)$. Thus the geometric fiber X_t of $X \to Y$ over $t \in Y(K)$ is the disjoint union of the curves $y^2 = x(x - 1)(x - t^{1/m})$ over all mth roots of t.

4.3. **Proof of Finiteness.** Assume for the moment the following two Lemmas; they will be proved in §4.4.

Lemma 4.3 (Big monodromy). *Consider the family of curves over $\mathbf{C} - \{0, 1\}$ whose fiber over $t \in \mathbf{C}$ is the union of the elliptic curves $E_z : y^2 = x(x-1)(x-z)$, over all m^{th} roots $z^m = t$. Then the action of monodromy*

$$(4.3) \qquad \pi_1(\mathbf{C} - \{0, 1\}, t_0) \longrightarrow \operatorname{Aut}\left(\bigoplus_{z^m = t_0} H^1_B(E_z, \mathbf{Q})\right)$$

has Zariski closure containing $\prod_z \operatorname{SL}(H^1_B(E_z, \mathbf{Q}))$.

Lemma 4.4 (Generic simplicity). *Let L be a number field and p a rational prime, larger than 2, and unramified in L. There are only finitely many $z \in L$ such that $z, 1 - z$ are both p-units, but for which the Galois representation of G_L on the Tate module $T_p(E_z) = H^1_{\text{et}}(E_{z,\bar{L}}, \mathbf{Q}_p)$ of the elliptic curve*

$$E_z : y^2 = x(x - 1)(x - z),$$

fails to be simple.

Of course much stronger results than Lemma 4.4 are known. The point here is that we prove this in a "soft" fashion, using the Torelli theorem as a substitute for more sophisticated arguments; although we use the specific feature of Hodge weights 0 and 1, the argument is robust enough to generalize (although with a little added complexity, see e.g. Lemma 6.3).

Proof of Lemma 4.2 assuming Lemmas 4.3 and 4.4. This argument is similar to the proof of Proposition 3.4, with added complication coming from the interaction of the fields K and L. Recall that we have fixed $t_0 \in U_{1,L}$ and we must verify the finiteness of the set of $t \in U_{1,L}$ with $t \equiv t_0$ modulo v.

By Lemmas 4.4 and 2.3, it is enough to verify the finiteness of the subset of such t where the pair $(K(t^{1/m}), \rho_t|G_{K(t^{1/m})})$ lies in a fixed isomorphism class; in particular $(K_v(t^{1/m}), \rho_t|G_{K_v(t^{1/m})})$ lies in a fixed isomorphism class.

Under the correspondence of p-adic Hodge theory, ρ_t restricted to $K_v(t^{1/m})$ corresponds to the filtered ϕ-module

$$(4.4) \qquad \left(H^1_{\text{dR}}(X_{t,K_v}/K_v) \text{ as } K_v(t^{1/m})\text{-module}, \text{Frobenius}, \text{filtration}\right),$$

where we equip $H^1_{\text{dR}}(X_{t,K_v}/K_v)$ with the structure of 2-dimensional vector space over $K_v(t^{1/m})$ that arises from the scheme structure of X_t over $K(t^{1/m})$.

Let us clarify this vector space structure over $K_v(t^{1/m})$, which is crucial to our argument. Although *a priori* a K-scheme, the factorization $X \to Y' \to Y$ induces on X_t the structure of $K(t^{1/m})$-scheme, i.e. arising from the morphism $X_t \to (Y')_t \simeq \mathrm{Spec} K(t^{1/m})$. Now the de Rham cohomology of X_t is the same whether we consider it as a $K(t^{1/m})$-variety or as a K-variety. If we consider it as K-variety, we can recover its structure of $K(t^{1/m})$-vector space by means of the natural map

$$K(t^{1/m}) = H^0_{\mathrm{dR}}(Y'_t/K) \to H^0_{\mathrm{dR}}(X_t/K).$$

The same picture works with K replaced by K_v everywhere.

(Similarly, there are two natural interpretations for "Frobenius" in (4.4), but they are equivalent: As just explained, we can consider the space H^1_{dR} as the de Rham cohomology of either a $K_v(t^{1/m})$-scheme, or of the associated K_v-scheme obtained simply by restricting the scalars. Both of these schemes have evident integral models, over $\mathcal{O}_v[x]/(x^m - t)$ and $\mathcal{O}_v$ respectively. Accordingly, the de Rham cohomologies can be identified with the crystalline cohomologies of the special fibers; these crystalline cohomologies are identified, in a fashion that respects the semilinear Frobenius endomorphisms.)

The Gauss–Manin connection for the family $X \to Y$ induces

$$(4.5) \qquad H^1_{\mathrm{dR}}(X_{t,K_v}/K_v) \simeq H^1_{\mathrm{dR}}(X_{t_0,K_v} K_v)$$

which, by compatibility of Gauss–Manin connection with the cup product, is compatible with their module structures over the corresponding H^0s. The corresponding identification of H^0s induces the standard identification $K_v(t^{1/m}) \simeq K_v(t_0^{1/m})$ and therefore the isomorphism (4.5) is compatible with structures of $K_v(t^{1/m}) \simeq K_v(t_0^{1/m})$-modules.

Therefore, under the identification of (4.5), the F^1-step of the filtration on $H^1_{\mathrm{dR}}(X_{t,K_v}/K_v)$ is identified with a $K_v(t_0^{1/m})$-line inside $H^1_{\mathrm{dR}}(X_{t_0,K_v}/K_v)$. Call this line $\Phi(t)$. The variation of this line gives a K_v-analytic period mapping

(4.6)

$$\Phi : \{t \in K_v, t \equiv t_0 \text{ modulo } v\} \longrightarrow K_v(t_0^{1/m})\text{-lines in } H^1_{\mathrm{dR}}(X_{t_0,K_v}/K_v) \xrightarrow{\simeq} \mathbf{P}^1_{K_v(t_0^{1/m})}$$

$$K_v\text{-subspaces in } H^1_{\mathrm{dR}}(X_{t_0,K_v}/K_v) \xrightarrow{\simeq} \mathrm{Gr}(2m, m)_{K_v}.$$

(The period mapping for the family $X \to Y$ *a priori* takes values in the bottom row, but we have just seen that it factors through the top row. See §3.3 for a more detailed discussion of the radius of convergence; in particular it defines a rigid analytic function on a domain containing $\{t \in K_v, t \equiv t_0 \text{ modulo } v\}$ i.e. the K_v-points in a residue disk.)

Therefore (applying the Gauss–Manin connection to identify (4.4) with similar data over t_0) the isomorphism class of the quadruple

$$\left(K_v(t_0^{1/m}), H^1_{\mathrm{dR}}(X_{t_0,K_v}/K_v) \text{ as } K_v(t_0^{1/m})\text{-module}, \Phi(t), \mathrm{Frob}_v \right)$$

is determined from (4.4) and therefore the triple

$$\left(H^1_{\mathrm{dR}}(X_{t_0,K_v}/K_v) \text{ as } K_v(t_0^{1/m})\text{-module}, \Phi(t), \mathrm{Frob}_v \right)$$

lies in a finite set of isomorphism classes for filtered ϕ-modules over $K_v(t_0^{1/m})$ (coming from the finitely many automorphisms of $K_v(t_0^{1/m})$ over K_v). Therefore, $\Phi(t)$ lies in a finite collection of orbits for

$$Z = \text{centralizer of } \mathrm{Frob}_v^{[K_v:\mathbf{Q}_p]} \text{ in } K_v(t_0^{1/m})\text{-linear automorphisms of } H^1_{\mathrm{dR}}(X_{t_0,K_v}/K_v).$$

Now we can apply Lemma 2.1 to the field extension $K_v(t_0^{1/m})/K_v$ and the K_v-linear automorphism $\mathrm{Frob}_v^{[K_v:\mathbf{Q}_p]}$ of $H^1_{\mathrm{dR}}(X_{t_0,K_v}/K_v)$. This gives us that

$$\dim_{K_v} Z \leqslant (\dim_{K_v(t_0^{1/m})} H^1_{\mathrm{dR}})^2 = 4.$$

Our analysis thus far has shown that the set of $t \in U_{1,L}$ such that $t \equiv t_0$ modulo v is contained in

$$\Phi^{-1}(\mathcal{Z}),$$

where Φ is the period map as in (4.6) and $\mathcal{Z} \subset \mathrm{Gr}_{K_v}(2m, m)$ has dimension at most 4. By Lemma 3.3, this set is finite so long as we verify an assertion about the complex period map, namely, that the dimension of the orbit of the algebraic monodromy group over $\mathbf{C}$ is strictly greater than 4. As in Lemma 3.3, we fix an embedding $K \hookrightarrow \mathbf{C}$ throughout the following discussion.

As mentioned, the vector space $V = H^1_{\mathrm{dR}}(X_{t_0}/K)$ has the natural structure of a 2-dimensional vector space over $K(t_0^{1/m})$. The splitting of $X_{t_0,\mathbf{C}}$ into geometric components induces a splitting

$$(4.7) \qquad\qquad V_{\mathbf{C}} = \bigoplus_{i=1}^{m} V_i,$$

where each V_i is a 2-dimensional complex vector space; moreover the Hodge filtration on $H^1_{\mathrm{dR}}(X_{t_0}/K) \otimes \mathbf{C}$ also splits along this decomposition. Lemma 4.3 shows that the algebraic monodromy group Γ contains $\prod_{i=1}^{m} \mathrm{SL}(V_i)$. The pertinent flag variety $\mathcal{H} \simeq \mathrm{Gr}(V, m)$ is the variety of m-dimensional subspaces in V; the splitting (4.7) induces a natural inclusion $\prod_{i=1}^{m} \mathbf{P}V_i \hookrightarrow \mathcal{H}_{\mathbf{C}}$. Therefore the the orbit Γh_0^ι is all of $\prod_{i=1}^{m} \mathbf{P}V_i$ and, in particular, has dimension $m \geqslant 8$. Lemma 3.3 now gives the desired finiteness.

In conclusion, assuming Lemmas 4.3 and 4.4, we have shown that the set described in (4.1) is finite. $\qquad\square$

4.4. **Big Monodromy and Generic Simplicity.** In this section we prove Lemmas 4.3 and 4.4.

Proof of Lemma 4.3. Write Γ for the Zariski closure in question. It preserves the splitting of (4.3), although not the individual summands. Then:

- Γ transitively permutes the factors on the right-hand side of (4.3), by considering the action of local monodromy near $t = 0$;
- $\Gamma \cap \mathrm{SL}(2)^m$ projects to $\mathrm{SL}(2)$ in each factor: indeed, this projection contains a finite-index subgroup of the algebraic monodromy group of the Legendre family.
- Γ contains an element of the form

$$(1, 1, \ldots, 1, u, 1, \ldots, 1)$$

where $u \in \mathrm{SL}(2)$ is a nontrivial unipotent element, as we see by considering the action of local monodromy near $t = 1$.

We now apply a slight variant of Lemma 2.12 to conclude that $\Gamma \supset \mathrm{SL}(2)^m$. $\qquad\square$

Proof of Lemma 4.4. Fix $z_0 \in L$ with the quoted p-integrality properties; in particular, E_{z_0} has good reduction at all primes of L above p.

It is enough to show the same finiteness when we restrict to the set

$$V_L = \{z \in L : z \equiv z_0 \text{ modulo } v, \text{ for all } v | p\}.$$

If $T_p(E_z)$ is reducible there exists a one-dimensional subrepresentation $W_z \subset T_p(E_z)$. By Lemma 2.10 (applied with $K = \mathbf{Q}$) there is a place w of L above p such that $F^1(W_z^{\mathrm{dR}}) =$

W_z^{dR}; here W_z^{dR} is the filtered L_w-vector space associated to W_z by p-adic Hodge theory over the p-adic field L_w.

Because the Newton and Hodge polygons of W_z^{dR} have the same endpoint, the slope of semilinear Frobenus acting on W_z^{dR} is equal to 1; by the same reasoning for $H_{\mathrm{dR}}^1(E_z/L_w)$, the sum of slopes for the semilinear Frobenius acting on $H_{\mathrm{dR}}^1(E_z/L_w)$ is 1, so it has another slope equal to 0.

In particular, the L_w-linear Frobenius $\mathrm{Frob}_w^{[L_w:\mathbf{Q}_p]}$ has distinct eigenvalues.

Also, the L_w-line W_z^{dR} must coincide with $F^1 H_{\mathrm{dR}}^1(E_z/L_w)$, so that the latter space is the slope-1 eigenline for the semilinear Frobenius Frob_w.

As in the discussion around (3.9), Gauss–Manin induces an identification

$$(4.8) \qquad H_{\mathrm{dR}}^1(E_{z_0}/L_w) \simeq H_{\mathrm{dR}}^1(E_z/L_w)$$

of L_w-vector spaces with semilinear Frobenius action. But the position of the Hodge line $F^1 H_{\mathrm{dR}}^1(E_z/L_w)$ varies w-adic analytically inside the disk V_L – here we use (4.8) to identify this line to a line inside the fixed space $H_{\mathrm{dR}}^1(E_{z_0}/L_w)$ – and the associated w-adic analytic function is nonconstant (by the – trivial – Torelli theorem for elliptic curves). It follows there are at most finitely many $z \in V_L$ for which $F^1 H_{\mathrm{dR}}^1(E_z/L_w)$ is the slope-1 Frobenius eigenline. Taking the union over possible w we still see that the exceptional set is finite. $\qquad\square$

5. Outline of the argument for Mordell's conjecture

The proof of the Mordell conjecture is substantially harder than the S-unit equation. To try to assist the reader, we summarize the proof here, and then elaborate on the ingredients over the next three sections.

First of all, we will make crucial use of the type of structure that occurred in §4.2, to which we give a name:

Definition 5.1. *An* abelian-by-finite family *over Y is a sequence of morphisms*

$$X \longrightarrow Y' \stackrel{\pi}{\longrightarrow} Y$$

where π is finite étale, and $X \to Y'$ is (equipped with the structure of) a polarized abelian scheme.

A good model *for such a family, over an S-integer ring $\mathcal{O} \subset K$, is a family $\mathcal{X} \to \mathcal{Y}' \to \mathcal{Y}$ of smooth, proper $\mathcal{O}$-schemes, satisfying the same conditions and also the assumptions at the start of §3.1, and recovering $X \to Y' \to Y$ on base change to K.*

Of course the polarization on $X \to Y'$ is an additional structure but for brevity we do not explicitly include it in the notation.

For any such abelian-by-finite family $X \to Y' \to Y$ take a complex point $y_0 \in Y(\mathbf{C})$ and consider the action of the topological fundamental group $\pi_1(Y(\mathbf{C}), y_0)$ on

$$H_B^1(X_{y_0}, \mathbf{Q}) \simeq \bigoplus_{\pi(\tilde{y}) = y_0} H_B^1(X_{\tilde{y}}, \mathbf{Q}),$$

where the sum is taken over $\tilde{y} \in Y'(\mathbf{C})$ lying over y_0. We say that the family has *full monodromy* if the Zariski closure of $\pi_1(Y, y_0)$, in its action on the right-hand side, contains the product of symplectic groups:

$$(5.1) \qquad \overline{(\text{image of } \pi_1(Y(\mathbf{C}), y_0))} \supset \prod_{\pi(\tilde{y}) = y_0} \mathrm{Sp}\left(H_B^1(X_{\tilde{y}}, \mathbf{Q}), \omega\right),$$

where the symplectic group is with reference to the form ω defined by the polarization.

The key reason to use abelian-by-finite families is that we can guarantee that the Galois orbits on any fiber of $Y' \to Y$, above a K-rational point of Y, are "large." In fact, what we need (see discussion in Introduction) is that most points in the fiber above $y_0 \in Y(K)$ cannot be defined over "small" extensions of K_v. To quantify the notions of large and small we introduce the following quantity:

Definition 5.2. *Let E be a G_K-set and v a place of K such that the G_K-action on E is unramified at v. Let*

$$(5.2) \qquad \mathrm{size}_v(E) = \frac{\textit{number of elements of } E \textit{ that belong to } \mathrm{Frob}_v\textit{-orbits of size} < 8}{\textit{number of elements of } E}$$

If E is a zero-dimensional K-scheme, we will write $\mathrm{size}_v(E)$ instead of $\mathrm{size}_v(E(\bar{K}))$.

Note that if $E \to E'$ is a morphism of G_K-sets, and all fibers have the same cardinality, then

$$(5.3) \qquad\qquad\qquad \mathrm{size}_v(E) \leqslant \mathrm{size}_v(E').$$

The next result is, in essence, a variant of Proposition 3.4, but it requires some careful indexing. It will be proved in §6.

Proposition 5.3. *Let Y be a curve over K of genus $g \geqslant 2$.*

Let $X \to Y' \xrightarrow{\pi} Y$ be an abelian-by-finite family over Y, with full monodromy (see Definition 5.1 and subsequent discussion). Let d be the relative dimension of $X \to Y'$. Suppose that $X \to Y' \xrightarrow{\pi} Y$ admits a good model over the ring $\mathcal{O}$ of S-integers of K. Let $v \notin S$ be a friendly place of K (Definition 2.7).

Let size_v be as in (5.2). Then the set

$$Y(K)^* := \left\{ y \in Y(K) : \mathrm{size}_v(\pi^{-1}(y)) < \frac{1}{d+1} \right\}$$

is finite.

In §7, we introduce a specific abelian-by-finite family $X_q \to Y_q' \xrightarrow{\pi} Y$ for each prime $q \geqslant 3$, referred to as the "Kodaira–Parshin family for the group $\mathrm{Aff}(q)$." Roughly, Y_q' is a Hurwitz space for $\mathrm{Aff}(q)$ and X_q is the Prym of the universal curve. It has the following properties:

(i) It has full monodromy (Theorem 8.1).
(ii) The relative dimension d_q of $X_q \to Y_q'$ is given by $d_q = (q-1)(g - \frac{1}{2})$.
(iii) For each $y_0 \in Y(K)$ there is a G_K-equivariant identification of $\pi^{-1}(y_0)$ with the conjugacy classes of surjections $\pi_1^{\mathrm{geom}}(Y - y_0, *) \twoheadrightarrow \mathrm{Aff}(q)$ that are nontrivial on a loop around y_0.

Note that we can identify π_1^{geom} with the profinite completion of a free group on $2g$ generators $x_1, x_1', \ldots, x_g, x_g'$ in such a way that the loop around y_0 corresponds to the conjugacy class of $[x_1, x_1'][x_2, x_2'] \ldots [x_g, x_g']$. Therefore, the set of surjections $\pi_1^{\mathrm{geom}}(Y - y_0, *) \twoheadrightarrow \mathrm{Aff}(q)$ nontrivial on a loop around y_0 is identified with the left-hand side of (2.3).

There is probably nothing very special about the use of $\mathrm{Aff}(q)$, but it is simple enough that we can compute everything explicitly.

Assuming these things we can prove:

Theorem 5.4. *Let Y be a curve over the number field K with genus $g \geqslant 2$. Then $Y(K)$ is finite.*

Proof. We apply Proposition 5.3 to the Kodaira–Parshin family with parameter q. What we will show is that we may choose q and the place v in such a way that v is friendly and

(5.4) $$\text{size}_v(\pi^{-1}(y)) < \frac{1}{d_q + 1} \quad \text{for all } y \in Y(K).$$

The key point is to use the mapping (5.5) below and the Weil pairing to give an upper bound on $\text{size}_v(\pi^{-1}(y))$.

We choose q with the following properties:

(i) $q - 1$ is not divisible by 4 or by any odd primes less than $8[K : \mathbf{Q}]$.
(ii) The Galois closure K' of K is linearly disjoint from $\mathbf{Q}(\zeta_{q-1})$ over $\mathbf{Q}$.
(iii) $\frac{8 \cdot 2^{g+1}}{(q-1)^g} < \frac{1}{(g-1/2)(q-1)+1}$.

This is possible by Dirichlet's theorem: we choose q such that q is not congruent to 1 mod ℓ for any prime ℓ that either divides the discriminant of K, or that is less than $8[K : \mathbf{Q}]$, and also q is not congruent to 1 mod 4. Then linear disjointness follows: for ramification reasons $K' \cap \mathbf{Q}(\zeta_{q-1}) = \mathbf{Q}$. Such a q can be chosen arbitrarily large; in particular it can be chosen to satisfy the third condition.

Now form the Kodaira–Parshin family $X = X_q \longrightarrow Y'_q \longrightarrow Y$ for the group $\text{Aff}(q)$ and choose a set S such that it has a good model over the ring of S-integers.

Next we show that there exists a place $v \notin S$ of K such that:

(i) v is friendly (in the sense of Definition 2.7)
(ii) $(q_v, q - 1) = 1$ (recall that q_v was the cardinality of the residue field at v)
(iii) For any odd prime factor r of $q - 1$, the class of q_v in $(\mathbf{Z}/r)^*$ has order at least 8.

Note that the latter two conditions depend only on the residue class of q_v modulo $q - 1$. We will produce v by the Chebotarev density theorem, applied to $\text{Gal}(K'(\zeta_{q-1})/\mathbf{Q})$. By hypothesis, K' and $\mathbf{Q}(\zeta_{q-1})$ are linearly disjoint over $\mathbf{Q}$, so the map

$$\text{Gal}(K'(\zeta_{q-1})/\mathbf{Q}) \longrightarrow \text{Gal}(K'/\mathbf{Q}) \times \text{Gal}(\mathbf{Q}(\zeta_{q-1})/\mathbf{Q}).$$

is an isomorphism.

If K has no CM subfield, choose $\sigma \in \text{Gal}(K'/\mathbf{Q})$ arbitrarily. Otherwise let E be the maximal CM subfield of K, and let E^+ the maximal totally real subfield; choose some $\sigma \in \text{Gal}(K'/E^+) \subseteq \text{Gal}(K'/\mathbf{Q})$ inducing the nontrivial automorphism of E over E^+.

By the Chinese Remainder Theorem, we can choose a residue class $a \in (\mathbf{Z}/(q-1))^*$ whose reduction modulo r is a primitive root for $(\mathbf{Z}/r)^*$ for every prime factor r of $(q-1)$.

By Chebotarev density, there is a place $\wp$ of $K'(\zeta_{q-1})$ such that the Frobenius $\text{Frob}_\wp$ is the element (σ, a) of $\text{Gal}(K'(\zeta_{q-1})/\mathbf{Q}) \simeq \text{Gal}(K'/\mathbf{Q}) \times (\mathbf{Z}/(q-1))^*$. Let p be the prime of $\mathbf{Q}$ below $\wp$; thus $p \equiv a$ modulo $q - 1$. The place v of K below $\wp$ has residue field of size $q_v = p^i$, with $i \leqslant [K : \mathbf{Q}]$; therefore, if r is an odd prime factor of $(q - 1)$, the order of q_v mod r is at least $\left\lceil \frac{r-1}{[K:\mathbf{Q}]} \right\rceil \geqslant 8$. For the last inequality we used property (i) of q.

If K admits a CM subfield then the place of E^+ below $\wp$ is inert in E, by choice of σ. This shows that there indeed exists v as desired.

Now consider the Kodaira–Parshin family $X = X_q \longrightarrow Y'_q \longrightarrow Y$ for the group $\text{Aff}(q)$ and write d_q for the relative dimension of $X \to Y$. For any $y \in Y(K)$ property (iii) of Kodaira–Parshin covers (page 26), and the surjection $\text{Aff}(q) \twoheadrightarrow \mathbf{F}_q^* \simeq \mathbf{Z}/(q-1)$, gives rise to a map of G_K-sets

(5.5) $$\pi^{-1}(y) \longrightarrow \underbrace{H^1_{\text{et}}(Y_{\bar{K}}, \mathbf{Z}/(q-1))}_{M}.$$

Let $\Upsilon \subseteq M$ be the image of the map. In explicit coordinates, the map (5.5) has been studied in Lemma 2.11 (see also remark after (iii) on page 26). Therefore, by Lemma 2.11, all fibers of the map have the same size. Therefore, in view of (5.3), it is enough to show that $\mathrm{size}_v(\Upsilon) < \frac{1}{d_q+1}$.

Now M has the structure of a $(2g)$-dimensional free module over $\mathbf{Z}/(q-1)$. On choosing an identification of M with $(\mathbf{Z}/(q-1))^{2g}$, the set Υ consists of those elements $(y_1, y_1', \ldots, y_g, y_g')$ such that the elements $y_1, y_1', \ldots, y_g, y_g'$ generate $(\mathbf{Z}/(q-1))$. (This is shown in the proof of Lemma 2.11.)

M is also equipped with a Galois-equivariant Weil pairing

$$\langle -, - \rangle : M \times M \to \mu_{q-1}^{\vee} := \mathrm{Hom}(\mu_{q-1}, \mathbf{Z}/(q-1)\mathbf{Z}).$$

The Weil pairing is perfect, i.e. the corresponding map $M \to \mathrm{Hom}(M, \mu_{q-1}^{\vee})$ is an isomorphism. The Frobenius at v induces, in particular, an automorphism $T : M \to M$ that satisfies

$$\langle Tv_1, Tv_2 \rangle = q_v^{-1} \langle v_1, v_2 \rangle.$$

We want to bound the number of elements of M belonging to T-orbits of size less than 8. These elements are contained in the union of the submodules $\ker(T^i - 1)$ for $1 \leqslant i \leqslant 8$. If $m_1, m_2 \in \ker(T^i - 1)$ then $(q_v^{-i} - 1)\langle m_1, m_2 \rangle = 0$. For every odd prime factor r of $q-1$ we know that q_v^i is not congruent to 1 modulo r; therefore $(q_v^i - 1)$ is relatively prime to r. Thus $2\langle m_1, m_2 \rangle = 0$ for any $m_1, m_2 \in \ker(T^i - 1)$.

Now if A is a finite abelian group endowed with a nondegenerate pairing $A \times A \to \mathbf{Q}/\mathbf{Z}$ then any subgroup $B \subset A$ such that $\langle B, B \rangle = 0$ has order at most $\sqrt{A}$. Applying this to $2M$ we find

$$\left| 2\ker(T^i - 1) \right| \leqslant \left(\frac{q-1}{2} \right)^g \implies \left| \ker(T^i - 1) \right| \leqslant 2^g (q-1)^g.$$

Hence, the number of elements of M contained in the union of the submodules $\ker(T^i - 1)$ for $1 \leqslant i \leqslant 8$ is at most $8 \cdot 2^g (q-1)^g$.

It remains to give an upper bound for the "size_v" of Υ, the image of (5.5). The number of generating $(2g)$-tuples in $\mathbf{Z}/N$ equals $\#(\mathbf{Z}/N)^* \times \mathbf{P}^{2g-1}(\mathbf{Z}/N)$, which equals $N^{2g} \cdot \prod_{p|N}(1 - p^{-2g}) \geqslant \frac{1}{2} N^{2g}$. So Υ has at least $\frac{1}{2}(q-1)^{2g}$ elements, of which at most $8 \cdot 2^g (q-1)^g$ belong to Frobenius orbits of size 8 or smaller. It follows that

$$\mathrm{size}_v(\pi^{-1}(y)) \overset{(5.3)}{\leqslant} \mathrm{size}_v(\Upsilon) \leqslant \frac{8 \cdot 2^g (q-1)^g}{\frac{1}{2}(q-1)^{2g}} = \frac{8 \cdot 2^{g+1}}{(q-1)^g} < \frac{1}{\underbrace{(g-1/2)(q-1)}_{d_q} + 1},$$

the last inequality by property (iii) of the prime q. This concludes the proof of (5.4).   □

6. RATIONAL POINTS ON THE BASE OF AN ABELIAN-BY-FINITE FAMILY

In this section we prove Proposition 5.3, which is in essence a variant of Proposition 3.4, and which we rewrite for the reader's convenience.

Proposition 5.3. *Let Y be a curve over K of genus $g \geqslant 2$.*

Let $X \to Y' \overset{\pi}{\to} Y$ be an abelian-by-finite family over Y, with full monodromy (see Definition 5.1 and subsequent discussion). Let d be the relative dimension of $X \to Y'$. Suppose that $X \to Y' \overset{\pi}{\to} Y$ admits a good model over the ring $\mathcal{O}$ of S-integers of K. Let $v \notin S$ be a friendly place of K (Definition 2.7).

Let size_v *be as in* (5.2). *Then the set*

$$Y(K)^* := \left\{ y \in Y(K) : \text{size}_v(\pi^{-1}(y)) < \frac{1}{d+1} \right\}$$

is finite.

Here's what happens in the proof. There are two central lemmas, Lemmas 6.1 and 6.2.

- The assumption that $\text{size}_v(\pi^{-1}(y)) < \frac{1}{d+1}$ guarantees that most points in the fiber $\pi^{-1}(y)$ are defined over fields of large degree over $\mathbf{Q}_p$. As discussed in §1.3, we will use the fact that an extension K_v of $\mathbf{Q}_p$ is of large degree to bound the centralizer of Frobenius for a variety defined over K_v.

   Some care is required with indexing since we only have *most* points; in particular, we need to identify the fibers over p-adically nearby points y. The discussion of indexing occupies the first part of the proof; the bound on the Frobenius centralizer is in the proof of Lemma 6.2.
- Lemma 6.1 handles the possible failure of semisimplicity (see discussion in §1.4). As in Lemma 4.4, we use constraints on Hodge weights coming from global representations (Lemma 2.9) to show that only finitely many fibers can give rise to non-semisimple Galois representations. This requires a general position argument in linear algebra (Lemma 6.4).

Proof. Through the proof, we denote by p the prime of $\mathbf{Q}$ below v; "Tate module" always means "p-adic Tate module," and "étale cohomology" means geometric étale cohomology taken with $\mathbf{Q}_p$ coefficients.

Recall also that we have fixed an algebraic closure $\overline{K}$ with Galois group G_K. Fix an extension of v to that field; the completion of $\overline{K}$ gives an algebraic closure $\overline{K_v}$ of K_v. In particular, if $L \subset \overline{K}$ is unramified at v, we obtain a Frobenius element $\text{Frob}_v \in \text{Gal}(L/K)$.

Fix $y_0 \in Y(K)^*$. It is sufficient to show that there are only finitely many points of $Y(K)^*$ that lie in the residue disk

$$\Omega_v = \{ y \in Y(K_v) : y \equiv y_0 \text{ modulo } v \},$$

which we are regarding as a K_v-analytic manifold.

For each $y \in Y(K)$, let E_y be the ring of regular functions on the zero-dimensional scheme $\pi^{-1}(y)$; this is an étale K-algebra and $\text{Hom}(E_y, \overline{K})$ is identified with the G_K-set $\pi^{-1}(y)_{\overline{K}}$ of preimages of y under π. By our assumptions, the G_K-set $\pi^{-1}(y)_{\overline{K}}$ is unramified at v. Write E_0 for E_{y_0}.

The fiber X_y of $X \to Y$ above $y \in Y(K)$ is *a priori* a K-scheme, but the factorization $X \to Y' \to Y$ gives it the structure of an E_y-scheme; in particular its de Rham cohomology $H^1_{\text{dR}}(X_y/K)$ has the structure of a free E_y-module. Moreover, the polarization on X induces an E_y-*bilinear* symplectic pairing

$$H^1_{\text{dR}}(X_y/K) \times H^1_{\text{dR}}(X_y/K) \longrightarrow E_y.$$

Write $E_{0,v} = E_0 \otimes_K K_v$, and $V_v := H^1_{\text{dR}}(X_{y_0}/K_v)$. Then V_v is a free $E_{0,v}$-module equipped with an ($E_{0,v}$-bilinear) symplectic form. Denote by $\mathcal{H}_v \subset \mathcal{G}_v$ the K_v-schemes defined by Weil restriction:

$$\mathcal{G}_v = \text{Res}_{K_v}^{E_{0,v}} \text{Gr}(V_v, g)$$

$$\mathcal{H}_v = \text{Res}_{K_v}^{E_{0,v}} \text{LGr}(V_v, \omega).$$

Here $\mathrm{Res}_{K_v}^{E_{0,v}}$ denotes Weil restriction of scalars, $\mathrm{Gr}(V_v, g)$ classifies free E_v-submodules of rank g inside V_v, and LGr classifies free rank-g submodules on which the symplectic pairing is trivial.

Then the period map at y_0 gives a K_v-analytic function

$$\Phi_v : \Omega_v \longrightarrow \mathcal{H}_v$$

(see §3.3 for a more detailed discussion of the radius of convergence; in particular it defines a rigid analytic function on a domain containing Ω_v, i.e. the K_v-points in a residue disk).

A priori, this period mapping is valued in a suitable Lagrangian Grassmannian of K_v-linear subspaces inside V_v, but, just as in the discussion of §4.3, each of these Lagrangian subspaces are actually $E_{0,v}$-stable, so that the period mapping actually takes values inside $\mathcal{H}_v$. Lemma 3.3, and the assumption of full monodromy, imply that $\Phi_v(\Omega_v)$ is Zariski-dense in $\mathcal{H}_v$.

To proceed further, as we discussed in the proof sketch, we need to carefully index the points above y. Firstly, E_y decomposes as a product of fields:

$$E_y = \prod_{y'} K(y')$$

where the product is over points y' of the scheme Y' lying above y. For any such y', the fiber $X_{y'}$ of $X \to Y'$ above y' is a d-dimensional abelian variety over the field $K(y')$; write $\rho_{y'}$ for the corresponding $2d$-dimensional p-adic Galois representation of the absolute Galois group of $K(y')$.

The base change $E_y \otimes_K K_v$ splits as a product of fields

$$(6.1) \qquad E_y \otimes_K K_v = \prod_{y',w} K(y')_w$$

indexed by pairs (y', w), where y' is a closed point of $\pi^{-1}(y)$ as above, and w is a place of $K(y')$ over v. In this situation we will say, for short, that (y', w) is above (y, v).

Write $X_{y',w}$ for the base change of X_y along $E_y \to K(y')_w$, and $\rho_{y',w}$ for the $G_{K(y')_w}$-representation on its étale cohomology. The de Rham cohomology $V_v = H^1_{\mathrm{dR}}(X_y/K_v)$ over K_v splits as a product

$$(6.2) \qquad V_v = \prod_{y',w} V_{y',w}, \ \ V_{y',w} = H^1_{\mathrm{dR}}(X_{y',w}/K(y')_w)$$

in a fashion that is compatible with the $E_y \otimes_K K_v$-module structure and (6.1). The dimension of each $V_{y',w}$ over $K(y')_w$ is the same, namely, $2d$.

Crystalline cohomology of the reduction modulo v (or, phrased differently, the Gauss–Manin connection for $Y' \to Y$) gives an isomorphism

$$(6.3) \qquad E_y \otimes_K K_v \xrightarrow{\sim} E_{0,v} = E_0 \otimes_K K_v$$

whenever y belongs to the residue disk Ω of y_0. In particular this induces a bijection

$$(6.4) \qquad (y', w) \text{ above } (y, v) \xleftrightarrow{\sim} (y'_0, w_0) \text{ above } (y_0, v)$$

since both sides are identified with the spectrum of the common algebra of (6.3). Moreover, the identification (6.3) is compatible with the Gauss-Manin isomorphism

$$(6.5) \qquad H^1_{\mathrm{dR}}(X_y/K_v) \xrightarrow{\mathrm{GM}} H^1_{\mathrm{dR}}(X_{y_0}/K_v).$$

If (y', w) corresponds to (y'_0, w_0) under this identification, then (6.3) and (6.5) induce

$$(6.6) \qquad K(y')_w \simeq K(y'_0)_{w_0}, \ \ H^1_{\mathrm{dR}}(X_{y',w}/K(y')_w) \simeq H^1_{\mathrm{dR}}(X_{y'_0,w_0}/K(y'_0)_w).$$

Also, (6.2) induces the splitting of the variety $\mathcal{H}_v$ as a product

$$\mathcal{H}_v = \prod_{(y_0', w)} \mathcal{H}_{(y_0', w)},$$

where the product is taken over (y_0', w) above (y_0, v), and where

$$\mathcal{H}_{(y_0', w)} = \operatorname{Res}_{K_v}^{K(y_0')_w} \operatorname{LGr}(V_{y_0', w}, \omega).$$

We have a similar decomposition $\mathcal{G}_v = \prod_{(y_0', w)} \mathcal{G}_{(y_0', w)}$.

If $y \in \Omega_v$, and if (y', w) above (y, v) corresponds to (y_0', w_0) under (6.4), then

(6.7) \qquad projection to $\mathcal{H}_{(y_0', w_0)}$ of $\Phi_v(y) = F^1 H_{\mathrm{dR}}^1(X_{y', w})$.

where we identify $F^1 H_{\mathrm{dR}}^1(X_{y', w})$ with a Lagrangian in the $K(y_0')_{w_0}$-vector space V_{y_0', w_0} using the Gauss–Manin connection (6.6). This result (6.7) comes down to the fact, already noted, that (6.4) and (6.5) are compatible.

We will establish the following two lemmas.

Lemma 6.1 (Generic Simplicity). *There is a finite subset $F \subset \Omega_v \cap Y(K)^*$ such that, for $y \in (\Omega_v \cap Y(K)^*) - F$, there exists (y', w) above (y, v) such that:*

(i) $[K(y')_w : K_v] \geqslant 8$
(ii) $\rho_{y'}$ *is simple as a $G_{K(y')}$-representation.*

Observe that for $y \in (\Omega_v \cap Y(K)^*) - F$, and y' above y, there are only finitely many possibilities for the isomorphism class of the field $K(y')$. Thus, by Lemma 2.3, there are only finitely many possibilities for the isomorphism class of the pair $(K(y)', \rho_{y'})$, and so also only finitely many possibilities for the isomorphism class of any pair $(K(y')_w, \rho_{y'}|_{K(y)_w'})$ arising from (y', w) as in Lemma 6.1. The proof of Proposition 5.3 will then follow from Lemma 6.1 above and Lemma 6.2 below. □

Lemma 6.2 (Galois representations really do vary in our family). *Fix a finite field extension K_v' of K_v, with $[K_v' : K_v] \geqslant 8$, and a Galois representation ρ' of the absolute Galois group of K_v'.*

There are only finitely many $y \in \Omega_v \cap Y(K)$ for which there exist (y', w) satisfying conditions (i) and (ii) of Lemma 6.1 and moreover the pair

$$(K(y')_w, \rho_{y', w}) \text{ is isomorphic to } (K_v', \rho')$$

i.e. there is an isomorphism $K(y)_w' \to K_v'$ carrying the isomorphism class of ρ' to that of $\rho_{y', w}$.

To prove Lemmas 6.1 and 6.2 we shall analyze the period mapping more carefully.

Proof of Lemma 6.2. Under the correspondence of p-adic Hodge theory, $\rho_{y', w}$ corresponds to the $K(y')_w$-vector space $H_{\mathrm{dR}}^1(X_{y'}/K(y')_w)$, together with its natural semilinear Frobenius operator ϕ, and the (two-step) filtration defined by $F^1 H_{\mathrm{dR}}(X_{y'}/K(y')_w)$.

Suppose that (y', w) corresponds to (y_0', w_0) under (6.4). Using the isomorphism (6.6) the triple just described corresponds to

$$(H_{\mathrm{dR}}^1(X_{y_0'}/K(y_0')_{w_0}), \phi_v = \text{semilinear Frobenius, projection of } \Phi_v(y) \text{ to } \mathcal{H}_{y_0', w_0}).$$

It is enough to show that the set of y, for which this triple belongs to a fixed isomorphism class, is finite.

Belonging to a fixed isomorphism class means that the projection of $\Phi_v(y)$ to $\mathcal{H}_{y_0', w_0}$ lies inside a single orbit for the action of the Frobenius centralizer $Z(\phi_v)$ on $\mathcal{G}_{y_0', w_0}$, and

so also a single orbit of $Z(\phi_v^{[K_v:\mathbf{Q}_p]})$ on $\mathcal{G}_{y_0',w_0}$. (In both cases, these centralizers are taken inside $K(y_0')_{w_0}$-linear automorphisms of $V_{y_0',w}$.)

Apply Lemma 2.1 to the field extension $K(y_0')_w/K_v$ to see that this Frobenius centralizer has K_v-dimension at most $(\dim_{K(y_0')_{w_0}} V_{y_0',w_0})^2 = 4d^2$.

As noted earlier, the period map Φ_v has Zariski-dense image (in the K_v-variety $\mathcal{H}_v$; therefore this remains true when projected to $\mathcal{H}_{y_0',w_0}$). Since $\dim_{K_v} \mathcal{H}_{y_0',w_0} = [K(y')_w : K_v] \cdot \frac{d(d+1)}{2} \geqslant 4d(d+1) > 4d^2$, Lemma 3.3 completes the proof of Lemma 6.2. $\qquad\square$

Proof of Lemma 6.1. Let us call $y \in Y(K)^* \cap \Omega_v$ "bad" when, for every (y',w) above (y,v) such that $[K(y')_w : K_v] \geqslant 8$, the representation $\rho_{y'}$ fails to be simple. We must show there are only finitely many bad $y \in Y(K)^* \cap \Omega_v$.

> **Sublemma:** If $y \in Y(K)^* \cap \Omega_v$ is bad, there exists:
> - (y', w) above (y,v), with $[K(y')_w : K_v] \geqslant 8$;
> - a nonzero proper Frobenius-stable subspace $W_{y',w}^{\mathrm{dR}}$ of $H_{\mathrm{dR}}^1(X_{y'}/K(y')_w)$ such that $\dim F^1 W_{y',w}^{\mathrm{dR}} \geqslant \dim(W_{y',w}^{\mathrm{dR}})/2$. (Here, and in the discussion below, dimensions are dimensions over $K(y')_w$.)

Proof of sublemma: Take a bad $y \in Y(K)^* \cap \Omega_v$. For each y' above y let $W_{y'}$ be a nonzero subrepresentation of $\rho_{y'}$ of minimal positive dimension. (It is therefore possible that $W_{y'}$ is all of $\rho_{y'}$). For each place w of $K(y')$ we define $W_{y',w}^{\mathrm{dR}}$ by applying p-adic Hodge theory to $W_{y'} \leqslant \rho_{y'}$; thus $W_{y',w}^{\mathrm{dR}}$ is a ϕ-stable submodule of $H_{\mathrm{dR}}^1(X_{y'}/K(y')_w)$.

Note that

$$(6.8) \qquad \dim W_{y',w}^{\mathrm{dR}} \leqslant d \text{ whenever } [K(y')_w : K_v] \geqslant 8.$$

Indeed because y' is bad, the supposition $[K(y')_w : K_v] \geqslant 8$ forces $\rho_{y'}$ to be non-simple; because it preserves (up to similitude) a bilinear form, we have $\dim W_{y'} \leqslant \frac{1}{2} \dim \rho_{y'}$, thus (6.8).

Now assume that, for each (y',w) above (y,v), satisfying $[K(y')_w : K_v] \geqslant 8$, we have

$$\dim F^1 W_{y',w}^{\mathrm{dR}} < \frac{1}{2} \dim W_{y',w}^{\mathrm{dR}}.$$

We will derive a contradiction, which will conclude the proof.

By Lemma 2.10, applied to $W_{y'}$ as a Galois representation of $K(y')$, we have

$$(6.9) \qquad \sum_{w|v} [K(y')_w : K_v] \frac{\dim F^1 W_{y',w}^{\mathrm{dR}}}{\dim W_{y',w}^{\mathrm{dR}}} = \frac{1}{2}[K(y') : K]$$

for any y' a closed point of $\pi^{-1}(y)$. Sum over y' above y; using (6.8) we get

$$(6.10)$$
$$\sum_{[K(y')_w:K_v]\geqslant 8} [K(y')_w : K_v] \left(\frac{1}{2} - \frac{1}{2d}\right) + \sum_{[K(y')_w:K_v]<8} [K(y')_w : K_v] \geqslant \frac{1}{2} \sum_{(y',w)} [K(y')_w : K_v]$$

Here all summations are over (y',w) above (y,v). Therefore,

$$(6.11) \qquad \sum_{[K(y')_w:K_v]<8} \frac{1}{2}[K(y')_w : K_v] \geqslant \frac{1}{2d} \sum_{[K(y')_w:K_v]\geqslant 8} [K(y')_w : K_v].$$

Let $e_1, \ldots, e_k$ be the cycle structure of Frob_v acting on the $\bar{K}$ points of $\pi^{-1}(y)$. The inequality above means that

$$\frac{1}{2} \sum_{i : e_i < 8} e_i \geqslant \frac{1}{2d} \sum_{i : e_i \geqslant 8} e_i,$$

which is to say that $\mathrm{size}_v(\pi^{-1}y) \geqslant \frac{1}{d+1}$. This contradicts the assumption that $y \in Y(K)^*$. $\qquad\square$

We now return to the proof of Lemma 6.1. Fix any (y_0', w) above (y_0, v) with $[K(y_0')_w : K_v] \geqslant 8$. Such a (y_0', w) exists because of the assumption that $y_0 \in Y(K)^*$. In view of the *Sublemma* and (6.7), it is enough to show that there are only finitely many $y \in Y(K) \cap \Omega_v$ such the projection of $\Phi_v(y)$ to $\mathcal{H}_{(y_0', w)}$ lies in the subvariety

$$\mathcal{H}^{\mathrm{bad}}_{(y_0', w)} \subset \mathcal{H}_{(y_0', w)}$$

defined as the Lagrangian, $K(y_0')_w$-subspaces $F \subset V_{y_0', w}$ (recall (6.2) for definition) for which there exists a Frobenius-stable subspace $W \subset V_{y_0', w}$, satisfying

$$(6.12) \qquad\qquad \dim(F \cap W) \geqslant \frac{1}{2} \dim(W),$$

By the lemmas that follow, $\mathcal{H}^{\mathrm{bad}}_{(y_0', w)}$ is contained in a proper closed K_v-subvariety of $\mathcal{H}_{(y_0', w)}$; we conclude as in the proof of Lemma 6.2. $\qquad\square$

Lemma 6.3. *Suppose L_w is a finite unramified extension of K_v of degree $r \geqslant 8$. Let (V, ω) be a symplectic L_w-vector space, with $\dim_{L_w} V = 2d$; let $\phi : V \to V$ be semilinear for the Frobenius automorphism of L_w/K_v and bijective.*

Then there is a Zariski-open

$$\mathcal{A} \subseteq \mathrm{Res}^{L_w}_{K_v} \mathrm{LGr}(V, \omega)$$

(where $\mathrm{LGr}(V, \omega)$ is the Lagrangian Grassmannian, and $\mathrm{Res}^{L_w}_{K_v}$ denotes Weil restriction of scalars from L_w to K_v) with the following property:

If $F \subset V$ is a Lagrangian L_w-subspace, corresponding to a point of $\mathcal{A}(K_v)$, there is no ϕ-invariant L_w-subspace W of V satisfying (6.12).

Proof. Just as in Lemma 2.1, $V \otimes_{K_v} \overline{K_v}$ splits into $2d$-dimensional spaces $V_1, \ldots, V_r$ indexed by embeddings $L_w \hookrightarrow \overline{K_v}$; we can order them so that ϕ induces isomorphisms $V_i \simeq V_{i+1}$ for $1 \leq i \leq r - 1$, and thus can identify them all with V_1 (we do not use the "cyclic" isomorphism $V_r \simeq V_1$).

The base extension $W \otimes_{K_v} \overline{K_v}$ of any ϕ-invariant L_w-subspace yields a subspace $\bigoplus W_i \leqslant \bigoplus V_i$, where each W_i corresponds to W_1 under the above identifications. Similarly, the base extension of a Lagrangian L_w-subspace $F \leqslant V$ gives an subspace $\bigoplus F_i \leqslant \bigoplus V_i$, where each F_i is Lagrangian. If (6.12) is satisfied, then $\dim(F_i \cap W_i) \geqslant \frac{1}{2} \dim(W_i)$ for each $1 \leqslant i \leqslant r$.

The next, and final, Lemma shows that the set of $(F_1 \ldots, F_r)$ for which such a W exists is a proper, Zariski-closed subset. Thus there is a Zariski-open set inside in

$$\left(\mathrm{Res}^{L_w}_{K_v} \mathrm{LGr}(V, \omega) \right) \times_{K_v} \overline{K_v}$$

such that, if F belongs to this Zariski-open, it has the property quoted in the statement. Taking the intersection of Galois conjugates of this set, we get the desired Zariski-open inside $\mathrm{Res}^{L_w}_{K_v} \mathrm{LGr}(V, \omega)$. $\qquad\square$

Lemma 6.4. *Let (V, ω) be a symplectic vector space over a field of characteristic zero with $\dim(V) = 2d$; write $\mathrm{LGr}(V, \omega)$ for the Grassmannian of Lagrangian subspaces. Let* E *be the set of r-tuples of Lagrangian subspaces*

$$(F_1, \ldots, F_r) \in \mathrm{LGr}(V, \omega)^r$$

for which there exists a proper nonzero subspace $W \subset V$ such that $\dim(F_j \cap W) \geqslant \frac{1}{2} \dim(W)$ for every j. If $r \geqslant 8$ then E *is contained in a proper, Zariski-closed subset of $\mathrm{LGr}(V, \omega)^r$.*

Proof. In fact our argument will show that $r \geqslant 5$ is enough.

First we argue that E is Zariski-closed. Consider the product $\mathrm{Gr}(V) \times \mathrm{LGr}(V, \omega)^r$ parametrizing tuples $(W, F_1, F_2, \ldots, F_r)$ such that each F_i is Lagrangian. For each i, the dimension $\dim F_i \cap W$ is (Zariski) upper semicontinuous; so the set $\tilde{\mathrm{E}}$ of tuples satisfying the conditions described is closed. Now E is the image of the closed set $\tilde{\mathrm{E}}$ under a proper map, so it is itself closed.

Since E is closed it's enough to produce a single tuple $(F_1, \ldots, F_r)$ not in E.

Take $e_1, \ldots, e_d, e'_1, \ldots, e'_d$ a standard symplectic basis for V, so $\langle e_i, e'_i \rangle = 1$, and $\langle e'_i, e_i \rangle = -1$, and all other pairings between basis vectors are zero. Let

$$
\begin{aligned}
F_1 &= \mathrm{span}(e_1, e_2, \ldots, e_d) \\
F_2 &= \mathrm{span}(e'_1, e'_2, \ldots, e'_d) \\
F_3 &= \mathrm{span}(e_1 + e'_1, e_2 + e'_2, \ldots, e_d + e'_d) \\
F_4 &= \mathrm{span}(e_1 + 2e'_1, e_2 + 4e'_2, \ldots, e_d + 2de'_d).
\end{aligned}
$$

Now each of these four spaces is maximal isotropic, and any two of them have trivial intersection.

Write $\pi_{12} : V \to F_1$ for the projection along the decomposition $V = F_1 \oplus F_2$, and similarly define $\pi_{21} : V \to F_2$. Both π_{12} and π_{21} are isomorphisms when restricted to either F_3 or F_4. Write $\Phi_{12;3} : F_1 \to F_2$ for the isomorphism

$$F_1 \overset{\pi_{12}^{-1}}{\longleftarrow} F_3 \overset{\pi_{21}}{\longrightarrow} F_2.$$

In explicit coordinates $\Phi_{12;3}$ takes e_i to e'_i, and the similar map $\Phi_{12;4}$ takes e_i to $2ie'_i$.

We claim that only finitely many W can satisfy the condition stated in the Lemma with respect to F_1, F_2, F_3, F_4. Suppose given such a W. Since $W \cap F_1$ and $W \cap F_2$ have trivial intersection with each other, and they each have dimension at least $\frac{1}{2} \dim(W)$, we have a direct sum decomposition

$$(6.13) \qquad\qquad\qquad W = (W \cap F_1) \oplus (W \cap F_2)$$

and an equality $\dim(W \cap F_1) = \dim(W \cap F_2) = \frac{1}{2} \dim W$. Similarly, we find that $\dim(W \cap F_3) = \dim(W \cap F_4) = \frac{1}{2} \dim W$.

Next π_{12} gives an isomorphism $F_3 \to F_1$; comparing dimensions, we see the restriction

$$\pi_{12} : W \cap F_3 \overset{\sim}{\longrightarrow} W \cap F_1$$

is an isomorphism as well. Similarly $\pi_{21} : W \cap F_3 \overset{\sim}{\longrightarrow} W \cap F_2$.

In particular, $\Phi_{12,3}$ carries $W \cap F_1$ isomorphically to $W \cap F_2$. The same reasoning applies to $\Phi_{12,4}$. Therefore, $W \cap F_1$ is stable under $\Phi_{12,4}^{-1} \Phi_{12,3}$, which shows that $W \cap F_1 \subseteq F_1$ is stable under the map $e_i \mapsto 2ie_i$.

There are then finitely many possibilities for $W \cap F_1$; then there are also finitely many possibilities for $W \cap F_2 = \Phi_{12,3}(W \cap F_1)$ and then by (6.13) finitely many possibilities for W; call them $W_1, \ldots, W_N$.

Now, for each W_i, the condition that $\dim(F_5 \cap W_i) \geqslant \frac{1}{2}\dim(W_i)$ cuts out a proper Zariski-closed subset of the Lagrangian Grassmannian parametrizing F_5; thus we may choose F_5 so that no W satisfies the dimension bound. $\square$

7. THE KODAIRA–PARSHIN FAMILY

The argument that we have given for Mordell's conjecture in Section 5 made use of a specific abelian-by-finite family, the Kodaira–Parshin family. In this section we explain how to construct this family, making use (in effect) of an algebraic version of the theory of Hurwitz spaces. We need this theory only in characteristic zero.

7.1. **Hurwitz spaces for curves.**

Proposition 7.1. *Let Y be a curve of genus at least 2 over a number field K, and let G be a center-free finite group. Then there is a K-curve Y' equipped with an étale map $\pi : Y' \to Y$, and a relative curve $Z \to Y'$, with the following properties:*

(i) "Y' parameterizes G-covers of Y branched at a single point": For $y \in Y(\bar{K})$, there is a bijection between $\pi^{-1}(y)$ and the set of G-conjugacy classes of surjections $\pi_1^{\mathrm{geom}}(Y-y,*) \twoheadrightarrow G$ nontrivial on a loop around y. Moreover, if $y \in Y(K)$, this identification is G_K-equivariant.

(ii) "Z gives the universal G-cover of Y branched at a single point": There is a morphism $Z \to Y' \times Y$ of relative curves over Y' (here, we are regarding $Y' \times Y$ as the trivial family of curves over Y', with fiber Y everywhere).

Moreover G acts on Z covering the trivial action on $Y' \times Y$. This action makes $Z \to Y' \times Y$ into a G-covering away from the graph of π. If we take the fiber of this morphism of relative curves above $y' \in Y'(\bar{K})$, the resulting map $Z_{y'} \to Y$ of curves is ramified exactly at $\pi(y')$. The induced homomorphism

$$\pi_1^{\mathrm{geom}}(Y - y', y_0) \to \mathrm{Aut}_G(Z_{(y',y_0)}) \cong G$$

is exactly (in the conjugacy class of) the surjection from (i) classified by y'.

There are several references on this matter that address much more general settings (e.g. [30, §3.22]) but since none of them give the precise statement we need, we will simply outline a direct proof, descending from the complex analytic analogue, in §7.3.

Now we apply this to the group $G = \mathrm{Aff}(q)$:

Definition 7.2. *Let Y be a curve of genus at least 2 over a number field K, and let q be a prime number. The* Kodaira-Parshin curve family *over Y with parameter q will be the sequence of morphisms*

$$(7.1) \qquad\qquad\qquad Z_q \longrightarrow Y'_q \longrightarrow Y,$$

obtained from Proposition 7.1 applied to the group $G = \mathrm{Aff}(q)$.

We now want to form an associated abelian-by-finite family to the Kodaira–Parshin curve family.

7.2. **Prym varieties.** We first describe the situation fiberwise:

Given a morphism $C_1 \to C_2$ of curves over an algebraically closed field, the associated *Prym variety* is the cokernel of the induced map $\mathrm{Pic}^0(C_2) \to \mathrm{Pic}^0(C_1)$ on Jacobians.

Now suppose that that the covering $C_1 \to C_2$ is Galois, with Galois group $\mathrm{Aff}(q)$, and ramified over exactly one point of C_2. The degree of this covering is $q(q-1)$. Rather than take its Prym directly, however, we prefer to use a reduced version. Namely, we can form

a smaller degree-q covering $C_1' \to C_2$ using the permutation action of $\mathrm{Aff}(q)$ on $\mathbf{Z}/q\mathbf{Z}$, and we are interested in the Prym variety of this associated covering:

$$(7.2) \qquad\qquad \mathrm{coker}(\mathrm{Pic}^0(C_2) \to \mathrm{Pic}^0(C_1'))$$

We emphasize again that this is *not* the Prym variety of $C_1 \to C_2$ but a "reduced" version of it where the role of C_1 has been replaced by C_1'.

We can reformulate this in terms of C_1, rather than the associated curve C_1'. Both $\mathrm{Pic}^0(C_2)$ and $\mathrm{Pic}^0(C_1')$ map to $\mathrm{Pic}^0(C_1)$, with finite kernel. The image of $\mathrm{Pic}^0(C_2)$ in $\mathrm{Pic}^0(C_1)$ is now the connected component of the $\mathrm{Aff}(q)$-invariants; similarly the image of $\mathrm{Pic}^0(C_1')$ in $\mathrm{Pic}^0(C_1)$ is the connected component of the invariants by the subgroup $H_q = (\mathbf{Z}/q\mathbf{Z})^*$, which is a point stabilizer in the permutation action of $\mathrm{Aff}(q)$ on $\mathbf{Z}/q\mathbf{Z}$. In summary, then, the Prym variety of $C_1' \to C_2$ is isogenous to cokernel of the map

$$\text{connected component of } \mathrm{Pic}^0(C_1)^{G_q} \to \text{connected component of } \mathrm{Pic}^0(C_1)^{H_q}.$$

This is an abelian variety of dimension $(2g-1) \cdot \frac{q-1}{2}$, isogenous to (7.2).

We may alternately describe this as follows: Form the idempotent

$$e := \frac{1}{\#H_q} \sum_{h \in H_q} h - \frac{1}{\#\mathrm{Aff}(q)} \sum_{g \in \mathrm{Aff}(q)} g \in \mathbf{Q}[\mathrm{Aff}(q)]$$

and let $e' = 1 - e$ be the complementary idempotent. Then $e'' := \#\mathrm{Aff}(q) \cdot e' \in \mathbf{Z}[\mathrm{Aff}(q)]$ acts on $\mathrm{Pic}^0(C_1)$, and moreover the connected component of its kernel is isogeneous to the Prym variety described above:

$$(7.3) \qquad \text{connected component of } \mathrm{Pic}^0(C_1)[e''] \xrightarrow{\mathrm{isog.}} (\text{Prym for } C_1' \to C_2).$$

Equation 7.3 gives a way to access the 'reduced' Prym variety (at least up to isogeny) that we can conveniently apply in our relative situation: In the situation described in Definition 7.2, $Z_q \to Y_q'$ is a relative curve over Y_q' and it admits a $\mathrm{Aff}(q)$-action, where $\mathrm{Aff}(q)$ acts trivially on the base. The relative Picard scheme of this curve is an abelian scheme over Y_q' equipped with a symmetric and fiberwise ample line bundle. Thus we may form

$$X_q = \text{ relative identity component of } \mathrm{Pic}^0_{Z_q \to Y_q'}[e''],$$

where $[e'']$ means the kernel of e'', and for the notion of "relative identity component," see [13, Proposition 15.6.4]. This X_q is an abelian scheme over Y_q', equipped with a symmetric and fiberwise ample line bundle; its fiber over any $y \in Y_q'(\bar{K})$ coincides with the construction on the left hand side of (7.3); in particular this fiber is isogenous to the reduced Prym variety of the associated $\mathrm{Aff}(q)$-covering $Z_y \to Y$.

Definition 7.3. *Notation as in the prior definition. The* Kodaira–Parshin family of Jacobians *over Y, associated to the group* $\mathrm{Aff}(q)$*, is the sequence of morphisms*

$$X_q \longrightarrow Y_q' \to Y,$$

where X_q is, as defined above, the reduced relative Prym of $Z \to Y_q' \times Y$, considered as a morphism of relative cuves over Y_q'.

This is an abelian-by-finite family, in the sense of Definition 5.1.

7.3. **Proof of Proposition 7.1.** We give the proof of Proposition 7.1. As we have mentioned this is largely for lack of a good reference which states precisely what we need; certainly much more general statements about Hurwitz schemes exist in the literature.

We start by supposing that Y is a proper smooth curve over $\mathbf{C}$; while we work over $\mathbf{C}$ we identify Y with its complex points.

For $y \in Y$ set $S(y)$ to be the set of conjugacy classes of surjective homomorphisms from $\pi_1(Y - y, *) \twoheadrightarrow G$, with the property that a loop around y has nontrivial image. Equivalently, $S(y)$ is the finite set of isomorphism classes of connected coverings of Y with Galois group G, branched precisely at y.

For y near y^* there is a natural identification $S(y) \cong S(y^*)$ since we can topologically identify (Y, y) and (Y, y^*). Thus the set $\coprod_{y \in Y(\mathbf{C})} S(y)$ has the structure of a Riemann surface Y' equipped with a covering map $e : Y' \to Y$. Explicitly, for each $y' \in Y'$, we have $y' \in S(e(y'))$, or in words: y' classifies a connected G-covering of Y branched at $y = e(y')$.

Moreover, the coverings indexed by the elements of $S(y)$ fit together to a morphism

$$f : Z \to Y' \times Y$$

of smooth complex manifolds; here G acts on Z, covering the trivial action on $Y' \times Y$. More explicitly:

- f is a covering map and a G-torsor when restricted to the complement of the analytic divisor

$$\Delta := \text{graph of } e \subset Y' \times Y$$

- the pullback of the above morphism along $y' \times Y \hookrightarrow Y' \times Y$ (for $y' \in Y'$) is isomorphic to the covering of Y classified by y'.

Near the preimage of Δ on Z the map looks in local coordinates like $(z, w) \mapsto (z, w^n)$ for suitable n.

Now everything can be algebraized, i.e. Z and Y' have unique structures of complex algebraic variety compatible with their analytic structures, and the G-action on Z as well as the morphisms $Z \to Y' \times Y$ and $Y' \to Y$ are algebraic. This is clear for Y'; also the the structure sheaf of Z defines a coherent analytic sheaf on $Y' \times Y$ which can be made algebraic by GAGA ([40, Theorem 3]); similarly the algebra structure on this coherent analytic sheaf comes from an algebra structure on the algebraic sheaf [40, Theorem 2].

We now switch to using the letters $Z, Y, \ldots$ for the complex algebraic varieties, rather than the associated analytic spaces. So we have defined a sequence of complex algebraic varieties

$$(7.4) \qquad Z \xrightarrow{f} Y' \times Y \xrightarrow{e \times \text{id}} Y \times Y$$

where f is étale away from the graph of e, and e is étale; the composite $Z \to Y^2$ is therefore étale away from the diagonal Δ. (Note that it is equivalent to check étale in the algebraic and analytic settings, see [1, XII, §3]).

Now suppose that Y is actually defined over a subfield $K \subset \mathbf{C}$; we denote by Y_K the corresponding K-scheme (similarly $(Y^2)_K$, etc.); we want now to descend everything in sight to K.

Lemma 7.4. *Write Z° for the preimage of $Y^2 - \Delta$ in Z.*

(1) The étale cover $F : Z^\circ \to Y^2 - \Delta$ can be uniquely extended to a cover $F_K : Z_K^\circ \to (Y^2 - \Delta)_K$. (In both cases, these étale covers are understood to be equipped with G-action.)

(2) Let $(y_1, y_0) \in Y(\bar{K})^2$, with $y_1 \neq y_0$. The geometric fiber

$$F_K^{-1}(y_1, y_0)/G$$

is identified with the set $S(y_0)$, as defined above, now using étale $\pi_1^{\mathrm{geom}}(Y - y_0, y_1)$. If $(y_1, y_0) \in Y(K)^2$ this identification is equivariant for G_K.

(3) The quotient Z_K°/G (which is étale over $(Y^2 - \Delta)_K$) extends uniquely to an étale cover of Y_K^2. This cover is isomorphic to one of the form $Y_K' \times Y_K \to Y_K^2$ for an étale cover $Y_K' \to Y_K$, such that Y' is the base change of Y_K' to $\mathbf{C}$.

Assume Lemma 7.4 (the proof, which will be given in a moment, will involve only the theory of étale π_1 and group theory). It produces a sequence $Z_K^\circ \to Y_K' \times Y_K \to Y_K^2$; we need to extend Z_K° to a K-structure on all of Z, and extend the first map accordingly.

Let $Z_K \to Y_K^2$ be the normalization of Y_K^2 inside the fraction field of Z_K°. Then Z_K is normal, and finite over Y_K^2. The base extension $Z_K \otimes_K \mathbf{C}$ is therefore also normal (the extension of a normal scheme along a field extension in characteristic zero is normal – see [44, Tag 037Z] or [20, Cor. 6.14.2]), and it is finite over Y^2. Consequently, $Z_K \otimes_K \mathbf{C}$ coincides with the normalization of Y^2 in the function field of Z°. This latter normalization is identified with Z, for Z is also normal and finite over Y^2.

The morphism $Z_K^\circ \to Z_K^\circ/G \to Y_K' \times Y_K$ now extends to $Z_K \to Y_K' \times Y_K$, and the other desired properties can be verified since they are true over $\mathbf{C}$.

Proof of Lemma 7.4. We do this by means of the theory of the étale fundamental group. We first formulate the basic point in purely group theoretic terms.

Let Γ, G be groups, with G finite center-free, and c a conjugacy class of morphisms in $\mathrm{Hom}(\Lambda, \Gamma)$ for some other group Λ; when we apply this, Γ will be a π_1 of a punctured curve, Λ will be the profinite completion of an infinite cyclic group, and c will come from monodromy around the puncture. Consider the set $S = S(\Gamma, c, G)$ of all surjective homomorphisms $\varphi : \Gamma \to G$, with the property that they are nontrivial when pulled back by c. There are natural commuting actions of Γ and G on S:

$$\gamma \cdot \varphi = \varphi \circ \mathrm{Ad}(\gamma)^{-1} \ (\gamma \in \Gamma), \ \ \varphi \cdot h = \mathrm{Ad}(h^{-1}) \circ \varphi \ (h \in G).$$

where we've written $\mathrm{Ad}(x)$ for the automorphism $g \mapsto xgx^{-1}$.

This Γ-action extends *uniquely* to an action (commuting with G) of any overgroup $\widetilde{\Gamma} \supset \Gamma$ in which Γ is normal and whose conjugation action preserves c. Indeed the extension is described by exactly the same formula; uniqueness comes from the fact that the stabilizer of $\varphi \in S(\Gamma, c, G)$ inside $\Gamma \times G^{\mathrm{op}}$ is given by

$$\{(\gamma \in \Gamma, h \in G) : h^{-1} = \varphi(\gamma)\},$$

and so φ is determined by its stabilizer in $\Gamma \times G^{\mathrm{op}}$. (We used that φ is surjective and that G is center-free.)

We apply this as follows. As above, fix two points $y_0 \neq y_1 \in Y(\mathbf{C})$; we will use $\mathbf{y} = (y_1, y_0)$ as a geometric basepoint for $Y \times Y$. Consider the sequence of pointed schemes:

$$(7.5) \qquad \underbrace{(Y - \{y_0\}, y_1)}_{\Gamma := \pi_1} \overset{p \mapsto (p, y_0)}{\longrightarrow} \underbrace{(Y^2 - \Delta, \mathbf{y})}_{\widetilde{\Gamma}^{\mathrm{geom}} := \pi_1} \overset{(y, y') \mapsto y'}{\longrightarrow} (Y, y_0)$$

and let $\Gamma, \widetilde{\Gamma}^{\mathrm{geom}}$ be defined as the geometric étale π_1 of the first and second spaces, at the specified basepoints. Now the long exact sequence for homotopy groups of a fibration gives rise to an exact sequence of *topological* fundamental groups; in the setting at hand this is short exact because the π_2 of $Y - \{y_0\}$ vanishes. The corresponding sequence of

geometric étale fundamental groups is obtained by profinite completion; it remains exact by the results of [38]. Therefore the first map above identifies Γ with a normal subgroup of $\tilde{\Gamma}^{\mathrm{geom}}$. It follows easily that, if we write

$$\tilde{\Gamma} = \pi_1((Y^2 - \Delta)_K, \mathbf{y}),$$

(arithmetic fundamental group) then the map $\Gamma \to \tilde{\Gamma}^{\mathrm{geom}}$ identifies Γ to a normal subgroup of $\tilde{\Gamma}$.

Now let $S = S(\Gamma, c, G)$ be as above, where c is the conjugacy class of maps $\widehat{\mathbf{Z}} \to \pi_1(Y - \{y_0\}, y_1) = \Gamma$ arising from the monodromy around y_0. The commuting $\Gamma \times G$ actions on S define a cover of $Y - \{y_0\}$, equipped with an action of G by automorphisms, whose fiber at y_1 is identified with S. This cover may be described as follows: it is the disjoint union of all the connected G-covers of Y branched precisely at y_0. In other words, it is the restriction of $Z \to Y^2 - \Delta$ to the fiber $\{y_0\} \times (Y - \{y_0\})$. From the uniqueness just described, the extension of this $\Gamma \times G$ action on S to an action of $\widetilde{\Gamma}^{\mathrm{geom}} \times G$ corresponds to the cover $Z \to Y^2 - \Delta$. Therefore, the (further) unique extension of the $\Gamma \times G$-action on S to $\widetilde{\Gamma} \times G$ gives the statement (1) in the Claim.

Statement (2) of the *Claim* (and the G_K-equivariance if y_1, y_0 are K-rational) follows for the specific (y_1, y_0) chosen above; however, since we showed that the K-structure on Z° is unique, it must also be true for any choice of (y_1, y_0).

For statement (3) we notice that the action of Γ on $S(\Gamma, c, G)/G$ is in fact trivial. Therefore the resulting action of $\widetilde{\Gamma}$ factors through the quotient $\pi_1(Y_K, y_0)$ arising from the last map of (7.5). This amounts to the third assertion. $\qquad\square$

8. The monodromy of Kodaira–Parshin families

8.1. **Introduction, Notation, Statement of Main Theorem.** In this section we consider surfaces in the classical topological category: by a "surface" we mean the complement of finitely many interior points inside a connected, orientable, compact two-dimensional manifold with boundary. Thus a surface can have both boundary and punctures. Throughout this section, the letters Y and Z will denote such a surface, and we will use y_0 to denote a base point on Y. For such a surface Y, $\mathrm{MCG}(Y)$ denotes the mapping class group of Y. To emphasize, Y could have "punctures" or boundary. The book of Farb and Margalit [17] is a reference on this material that contains all the results we will use. When we discuss homology or cohomology, the coefficients are always assumed to be the rational numbers $\mathbf{Q}$ unless stated otherwise.

We first reformulate the statement to be proven.

8.2. **Covers and their homology.** Let Y be a surface (possibly with punctures or boundary). An $\mathrm{Aff}(q)$-cover of Y is, by definition, a connected surface Z together with a degree q covering map

$$\pi : Z \longrightarrow Y$$

whose monodromy representation on a general fiber is equivalent to the action of $\mathrm{Aff}(q)$ on $\mathbf{F}_q$ (i.e. we can label points in the fiber by $\mathbf{F}_q$ in such a way that the monodromy representation has image $\mathrm{Aff}(q)$). We will often abuse notation and refer to this cover simply as Z, i.e., regard the map π as implicit.

After choice of basepoint $y_0 \in Y$, such a cover determines an $\mathrm{Aff}(q)$-conjugacy class[5] of maps

$$(8.1) \qquad\qquad \pi_1(Y, y_0) \twoheadrightarrow \mathrm{Aff}(q).$$

We define two $\mathrm{Aff}(q)$-covers (Z_1, π_1) and (Z_2, π_2) to be *isomorphic* when there is a homeomorphism $Z_1 \simeq Z_2$ commuting with the projections to Y; equivalently, when the associated conjugacy classes of π_1-representations (8.1) coincide.

If we have fixed a $\mathrm{Aff}(q)$-cover $Z \to Y$, we denote by $\mathrm{Cov} : \pi_1 \to \mathrm{Aff}(q)$ any homomorphism in the conjugacy class of (8.1). For $\eta \in \pi_1$ we can unambiguously talk about the *cycle decomposition* of $\mathrm{Cov}(\eta)$ in $\mathrm{Sym}(\mathbf{F}_q)$, which we regard as a partition of the positive integer q; this cycle decomposition is conjugation-invariant.

Given any covering map $\pi : Z \to Y$, the pullback and pushforward on homology define a splitting

$$H_1(Z, \mathbf{Q}) = \pi^* H_1(Y, \mathbf{Q}) \oplus \underbrace{H_1^{\mathrm{Pr}}(Z, Y; \mathbf{Q})}_{\ker(\pi_* : H_1(Z) \to H_1(Y)).}$$

Henceforth we will drop the coefficients $\mathbf{Q}$ from the notation. The symbol Pr stands for primitive; alternatively, $H_1^{\mathrm{Pr}}(Z, Y)$ is the homology of a Prym variety.

Now $H_1(Z, \mathbf{Q})$ and $H_1(Y, \mathbf{Q})$ are both equipped with skew-symmetric pairings, the intersection pairings. The map π^* scales the pairing by the degree q of the covering of $Z \to Y$. If the pairing on $H_1(Z, \mathbf{Q})$ is nondegenerate, we may identify the primitive homology with the orthogonal complement to $\pi^* H_1(Y, \mathbf{Q})$ in $H_1(Z, \mathbf{Q})$, and in particular this primitive homology inherits a skew-symmetric pairing. In our case, Z and Y will both be compact surfaces punctured at a single point, and therefore the intersection pairings on $H_1(Z, \mathbf{Q})$ and $H_1(Y, \mathbf{Q})$ are perfect.

8.2.1. *The mapping class group and its action on homology; the map* Mon. Clearly the diffeomorphism group of Y acts on the finite set of isomorphism classes of $\mathrm{Aff}(q)$-covers of Y, and this action factors through the mapping class group $\mathrm{MCG}(Y)$. In algebraic terms, this action is induced from the map $\mathrm{MCG}(Y) \longrightarrow \mathrm{Out}(\pi_1(Y, y_0))$.

Let $\mathrm{MCG}(Y)_Z$ denote the stabilizer of (Z, π) for this action. Since $\mathrm{Aff}(q)$ has trivial centralizer in $\mathrm{Sym}(\mathbf{F}_q)$, such elements lift uniquely to mapping classes on Z, i.e. there is a homomorphism

$$\mathrm{MCG}(Y)_Z \longrightarrow \mathrm{MCG}(Z).$$

Namely, fixing a representative $\alpha : Y \to Y$, there is a unique $f : Z \to Z$ that renders the diagram

$$(8.2) \qquad\qquad \begin{array}{ccc} Z & \xrightarrow{\;f\;} & Z \\ \downarrow{\scriptstyle\pi} & & \downarrow{\scriptstyle\pi} \\ Y & \xrightarrow{\;\alpha\;} & Y \end{array}$$

commutative. Sending the mapping class of α to the mapping class of f defines the desired homeomorphism.

This construction gives rise to actions of $\mathrm{MCG}(Y)_Z$ on $H_1(Z)$ and $H_1^{\mathrm{Pr}}(Z, Y)$. This latter action is the *monodromy map*

$$\mathrm{Mon} : \mathrm{MCG}(Y)_Z \to \mathrm{Sp}(H_1^{\mathrm{Pr}}(Z, Y)).$$

---

[5]A priori, the map is defined up to conjugation by the normalizer of $\mathrm{Aff}(q)$ in $\mathrm{Sym}(\mathbf{F}_q)$. This normalizer is equal to $\mathrm{Aff}(q)$.

8.2.2. *The main theorem.* Fix a surface Y of genus $g \geqslant 2$, a point $y \in Y$, a prime $q \geqslant 3$; as before, $\mathrm{Aff}(q)$ denotes the group of affine-linear transformations of $\mathbf{F}_q$.

We consider $\mathrm{Aff}(q)$-covers Z° of $Y - \{y\}$ such that the monodromy around y is nontrivial (hence a q-cycle); the compactification of such a cover is a surface Z of genus $gq - \frac{q-1}{2}$. We call such Z *singly ramified* $\mathrm{Aff}(q)$-covers of Y. The notation hides the dependence on the point y, which will remain fixed. There are, up to isomorphism, only finitely many such Z; choose a representative for each isomorphism class and call them $Z_1, Z_2, \ldots, Z_N$, and let $\mathrm{Cov}_1, \mathrm{Cov}_2, \ldots, \mathrm{Cov}_N : \pi_1(Y - \{y\}) \to \mathrm{Aff}(q)$ be representatives for the associated monodromy mappings.

Let $\mathrm{MCG}(Y - \{y\})_0$ denote the intersection of the groups $\mathrm{MCG}(Y - \{y\})_{Z_i}$. The individual monodromy maps attached to the covers Z_i combine to give a map

$$(8.3) \qquad \mathrm{Mon} : \mathrm{MCG}(Y - \{y\})_0 \to \prod_{i=1}^{N} \mathrm{Sp}(H_1^{\mathrm{Pr}}(Z_i, Y)).$$

The mapping class group of a punctured surface fits in the Birman exact sequence [17, Theorem 4.6]

$$(8.4) \qquad 0 \to \pi_1(Y, y) \to \mathrm{MCG}(Y - \{y\}) \to \mathrm{MCG}(Y) \to 0.$$

Let $\pi_1(Y, y)_0$ denote the inverse image of $\mathrm{MCG}(Y - \{y\})_0$ in $\pi_1(Y, y)$; the inclusion $\pi_1(Y, y)_0 \subseteq \pi_1(Y, y)$ is of finite index.

The restriction of (8.3) to the subgroup $\pi_1(Y, y)_0$ describes the monodromy of a Kodaira-Parshin family, as in Definition 7.3. We review this connection in more detail in §8.2.3. The following statement is equivalent to the large monodromy property of Kodaira–Parshin families, stated without proof as point (i) before Theorem 5.4.

Theorem 8.1. *Let notation be as above; in particular, $Z_1, \ldots, Z_N$ are a set of representatives for isomorphism classes of singly ramified $\mathrm{Aff}(q)$-covers of Y. Then the map*

$$(8.5) \qquad \mathrm{Mon} : \pi_1(Y, y)_0 \to \prod_{i=1}^{N} \mathrm{Sp}(H_1^{\mathrm{Pr}}(Z_i, Y))$$

has Zariski-dense image.

We briefly outline the proof. We give in §8.4 a "normal form" for each $\mathrm{Aff}(q)$-cover. Using the sequence (8.4), we reduce to showing a similar assertion with $\pi_1(Y, y)_0$ replaced by $\mathrm{MCG}(Y - \{y\})_0$. This allows us to use Dehn twists. Using our normal form for $\mathrm{Aff}(q)$-covers, and constructing a suitable system of curves to Dehn-twist around, we can see that the monodromy surjects onto each factor $\mathrm{Sp}(H_1^{\mathrm{Pr}}(Z_i, Y))$. A version of Goursat's lemma completes the proof.

When considering the general problem (replacing $\mathrm{Aff}(q)$ or cyclic covers by G-covers) the primitive homology must be further decomposed according to the representation theory of G. Looijenga [28] has proven a similar result for cyclic covers of surfaces without monodromy; in fact, Looijenga determines the exact image of Mon in this situation. See also [19, Theorem 1.6] for an analogous result for unramified covers of a closed surface, and [37] for covers whose covering group is the Heisenberg group.

8.2.3. *Application to Theorem 5.4.* For clarity we now write out why Theorem 8.1, in the form stated, above, implies what is used in Theorem 5.4: namely, the Kodaira-Parshin family $X_q \to Y'_q \xrightarrow{\pi} Y$ for the group $\mathrm{Aff}(q), q \geq 3$, has full monodromy.

In stages:

- We begin, as in §7.1 with a family $Z_q \to Y_q' \to Y_q$, with $Z_q \to Y_q'$ a relative curve. (The Kodaira-Parshin family was constructed by applying a Prym construction to this, as explicated in Definition 7.3).
- Fix $y \in Y_q(\mathbf{C})$. The fiber of $Y_q' \to Y_q$ above $y \in Y_q(\mathbf{C})$ is identified with the isomorphism classes of singly ramified $\mathrm{Aff}(q)$-covers of $Y(\mathbf{C})$, branched at y. This follows from property (i) of Proposition 7.1.
- Fix $y' \in Y_q'(\mathbf{C})$ above $y \in Y_q(\mathbf{C})$. Let $Z_{q,y'}$ be the fiber of $Z_q \to Y_q'$ over y'.

  By construction, $Z_{q,y'} \to Y(\mathbf{C})$ is a singly branched $\mathrm{Aff}(q)$-cover, and we can form the degree q cover associated to the action of $\mathrm{Aff}(q)$ on $\mathbf{Z}/q\mathbf{Z}$, i.e. $Z_{q,y'}(\mathbf{C}) \times_{\mathrm{Aff}(q)} \mathbf{Z}/q\mathbf{Z}$. By our definitions above, we have

$$Z_{q,y'}(\mathbf{C}) \times_{\mathrm{Aff}(q)} \mathbf{Z}/q\mathbf{Z} \simeq Z_i$$

  (for some unique i in $\{1, 2, \ldots, N\}$) as Riemann surfaces over $Y(\mathbf{C})$.
- The construction of Kodaira–Parshin families then gives rise, as in (7.3), to an isogeny

$$\text{fiber } X_{q,y'} \text{ of } X_q \text{ above } y' \longrightarrow \mathrm{Prym}(Z_i \to Y(\mathbf{C}))$$

  This isogeny induces an isomorphism of the rational homology groups:

$$H_1^{\mathrm{Pr}}(Z_i, Y; \mathbf{Q}) \simeq \text{first homology of } X_{q,y'}(\mathbf{C}) \text{ with rational coefficients.}$$

- This identification is compatible with monodromy, and so Theorem 8.1 translates to definition (5.1) of full monodromy.

8.3. Dehn twists and liftable curves.

We say that e is a *simple closed curve* in a surface Y if it is the image of a smooth embedding $S^1 \to Y$; a simple closed curve has no self-intersection. For us, a simple closed curve will always come with an orientation, namely, the orientation induced from a fixed orientation on S^1. If $y \in Y$ is a point, then we say $\eta \in \pi_1(Y, y)$ is *represented by a simple closed curve* if there is a loop e in Y, based at y and representing the class $\eta \in \pi_1(Y, y)$, which is a simple closed curve. We may say (somewhat imprecisely) that η "is" a simple closed curve.

If e is a simple closed curve in $Y - \{y\}$, the Dehn twist D_e about e acts on $H_1(Y)$ by the transvection T_e; indeed, we can regard D_e as an element of $\mathrm{MCG}(Y - \{y\})$. We want to study how this lifts to an $\mathrm{Aff}(q)$-cover $Z \to Y$: Let n_e be the order of the image of e in $\mathrm{Aff}(q)$. Then $D_e^{n_e}$ lifts to an automorphism of the cover Z, as we now describe. Suppose the image of e under $\pi_1(Y, y_0) \to \mathrm{Aff}(q) \to \mathrm{Sym}(\mathbf{F}_q)$ has cycle structure $(d_1, \ldots, d_k)$. The preimage of e under $Z \to Y$ is a disjoint union of circles $e_1, \ldots, e_k$, with the circles e_i in natural bijection with the cycles in the permutation. Then $D_e^{n_e}$ lifts to the product of commuting Dehn twists

$$\prod_i D_{e_i}^{n_e/d_i}$$

on Z.

In our cases, the only possibilities for cycle structure are as follows:

- If e maps to an element of $\mathrm{Aff}(q)$ that is not in $\mathbf{F}_q^+$, i.e. has nontrivial image $a \in \mathbf{F}_q^*$, then $(d_1, \ldots, d_k) = (1, \mathrm{ord}_q(a), \mathrm{ord}_q(a), \ldots, \mathrm{ord}_q(a))$.
- If e maps to a nonzero element of $\mathbf{F}_q^+$ then $(d_1, \ldots, d_k) = (q)$.
- If e maps to the identity element of $\mathrm{Aff}(q)$, then $(d_1, \ldots, d_k) = (1, \ldots, 1)$.

Now we note that:

Lemma 8.2. *Let e be a simple closed curve in $Y - \{y\}$. Then the classes of the preimages $[e_1], \ldots, [e_k]$ in the homology of Z are linearly independent; projected to $H_1^{\mathrm{Pr}}(Z, Y)$, their span has dimension $k - 1$.*

Proof. Y admits the structure of a CW complex with one 2-cell such that e belongs to the 1-skeleton. The inclusion of this 1-skeleton into $Y - y$ is a homotopy equivalence. Correspondingly the inclusion of the preimage (in Z) of this 1-skeleton into the preimage (in Z) of $Y - y$ is also a homotopy equivalence. Note also that the inclusion of $Y - y$ into Y induces an isomorphism on H_1, with a similar statement for $Z - \pi^{-1}(y)$.

These remarks allow us to reduce the Lemma to corresponding assertions for a covering of a finite graph, which are clear. $\qquad\square$

The action of $\mathrm{Mon}(D_e^{n_e})$ on $H_1^{\mathrm{Pr}}(Z, Y)$ is a unipotent transformation u such that the image of $u - 1$ is exactly the span of the classes of the circles e_i. By the Lemma just proven, this has dimension $k - 1$; correspondingly the fixed space $\mathrm{Mon}(D_e^{n_e})$ on $H_1^{\mathrm{Pr}}(Z, Y)$ has codimension $k - 1$.

We record the following consequence:

Lemma 8.3. *Suppose $Z \to Y$ is an $\mathrm{Aff}(q)$-cover. Let e be a simple closed curve in Y, and take M such that $D_e^M \in \mathrm{MCG}(Y)_Z$. Then the rank of $\mathrm{Mon}(D_e^M) - \mathrm{Id}$ acting on $H_1^{\mathrm{Pr}}(Z, Y)$ determines the conjugacy class of $\mathrm{Cov}(e)$ in the symmetric group $\mathrm{Sym}(\mathbf{F}_q)$.*

A particularly important case is when e is a simple closed curve in Y such that $\mathrm{Cov}(e)$ maps to a generator for $\mathbf{F}_q^*$ under the natural map $\mathrm{Aff}(q) \to \mathbf{F}_q^*$. We call such a e a *liftable curve* (for the $\mathrm{Aff}(q)$-cover $Z \to Y$). Its preimage in Z splits into a union of simple closed curves e^+ of degree 1 over e, and e^- of degree $q - 1$ over e. For liftable e we write

$$\widetilde{e} := \text{ projection of the class of } e^+ \text{ to primitive homology.}$$

According to our discussion above, D_e induces a transvection on $H_1^{\mathrm{Pr}}(Z, Y)$, with center $\widetilde{e}$.

Write $\cdot$ for the intersection pairing on homology. Given liftable curves A and B, we have

$$(8.6) \qquad \widetilde{A} \cdot \widetilde{B} = (A^+ \cdot B^+) - \frac{1}{q} A \cdot B.$$

Indeed, identifying primitive homology with the kernel of the pushforward, we have $q\widetilde{A} = qA^+ - \pi^* A$, and so the intersection pairing of $q\widetilde{A}$ with $q\widetilde{B}$ is

$$(qA^+ - \pi^* A) \cdot (qB^+ - \pi^* B) = q^2 (A^+ \cdot B^+) - 2q(A \cdot B) + q(A \cdot B),$$

as desired.

8.4. A normal form for an Aff_q-cover. Again, take Z a singly ramified $\mathrm{Aff}(q)$-cover of Y. We will describe the cover $Z \to Y$ in a normal form by cutting Y carefully, using essentially the fact that $\mathrm{Aff}(q)$ is solvable. The end result is roughly that the covering $Z \to Y$ can be expressed as the sum of a trivial cover of a genus $g - 1$ surface and a nontrivial cover on a torus.

Choose a basepoint $y_0 \in Y$. The map $\mathrm{Cov} : \pi_1(Y - \{y\}, y_0) \to \mathrm{Aff}(q)$ specifying the cover $Z \to Y$ induces a map

$$H_1(Y, \mathbf{Z}) \cong H_1(Y - \{y\}, \mathbf{Z}) \to \mathbf{F}_q^*$$

on abelianizations. The group $\mathbf{F}_q^*$ is cyclic and $H_1(Y, \mathbf{Z})$ is free. If we choose a surjection $\mathbf{Z} \twoheadrightarrow \mathbf{F}_q^*$, then the map on abelianizations lifts to a map

$$H_1(Y, \mathbf{Z}) \to \mathbf{Z} \to \mathbf{F}_q^*.$$

We can choose this map so that $H_1(Y, \mathbf{Z}) \to \mathbf{Z}$ is surjective, so it is given by intersecting with a primitive integral homology class α_1.

Choose a simple closed curve, which we also call α_1, representing this class. (Indeed, any primitive integral homology class is represented by a simple closed curve: [17, Proposition 6.2].) In fact, choose two such curves, α_1^+ and α_1^-, which pass "close by" but on either side of the ramification point y, and are parallel to one another. Note that, since our cover is ramified at y, and the monodromy at y is a nontrivial element of in $\mathbf{F}_q^+$, $\mathrm{Cov}(\alpha_1^+)$ and $\mathrm{Cov}(\alpha_1^-)$ cannot both be trivial.
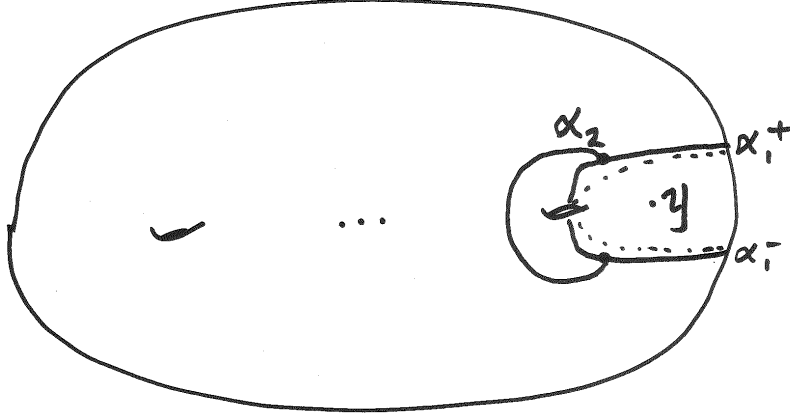

FIGURE 1. The curves $\alpha_1^{\pm}$ and α_2 on Y.

Cutting Y along the curves $\alpha_1^{\pm}$ and discarding the connected component of y, we obtain a surface Y^1 with two boundary components. Let Z^1 be the pullback of our covering to Y^1. The map $\mathrm{Cov} : \pi_1(Y^1, y_0) \to \mathrm{Aff}(q)$ has image contained in $\mathbf{F}_q^+ \subseteq \mathrm{Aff}(q)$ by our choice of α; so it factors through $H_1(Y^1, \mathbf{Z})$.

The boundary components (with orientations defined by an outward normal) define classes $b_+, b_- \in H_1(Y^1, \mathbf{Z})$; these classes satisfy $b_+ + b_- = 0$ because their sum is the boundary of Y^1. We saw above that b_+ and b_- cannot both have trivial image in $\mathbf{F}_q^+$; so $\mathrm{Cov}(b_+) = -\mathrm{Cov}(b_-)$ must be nontrivial. Conjugating by a suitable element of $\mathrm{Aff}(q)$ as necessary, we may as well suppose that $b_+ \in H_1(Y^1, \mathbf{Z})$ maps to $1 \in \mathbf{F}_q^+$.

For a surface such as Y^1 with boundary ∂Y^1, Poincaré duality takes the form of a perfect pairing between absolute and relative homology:

$$(8.7) \qquad\qquad H_1(Y^1, \partial Y^1; \mathbf{Z}) \times H_1(Y^1; \mathbf{Z}) \to \mathbf{Z}$$

The map $H_1(Y^1, \mathbf{Z}) \to \mathbf{F}_q^+$ lifts to a map $H_1(Y^1, \mathbf{Z}) \to \mathbf{Z}$; since b_+ is a primitive element of $H_1(Y^1, \mathbf{Z})$, we can choose such a lift taking b_+ to 1. This lift is of the form $x \mapsto \langle x, \alpha_2 \rangle$ for a relative homology class $\alpha_2 \in H_1(Y^1, \partial Y^1; \mathbf{Z})$. Therefore α_2 intersects the boundary components with multiplicity $+1$ and -1.

The following lemma readily implies that α_2 can be represented by a simple curve, the image of an immersion $e : [0, 1] \to Y^1$ that meets ∂Y^1 only at the endpoints, which we also call α_2. Indeed, it implies $MCG(Y)$ acts transitively on that subset of $H_1(Y, \partial Y) \simeq \mathrm{Hom}(H_1(Y, \mathbf{Z}), \mathbf{Z})$ consisting of elements whose pairing with a fixed boundary circle is

1. It follows that we can find a mapping class carrying the homology class of e to the homology class α_2, as desired.

Lemma 8.4. *Suppose Y is a surface of genus g with 2 boundary components, so $V = H_1(Y, \mathbf{Z})$ is a free $\mathbf{Z}$-module of rank $2g + 1$. We regard it as equipped with a (degenerate) alternating form via $H_1(Y) \to H_1(Y, \partial Y)$ and the duality pairing (8.7); the radical of this form is the rank-1 submodule V^0 generated by b, the class of one of the two boundary components of Y.*

Let $\mathrm{Sp}(V, b)$ denote the group of automorphisms of V preserving the bilinear form and fixing b. Then the natural map $MCG(Y) \to \mathrm{Sp}(V, b)$ is surjective.

Proof. The group $\mathrm{Sp}(V, b)$ fits into an exact sequence

$$1 \to \mathrm{Hom}(V/V^0, V^0) \to \mathrm{Sp}(V, b) \to \mathrm{Sp}(V/V^0) \to 1,$$

where the left-hand map is given by $f \mapsto 1 + f$.

Now one obtains a closed surface from Y by capping off both boundary components. The mapping class group $MCG(Y)$ surjects onto the mapping class group of this closure [17, Prop 3.19]. Therefore (by the surjectivity of the symplectic representation for a closed surface [17, Theorem 6.4]) it surjects onto $\mathrm{Sp}(V/V^0)$.

Now let $v \in V$ be a class, not in V^0, which is represented by a simple closed curve in Y; and let b be one of the two boundary components of Y. We can represent $v + b$ by a simple closed curve as well (possibly after replacing b with $-b$). Thus the image of $MCG(Y)$ contains the transvections T_v and T_{v+b}. The composition $T_{v+b}T_v^{-1}$ is a nontrivial element of $\mathrm{Sp}(V, b)$, coming from the element

$$x \mapsto \langle x, v \rangle b \in \mathrm{Hom}(V/V^0, V^0).$$

These generate $\mathrm{Hom}(V/V^0, V^0)$ so the result follows. $\qquad\square$

Cut Y^1 along α_2, and let Y^2 be the resulting surface; it is a surface of genus $g - 1$ with one boundary component. The pullback of the cover $Z \to Y$ to Y^2 *splits*, i.e., becomes a disjoint union of q copies of Y^2. We can recover Y from Y^2 by gluing to Y^2 a torus with one boundary component. Thus our discussion has shown that it is possible to put any $\mathrm{Aff}(q)$-cover $Z \to Y$ into a normal form: a connected sum of a trivial cover of a genus $g - 1$ surface and a nontrivial cover of a torus.

To summarize: Let S_{g-1} be a genus-$(g - 1)$ surface and let T be a torus. Fix a small open disk D in S_{g-1} and D' in T, and set

(8.8) $$\mathsf{S}_{g-1}^\circ = \mathsf{S}_{g-1} - D, \ \ \mathsf{T}^\circ = \mathsf{T} - D',$$

so these are, respectively, a surface of genus $g - 1$ with one boundary component and a torus with one boundary component. We identify Y with the genus-g surface obtained by gluing S_{g-1}° to T° along an identification $\partial D' \simeq \partial D$. (In relation to the discussion just given, S_{g-1}° is homotopy-equivalent to Y^2.)

Proposition 8.5 (Normal form for $\mathrm{Aff}(q)$-covers). *Let Z be a singly ramified $\mathrm{Aff}(q)$-cover of Y. Then we may write Y as a connected sum:*

$$Y = \mathsf{S}_{g-1} \# \mathsf{T},$$

where S_{g-1} is a genus-$(g - 1)$ surface and T is a genus-1 surface, satisfying the following properties (with notation as above).

- *The ramification point y belongs to the interior of T°,*
- *the cover $Z \to Y$ splits over S_{g-1}°.*
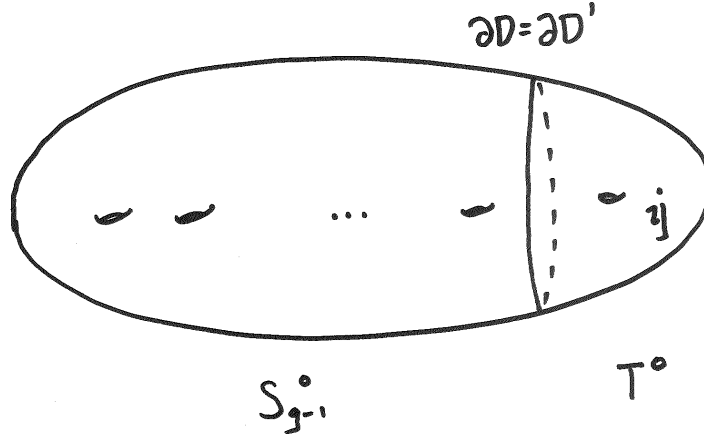
FIGURE 2. Y as a connected sum.

- *the cover $Z \to Y$, when restricted to T°, extends over T, i.e. has trivial monodromy around the boundary circle of T°.*
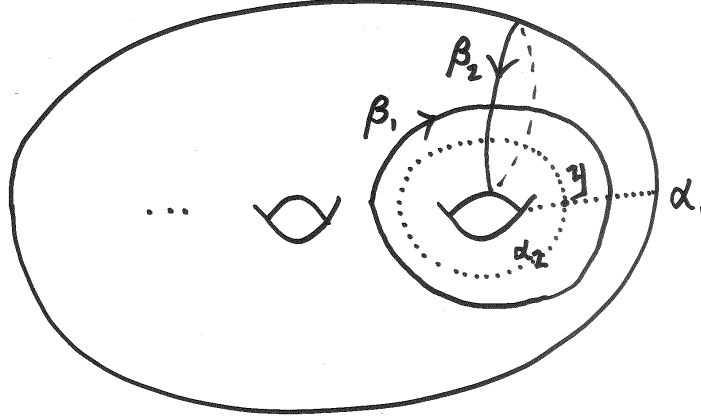- *With respect to a standard basis for $\pi_1(\mathsf{T} - y, *)$, a free group on two generators β_1, β_2, the monodromy of the cover sends*
  - *β_1 to an element of $\mathrm{Aff}(q)$ projecting to a generator for $\mathbf{F}_q^*$, and*
  - *β_2 to an nonzero element of $\mathbf{F}_q^+$.*

Here β_1 is a curve which crosses α_1 once and does not cross α_2; and similarly for β_2.

FIGURE 3. The curves β_1 and β_2. (The basepoint is the intersection of β_1 and β_2.)

Thus Z consists of q copies of S_{g-1}° glued to a degree-q cover $\widetilde{\mathsf{T}^\circ}$ of T° along q boundary circles. In the sequel we will use $\widetilde{\mathsf{S}_{g-1}^\circ}$ for the cover of S_{g-1}° induced by Z.

8.5. **Proof of Theorem 8.1.** We must show (8.5) has Zariski-dense image. We will perform a series of reductions; the main steps are Lemmas 8.7, 8.9, and 8.10.

Lemma 8.6. *The image of $\pi_1(Y, y)_0$ (see after (8.4) for definition) under the monodromy map*

$$\pi_1(Y, y)_0 \to \mathrm{Sp}(H_1^{\mathrm{Pr}}(Z_i, Y))$$

to any factor $\mathrm{Sp}(H_1^{\mathrm{Pr}}(Z_i, Y))$ of the right-hand side of (8.5) is not contained in the center of Sp.

Proof. We leave the simple topological proof to the reader.[6] □

Because $\pi_1(Y, y)_0$ is normal inside $\mathrm{MCG}(Y - \{y\})_0$ and the symplectic groups are almost simple, Theorem 8.1 follows from the subsequent Lemma:

Lemma 8.7. *The monodromy map restricted to $\mathrm{MCG}(Y - \{y\})_0$,*

$$\mathrm{Mon} : \mathrm{MCG}(Y - \{y\})_0 \to \prod_{i=1}^{N} \mathrm{Sp}(H_1^{\mathrm{Pr}}(Z_i, Y)),$$

has Zariski-dense image.

In turn, using Lemma 2.12, this will follow from Lemmas 8.8 and 8.9.

Lemma 8.8 (Distinct covers are distinguished by monodromy around a simple closed curve.)**.** *For two non-isomorphic $\mathrm{Aff}(q)$-covers Z_1, Z_2 there exists a simple closed curve η in Y such that the cycle decompositions of the monodromy around η in Z_1 and Z_2 are different.*

Proof. Two coverings Z_1, Z_2 define two maps $\pi_1(Y - y) \to \mathrm{Aff}(q)$. Suppose, first of all, that their projections to $\mathbf{F}_q^*$ have different kernels (i.e. are not related by an automorphism of $\mathbf{F}_q^*$). We may find a primitive homology class whose images under the two maps $f_1, f_2 : H_1(Y, \mathbf{Z}) \to \mathbf{F}_q^*$ have different orders in $\mathbf{F}_q^*$. Indeed, there is a basis $e_1, \ldots, e_r$ for $H_1(Y, \mathbf{Z})$ such that the kernel of f_1 equals $(q-1)e_1, e_2, \ldots, e_r$; not all of $e_2, \ldots, e_r$ can be in the kernel of f_2, and so at least one of these latter classes suffice. Represent this primitive homology class by a simple closed curve to construct η.

Otherwise, the coverings Z_1, Z_2 define maps $\pi_1(Y - y) \to \mathbf{F}_q^*$ having the same kernel. Accordingly, in the algorithm to convert an $\mathrm{Aff}(q)$-cover into a normal form described in §8.4, we can cut Y along the *same* curve α_1, as in §8.4, for both Z_1 and Z_2. We obtain, as before, a surface Y^1 with two boundary components; the covers Z_1, Z_2 define two maps

$$g_1, g_2 : H_1(Y^1, \mathbf{Z}) \longrightarrow \mathbf{F}_q^+.$$

If g_1 is not proportional to g_2, we can find a primitive homology class for $H_1(Y^1, \mathbf{Z})$ which is in the kernel of one map but not the other. Represent this primitive homology class by a simple closed curve to construct η.

Otherwise g_1 and g_2 are proportional, so the two maps $\pi_1(Y^1) \to \mathbf{F}_q^+$ have the same kernel. Therefore, we can cut Y_1 along the same curve α_2 for both Z_1 and Z_2. So we get a decomposition of Y as a connected sum $Y = \mathsf{S}_{g-1} \# \mathsf{T}$ as above, such that both Z_1 and Z_2 become trivial on S_{g-1}.

Let β_1 and β_2 be curves on T as in the end of §8.4. Then both maps $\pi_1(Y-y) \to \mathrm{Aff}(q)$ send β_1 to an element of $\mathrm{Aff}(q)$ projecting to a generator for $\mathbf{F}_q^*$; and they both send β_2

---

[6]One can also give an algebro-geometric argument, as follows. Suppose to the contrary. Now, as in §7, there is an associated finite covering $Y' \to Y$ such that the various Z_i fit together into a curve fibration $\mathsf{Z} \to Y'$. If (a) were false, the theorem of the fixed part means that the Hodge structure of the fibers of $\mathsf{Z} \to Y'$ are constant, at least over one component of Y'. By Torelli, this means that all the fibers are actually isomorphic. This contradicts de Franchis's theorem.

to an element of $\mathbf{F}_q^+$. Each of $\mathrm{Cov}_1(\beta_1)$ and $\mathrm{Cov}_2(\beta_1)$ has a unique fixed point in $\mathbf{F}_q$; up to conjugation, we may suppose this fixed point is 0. By a further conjugation we may assume that $\mathrm{Cov}_1(\beta_2) = \mathrm{Cov}_2(\beta_2) = 1 \in \mathbf{F}_q^+$.

So we can write

$$\mathrm{Cov}_1(\beta_1)\colon x \mapsto c_1 x$$

and

$$\mathrm{Cov}_2(\beta_1)\colon x \mapsto c_2 x.$$

If Z_1 and Z_2 are not isomorphic covers, we must have $c_1 \neq c_2$.

There is a map

$$\pi_1(Y - y) \longrightarrow \pi_1(\mathsf{T} - y)$$

which is obtained (in the notation of (8.8)) by collapsing S_{g-1}° to a point; this gives a map from $Y - y$ to a surface that is homotopy equivalent to $\mathsf{T} - y$.

There exists a simple closed curve $\eta \in \pi_1(Y - y)$ mapping to $\beta_1\beta_2\beta_1^{-1}\beta_2^{q-c_1}$ under this map $\pi_1(Y - y) \to \pi_1(\mathsf{T} - y)$: see Figure 8.5 and its caption.

Then $\mathrm{Cov}_1(\eta)$ is trivial but $\mathrm{Cov}_2(\eta)$ is not trivial. This concludes the proof.   □
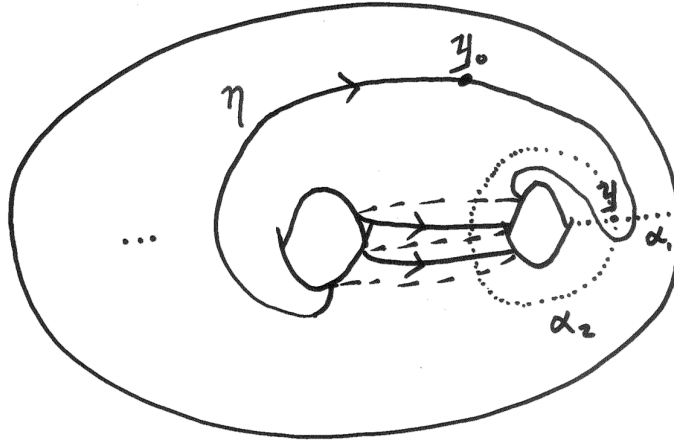

FIGURE 4. The curve η.
How to read the picture: Follow along the path η, starting at the basepoint y_0. Write down a word in the symbols β_1 and β_2 as follows. Every time η crosses α_1, write β_1 or β_1^{-1}, depending whether the crossing was in the positive or negative direction. Every time η crosses α_2, write β_2 or β_2^{-1}. The resulting word is the image of η under the map $\pi_1(Y - y, y_0) \to \pi_1(\mathsf{T} - y, y_0)$, which we readily see is $\beta_1\beta_2\beta_1^{-1}\beta_2^2$.

Lemma 8.9. *The monodromy map* $\mathrm{Mon} : \mathrm{MCG}(Y)_{Z_i} \longrightarrow \mathrm{Sp}(H_1^{\mathrm{Pr}}(Z_i, Y))$ *has Zariski-dense image.*

We are now reduced to proving Lemma 8.9. Let $Z = Z_i$ for some fixed i. By the construction of Dehn twists from liftable curves (see discussion at end of §8.3), as well as Lemma 2.14 on generation by transvections, it is enough to show:

Lemma 8.10. *There exists a collection of liftable curves* $A_1, \ldots, A_N$ *on Y such that:*
  (a) the $\widetilde{A}_i$ span the primitive homology $H_1^{\mathrm{Pr}}(Z, Y)$;
  (b) the graph obtained by connecting A_i, A_j when $\widetilde{A}_i \cdot \widetilde{A}_j \neq 0$ is connected.

8.6. **Proof of Lemma 8.10.** We put the singly ramified $\mathrm{Aff}(q)$-cover $Z \to Y$ in a normal form, as explained in §8.4. Recall notation (D, D', S_{g-1}, T, and so forth) from the end of §8.4. We will produce the curves A_i by concatenating curves on T° and curves on S_{g-1}°.

Fix a point $p \in \partial D \cong \partial D'$. Fix a labelling of the points of Z above p by $\mathbf{F}_q$, compatible with the usual action of $\mathrm{Aff}(q)$ for some fixed homomorphism

$$\mathrm{Cov} : \pi_1(Y - y, p) \longrightarrow \mathrm{Aff}(q).$$

Recall that the cover $Z \to Y$ splits over S_{g-1}°; the labelling above p therefore permits us also to label the components of S_{g-1}° by $\mathbf{F}_q^+$.

Lemma 8.11. *There exist $q + 1$ simple closed curves $\{\gamma_j : 0 \leqslant j \leqslant q\}$ on T°, beginning and ending at p, not passing through y, and intersecting $\partial D'$ only at its endpoints, such that:*

   (i) For each j, the monodromy $\mathrm{Cov}(\gamma_j)$ projects under $\mathrm{Aff}(q) \to \mathbf{F}_q^*$ to the same fixed generator of $\mathbf{F}_q^*$;
   (ii) The monodromy of γ_j, defining a map $\mathbf{F}_q \to \mathbf{F}_q$, fixes exactly j modulo q.
   (iii) The (unique) lifts γ_j^+ to simple closed curves on $\widetilde{\mathsf{T}^\circ}$ span the homology of $\widetilde{\mathsf{T}^\circ}$ modulo the homology of its boundary.
   (iv) Each γ_j has the same orientation near p, i.e., either the outgoing branch is "above" the incoming branch for all j, or vice versa.
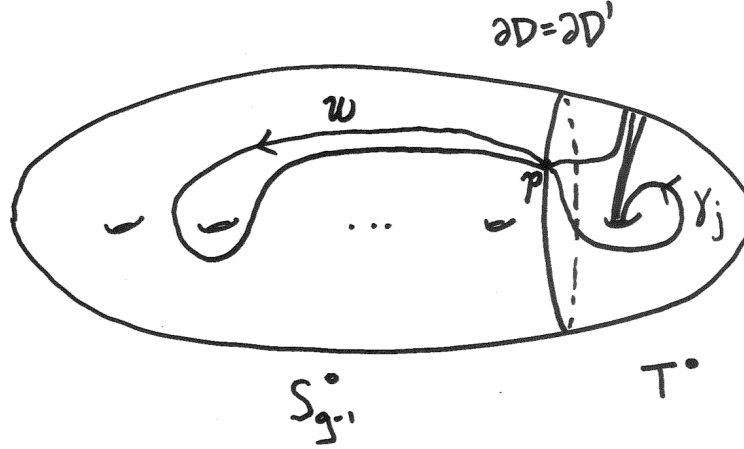
FIGURE 5. The curves γ_j and w on Y.

Proof. Take an explicit basis β_1, β_2 of homology of T, such as was described in §8.4; conjugating if necessary we can suppose that the monodromy of β_1 is $x \mapsto gx$ (for $g \in \mathbf{F}_q^*$ a generator) and the monodromy of β_2 is $x \mapsto x + 1$. We can choose this basis in such a way that all powers $\beta_1 \beta_2^j$ with j non-negative are represented by simple closed curves on T°, which start and end at p.

The monodromy around $\beta_1 \beta_2^j$ is given by $x \mapsto g(x + j)$, which fixes $\frac{gj}{1-g} \in \mathbf{F}_q^+$. Write $[\ell]$ for the unique representative of $\ell \in \mathbf{F}_q$ that lies in $[0, q - 1]$, and put

$$j^* = \begin{cases} [\frac{gj}{1-g}], & j \neq q \\ q, & j = q. \end{cases}$$

The map $j \mapsto j^*$ gives a bijection from $[0, q]$ to itself. Now put

$$\gamma_{j^*} = \beta_1 \beta_2^j \quad (j \in [0, q]).$$

Conditions (i) and (ii) are clearly satisfied. To check (iii) we must verify that the associated homology classes span homology of $\widetilde{\mathsf{T}^\circ}$ modulo its boundary. One could verify this by computing an explicit CW-complex for $\widetilde{\mathsf{T}^\circ}$; we present an alternative group-theoretic proof. It is sufficient to show that

(8.9)   the homology clases of the lifts of γ_j span $H_1(\widetilde{\mathsf{T} - y})$.

Here and in what follows we make use of the fact that our $\mathrm{Aff}(q)$-cover extends over T, and thus write (e.g.) $\widetilde{\mathsf{T}}$. To see that (8.9) indeed implies (iii), consider the diagram

$$
\begin{array}{ccc}
H_1(\widetilde{\mathsf{T}^\circ - y}) & \longrightarrow & H_1(\widetilde{\mathsf{T}^\circ})/H_1(\partial\widetilde{\mathsf{T}^\circ}) \\
\downarrow & & \downarrow {\scriptstyle \sim} \\
H_1(\widetilde{\mathsf{T} - y}) & \xrightarrow{\ f\ } & H_1(\widetilde{\mathsf{T}})/H_1(\widetilde{\mathsf{D}}) = H_1(\widetilde{\mathsf{T}})
\end{array}
$$

where f is surjective because in fact $H_1(\widetilde{\mathsf{T} - y}) \simeq H_1(\widetilde{\mathsf{T}})$: the preimage of y is a single point.

Let $\tilde{p}$ be the point above p corresponding to $0 \in \mathbf{F}_q$. Projection to T identifies $\pi_1(\widetilde{\mathsf{T} - y}, \tilde{p})$ with the subgroup $\mathsf{H} \leqslant \langle \beta_1, \beta_2 \rangle$ defined by

$$\mathsf{H} = \text{stabilizer of } 0 \in \mathbf{F}_q.$$

Therefore the first homology of $\widetilde{\mathsf{T} - y}$ is the abelianization of H. Under this correspondence, the homology class of the lift of γ_j corresponds to the image in H^{ab} of $(\beta_2^{-j^*})\beta_1\beta_2^j(\beta_2^{j^*}) \in \mathsf{H}$.

We must therefore show that the elements $\beta_2^{-j^*}\beta_1\beta_2^{j+j^*}$ actually generate H^{ab}. Note that among these elements are β_1 and (a conjugate of) $\beta_1\beta_2^q$, so it is enough to show that

(8.10) β_2^q and $\beta_2^{-j^*}\beta_1\beta_2^{j+j^*} \quad (0 \leqslant j \leqslant q-1)$

generate H^{ab}. However, a set of left coset representatives for H are given by $1, \beta_2, \ldots, \beta_2^{q-1}$; according to Schreier's algorithm a generating set for H is given by

$$\beta_2^q, \beta_2^{-[gj]}\beta_1\beta_2^j, j \in [0, q-1]$$

Now considered modulo q the set of pairs $(-[gj], j) \equiv (-gj, j)$ appearing here coincide with the pairs $(-j^*, j+j^*) \equiv (-\frac{gj}{1-g}, \frac{j}{1-g})$ appearing in (8.10). So the elements of (8.10) even generate H not just its abelianization. $\qquad\square$

We return to the proof of Lemma 8.10. For each primitive homology class in S°_{g-1} we fix a representative which is a simple closed curve on S°_{g-1} beginning and ending at p. Let W be the resulting collection of simple closed curves. For each $w \in W$ at least one of the two homotopy classes

(8.11) $A(w, j) = \gamma_j \cdot w^{\pm 1} \in \pi_1(Y, p)$

is representable by a simple closed curve on Y. The choice of sign depends only on w and does not depend on j, in view of property (iv) of the curves γ_j. For a picture of the curve $A(w, j)$, see Figure 8.6.

The image of this curve in $\mathrm{Aff}(q)$ projects to a generator of $\mathbf{F}_q^*$; therefore it is "liftable" in the sense of §8.3. Recall also from §8.3 the notation e^+ for the degree-1 lift of a liftable curve e. The lift of $A(w, j)$ has homology class given by

$$[A(w,j)^+] = [\gamma_j^+] \pm [w_j],$$

where w_j means that we lift w to a closed loop on the jth preimage of S_{g-1}° inside Z; the sign above is the same as in (8.11).

We have

$$[A(w,j)^+] - [A(w',j)^+] = \epsilon[w_j] + \epsilon'[(w')_j] \ (\epsilon, \epsilon' \in \pm 1)$$

and we see readily that these classes – as both w, w' vary through W – span the homology of the jth preimage of S_{g-1}° in the cover Z.

The boundary of $\widetilde{\mathsf{S}_{g-1}^\circ}$ is a union of q circles. Considering the Mayer–Vietoris sequence

$$H_1(S^1)^q \to H_1(\widetilde{\mathsf{S}_{g-1}^\circ}) \oplus H_1(\widetilde{\mathsf{T}^\circ}) \twoheadrightarrow H_1(Z).$$

and using the fact that the $[\gamma_j]$ span $H_1(\widetilde{\mathsf{T}^\circ})$ modulo its boundary (Lemma 8.11 (iii)), we see that the $[A(w,j)^+]$ span $H_1(Z)$.

As before, we define

$$\widetilde{A(w,j)} = \text{projection of } [A(w,j)^+] \text{ to primitive homology.}$$

so that the homology classes $\widetilde{A(w,j)}$ span $H_1^{\mathrm{Pr}}(Z, Y)$. This completes the proof of part (a) of Lemma 8.10.

To prove part (b), we need to compute some intersection numbers. We note that the intersection number between any γ_j^+ and any w_k is trivial. Thus

$$[A(w_1, j)] \cdot [A(w_2, k)] = [\gamma_j] \cdot [\gamma_k] \pm [w_1] \cdot [w_2]$$

$$[A(w_1, j)^+] \cdot [A(w_2, k)^+] = [\gamma_j^+ \cdot \gamma_k^+] \pm \delta_{jk}[w_1] \cdot [w_2],$$

where δ_{jk} is the Kronecker δ symbol, and, in both instances, the sign that appears ithe product of the sign for w_1 and the sign for w_2. Upon projecting to primitive homology, (8.6) gives

$$\widetilde{A(w_1,j)} \cdot \widetilde{A(w_2,k)} = \pm(\delta_{jk} - q^{-1})[w_1] \cdot [w_2] + ([\gamma_j^+] \cdot [\gamma_k^+] - q^{-1}[\gamma_j \cdot \gamma_k]).$$

The connectedness of the "intersection graph" follows from this. It is enough to show that given (w_1, j) and (w_2, ℓ) there exists (w_3, k) with both intersection numbers nonzero. For this, we note that the factor $(\delta_{jk} - q^{-1})$ is never zero, so we simply choose w_3 so that $[w_1] \cdot [w_3]$ and $[w_2] \cdot [w_3]$ are sufficiently large: this is possible because $[w_1], [w_2] \neq 0$, the intersection pairing on $H_1(\mathsf{S}_{g-1}^\circ)$ is perfect, and we can choose w_3 such that $[w_3]$ is any given primitive homology class. This proves Lemma 8.10 and Theorem 8.1. $\qquad\square$

9. Transcendence of period mappings; the Bakker–Tsimerman theorem

It is desirable to extend the method to settings where the base Y is higher-dimensional, thus feasibly leading to finiteness results for integral points on Y. We will study the example when $X \to Y$ is the moduli space of smooth hypersurfaces in $\mathbf{P}^m$; then integral points on Y correspond to integral homogeneous polynomials $P(x_0, \ldots, x_m)$ of degree d whose discriminant $(\mathrm{disc}\ P) \in \mathcal{O}^*$.

(A natural family of generalizations of this example is given by considering the integral points on $\mathbf{P}^m - Z^\vee$, where $Z \subset \mathbf{P}^m$ is a smooth subvariety, and $Z^\vee$ is the dual projective

variety to Z: there is a natural smooth projective family over $\mathbf{P}^m - Z^\vee$, namely, the family of smooth hyperplane sections of Z.)

9.1. **The Ax-Schanuel theorem of Bakker and Tsimerman.** Suppose that we are given a smooth proper map $X \to Y$ of relative dimension d over the complex numbers (we identify complex algebraic varieties with their complex points). The primitive cohomology of each fiber $H^d(X_y, \mathbf{C})^{\mathrm{prim}}$ carries a polarized Hodge structure. Let $\mathfrak{H}$ be the associated period domain which classifies polarized Hodge structures with the same numerical data as this primitive cohomology, so we have an analytic period map

$$\Phi : \widetilde{Y} \longrightarrow \mathfrak{H}$$

where $\widetilde{Y}$ is the universal cover of $Y(\mathbf{C})$. This $\mathfrak{H}$ is open (for the analytic topology) in a certain complex flag variety $\mathfrak{H}^*$, which parameterizes isotropic flags with a given dimensional data inside a certain orthogonal or symplectic complex vector space.

Bakker and Tsimerman [2] have proven the following analogue of the Ax–Schanuel theorem. It is a very strong statement about the transcendence of Φ.

To simplify the statement, we assume that the monodromy mapping

$$\pi_1(Y) \to \mathrm{Aut}(H^d(X_y, \mathbf{C})^{\mathrm{prim}})$$

has image whose Zariski closure contains the full special orthogonal or symplectic group, stabilizing the intersection form. (This restriction, which guarantees that the image $\Phi(\widetilde{Y})$ is Zariski-dense in $\mathfrak{H}^*$, is not important, and in [2] the theorem is formulated for an arbitrary Mumford-Tate domain as target.)

Theorem 9.1. *(Theorem of Bakker and Tsimerman.) Suppose that $V \subset Y \times \mathfrak{H}^*$ is algebraic. Write W for the image of $\widetilde{Y}$ in $Y \times \mathfrak{H}$. Suppose that $U \subset V \cap W$ is irreducible analytic such that*

$$\mathrm{codim}_{Y \times \mathfrak{H}^*}(U) < \mathrm{codim}_{Y \times \mathfrak{H}^*}(V) + \mathrm{codim}_{Y \times \mathfrak{H}^*}(W),$$

where all the codimensions are taken inside $Y \times \mathfrak{H}^*$. Then the projection of U to Y is contained in a proper ("weak Mumford-Tate") subvariety.

In particular this has the following corollary:

Corollary 9.2 (Transcendence property of period mappings)**.** *With notation as above, suppose that $Z \subset \mathfrak{H}^*$ is an algebraic subvariety, and*

$$(9.1) \qquad\qquad\qquad \mathrm{codim}_{\mathfrak{H}^*}(Z) \geqslant \dim(Y).$$

Then any irreducible component of $\Phi^{-1}(Z)$ is contained inside the preimage, in $\widetilde{Y}$, of the complex points of a proper subvariety of Y.

Proof. Let Q be an irreducible component as in the statement of the corollary.

Let $V = Y \times Z$. The intersection W^Z of W with $Y \times (Z \cap \mathfrak{H})$, intersction taken in $Y \times \mathfrak{H}$, is an analytic set. Moreover, the image of Q under the analytic map $\widetilde{Y} \to Y \times \mathfrak{H}$ is contained in W^Z. Therefore, the image of Q is contained in some irreducible component of W^Z, call it U:

$$U = \text{an irreducible component of } W \cap (Y \times (Z \cap \mathfrak{H})).$$

We apply Theorem 9.1 with this chioce of U, V, W. Then

$$\mathrm{codim}_{Y \times \mathfrak{H}^*} V = \mathrm{codim}_{\mathfrak{H}^*} Z \text{ and } \mathrm{codim}_{Y \times \mathfrak{H}^*} W = \dim \mathfrak{H}^*.$$

so

$$\mathrm{codim}_{Y \times \mathfrak{H}^*} W + \mathrm{codim}_{Y \times \mathfrak{H}^*} V = \dim \mathfrak{H}^* + \mathrm{codim}_{\mathfrak{H}^*} Z \overset{(9.1)}{\geqslant} \dim \mathfrak{H}^* + \dim(Y).$$

This shows $\dim U = 0$, unless the projection of U to Y is contained in a proper weak Mumford–Tate subvariety. This implies the same property for Q, as desired. $\qquad\square$

9.2. Transferring transcendence to a p-adic setting.
Theorem 9.1 can be transferred to the p-adic setting, which is where we use it:

With notation as above, suppose additionally that $X \to Y$ is defined over $\mathbf{Z}[S^{-1}]$. Fix $p \notin S$ and $y_0 \in Y(\mathbf{Z}_p)$. As before we can form the p-adic period map

$$(9.2) \qquad \Phi_p : \underbrace{\text{residue disk around } y_0 \text{ in } Y(\mathbf{Q}_p)}_{U_p} \longrightarrow \mathfrak{H}^*_{\mathbf{Q}_p},$$

where $\mathfrak{H}^*_{\mathbf{Q}_p}$ is the base-change of $\mathfrak{H}^*$ to $\mathbf{Q}_p$; the map above is p-adic analytic, i.e., it is given in suitable coordinate charts by power series absolutely convergent on the residue disk.

Now suppose that we give ourselves a $\mathbf{Q}_p$-algebraic subvariety $Z \subset \mathfrak{H}^*_{\mathbf{Q}_p}$ satisfying the dimensional condition (9.1), i.e. the codimension of Z is greater than or equal to the dimension of Y.

Lemma 9.3. *Let U_p be as in (9.2), i.e. $\{y \in Y(\mathbf{Z}_p) : y \equiv y_0 \text{ modulo } p\}$. The set*

$$\Phi_p^{-1}(Z)$$

is not Zariski dense in Y.

Note that Φ_p is defined only on the residue disk U_p.

Proof. [7]

It will be convenient to have the freedom to vary y_0 later in the argument. To that end, note that the statement above depends only on U_p; after all, at the level of points, Φ_p is the map sending $y \in U_p$ to the induced Hodge filtration on the primitive crystalline cohomology of the special fiber of X_y.

By [29, Thm. 7.6] (or by the discussion of §3.3) the image of $\Phi_p(U_p)$ is contained in a residue disk on $\mathfrak{H}^*_p$ containing $\Phi_p(y_0)$, in particular, in some affine open set $\mathrm{Spec}\, A_p$ of $\mathfrak{H}^*_p$ containing $\Phi_p(y_0)$. We may suppose that $Z \subset \mathfrak{H}^*_{\mathbf{Q}_p}$ is defined locally by equations $F_i = 0$, where we suppose $F_i \in A_p$, i.e. the F_i are regular functions on this affine open set.

Consider now

$$G_i = F_i \circ \Phi_p.$$

These are defined by power series converging absolutely in U_p, i.e. in a suitable choice of local coordinates, G_i lies in a Tate algebra

$$R = \mathbf{Q}_p \left\langle \frac{x_1}{p}, \ldots, \frac{x_N}{p} \right\rangle$$

of formal power series convergent on a disk of p-adic radius $|p|$. In these coordinates U_p corresponds to $(x_1, \ldots, x_N) \in (p\mathbf{Z}_p)^N$. We want to show that the common zero-locus, inside U_p, of the G_i is contained in (the $\mathbf{Q}_p$-points of) an algebraic set. As a preliminary reduction, we will reduce to considering a single "irreducible component" of this common zero locus.

---

[7]The anonymous referee pointed out that one could also deduce the p-adic transcendence result from the complex transcendence result using the Seidenberg embedding theorem, as in [32, Section 2.5].

Fix a suitable open affine set $\operatorname{Spec} B_p \subset Y_{\mathbf{Q}_p}$ "containing the residue disc of y_0." (More precisely, we may fix an open affine neighbourhood in Y, considered now as $\mathbf{Z}_p$-scheme, of the image of the $\mathbf{Z}_p$-valued point $y_0 : \operatorname{Spec} \mathbf{Z}_p \to Y$, and take its generic fiber.) Then there is a morphism from B_p to R. Our result will now follow from the

> *Claim:* Let $\mathfrak{p}$ be a prime ideal of R, vanishing at some point of U_p.
> Suppose that $\mathfrak{p}$ is minimal among prime ideals containing $\langle G_1, \ldots, G_n \rangle$.
> Then $\mathfrak{p}$ contains (the image in R of) a regular function H, i.e. a function
> H belonging to B_p as above.

To see why this implies the statement of the lemma, assume the *Claim.* There are only finitely many such minimal primes as in the statement. Call them $\mathfrak{p}_1, \ldots, \mathfrak{p}_t$. Let $H_j \in \mathfrak{p}_j$ the function constructed according to the claim above. Then the vanishing locus of $\prod_j H_j$ contains the common vanishing locus of the G_i: if y lies in this common zero-locus, it lies in the vanishing locus of some $\mathfrak{p}_j$, and then $H_j(y) = 0$.

We now prove the *Claim.* The ideal $\mathfrak{p}$ vanishes at some point at U_p by assumption; choose such a point y_0.

We now transfer the question to the complex numbers. We fix an isomorphism $\sigma : \overline{\mathbf{Q}_p} \simeq \mathbf{C}$, which gives in particular an embedding $\sigma : \mathbf{Q}_p \hookrightarrow \mathbf{C}$. Then y_0 gives rise to a complex point $y_0^\sigma \in Y(\mathbf{C})$, and the de Rham cohomology of $X_{y_0}^\sigma$ is obtained from that of X_{y_0} via σ:

$$(9.3) \qquad H_{\mathrm{dR}}^*(X_{y_0}) \otimes_{(\mathbf{Q}_p, \sigma)} \mathbf{C} = H_{\mathrm{dR}}^*(X_{y_0^\sigma}/\mathbf{C}).$$

We may regard the period map Φ_p as taking values in the Grassmannian $\mathfrak{H}_{\mathbf{Q}_p}^*$ for the left-hand de Rham cohomology. Also let $U_{\mathbf{C}}$ be a small complex neighbourhood of y_0^σ and let $\Phi_{\mathbf{C}} : U_{\mathbf{C}} \to \mathfrak{H}_{\mathbf{C}}^*$ be the complex period mapping, which we regard as taking values in the associated complex variety $\mathfrak{H}_{\mathbf{C}}^* := \left(\mathfrak{H}_{\mathbf{Q}_p}^*\right)^\sigma$. This complex variety parameterizes certain flags inside the right-hand space of (9.3). Note the identification $\Phi_{\mathbf{C}}(y_0^\sigma) = \Phi_p(y_0)^\sigma$.

Now Z gives rise to an algebraic subvariety $Z^\sigma \subset \mathfrak{H}_{\mathbf{C}}^*$ and this subvariety again satisfies condition (9.1). The functions F_i are regular on an open affine containing $\Phi_p(y_0)$; correspondingly we obtain F_i^σ on an affine open in $\mathfrak{H}_{\mathbf{C}}^*$ containing $\Phi_{\mathbf{C}}(y_0^\sigma)$, which locally cut out Z^σ.

Ignoring convergence for a moment, regard the G_i in the completed local ring of $Y_{\mathbf{Q}_p}$ at y_0. This is a formal power series ring over $\mathbf{Q}_p$, and σ induces an injection from this completed local ring to the corresponding completed local ring of $Y_{\mathbf{C}}$ at y_0^σ; call this map $G \mapsto G^\sigma$. Then we have in fact

$$(9.4) \qquad G_i^\sigma = \text{ power series expansion of } F_i^\sigma \circ \Phi_{\mathbf{C}} \text{ at } y_0^\sigma.$$

This follows from just the same analysis of §3.3, or phrased informally, from the fact that the complex and p-adic period map satisfy the same differential equation.

It follows from (9.4) that the G_i^σ, *a priori* complex formal power series, are in fact convergent in a small complex neighbourhood of y_0^σ; their common vanishing locus for a sufficiently small such neighbourhood V coincides with $\Phi_{\mathbf{C}}^{-1}(Z^\sigma) \cap V$.

Corollary 9.2, applied to $Z^\sigma \subset \mathfrak{H}_{\mathbf{C}}^*$, shows that $\Phi_{\mathbf{C}}^{-1}(Z^\sigma) \cap V \subset Y_{\mathbf{C}}$ is not Zariski dense in $Y_{\mathbf{C}}$. Indeed, after analytically continuing $\Phi_{\mathbf{C}}$ from V to a universal cover of $Y_{\mathbf{C}}$, there are only finitely many irreducible components of $\Phi_{\mathbf{C}}^{-1}(Z^\sigma)$ which intersect V (by local finiteness of irreducible components of an analytic set). We can apply Corollary 9.2 to each of them to conclude that the common zero-locus of G_i^σ on V is contained in the zero locus of some algebraic function G (i.e., G arises from a regular function on a Zariski-open subset of $Y_{\mathbf{C}}$ containing V).

Consider the ring $R_{\mathbf{C}} = \mathbf{C}\{x_1, \ldots, x_n\}$ of formal power series that are convergent in some neighbourhood of 0. Given an ideal I of this ring, we can associate a germ $V(I)$ of an analytic set at the origin. The locally analytic Nullstellensatz [10, §3.4] asserts that the ideal of functions vanishing along this germ is precisely the radical $\sqrt{I}$ of I.

We apply this with $R_{\mathbf{C}}$ the ring of germs of holomorphic functions near $y_0^{\sigma} \in Y_{\mathbf{C}}$, taking I to be the ideal generated by the G_i^{σ}. Then $\sqrt{I}$ is the ideal of functions vanishing on $V(I)$ and in particular contains G. Thus $G^m \in I$ for some $m \geqslant 1$.

Therefore the ideal spanned by G_i^{σ} inside the ring of locally convergent power series contains the image of an algebraic function, i.e. a regular function on some Zariski-open subset of $Y_{\mathbf{C}}$ containing y_0^{σ}. The same is then *a fortiori* true if we replace "locally convergent" by "formal," and this latter assertion can be carried back, via σ^{-1}, to $Y_{\overline{\mathbf{Q}_p}}$. Thus, there is a regular function H, in a neighbourhood of y_0 on $Y_{\overline{\mathbf{Q}_p}}$, belonging to the ideal

$$(9.5) \qquad\qquad H \in \langle G_1, \ldots, G_k \rangle$$

generated by the G_i in the completed local ring $\widehat{\mathcal{O}}$ of $Y_{\overline{\mathbf{Q}_p}}$ at y_0.

By taking a norm we may suppose that H in fact arises from a regular function in a neighbourhood of y_0 on $Y_{\mathbf{Q}_p}$. Without loss of generality (multiplying by a suitable denominator if necessary), we may suppose that H is regular on the chosen open affine around y_0, i.e., $H \in B_p$. Note that $B_p \otimes \overline{\mathbf{Q}_p}$ surjects on to each quotient $\widehat{\mathcal{O}}/\mathfrak{m}_{\widehat{\mathcal{O}}}^t$ (where $\mathfrak{m}_{\widehat{\mathcal{O}}}$ is the maximal ideal). Therefore, for each $t \geqslant 1$, there are $Z_1, \ldots, Z_k \in B_p \otimes \overline{\mathbf{Q}_p}$ such that

$$(9.6) \qquad\qquad H \in \sum Z_i G_i + \mathfrak{m}_{\widehat{\mathcal{O}}}^t.$$

By linear algebra we see that we can even choose $Z_i \in B_p$.

The function H then defines a rigid-analytic function on the residue disk of y_0. Thus H and G_i both lie inside the Tate algebra R previously defined. Recall that we have fixed a prime ideal $\mathfrak{p}$ of R, contained in the maximal ideal $\mathfrak{m}$ associated to y_0, and containing the ideal J generated by the G_i inside R.

Now (9.6) implies that

$$H \in J + \mathfrak{m}^t$$

for every $t \geqslant 1$. Then the image of H in $R/\mathfrak{p}$ lies in the intersection $\bigcap_{t \geqslant 1} \mathfrak{m}^t$. Krull's intersection theorem, applied to the Noetherian integral domain $R/\mathfrak{p}$, implies that the intersection of powers of $\mathfrak{m}$ is trivial. Therefore $H \in \mathfrak{p}$, as desired. $\qquad\square$

10. BOUNDS ON POINTS WITH GOOD REDUCTION

Let $\pi : X \to Y$ be a smooth proper morphism over $\mathbf{Z}[S^{-1}]$, whose fibers are geometrically connected of relative dimension d. The goal of this section is to bound $Y(\mathbf{Z}[S^{-1}])$ by means of the same general techniques we have used elsewhere in the paper, i.e., by studying the variation of p-adic Galois representations of the fibers. We refer the reader to the Introduction (§1) for a discussion of the methods and how they compare with the curve case; the main difference in this general setting is that the linear algebra arguments required to avoid semisimplicity are much more elaborate, and are discussed in §11.

10.1. Fix $y_0 \in Y(\mathbf{C})$, with fiber X_0 and set $\mathsf{V}_0 = H^d(X_{0,\mathbf{C}}, \mathbf{Q})^{\mathrm{prim}}$. This is equipped with an intersection form $\langle -, - \rangle$. Assume that the image of

$$(10.1) \qquad\qquad \pi_1(Y_{\mathbf{C}}, y_0) \to \mathrm{Aut}\,(\mathsf{V}_0 \otimes \mathbf{C}, \langle -, - \rangle)$$

has Zariski closure containing the identity component of the right-hand group.

The Hodge structure on V_0 induces a weight-zero Hodge structure on

(10.2) $$\text{Lie GAut}(V_0 \otimes \mathbf{C}, \langle -, - \rangle) \simeq \mathbf{C} \oplus \text{Sym}^2 V_0 \text{ or } \mathbf{C} \oplus \wedge^2 V_0,$$

according to the parity of $\langle -, - \rangle$. We will refer to this as the *adjoint Hodge structure* to distinguish it from the Hodge structure on $V_0 \otimes \mathbf{C}$.

Let h^p be the dimension of the Hodge component $(p, -p)$ in the adjoint Hodge structure.[8] For any $E \in \mathbf{Z}_{\geqslant 0}$ that is at most the dimension of the adjoint Hodge structure, let

$$T(E) = \text{sum of the topmost } E \text{ Hodge numbers.}$$

Here the *Hodge numbers* are the list of ps for which $h^p \neq 0$, each written with multiplicity h^p; thus, for example, if $p_{\max}$ is the largest p for which $h^p \neq 0$, then $T(1) = p_{\max}$, and if $h^{p_{\max}} > 1$ then $T(2) = 2p_{\max}$.

We can extend T to be a continuous piecewise linear function $[0, \sum_{j \in \mathbf{Z}} h^j] \to \mathbf{R}_{\geqslant 0}$ such that $T(0) = 0$, and with derivative specified as

(10.3) $$T'(x) = \begin{cases} p_{\max} & \text{for } x \in (0, h^{p_{\max}}), \\ p_{\max} - 1 & \text{for } x \in (h^{p_{\max}}, h^{p_{\max}} + h^{p_{\max}-1}), \\ \text{and so forth.} \end{cases}$$

The transcendence property of period mappings is an essential ingredient in the following theorem. It says that integral points on the base are not Zariski dense whenever the adjoint Hodge structure is quite "spread out," that is to say, whenever the contribution of large $|p|$ to the total dimension $\sum h^p$ is large.

Theorem 10.1. *Let $\pi : X \to Y$ be a smooth proper morphism over $\mathbf{Z}[S^{-1}]$, whose fibers are geometrically connected of relative dimension d. With notation as above, suppose that the monodromy representation has large image, i.e. that (10.1) is satisfied, and moreover that*

(10.4) $$\sum_{p>0} h^p \geqslant h^0 + \dim(Y)$$

and

(10.5) $$\sum_{p>0} p h^p > T\left(h^0 + \dim(Y)\right) + T\left(\frac{3}{2} h^0 + \dim(Y)\right).$$

Then $Y(\mathbf{Z}[S^{-1}])$ is not Zariski dense in Y.

If we assume, moreover, that the monodromy representation for any subvariety $Y' \subset Y$ continues to have large image[9] (see (10.1)), then in fact $Y(\mathbf{Z}[S^{-1}])$ is finite.

Roughly speaking, a condition of type (10.4) is easily seen to be necessary for our method: with reference to the discussion of §1.2 we want Y to be transverse to all orbits of a certain group $Z(\phi)$ on a flag variety; the dimension of the flag variety is $\sum_{p>0} h^p$, and in our argument we shall bound the dimension of $Z(\phi)$ above by h^0. Equation 10.5 is in practice a much more restrictive condition and is needed to control semisimplification.

The combinatorial machinations that give rise to inequality 10.5 could probably be greatly optimized. We aimed to give a treatment that was fairly short, at some cost to the sharpness of the results. Informally speaking, the condition says that the Hodge diamond of Y is not very concentrated near the middle.

---

[8]As in Section 2.5, we are abusing the symbol p to refer to the indexing on a Hodge filtration.

[9]This is an unrealistically strong assumption. We include this statement simply to make clear the importance of this problem – controlling monodromy drop along subvarieties – for our method.

10.2. **Application to hypersurfaces.** We will now outline the proof of the following statement:[10]

Proposition 10.2. *There exists n_0 and a function $D_0(n)$ such that both* (10.4) *and* (10.5) *apply to $X \to Y$ the universal family of hypersurfaces in $\mathbf{P}^n$ of degree d, so long as $n \geqslant n_0$ and $d \geqslant D_0(n)$.*

Numerical experiments suggest that $n_0 \approx 60$ will do. Note that this family indeed has large monodromy image by [4].

We must emphasize that, in this case, the dimension of Y is *very large*, and so the statement that $Y(\mathbf{Z}[S^{-1}])$ is not Zariski dense is very modest indeed; but it seems to us an interesting first step, and potentially one can then iterate the argument by replacing Y by the Zariski closure of integral points. As suggested by the last line of the Theorem, it becomes relevant to analyze the following question:

> What is the smallest possible codimension of a subvariety $Y' \subset Y$ along which the monodromy drops?

In outline, the proof of Proposition 10.2 is as follows. It can be verified (we will omit the proof) that the middle Hodge numbers h^{pq} of a degree-d hypersurface inside $\mathbf{P}^n$ satisfy

$$(10.6) \qquad h^{pq}(d) \sim \frac{d^n}{n!} A(n,p)$$

where $p + q = n - 1$, and $A(n,p)$ is the *Eulerian number*: the number of permutations $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ with the property that $\sigma(i+1) > \sigma(i)$ for precisely p values of i. (Here we fix the dimension n of the ambient projective space, and the meaning of $\sim$ is that the ratio approaches 1 as $d \to \infty$.) Now consider $\alpha_p := \frac{1}{n!} A(n,p)$, which defines a probability distribution on $p \in \{0, \ldots, n-1\}$. The conclusion will be deduced, in essence, from the fact that α_p is well approximated by a binomial distribution with mean $n/2$ and variance $n/12$. We now describe the details.

First, consider the Hodge numbers h^p for the adjoint Hodge structure. Since the dimension of the symmetric or adjoint square of a k-dimensional vector space equals $\frac{k^2 \pm k}{2}$, we have

$$2h^p = \sum_{p_1 + p_2 = p + (n-1)} h^{p_1, q_1} h^{p_2, q_2} \pm h^{(p+n-1)/2, (-p+n-1)/2}$$

where in all cases $p_1 + q_1 = p_2 + q_2 = n - 1$. In particular, we deduce that

$$(10.7) \qquad h^p \sim d^{2n} \frac{1}{2} \underbrace{\sum_{p_1 - p_2 = p} \alpha_{p_1} \alpha_{p_2}}_{\beta_p}.$$

Next, note that the dimension of the moduli space of degree d hypersurfaces in $\mathbf{P}^n$ is given by

$$\binom{n+d}{d-1} - 1 = \frac{d(d+1) \ldots (d+n)}{(n+1)!} - 1 \sim \frac{d^{n+1}}{(n+1)!}$$

where the meaning of $\sim$ is as before. In particular, for any fixed $n \geqslant 2$,

$$(10.8) \qquad \lim_{d \to \infty} \frac{\dim Y}{h^0} = 0,$$

where h^0 is the dimension of the zeroth Hodge number for the adjoint structure.

---

[10]In Section 10.2 *only*, the symbol d represents the degree of a hypersurface, and $n - 1$ its dimension.

Let $X(n)$ (or just X for short) be the random variable which sends a uniformly distributed random permutation σ of $\{1, \ldots, n\}$ to the number of i for which $\sigma(i+1) > \sigma(i)$, subtract $\frac{n-1}{2}$. Write $y_i (1 \le i \le n-1)$ for the random variable, on the same space, with value $1/2$ if $\sigma(i+1) > \sigma(i)$, and $-1/2$ if $\sigma(i+1) < \sigma(i)$. Thus $X = \sum y_i$ and the expectation $\mathbb{E}(X)$ is zero. The variance of X is then given by

$$(10.9) \qquad \mathrm{Var}(X) = \sum_{i,j} \mathbb{E}(y_i y_j) = \underbrace{\frac{n-1}{4}}_{i=j} - \underbrace{2\frac{n-2}{12}}_{|i-j|=1} = \frac{n+1}{12}.$$

Now let $X'(n)$ be the random variable obtained by convolving $X(n)$ with itself, i.e. with adding together two copies of $X(n)$. Then

$$(10.10) \qquad \mathrm{Var}(X') = 2\mathrm{Var}(X) = \frac{n+1}{6}, \text{ and the probability that } (X' = p) = \beta_p,$$

where β_p is as in (10.7). Moreover, it is also known (see [9] for discussion and references to the literature) that as $n \to \infty$,
(10.11)

$X(n)/\sqrt{n}$ converges in distribution to a normal distribution with variance $1/12$.

and it follows then that $X'(n)/\sqrt{n}$ converges in distribution to a normal distribution with variance $1/6$. It follows in particular that

$$(10.12) \qquad \sum_{p>0} p\beta_p > A\sqrt{n}$$

for some absolute $A > 0$. We also need:

Lemma 10.3. *For sufficiently large n, we have $\beta_0 < \frac{40}{\sqrt{n}}$.*

Proof. The sequence β_p is symmetric and log-concave. The symmetry follows readily from the definition, whereas the second statement follows from the classical fact that the Eulerian numbers are log-concave. (See, for example, [21, Thms 1.4, 3.3].)

Let $c = \frac{1}{40}$. This number is chosen to be less than the density of the normal distribution with mean zero and variance $1/6$, at the point 1.1. From the convergence in distribution of X', it follows that for all large enough n there exists $P > \sqrt{n}$ with the property that $\beta_P > \frac{c}{\sqrt{n}}$.

We show that $\beta_0 \le \frac{c^{-1}}{\sqrt{n}}$. Suppose not; then log-concavity means

$$\beta_p > \frac{(1/c)^{1-p/P} c^{p/P}}{\sqrt{n}}$$

for all $p \in [0, P]$. In particular, this implies that $\beta_p > \frac{1}{\sqrt{n}}$ whenever $|p| \le P/2$. This contradicts $\sum \beta_p = 1$ for large enough n. $\qquad \square$

Proof. (of Proposition 10.2): In what follows, write "for big enough n and d" as an abbreviation for "for $n \ge n_0$ and $d \ge D_0(n)$, for some function D_0 of n."

There are two conditions to be checked, (10.4) and (10.5). That the former condition holds for big enough n and d follows from (10.8), (10.7) and the convergence in distribution of $X'(n)/\sqrt{n}$. It remains then to verify that (10.5) holds for big enough n and d.

Write $T(y)$ for the sum of the topmost y adjoint Hodge numbers and H for the total dimension of the adjoint Hodge structure. We claim that

$$2T(2h^0) < \sum_{p>0} ph^p,$$

for big enough n and d. That statement readily implies the desired conclusion, in view of (10.8).

By Lemma 10.3, for sufficiently large n we have $\beta_0 < \frac{40}{\sqrt{n}}$. Therefore, for d sufficiently large (depending on n) we have $h^0 < 40H/\sqrt{n}$. On the other hand, by (10.12), the right-hand side $\sum_{p>0} ph^p$ is bounded below by a constant multiple of $H\sqrt{n}$, for big enough d and n. Therefore, it is enough to verify that, for fixed positive constants c, δ, we have the inequality

$$(10.13) \qquad\qquad T\left(\frac{cH}{\sqrt{n}}\right) \leqslant \delta H \sqrt{n}$$

for big enough n and d.

Let $\epsilon = \frac{\delta}{2c}$. Separate the contribution of Hodge numbers above and below ϵn to T; we get:

$$T\left(\frac{cH}{\sqrt{n}}\right) \leqslant (\epsilon n)\frac{cH}{\sqrt{n}} + \sum_{p > \epsilon n} ph^p$$

Now the first quantity is bounded by $\frac{1}{2}\delta H\sqrt{n}$. The second quantity equals

$$\sum_{p>\epsilon n} ph^p = H \sum_{p>\epsilon n} p\beta_p + H \sum_{p>\epsilon n} p\left(\frac{h_p}{H} - \beta_p\right).$$

There is a function D_1 such that, for $d \geqslant D_1(n)$, the second term is at most H. Also, using the variance bound $\sum p^2 \beta_p = \frac{n+1}{6}$, the first term is at most $H\epsilon^{-1}$. Thus,

$$T\left(\frac{cH}{\sqrt{n}}\right) \leqslant \frac{1}{2}\delta H\sqrt{n} + H(1 + \epsilon^{-1}),$$

and the latter term is certainly bounded above by $\frac{1}{2}\delta H\sqrt{n}$ for $n \geqslant n_0$. This concludes the proof of (10.13), so also of our Proposition.

$\square$

10.3. **Setup for the proof of Theorem 10.1.** In what follows, ℓ denotes an arbitrary prime number not belonging to the fixed set S.

Working in the complex analytic category, let $\mathsf{V} = \mathbf{R}^d\pi_*\mathbf{Q}$. It is a local system of $\mathbf{Q}$-vector spaces on $Y(\mathbf{C})$ (and it admits an integral structure); the V_0 defined in §10.1 is its fiber above y_0.

Let $\mathbf{G}$ be the connected automorphism group of the intersection form on V_0, a semisimple $\mathbf{Q}$-group; also let

$$\mathbf{G}' = \mathrm{GAut}(\mathsf{V}_0, \langle -, - \rangle),$$

the corresponding generalized automorphism group, where we permit to scale the form $\langle -, - \rangle$.

Passing to ℓ-adic étale cohomology, there is a monodromy mapping $\pi_1^{\mathrm{arith}}(Y, y_0) \longrightarrow \mathbf{G}'(\mathbf{Q}_\ell)$ and the section associated to an integral point $y \in Y(\mathbf{Z}[S^{-1}])$ gives a representation

$$\rho_{y,\ell} : G_\mathbf{Q} \longrightarrow \mathbf{G}'(\mathbf{Q}_\ell).$$

This describes the Galois action on the primitive geometric étale cohomology of the fiber X_y in degree d (after using an isomorphism $\mathsf{V}_0 \otimes \mathbf{Q}_\ell \simeq \mathsf{V}_y \otimes \mathbf{Q}_\ell$).

In what follows we will freely use certain results about Galois representations into $\mathbf{G}'$ which are parallel to certain known results about GL_n-valued representations; we refer to §2.3 for further discussion of these points.

We denote by $\rho_{y,\ell}^{\mathrm{ss}}$ the semisimplification of $\rho_{y,\ell}$ relative to $\mathbf{G}'$ (see §2.3). By Faltings' finiteness theorem (Lemma 2.6) there are only finitely many possibilities for the $\mathbf{G}'(\mathbf{Q}_\ell)$-conjugacy class of $\rho_{y,\ell}^{\mathrm{ss}}$.

We must understand the variation of the representation ρ_y with y; as usual, we will study this using the period mapping. We begin with the complex Hodge structures.

The Hodge structure on V_y, the fiber of V at y, is given by a self-dual filtration

$$\text{(10.14)} \qquad\qquad \mathsf{V}_y = F^0 \mathsf{V}_y \supset \cdots \supset F^i \mathsf{V}_y \supset \ldots$$

and in this way we can regard the period mapping as

$$\text{(10.15)} \qquad \text{universal cover of } Y_{\mathbf{C}} \longrightarrow \text{Mumford-Tate domain for } \mathbf{G}',$$

where the Mumford-Tate domain in question is understood to be the space of self-dual filtrations on V_0 with the same dimensional data as the Hodge filtration on V_0.

Also, the Hodge structure on V_0 gives rise to a morphism

$$\varphi_0 : S^1 \longrightarrow \mathbf{G}'(\mathbf{C}).$$

For each $y \in Y(\mathbf{Z}[S^{-1}])$, we may reduce modulo ℓ and consider the crystalline Frobenius of the reduction $X_{y,\mathbf{F}_\ell} := X_y \times_{\mathbf{Z}[S^{-1}]} \mathbf{F}_\ell$. This determines a transformation of the (primitive) crystalline cohomology

$$\mathrm{F}_y^{\mathrm{crys},\ell} \in \mathrm{Aut}\, H_{\mathrm{crys}}^d(X_{y,\mathbf{F}_\ell})^{\mathrm{prim}}.$$

The characteristic polynomial of this endomorphism is determined by the ζ-function of $X_{y,\mathbf{F}_\ell}$, and it can be deduced (see [22]) that its eigenvalues coincide with the eigenvalues of ℓ-Frobenius on p-adic absolute étale cohomology for any prime $p \neq \ell$.

In the coming subsections we will prove the following two Lemmas:

Lemma 10.4 (Frobenius centralizer small, for some ℓ below an absolute bound). *There exists an integer L with the following property:*

> *For any $y \in Y(\mathbf{Z}[S^{-1}])$, there exists a prime $\ell \leqslant L, \ell \notin S$ such that the semisimplification of $\mathrm{F}_y^{\mathrm{crys},\ell}$ (and so also the crystalline Frobenius itself) satisfies*

$$\text{(10.16)} \qquad\qquad \dim Z\big(\big[\mathrm{F}_y^{\mathrm{crys},\ell}\big]^{\mathrm{ss}}\big) \leqslant \dim Z_{\mathbf{G}'(\mathbf{C})}(\varphi_0).$$

On the left hand side, we take the centralizer inside $\mathrm{GAut}(H_{\mathrm{crys}}^{d,\mathrm{prim}})$, *to which the crystalline Frobenius – and so also its semisimplification – belongs.*

Lemma 10.5 (Not Zariski dense. This is where semisimplicity gets taken care of.). *Given a prime $\ell \notin S$ and $y_0 \in Y(\mathbf{Z}[\frac{1}{S}])$ with the property that the centralizer of crystalline Frobenius $\mathrm{Frob}_{y_0}^{\mathrm{crys},\ell}$ is at most the dimension of $Z_{G'(\mathbf{C})}(\varphi_0)$, the set*

$$\text{(10.17)} \qquad\qquad \{y \in Y(\mathbf{Z}[\tfrac{1}{S}]) : y \equiv y_0 \text{ modulo } \ell, \rho_{y,\ell}^{\mathrm{ss}} \simeq \rho_{y_0,\ell}^{\mathrm{ss}}\}$$

is not Zariski dense. (Here $\simeq$ means that the representations are $\mathbf{G}'$-conjugate).

Assuming these Lemmas, let us conclude the proof of Theorem 10.1. With L as in Lemma 10.4 let $N = \prod_{\ell \leqslant L, \ell \notin S} \ell$. Now each $y \in Y(\mathbf{Z}[\frac{1}{S}])$ gives a collection of representations $\rho_y^{\mathrm{ss}} : G_{\mathbf{Q}} \to \mathbf{G}'(\mathbf{Q}_\ell)$, one for each ℓ dividing N. For each ℓ dividing N, let $\mathcal{G}_\ell$ be the set of representations $G_{\mathbf{Q}} \to \mathbf{G}'(\mathbf{Q}_\ell)$ that arises as some ρ_y^{ss}. This is a finite set (modulo conjugacy) by Lemma 2.6 applied to $\mathbf{G}' \subset \mathrm{GL}(\mathsf{V}_0)$; note that it is straightforward to verify that the integrality of characteristic polynomial of Frobenius passes from the whole cohomology to the primitive cohomology.

Call a pair (y, ℓ) as in Lemma 10.4 *good* if it satisfies (10.16). For each ℓ, Lemma 10.5 and the finiteness of $\mathcal{G}_\ell$ guarantee that the set of y for which (y, ℓ) is good is not Zariski dense. Taking the union over $\ell \leq L$ and applying Lemma 10.4, we see that $Y(\mathbf{Z}[S^{-1}])$ is itself not Zariski dense.

10.4. **Proof of Lemma 10.4**. Fix a prime $p \notin S$ and let $\rho_{y,p} : G_{\mathbf{Q}} \to \mathbf{G}'(\mathbf{Q}_p)$ be the p-adic Galois representation at y, as above. We have observed that there are only finitely many possibilities for $\rho_{y,p}^{\mathrm{ss}}$ (here, and below, the semisimplification is taken inside $\mathbf{G}'$).

Let $\mathbf{H}$ be the Zariski closure of $\rho_{y,p}^{\mathrm{ss}}(G_{\mathbf{Q}})$, with identity component $\mathbf{H}^\circ$. It is a reductive group (because we took the semisimplification, see §2.3 and references therein). Call an element in $\mathbf{H}^\circ(\overline{\mathbf{Q}_p})$ *very regular* if it is semisimple and:

> (*) its centralizer inside $\mathrm{Aut}(\mathsf{V}_0 \otimes \overline{\mathbf{Q}_p})$ has minimal dimension amongst all semisimple elements of $\mathbf{H}^\circ(\overline{\mathbf{Q}_p})$.

Choose a maximal torus $\mathbf{T}_0 \subset \mathbf{H}^\circ$, and let Φ be the set of nontrivial characters $\mathbf{T}_0 \to \mathbf{G}_m$ arising from the conjugation action of $\mathbf{T}_0$ on the Lie algebra of $\mathrm{Aut}(\mathsf{V}_0 \otimes \overline{\mathbf{Q}_p})$. For $t \in \mathbf{T}_0(\overline{\mathbf{Q}_p})$ the dimension of the centralizer of t, in $\mathrm{Aut}(\mathsf{V}_0 \otimes \overline{\mathbf{Q}_p})$, is the dimension of the centralizer of $\mathbf{T}_0$ in $\mathrm{Aut}(\mathsf{V}_0 \otimes \overline{\mathbf{Q}_p})$, plus the number of roots $\alpha \in \Phi$ with $\alpha(t) = 1$ (counted with multiplicity). The condition (*) for an element $t \in \mathbf{T}_0(\overline{\mathbf{Q}_p})$ amounts to asking that $\alpha(t) \neq 1$ for all $\alpha \in \Phi$. In particular:

- Any very regular element is regular inside $\mathbf{H}^\circ$, and
- Condition (*) implies the same condition with $\mathrm{Aut}(\mathsf{V}_0 \otimes \overline{\mathbf{Q}_p})$ replaced by $\mathbf{G}'$.

The set of very regular elements is a nonempty Zariski-open subset of $\mathbf{H}^\circ$ (so also of $\mathbf{H}$). Indeed, the function $f = \prod_{\alpha \in \Phi}(\alpha(t) - 1)$ defines a regular function on $\mathbf{T}_0$ which is invariant under the Weyl group. Therefore f extends to a regular function on $\mathbf{H}^\circ$, and the set of very regular elements is the locus where $f \neq 0$ (this forces semisimplicity).

It follows, then, that the set of very regular elements in $\mathbf{H}(\mathbf{Q}_p)$ is the complement of a proper Zariski-closed set. The preimage of the very regular set under $\rho_{y,p}^{\mathrm{ss}} : G_{\mathbf{Q}} \to \mathbf{H}(\mathbf{Q}_p)$ is nonempty, because $\rho_{y,p}^{\mathrm{ss}}(G_{\mathbf{Q}})$ is Zariski-dense in $\mathbf{H}$. This preimage is also topologically open, since the very regular set is open. By the Chebotarev density theorem, then, we may choose some ℓ such that

$$(10.18) \qquad \rho_{y,p}^{\mathrm{ss}}(\mathrm{Frob}_\ell) \text{ is a very regular element of } \mathbf{H}^\circ.$$

Because there are only finitely many possibilities for $\rho_{y,p}^{\mathrm{ss}}$, this ℓ can be taken to be bounded above by L that depends only on $S, p, \dim(\mathsf{V})$.

On the other hand, it is known that:

> the Zariski closure of $\rho_{y,p}(G_{\mathbf{Q}_p})$ (this is an algebraic subgroup of $\mathbf{G}'$) contains a group $\mathbf{S}$ defined over $\overline{\mathbf{Q}_p}$, with the following property: with respect to a suitable isomorphism $\overline{\mathbf{Q}_{\mathfrak{p}}} \simeq \mathbf{C}$, the group $\mathbf{S}$ becomes isomorphic to the Hodge torus, i.e. to the Zariski closure of the image of φ_0 in $\mathbf{G}'(\mathbf{C})$.

A result of rather similar nature to the quoted statement was proved by Sen [39] using Hodge–Tate decomposition (Sen's result pertains to the target group GL_n). It can be deduced using a remarkable result of Wintenberger [46] about functorially splitting the Hodge filtration for Fontaine–Laffaille modules. This is carried out by Pink [33, §2]; this latter method also readily adapts to $\mathbf{G}'$ target.[11]

---

[11] We outline how this is done. We may describe the Zariski closure $\mathbf{Z}$ in question as the Tannakian group associated to the neutral Tannakian category of $G_{\mathbf{Q}_p}$-modules generated by $\mathsf{V}_y \otimes \mathbf{Q}_p$ (i.e., the automorphisms

Thus

$$(10.19) \qquad \dim Z_{\mathbf{G}'}(\mathbf{S}) = \dim Z_{\mathbf{G}'(\mathbf{C})}(\varphi_0).$$

Moreover, a $\mathbf{G}(\overline{\mathbf{Q}_p})$-conjugate of $\mathbf{S}$ – call it $\mathbf{S}'$ – is also contained in the Zariski closure of the image of $\rho_{y,p}^{\mathrm{ss}}(G_{\mathbf{Q}})$. Indeed, choose a parabolic $\mathbf{Q} \leqslant \mathbf{G}'$ containing the image of $\rho_{y,p}$ and minimal for that property; then $\rho_{y,p}^{\mathrm{ss}}$ is obtained by projecting $\rho_{y,p}$ to a Levi factor of $\mathbf{Q}$, and in particular the Zariski closure of the image of $\rho_{y,p}^{\mathrm{ss}}$ certainly contains the projection of the Zariski closure of the image of $\rho_{y,p}$. Now apply Lemma 2.5.

Now we have

$$[\rho_{y,p}(\mathrm{Frob}_\ell)]^{\mathrm{ss}} \overset{\mathrm{Lemma\ 2.4}}{\sim} \left[\rho_{y,p}^{\mathrm{ss}}(\mathrm{Frob}_\ell)\right]^{\mathrm{ss}} \overset{(10.18)}{=} \rho_{y,p}^{\mathrm{ss}}(\mathrm{Frob}_\ell)$$

where $\sim$ denotes $\mathbf{G}'(\overline{\mathbf{Q}_p})$-conjugacy. By (10.18), the definition of "very regular" element, and the discussion that follows it, the centralizer of this element inside $\mathbf{G}'$ is as small as possible, amongst semisimple elements in $\mathbf{H}^{\circ}(\overline{\mathbf{Q}_p})$. In particular, this centralizer is at most as large as the centralizer of $\mathbf{S}'$ on $\mathbf{G}'$, and so

$$(10.20) \quad \dim Z_{\mathbf{G}'}\left[\rho_{y,p}(\mathrm{Frob}_\ell)\right]^{\mathrm{ss}} \leqslant \dim Z_{\mathbf{G}'}(\mathbf{S}') = \dim Z_{\mathbf{G}'}(\mathbf{S}) = \dim Z_{\mathbf{G}'(\mathbf{C})}(\varphi_0).$$

We now transfer this to the corresponding assertion for the crystalline Frobenius $\mathrm{Frob}_\ell^{\mathrm{crys}}$. We know that the crystalline ℓ-Frobenius on the ℓ-adic vector space $H_{\mathrm{crys}}^d(X_{y,\mathbf{F}_\ell})$ and the usual ℓ-Frobenius on the p-adic geometric étale cohomology of X_y have the same characteristic polynomial. The same is true for primitive parts. Thus $\rho_{y,p}(\mathrm{Frob}_\ell)^{\mathrm{ss}}$ and $(\mathrm{Frob}_\ell^{\mathrm{crys}})^{\mathrm{ss}}$ both have the same characteristic polynomial; also they both scale the bilinear forms by ℓ.

Split $\mathsf{V}_0 \otimes \mathbf{Q}_p = \bigoplus \mathsf{V}_\lambda$ into eigenspaces for $\rho_{y,p}(\mathrm{Frob}_\ell)^{\mathrm{ss}}$. The bilinear form gives a perfect pairing between each V_λ and $\mathsf{V}_{\ell\lambda^{-1}}$ (interpreted as a self-pairing when $\lambda^2 = \ell$); the centralizer of $\rho_{y,p}(\mathrm{Frob}_\ell)^{\mathrm{ss}}$ in $\mathbf{G}'$ is the set of g stabilizing each V_λ and respecting these pairings. In particular the centralizer dimension is determined by the function $\lambda \mapsto \dim(\mathsf{V}_\lambda)$; the same analysis applies for $(\mathrm{Frob}_\ell^{\mathrm{crys}})^{\mathrm{ss}}$. We deduce that

$$(10.21) \quad \dim Z_{\mathrm{GAut}}([\mathrm{Frob}_\ell^{\mathrm{crys}}]^{\mathrm{ss}}) = \dim Z_{\mathbf{G}'}(\rho_{y,p}(\mathrm{Frob}_\ell)^{\mathrm{ss}}) \overset{(10.20)}{\leqslant} \dim Z_{\mathbf{G}'(\mathbf{C})}(\varphi_0),$$

concluding the proof of the Lemma. $\qquad\square$

10.5. **Proof of Lemma 10.5.** We must analyze the set

$$(10.22) \qquad \{y \in Y(\mathbf{Z}[\tfrac{1}{S}]) : y \equiv y_0 \text{ modulo } p, \rho_{y,p}^{\mathrm{ss}} \simeq \rho_{y_0,p}^{\mathrm{ss}}\}$$

(we have switched from ℓ to p for typographical simplicity). Here we are assuming that the centralizer of crystalline Frobenius $\mathrm{Frob}_{y_0}^{\mathrm{crys},p}$, inside the group GAut of generalized automorphisms of the intersection pairing, has dimension at most the dimension of $Z_{G'(\mathbf{C})}(\varphi_0)$,

---

of the natural fiber functor). By the theory of Fontaine–Laffaille, there is another fiber functor on this category, arising from passing to filtered ϕ-modules; in particular, this gives rise to another Tannakian group $\mathbf{Z}'$, which acts on the (primitive part of the) de Rham cohomology of $X_y \times_{\mathbf{Q}} \mathbf{Q}_p$. These two fiber functors become isomorphic over $\overline{\mathbf{Q}_p}$ (cf. [12, §3]); in particular there is an isomorphism of $\mathsf{V}_y \otimes \overline{\mathbf{Q}_p}$ with the de Rham cohomology of $X_y \otimes_{\mathbf{Q}} \overline{\mathbf{Q}_p}$, which can be taken to preserve the respective intersection forms, and which carries $\mathbf{Z}_{\overline{\mathbf{Q}_p}}$ to $\mathbf{Z}'_{\overline{\mathbf{Q}_p}}$.

The Hodge filtration gives this fiber functor the structure of a filtered fiber functor; it gives a parabolic subgroup $\mathbf{P}' \subset \mathbf{Z}'$. Now Wintenberger's canonical splitting of the Hodge filtration provides a character $\varphi_W : \mathbf{G}_m \to \mathbf{P}'$.

Now pass to $\mathbf{C}$ by means of an isomorphism $\overline{\mathbf{Q}_p} \simeq \mathbf{C}$; then $\mathbf{Z}'_{\mathbf{C}}$ acts on the cohomology of $X_y \otimes_{\mathbf{Q}} \mathbf{C}$, as does φ_0. We claim that φ_0 and $\varphi_W|_{S^1}$ are conjugate inside $\mathrm{GAut}(H^d(X_y \otimes_{\mathbf{Q}} \mathbf{C})^{\mathrm{prim}})$; but they both preserve the Hodge filtration and induce the same scalar on the successive quotients; the conjugacy then follows by Lemma 2.5.

Now let us unwind the condition in (10.22), namely, that the semisimplified p-adic Galois representations for y and for y_0 are isomorphic. Recall that semisimplification is taken relative to the ambient group $\mathbf{G}'(\mathbf{Q}_p)$. The representation $\rho_{y,p}$ is realized on $H^d(X_y, \mathbf{Q}_p)^{\mathrm{prim}}$, and similarly for y_0. The semisimplification of $\rho_{y_0,p}$ (in the ambient group $\mathbf{G}'$) is obtained by taking a maximal self-dual flag of $\rho_{y_0,p}$-stable subspaces

$$0 \subset \mathfrak{f}^1 \subset \mathfrak{f}^2 \subset \cdots \subset \mathfrak{f}^\mathsf{m} \subset \underbrace{(\mathfrak{f}^\mathsf{m})^\perp}_{\mathfrak{f}^{\mathsf{m}+1}} \subset \cdots \subset H^d(X_{y_0}, \mathbf{Q}_p)^{\mathrm{prim}}$$

with the property that the representation on each graded piece is irreducible. (For the middle graded piece, i.e. the piece $\mathfrak{f}^{\mathsf{m}+1}/\mathfrak{f}^\mathsf{m}$, we interpret "irreducible" to mean that there is no *isotropic* invariant subspace, see §2.3 for explanation. We also permit the possibility that $\mathfrak{f}^\mathsf{m} = \mathfrak{f}^{\mathsf{m}+1}$.)

Since ρ_y^{ss} and $\rho_{y_0}^{\mathrm{ss}}$ are isomorphic, it means that there exist such flags $\mathfrak{f}_y$ and $\mathfrak{f}_0$ for both y and y_0 such that the $G_\mathbf{Q}$-representations on $\bigoplus_j \mathrm{gr}_j^\mathfrak{f}$ are isomorphic. In fact, we can arrange even that this is true for every j individually, and that the isomorphism preserves the intersection form for $j = \mathsf{m}$: this follows by using the last sentence of Lemma 2.6. [12]

Now the functors of p-adic Hodge theory carry $H^d(X_{y_0}, \mathbf{Q}_p)$ to $H^d_{\mathrm{dR}}(X_{y_0}, \mathbf{Q}_p)$ and similarly for y. Moreover, the intersection form

$$H^d(X_{y_0}, \mathbf{Q}_p) \otimes H^d(X_{y_0}, \mathbf{Q}_p) \longrightarrow H^{2d}(X_{y_0}, \mathbf{Q}_p) \, (\simeq \mathbf{Q}_p(-d))$$

is carried to the intersection form $H^d_{\mathrm{dR}}(X_{y_0}, \mathbf{Q}_p) \otimes H^d_{\mathrm{dR}}(X_{y_0}, \mathbf{Q}_p) \longrightarrow H^{2d}_{\mathrm{dR}}(X_{y_0}, \mathbf{Q}_p)$. These assertions remain valid for the primitive parts of cohomology.

The flags $\mathfrak{f}_y$ and $\mathfrak{f}_0$ are in particular $G_{\mathbf{Q}_p}$-invariant, and, under the correspondence of p-adic Hodge theory, these flags $\mathfrak{f}_y, \mathfrak{f}_0$ correspond to self-dual flags $\mathfrak{f}_y^{\mathrm{dR}}$ and $\mathfrak{f}_0^{\mathrm{dR}}$ inside the associated "de Rham" vector spaces:

$$\mathfrak{f}_y^{\mathrm{dR}} \text{ in } H^d_{\mathrm{dR}}(X_y)^{\mathrm{prim}} \text{ and } \mathfrak{f}_0^{\mathrm{dR}} \text{ in } H^d_{\mathrm{dR}}(X_{y_0})^{\mathrm{prim}}.$$

Moreover, under the correspondence of p-adic Hodge theory, the filtered ϕ-modules

$$(\mathfrak{f}_y^{\mathrm{dR}})^{m+1}/(\mathfrak{f}_y^{\mathrm{dR}})^m \text{ and } (\mathfrak{f}_0^{\mathrm{dR}})^{m+1}/(\mathfrak{f}_0^{\mathrm{dR}})^m$$

correspond, respectively, to the Galois representations of $G_{\mathbf{Q}_p}$ on $\mathfrak{f}_y^{m+1}/\mathfrak{f}_y^m$ and $\mathfrak{f}_0^{m+1}/\mathfrak{f}_0^m$. These Galois representations are isomorphic, so the filtered ϕ-modules just mentioned above are also isomorphic. For $m = \mathsf{m}$, the middle degree, the isomorphism of Galois representations can be taken to preserve the bilinear form, and so the same is true for the isomorphism of filtered ϕ-modules.

The map sending y to the Hodge filtration on X_y defines a period map

$$\Phi_p : \text{residue disk at } y_0, \text{ modulo } p \longrightarrow p\text{-adic period domain } \mathfrak{H}_p$$

where $\mathfrak{H}_p$ is now the set of self-dual flags inside $V := H^d_{\mathrm{dR}}(X_{y_0}, \mathbf{Q}_p)^{\mathrm{prim}}$ with the same dimensional data as the Hodge filtration on $H^d_{\mathrm{dR}}(X_{y_0})^{\mathrm{prim}}$. Write ϕ for the Frobenius map on V. Our analysis above shows that the the set $\{y \in Y(\mathbf{Z}[\frac{1}{S}]) : y \equiv y_0 \text{ modulo } p, \rho_{y,p}^{\mathrm{ss}} \simeq \rho_{y_0,p}^{\mathrm{ss}}\}$ is contained in a finite union of sets of the following type:

$$\Phi_p^{-1}(\mathfrak{S}),$$

---

[12] In more detail: In our reasoning to date, instead of using the finiteness of conjugacy classes of possible $\rho_y^{\mathrm{ss}} : G_\mathbf{Q} \to \mathbf{G}'$, we could instead use the stronger finiteness provided by the last sentence of Lemma 2.6. Namely, we fix for each y a parabolic subgroup $\mathbf{Q}_y$ containing the image of ρ_y, such that the projection of ρ_y to its Levi gives the semisimplification, and then use the finiteness up to conjugacy of possible pairs $(\mathbf{Q}_y, \rho_y^{\mathrm{ss}})$.

where $\mathfrak{S} \subset \mathfrak{H}_p$ is the space of filtrations F on $V = H^d_{\mathrm{dR}}(X_{y_0}, \mathbf{Q}_p)^{\mathrm{prim}}$ with the property that there exists *another* self-dual filtration $\mathfrak{f}$, the "semisimplification filtration":

$$0 = \mathfrak{f}^0 \subset \mathfrak{f}^1 \subset \mathfrak{f}^2 \subset \cdots \subset \mathfrak{f}^{\mathsf{m}} \subset \underbrace{(\mathfrak{f}^{\mathsf{m}})^{\perp}}_{\mathfrak{f}^{\mathsf{m}+1}} \subset \cdots \subset \mathfrak{f}^{2\mathsf{m}+1} = V$$

with the following properties:

- $\mathfrak{f}$ is ϕ-stable.
- The filtration induced by F on each graded piece $\mathrm{gr}^{\mathfrak{f}}_j$ has weight equal to $d/2$. (This follows because it arises from applying p-adic Hodge theory to the restriction of a global Galois representation that is pure of weight d, using Lemma 2.9.)
- We have an isomorphism of filtered ϕ-modules

$$(\mathrm{gr}^{\mathfrak{f}}_j, \text{filtration induced by } F) \simeq (\mathrm{gr}^{\mathfrak{f}_0}_j, \text{filtration induced by } F_0)$$

  (i.e., an isomorphism of vector spaces respecting Hodge filtration and Frobenius). In particular, the left-hand side of the above equation lies in a fixed isomorphism class.

  On the right hand side F_0 is the filtration at y_0. In the case of the middle graded piece $j = \mathsf{m}$, the isomorphism above may be taken, moreover, to preserve the bilinear forms on both sides.

The following Proposition 10.6 implies that the codimension of the set $\mathfrak{S}$ above is at least equal to the dimension of Y. Given this Proposition, Lemma 10.5 now follows from Lemma 9.3 (the p-adic transcendence of period mappings).

Proposition 10.6. *Suppose V is a vector space over the field K equipped with a bilinear form $\langle -, - \rangle$ and a linear automorphism $\phi \in \mathrm{GAut}(V)$.*

Suppose $A_1, \ldots, A_{\mathsf{m}}$ is a collection of K-vector spaces, each equipped with a decreasing filtration and a linear automorphism $\phi_i : A_i \to A_i$. We suppose the final space A_{m} is equipped with a bilinear form $\langle -, - \rangle$.

Consider all self-dual filtrations

$$V = F^0 V \supset F^1 V \supset \cdots \supset F^d V \supset F^{d+1} V = \{0\}$$

on V, where we fix the dimensions of each F^i.

Call such a filtration F "bad" if there exists another self-dual filtration $\mathfrak{f}$ on V

$$0 = \mathfrak{f}^0 \subset \mathfrak{f}^1 \subset \cdots \subset \mathfrak{f}^{\mathsf{m}} \subset \mathfrak{f}^{\mathsf{m}+1} \subset \cdots \subset \mathfrak{f}^{2\mathsf{m}+1} = V.$$

such that the following conditions hold.

- (a) *$\mathfrak{f}$ is ϕ-stable.*
- (b) *The weight of the filtration induced by F on each graded piece $\mathrm{gr}^k_{\mathfrak{f}}$ equals $d/2$, i.e. the weight of the filtration F on V.*
- (c) *There exists an isomorphism of filtered ϕ-modules:*

$$\left(\mathrm{gr}^j_{\mathfrak{f}} V, \text{filtration induced by } F \right) \simeq A_j$$

  for each $j \leq \mathsf{m}$, and in the middle dimension $j = \mathsf{m}$ this also preserves bilinear forms.

Define the Hodge numbers h^p as the dimension of $\mathrm{gr}^p_F \mathrm{LieGAut}(V)$; let $T(y)$ be the sum of the topmost y Hodge numbers, extended by linearity as in (10.3).

Put $z = \dim Z(\phi^{\mathrm{ss}})$, the dimension of the centralizer of the semisimple part of ϕ in $\mathrm{GAut}(V)$.

If e is a positive integer such that

(10.23) $$\text{number of positive Hodge numbers} \geqslant z + e$$

and

(10.24) $$\text{sum of all positive Hodge numbers} > T(z+e) + T\left(\frac{h^0}{2} + z + e\right),$$

then the codimension of the space of bad filtrations is greater than or equal to e.

To be clear, we apply this with:

- $K = \mathbf{Q}_p$ and $V = H_{\mathrm{dR}}^d(X_{y_0}, \mathbf{Q}_p)^{\mathrm{prim}}$ for some fiber of the family of Theorem 10.1;
- The filtration F comes from the Hodge filtration on X_y, where y lies in the residue disk of y_0.
- $\mathfrak{f}$ is another filtration which comes from a potential failure of the global Galois representation at y to be semisimple; the passage to the graded $\mathrm{gr}_{\mathfrak{f}}$ affects semisimplification of the Galois representation.
- Condition (b) comes eventually from global purity.
- We have $z \leqslant h^0$ by assumption (this came from Lemma 10.4) and we take $e = \dim(Y)$.

The statement of Proposition 10.6 is complicated – it is an analogue, in our current setting, of Lemma 6.3. We'll offer some vague motivation here. Proposition 10.6 asks: for which filtrations F on V do $A_1, \ldots, A_{\mathsf{m}}$ form a composition series for V as filtered ϕ-modules? Of course, filtered ϕ-modules are in general far from simple, and the choice of F often amounts to a choice of extension class. Based on this, one might expect that the space of bad filtrations is large, perhaps even Zariski dense in the flag variety. This does not happen here because of the condition that the weight of the filtration on each A_i equal $d/2$. This equal-weight condition generalizes Equation (6.12). Requiring the subobjects $\mathfrak{f}^j V$ giving the composition series to have large intersection with pieces of the filtration turns out to impose strong conditions on the filtration F. This is the content of Proposition 10.6.

11. COMBINATORICS RELATED TO REDUCTIVE GROUPS

It remains to prove Proposition 10.6 from the prior section. This is "just" a problem in linear algebra but it is a notational mess. We analyze it using some simple ideas about root systems. Although we work in the generality of an arbitrary reductive group, to help the exposition we will often explicate the discussion in the case of GL_n. One other reason we chose to work in this generality is that analysis of this type is likely necessary when carrying out a similar analysis for more general monodromy groups.

Since Proposition 10.6 is geometric, concerning the dimensions of certain algebraic sets, we can and will suppose that the base field K is algebraically closed. We will therefore permit ourselves to identify algebraic groups with their K-points; they will be correspondingly denoted by usual letters P, G etc., rather than boldface letters as we have done previously.

There is a correspondence between filtrations and parabolic subgroups. We have a question about the interaction of two filtrations $\mathfrak{f}$ and F; we're going to convert it to a question about the interaction of two parabolic subgroups P and Q.

One important warning: As defined $\mathfrak{f}$ is an increasing filtration, whereas F is decreasing. However, in actual fact, the indexing of $\mathfrak{f}$ is irrelevant. All that will matter throughout is

the stabilizer of $\mathfrak{f}$; we could re-index it to be a decreasing filtration and nothing at all would change. On the other hand, the indexing of F *does* matter, and thus we will need to keep track of extra data beyond its stabilizer.

Tracing back the origins of Proposition 10.6, $\mathfrak{f}$ comes from the semisimplification filtration on a global Galois representation, and F from the Hodge filtration. The following informal dictionary may be helpful, at least in interpreting the material from §11.3 onward:

- The parabolic denoted P should be thought of as the stabilizer of the semisimplification filtration $\mathfrak{f}$.
- The Levi quotient M of P corresponds to the associated graded for $\mathfrak{f}$; globally, the semisimplification of the Galois representation takes values in M.
- The parabolic Q should be thought of as the stabilizer of the Hodge filtration F.

The argument can be informally summarized like this:

- First of all, we bound the number of possibilities for $\mathfrak{f}$, using the fact that it is ϕ-stable. This uses the fact that the centralizer of ϕ^{ss} is not too large and happens in (11.18). After this point, it is enough to work with a given $\mathfrak{f}$ and P.
- Having fixed $\mathfrak{f}$ and P, we break up the space of possible F into P-orbits. The set of F satisfying the weight condition (b) of Proposition 10.6, is a union of P orbits. We need to show that no P orbit of small codimension occurs in this set.
- To illustrate the idea, we will just explain why the *open* P orbit doesn't occur. Suppose F satisfies the weight condition (b) of Proposition 10.6. We show then that PQ/Q is not open in G/Q.

  We find a maximal torus $T \subset P \cap Q$ and a character $\nu : \mathbb{G}_m \to T$ which defines the filtration F. In particular, Q consists of non-negative root spaces for ν. The weight condition will imply that

$$(11.1) \qquad \sum_{\gamma \in \Sigma - \Sigma_P} \langle \nu, \gamma \rangle = 0,$$

  the sum being taken over roots Σ for T that correspond to root spaces outside P.

  By using the assumed numerology of Hodge numbers, not too many of these $\langle \nu, \gamma \rangle$ can be zero. In particular, (11.1) implies that $\langle \nu, \gamma \rangle < 0$ for at least one $\gamma \in \Sigma - \Sigma_P$.

  That means there is at least one such root $\gamma \in \Sigma - \Sigma_P$ that doesn't belong to the Lie algebra of Q; equivalently,

$$\mathrm{Lie}(Q) + \mathrm{Lie}(P) \neq \mathrm{Lie}(G),$$

  which implies the desired conclusion.

11.1. **Filtrations on reductive groups.** Let G be a reductive group over an algebraically closed field K.

A (rational) cocharacter $\lambda : \mathbf{G}_m \dashrightarrow G$ is simply a co-character that is allowed to be defined on a finite cover of $\mathbf{G}_m$. It determines a parabolic P_λ, whose Lie algebra is the sum of non-negative weight spaces for λ; the centralizer of λ is therefore a Levi factor for this parabolic. A "filtration" for G will be, by definition, an equivalence class of such rational cocharacters λ, where $\lambda \sim \lambda'$ if λ' is conjugate to λ under P_λ (or equivalently under the unipotent radical of P_λ).

Example 1. *Filtrations.*

- *A filtration on $G = \mathrm{GL}(V)$ is the same as a (decreasing) filtration $F^\bullet V$ on V, where the indices are indexed by rational numbers. Specifically we set*

(11.2) $\qquad F^p V = \text{ sum of all weight spaces for } \lambda \text{ on } V \text{ with weights} \geqslant p$

  The associated parabolic P_λ is precisely the stabilizer of this filtration.

  Note that $F^\bullet V$ determines λ up to the equivalence described above: any two rational characters $\mathbf{G}_m \to P$ with the same projection to a Levi quotient are actually P-conjugate by Lemma 2.5.
- *If V is equipped with a bilinear form $\langle -, - \rangle$, then a filtration on $\mathrm{GAut}(V, \langle -, - \rangle)$ is the same as a self-dual filtration on V, again via the formula (11.2); more precisely, if the filtration F corresponds to a character $\chi : \mathbf{G}_m \dashrightarrow \mathrm{GAut}(V)$ for which $\chi(x)$ scales the form by x^r, then*

(11.3)
  $F^p V$ and $F^{r-p+\epsilon} V$ *are orthogonal complements of one another (for sufficiently small ϵ).*

A map $G_1 \to G_2$ of reductive groups induces, obviously, a map from filtrations for G_1 to filtrations for G_2. Thus a filtration on G determines a filtration on the underlying space of any G-representation. If $G = \mathrm{GL}_n$, this corresponds to the usual way in which a filtration on V induces (e.g.) a filtration on $V \otimes V, V^*$, etc.

Indeed, for a general group G, to give a filtration of G is the same as giving a filtration functorially on every representation of G: this is part of the theory of filtered fibered functors, see [36, Section IV.2.1].

For any reductive group S write

$$\mathfrak{a}_\mathbf{S} := X_*(Z_S) \otimes \mathbf{Q}$$

where Z_S is the center. (As usual, we write X_* for cocharacters and X^* for characters.) This space is canonically in duality with $X^*(S) \otimes \mathbf{Q}$. If F is a filtration on S the projection of the associated cocharacter to the torus quotient of S defines a class in $\mathfrak{a}_S$. We call this the weight of F:

$$\mathrm{wt}(F) \in \mathfrak{a}_S.$$

Example 2. *Weights of filtrations.*

- *For $\mathrm{GL}(V)$, $\mathfrak{a}_{\mathrm{GL}(V)}$ is a one-dimensional $\mathbf{Q}$-vector space. We identify it with $\mathbf{Q}$ by identifying the character $t \in \mathbf{G}_m \mapsto t\mathrm{Id}_V$ with $1 \in \mathbf{Q}$. With this identification, the weight of the filtration on $\mathrm{GL}(V)$, corresponding to $F^p V$ as in (a) above, is $\frac{\sum_p p \dim(F^p / F^{p+1})}{\dim V}$; thus this definition coincides with our previous definition (2.2).*
- *For $\mathrm{GAut}(V)$, we can make the same identification of $\mathfrak{a}$ with $\mathbf{Q}$ as for GL. With this identification, the weight of the filtration described before (11.3) is necessarily equal to $r/2$, one-half of the integer by which the associated character scales the form.*

We can alternately describe filtrations using parabolics: For $\lambda : \mathbf{G}_m \to G$ the projection of λ to the Levi quotient M_λ of the parabolic P_λ is central in M_λ; thus we get a class $\bar{\lambda}$ in $\mathfrak{a}_{M_\lambda}$. The pair $(P_\lambda, \bar{\lambda} \in \mathfrak{a}_{M_\lambda})$ depends only on the filtration associated to λ, and moreover completely determines that filtration, because of Lemma 2.5. In fact, any pair $(P, e \in \mathfrak{a}_M)$ of a parabolic and a "strictly positive" element of $\mathfrak{a}_M$, i.e. positive on all roots in the unipotent radical of P, arises from a filtration.

11.2. **Levi subgroups.** Now suppose that N is a Levi subgroup of G. The center of N then contains the center of G. In this way we obtain a map

$$(11.4) \qquad\qquad \mathfrak{a}_G \longrightarrow \mathfrak{a}_N$$

which is naturally split: A character of G, i.e. a homomorphism $G \to \mathbf{G}_m$, can be pulled back to a character of N. The resulting map

$$\underbrace{X^*(G) \otimes \mathbf{Q}}_{\simeq \mathfrak{a}_G^*} \longrightarrow \underbrace{X^*(N) \otimes \mathbf{Q}}_{\simeq \mathfrak{a}_N^*}$$

gives rise to a splitting of (11.4).

Example 3. If $\dim V_i = n_i$ then $\mathrm{GL}(V_1) \times \mathrm{GL}(V_2)$ is a Levi subgroup of $\mathrm{GL}(V_1 \oplus V_2)$. We identify $\mathfrak{a}_N = \mathbf{Q}^2$ as in the previous example; then $\mathfrak{a}_G$ is embedded as the subspace $(1, 1)$ and the complementary subspace is spanned by $(-\dim(V_2), \dim(V_1))$.

11.3. **The induced filtration on a Levi subgroup.** If V is a vector space equipped with filtrations $F^\bullet$ and $\mathfrak{f}^\bullet$, then $F^\bullet$ induces a filtration on $\mathrm{gr}_*^{\mathfrak{f}} V$. We need to analyze this induced filtration carefully when $F^\bullet$ the Hodge filtration and $\mathfrak{f}^\bullet$ the semisimplification filtration.

It is convenient to again express this abstractly: For any reductive group G and any parabolic P, a filtration F on G induces a filtration F_M on the Levi quotient M of P. (With reference to the example above, P corresponds to the filtration $\mathfrak{f}^\bullet$, and M to the associated graded). To explain this we require the following Lemma:

Lemma 11.1. *Let $\chi : \mathbf{G}_m \to G$ be a character defining the parabolic subgroup Q. Let Q act transitively on an algebraic variety Y. Then all fixed points of χ on Y are conjugate under the centralizer N of χ.*

Proof. We have a Levi decomposition $Q = NV$, with V the unipotent radical. Suppose that $y_0 \in Y$ is χ-fixed. It is enough to verify that y_0 is the unique point in $V y_0$ that is χ-fixed. Let V_0 be the stabilizer of y_0 inside V. For $x \in \mathbf{G}_m$,

$$\chi(x) \cdot (v y_0) = (\chi(x) v \chi(x)^{-1}) y_0,$$

and thus the χ-fixed points on $V y_0$ correspond to the fixed points for $\chi(x)$-conjugation on V/V_0.

But all the weights of this $\mathbf{G}_m$-action on V are positive, i.e. the limit of $\chi(x) v \chi(x)^{-1}$ as $x \to 0$ is equal to the identity. Therefore the only fixed point on V/V_0 for conjugation by $\chi(\mathbf{G}_m)$ is the identity coset. $\qquad\square$

Before we formulate the induced filtration in terms of parabolics, we recall some linear algebra associated to two parabolics. Suppose that P, Q are parabolic subgroups of G, where Q is the stabilizer of a filtration F. It is known that P and Q contain a common maximal torus T and that $P \cap Q$ is connected; this, together with everything else we will use is contained in [14, Chapter 2]. We will briefly summarize what we need.

Fixing T as above, we get Levi decompositions of P and Q such that both Levi factors contain T:

$$(11.5) \qquad\qquad P = MU \text{ and } Q = NV$$

We have a factorization

$$(11.6) \qquad\qquad P \cap Q = (M \cap Q) \cdot (U \cap Q).$$

In particular, this implies that the projection of $P \cap Q$ to M along $P \twoheadrightarrow M$ is just $M \cap Q$.

To verify this factorization, we note that $(M \cap Q)$ normalizes $(U \cap Q)$, and also that it is easy to verify the corresponding splitting at the level of Lie algebras; since $P \cap Q$ is connected, this factorization also follows.

Lemma 11.2. *(Induced filtration on the Levi factor of a parabolic.) Let F be a filtration for the group G. There exists a representative $\chi_P : \mathbf{G}_m \to G$ for the filtration F with the property that χ is valued in P. Moreover, any two such representatives are conjugate under $P \cap Q$.*

For each such representative χ_P, the projection of χ_P to the Levi quotient M of P defines a filtration on M which is independent of the choice of χ_P.

Proof. Let $\chi : \mathbf{G}_m \to G$ represent the filtration, and let Q be the associated parabolic. The intersection $P \cap Q$ contains a maximal torus T of G and we may certainly conjugate χ so it is valued in T, so in P; this proves the existence statement.

For uniqueness fix χ_P, which we may now suppose to be valued in $P \cap Q$. Now the image of $\mathrm{Ad}(q)^{-1}\chi_P$ is in P if and only if

$$\chi_P(\mathbf{G}_m) \subset \mathrm{Ad}(q)P,$$

i.e. qP lies in the set of fixed points of $\chi_P(\mathbf{G}_m)$ on QP/P. These fixed points are all conjugate under N, the centralizer of χ_P, as we have seen above; thus $qP \in NP$, so that $q \in N(Q \cap P)$. Thus the characters $\mathrm{Ad}(q)^{-1}\chi_P$ lie in a single $(P \cap Q)$-orbit.

It remains to prove the final statement. Choose $\chi_P, \chi_P' : \mathbf{G}_m \to P$, as above, both P-valued representatives for the filtration F. We have $\chi_P' = \mathrm{Ad}(g)\chi_P$ for some $g \in P \cap Q$, so $\overline{\chi_P'} = \mathrm{Ad}(\bar{g})\overline{\chi_P}$, where bars denote projection to the Levi quotient of P. To see that these two characters define the same filtration we need to verify that

$$\bar{g} \in Q_{\overline{\chi_P}}.$$

This follows from the remark after (11.6): extend the image of χ_P to a maximal torus inside $P \cap Q$; then, with the corresponding choice M of Levi subgroup for P, we have $\overline{\chi_P} = \chi_P$ and $Q_{\overline{\chi_P}} = Q \cap M$.

$\square$

Example 4. *Induced filtration on the associated graded.*
- *Consider the case of $G = \mathrm{GL}(V)$. Suppose given a decreasing filtration $F^\bullet V$ (with associated parabolic Q) and another parabolic P; we fix an increasing filtration $\mathfrak{f}^\bullet V$ with stabilizer P. We show that the construction above gives precisely the filtration induced by F on the associated graded to $\mathfrak{f}$.*

  As above, we can represent the character for F by a character χ preserving the filtration $\mathfrak{f}$. Then, writing $\bar{F}$ for the induced filtration:

  $$\bar{F}^j(\mathfrak{f}^k/\mathfrak{f}^{k-1})$$

  is the sum of all eigenspaces with weights $\geqslant j$; this is the image of the corresponding space in $\mathfrak{f}^k$, that is to say,

  $$\bar{F}^j(\mathfrak{f}^k/\mathfrak{f}^{k-1}) = \text{image of } F^j \cap \mathfrak{f}^k \text{ in } \mathfrak{f}^k/\mathfrak{f}^{k+1}.$$

- *We now modify the example above by taking G to be $\mathrm{GAut}(V, \langle -, - \rangle)$ for some symmetric or skew-symmetric nondegenerate bilinear pairing $\langle -, - \rangle$. Now suppose that P is a parabolic subgroup of G, stabilizing the self-dual increasing filtration*

  $$0 = \mathfrak{f}^0 \subset \mathfrak{f}^1 \subset \cdots \subset \mathfrak{f}^m \subset \mathfrak{f}^{m+1} \subset \cdots \subset \mathfrak{f}^{2m+1} = V.$$

Just as before, F induces a filtration $\bar{F}$ on each graded piece $\mathfrak{g}^j = \mathfrak{f}^j/\mathfrak{f}^{j-1}$. The associated Levi subgroup is isomorphic to

$$\mathrm{GL}(\mathfrak{g}^1) \times \cdots \times \mathrm{GL}(\mathfrak{g}^m) \times \mathrm{GAut}(\mathfrak{g}^{m+1}),$$

where we regard the last factor as $\mathbf{G}_m$ even if $\mathfrak{f}^m = \mathfrak{f}^{m+1}$, and the corresponding filtration on each factor is the one induced by $\bar{F}$.

11.4. **Balanced filtrations and parabolic subgroups.**

As above, let G be a reductive group over a field K, and let:

- F be a filtration of G associated with the parabolic subgroup Q,
- P a parabolic subgroup of G, with Levi quotient M.

We say that F is *balanced* with respect to P if $\mathrm{wt}(F) \in \mathfrak{a}_G$ is carried, under the embedding $\mathfrak{a}_G \to \mathfrak{a}_M$, to the weight $\mathrm{wt}(F_M)$ of the filtration induced on the Levi quotient. Here $\mathfrak{a}_G \hookrightarrow \mathfrak{a}_M$ is as in §11.2.

Example 5. *Balanced filtrations.*

- *If $G = \mathrm{GL}(V)$, and P is associated to the increasing filtration $\mathfrak{f}^q V$, then "balanced" says that, for every q, the filtration that F induces on $\mathfrak{f}^q/\mathfrak{f}^{q+1}$ has the same weight as the filtration F on V.*
- *The same assertion holds for $\mathrm{GAut}(V)$, where now F and $\mathfrak{f}$ are self-dual filtrations.*

Note that if we choose a cocharacter $\chi_P : \mathbf{G}_m \dashrightarrow P$ representing F, the condition of being "balanced" implies that, for any character ψ of P trivial on the center of G,

$$(11.7) \qquad\qquad\qquad \langle \psi, \chi_P \rangle = 0.$$

Now define

$$X(F) = \{G\text{-conjugates of } F \text{ that are balanced with respect to } P\},$$

so that $X(F)$ is a P-stable subvariety of G/Q and is equipped with a map

$$(11.8) \qquad\qquad\qquad X(F) \longrightarrow \{\text{filtrations of } M\}$$

via the rule $F \mapsto F_M$. We may regard this, in an evident way, as a "constructible" map between algebraic varieties (i.e. its graph is a constructible set) and thus we can reasonably speak of dimension of fibers.

We will analyze (11.8) by breaking $X(F)$ into P-orbits. Consider for a moment $F \mapsto F_M$ as a map

$$(11.9) \qquad \text{filtrations } P\text{-conjugate to } F \longrightarrow \text{filtrations on Levi quotient of } P$$

where both sides are P-varieties. The left hand side is identified with $P/(P \cap Q)$, and – if we choose a maximal torus of $P \cap Q$ containing the image of a character defining F, and take the corresponding Levi decomposition $P = MU$ – the image is identified with $M/(M \cap Q)$. From this and (11.6) we find that each fiber of (11.9) has dimension

$$(11.10) \qquad\qquad\qquad \dim(U) - \dim(Q \cap U)$$

where U is the unipotent radical of P, and Q the stabilizer of F.

11.5. **Double cosets of parabolic subgroups.** Fix, as before, F a filtration of G associated with the parabolic subgroup Q, and P a parabolic subgroup of G, with Levi quotient M. Continue with notation $X(F)$ as above. We will be concerned with estimating the size of the fibers of (11.8).

Fix a Borel B contained in P and a maximal torus $T \subset B$. Since the variety $X(F)$ depends only on the G-orbit of F, we may harmlessly replace F by a G-conjugate; in particular we may suppose that F is defined by a co-character $\mu : \mathbf{G}_m \to T$ that is positive with respect to B, i.e. $B \subset Q$.

Let $\Sigma \supset \Sigma^+$ be the set of roots of T on G and on B, respectively; one therefore gets notions of simple and positive roots. Let Σ_P, Σ_Q be the set of roots of T on P and Q. Therefore, $\Sigma^+ \subset \Sigma_P$ and $\Sigma^+ \subset \Sigma_Q$. Let Δ_P be the subset of simple roots α for which $-\alpha \in \Sigma_P$, and similarly define Δ_Q; thus P and Q correspond to the subsets Δ_P, Δ_Q of the set of simple roots. Note that, since μ defines the parabolic subgroup Q, Σ_Q is the set of roots having nonnegative pairing with μ, and in particular μ is orthogonal to all roots for Δ_Q:
$$\langle \mu, \beta \rangle = 0, \ \ \beta \in \Delta_Q.$$

Recall the "adjoint" Hodge numbers associated to Lie $\mathrm{GAut}\,(\mathsf{V}_0 \otimes \mathbf{C}, \langle -, - \rangle)$, introduced in Section 10. The following proposition uses an abstraction of that notion:

Proposition 11.3. *Let the "Hodge numbers" be the multi-set of integers of the form $\langle \mu, \gamma \rangle$ with $\gamma \in \Sigma$, adding multiplicity $\dim(T)$ to the multiplicity of zero. For $i \neq 0$ let a_i be the number of roots $\gamma \in \Sigma$ with $\langle \mu, \gamma \rangle = i$, so that a_i is the multiplicity of i as a Hodge number and $\sum_{i>0} a_i = \dim(G/Q)$; we take a_0 the dimension of the Levi factor of Q.*

Suppose $e \leq \dim(G/Q)$ is a positive integer such that

(11.11)   *sum of all positive Hodge numbers > sum of top e Hodge numbers*

$$+ \text{ sum of top } \left(\tfrac{a_0}{2} + e \right) \text{ Hodge numbers,}$$

Then the codimension inside G/Q of any fiber of the mapping (11.8)

$$X(F) \to \text{filtrations of } M$$

is greater than e.

Proof. We are going to analyze this P-orbit by P-orbit. Note that we have $G = P W_{PQ} Q$, where W_{PQ} is the subset of the Weyl group W defined via

(11.12) $$W_{PQ} = \{w \in W : \ w^{-1}\Delta_P > 0, w\Delta_Q > 0\}.$$

Indeed, it is enough to see (by the Bruhat decomposition) that $W = W_P \cdot W_{PQ} \cdot W_Q$, where W_P and W_Q are generated by simple reflections corresponding to Δ_P and Δ_Q. Writing as usual $\ell(w) = \#\{\alpha > 0 : w\alpha < 0\}$ for the length of a Weyl element, any minimal-length representative in a fixed double coset $W_P \cdot w \cdot W_Q$ belongs to W_{PQ}: for $\alpha \in \Delta_P$, the element $s_\alpha w$ has shorter length than w if $w^{-1}\alpha < 0$. Similarly, for $\beta \in \Delta_Q$, we know that ws_β has shorter length than w if $w\beta < 0$.

For each $w \in W_{PQ}$ we have either $PwQ/Q \subset X(F)$, or $PwQ/Q \cap X(F) = \emptyset$. Call w *bad* in the former case. For each bad $w \in W_{PQ}$ let $X(F)_w$ be the corresponding locally closed subvariety of $X(F)$, i.e.

$$X(F)_w = X(F) \cap ((PwQ)/Q).$$

Thus $X(F) = \coprod X(F)_w$, the union taken over bad w. Assume, by way of contradiction, that there exists some bad w such that a fiber of

(11.13) $$X(F)_w \to \text{filtrations of } M$$

has codimension inside G/Q that is $\leqslant e$.

That w is bad means that the filtration defined by the co-character $w\mu$ is balanced with reference to P. This means in particular that

$$\tag{11.14} \sum_{\gamma \in \Sigma - \Sigma_P} \langle w\mu, \gamma \rangle = 0.$$

In fact $\sum_{\gamma \in \Sigma - \Sigma_P} \gamma$ computes the modular character of the parabolic subgroup P: it is the negative of the character by which P acts on the determinant of its unipotent radical, and then use (11.7).

For this (bad) w, write

$$X = \{\beta \in \Sigma - \Sigma_P : w^{-1}\beta > 0\} = \{\beta \in \Sigma - \Sigma_P : -w^{-1}\beta \in \Sigma - \Sigma_Q\}$$

(using Lemma 11.4, see below) and let X' be the complement of X inside $\Sigma - \Sigma_P$.

Each fiber of (11.13) has, by (11.10), dimension

$$\dim(U) - \dim(\mathrm{Ad}(w)Q \cap U) = \#\{\alpha \in \Sigma - \Sigma_P : -w^{-1}\alpha \in \Sigma - \Sigma_Q\} = \#X.$$

(see Lemma 11.4). This is equal to the length $\ell(w)$, although we won't make explicit use of it. Therefore our assumption means $\#X \geqslant \dim(G/Q) - e$. Then, since $\#X' = \dim(G/P) - \#X$, we have

$$\tag{11.15} \#X' \leqslant \dim(G/P) - \dim(G/Q) + e$$

$$= \dim(Q) - \dim(P) + e \leqslant \dim(Q/B) + e \leqslant \frac{a_0}{2} + e.$$

Also, by (11.14),

$$\tag{11.16} \sum_{\beta \in X} \langle \mu, w^{-1}\beta \rangle = \sum_{X'} -\langle \mu, w^{-1}\beta \rangle.$$

All entries on the left hand side are *strictly* positive because $w^{-1}\beta$ is the negative of an element of $\Sigma - \Sigma_Q$. All entries on the right-hand side are non-negative (because $B \subset Q$ the cocharacter μ is non-negative on positive roots.) Now X has size $\geqslant \dim(G/Q) - e$, so the image $-w^{-1}(X)$ omits at most e roots inside $\Sigma - \Sigma_Q$. Therefore, the left-hand side of (11.16) is *at least*

sum of all positive Hodge numbers $-$ sum of the topmost e Hodge numbers.

On the other hand, the right-hand side of (11.16) is *at most* the sum of the top $(a_0/2 + e)$ Hodge numbers. (Here we have used that, since $e \leq \dim(G/Q)$, the top e Hodge numbers are all positive.) So we get a contradiction to (11.16) under the stated hypothesis. $\square$

We used the following Lemma:

Lemma 11.4. *Let Σ be the set of all roots, and take $w \in W_{PQ}$ (see (11.12)).*

   (i) For $\beta \in \Sigma - \Sigma_Q$, we have $w\beta > 0 \iff -w\beta \in \Sigma - \Sigma_P$.
   (ii) For $\alpha \in \Sigma - \Sigma_P$, we have $w^{-1}\alpha > 0 \iff -w^{-1}\alpha \in \Sigma - \Sigma_Q$.
   (iii) The map $x \mapsto -w(x)$ induces a bijection of these sets:

$$\tag{11.17} \{\beta \in \Sigma - \Sigma_Q : w\beta > 0\} \longrightarrow \{\alpha \in \Sigma - \Sigma_P : w^{-1}\alpha > 0\}$$

   The size of this set is precisely the length $\ell(w)$.

Proof. Take $\beta \in \Sigma - \Sigma_Q$ with $w\beta > 0$. If $-w(\beta)$ were in Σ_P, then β is a positive linear combination of roots in $w^{-1}\Delta_P$, contradicting the negativity of β.

This shows the $\implies$ direction of (i) and the $\implies$ direction of (ii) is similar. The reverse directions for (i) and (ii) are clear. For example, if $-w\beta$ is in $\Sigma - \Sigma_P$, then $w\beta > 0$ because

all roots in $\Sigma - \Sigma_P$ are negative. Now it is clear that the maps w and w^{-1} give inverse bijections in (11.17). $\qquad\square$

11.6. **Conclusion of the argument.** We now return to the situation of Proposition 10.6. Let $G = \mathrm{GAut}(V, \langle -, - \rangle)$.

We translate the problem into reductive group language. Let F_0 be a fixed self-dual filtration on V; we will consider those filtrations F that are conjugate to F_0 under G. Let Q be the stabilizer of F_0 in G, with Levi quotient N. Reformulating Proposition 10.6 (replacing $\mathfrak{f}$ from the Proposition with the parabolic subgroup which is its stabilizer): we must estimate the codimension of $g \in G/Q$ such that, writing $F = gF_0$, there exists another parabolic subgroup $P \leqslant G$ such that:

- (a)' (from property (a) of Proposition 10.6): $\phi \in P$;
- (b)' (from property (b) of Proposition 10.6): F is balanced with respect to P, cf. the example of §11.4.
- (c)' (from property (c) of Proposition 10.6): The G-conjugacy class of (P, ϕ_M, F_M) is fixed, where ϕ_M is the projection of ϕ to the Levi quotient M of the parabolic P.[13]

First of all, we reduce to the case when ϕ is semisimple. Firstly, $\phi \in P \implies \phi^{ss} \in P$ and, supposing that $\phi \in P$, then also $(\phi^{ss})_M = (\phi_M)^{ss}$ (the subscript M denotes projection to M). Now if (P, ϕ_M, F_M) and $(P', \phi_{M'}, F_{M'})$ are conjugate, so that there is $g \in G$ with $\mathrm{Ad}(g)P = P'$ and $\mathrm{Ad}(g) : M \to M'$ carries ϕ_M to $\phi_{M'}$, then $\mathrm{Ad}(g) : M \to M'$ also carries $(\phi_M)^{ss} = (\phi^{ss})_M$ to $(\phi_{M'})^{ss} = (\phi^{ss})_{M'}$. In other words, if we replace ϕ by ϕ^{ss} then the codimension of the set described above will only decrease. We do this, and can therefore assume that ϕ is semisimple.

We will first show that

$$(11.18) \qquad \text{(dimension of possible pairs } (P, F_M)) \leq z = \dim Z(\phi),$$

the dimension of the centralizer of ϕ in G. (Note that, because of our reduction above, z corresponds to the dimension of the centralizer of ϕ^{ss}, for the original choice of ϕ.)

The set of P containing a given semisimple ϕ is a finite union of orbits of $Z(\phi)$, as we see by infinitesimal computations. It suffices, therefore, to examine a single $Z(\phi)$-orbit on the space of P. Fix P_1 in this orbit. The dimension of $Z(\phi) \cdot P_1$ equals

$$(11.19) \qquad \dim Z(\phi) - \dim Z_{P_1}(\phi)$$

Next, if we fix $P \in Z(\phi) \cdot P_1$, the collection of filtrations $\mathcal{F}$ on its Levi factor M for which $(P, \phi_M, \mathcal{F})$ belongs to a fixed G-isomorphism class corresponds to a finite collection of orbits of $Z_M(\phi_M)$ on the space of filtrations on M. Now ϕ is P-conjugate to ϕ_M by (2.1) so that $\dim Z_M(\phi_M) \leqslant \dim Z_P(\phi)$. It follows that the dimension of the space of possible filtrations on M, for P fixed, is at most $\dim Z_P(\phi) = \dim Z_{P_1}(\phi)$. Adding this to (11.19) we deduce (11.18).

We may now conclude the proof. Suppose e is as in (10.24), so that both conditions are satisfied:

$$\text{number of positive Hodge numbers} \geqslant z + e$$

$$\text{sum of all positive Hodge numbers} > T(z + e) + T\left(\frac{h^0}{2} + z + e\right).$$

Recall that $X(F) \subset G/Q$ is the set of filtrations that are G-conjugate to F and are balanced with respect to P; We may apply Proposition 11.3, but taking the e of that Proposition to

---

[13]We say here that (P, ϕ_M, F_M) is conjugate to $(P', \phi_{M'}, F_{M'})$ when there is $g \in G$ such that $\mathrm{Ad}(g)P = P'$, and the induced isomorphism of Levi quotients carries (ϕ_M, F_M) to $(\phi_{M'}, F_{M'})$.

be $z + e$ in the discussion above. (Note that the first displayed equation above guarantees, in the notation of Proposition 11.3, that $z + e \leqslant \dim(G/Q)$, as needed to apply it.) Thus, if we fix P, the codimension inside G/Q of any fiber of

$$X(F) \to \text{filtrations on } M$$

is at least $z + e$.

However, we saw above that the dimension of possibilities for $(P, \text{filtration on } M)$ is at most z. Therefore, the total codimension of the set of $g \in G/Q$ satisfying (a)', (b)', (c)' is at least e, concluding the proof. $\qquad\square$

12. BOUNDING FROBENIUS VIA POINT COUNTS

We remark on an alternative approach to bounding the size of the Frobenius centralizer, i.e. the step that was achieved in the previous argument by Lemma 10.4. It is likely that in some ranges this gives rise to better numerical bounds:

Lemma 12.1. *Let Y be a smooth hypersurface of degree d and dimension $n \geqslant 2$, defined over the finite field k with q elements; let $b = \dim H^n_{\text{prim}}(Y_{\bar{k}}, \mathbf{Q}_\ell)$. Then the centralizer Z of the semisimplified Frobenius, acting on $H^n_{\text{prim}}(Y_{\bar{k}}, \mathbf{Q}_\ell)$, has dimension at most $3b^2/N$, where N is the largest integer for which $q^{(n/2+1)N} < b/3$.*

Proof. To avoid confusion between $i = \sqrt{-1}$ and as an index we write $e(\alpha) := \exp(2\pi i \alpha)$.

Let the Frobenius eigenvalues on $H^n_{\text{prim}}(Y_{\bar{k}}, \mathbf{Q}_\ell)$ be given by

$$\lambda_1 = q^{n/2}e(\theta_1), \ldots, \lambda_b = q^{n/2}e(\theta_b),$$

and let μ be the measure on S^1 given by $\sum_{i=1}^{b} \delta_{\theta_i}$. If the multiplicities of the θ_i are $m_1, \ldots, m_r$, with $\sum m_i = b$, then $\dim Z = \sum m_i^2$.

If g is any non-negative real-valued function on S^1 we have $\int g(t-\theta)d\mu(\theta) = \sum_s g(t-\theta_s)$, and so

$$(12.1) \qquad \int_t dt \left| \int g(t-\theta)d\mu(\theta) \right|^2 \geqslant \dim Z \cdot \|g\|_{L^2}^2$$

which bounds from above the dimension of the centralizer; this estimate is most effective if the support of g is concentrated near 0. Here, and in what follows, the measure is the Haar probability measure on S^1.

If k' is the field extension of k of degree j, the number of points of $Y(k')$ is given by

$$|Y(k')| = \sum_{\ell=0}^{n} q^{\ell j} + (-1)^n q^{nj/2} \sum_{s=1}^{b} e(j\theta_s).$$

Since this lies between 0 and the size of $\mathbf{P}^{n+1}(k')$, i.e. between 0 and $\sum_{\ell=0}^{n+1} q^{\ell j}$, we see that

$$(12.2) \qquad \left| \sum_s e(j\theta_s) \right| \leqslant q^{(n/2+1)j}.$$

Let

$$g_N(t) = \left(\sum_{r=-N}^{N} e(rt) \right)^2 = \sum_{r=-2N}^{2N} (2N+1-|r|)e(rt),$$

a function on S^1. Note that $\|g_N\|_{L^2}^2 = (2N+1)^2 + 2\sum_{i=1}^{2N} i^2$. We have

$$\int g_N(t-\theta)d\mu(\theta) = \sum_{r=-N}^{N}(2N+1-|r|)\sum_{s=1}^{b}e(r(t-\theta_s)).$$

Using (12.2), we see that this is bounded in absolute value by

$$(2N+1)\left[b + 2\sum_{r=1}^{N} q^{(n/2+1)|r|}\right] \leqslant (2N+1)\left(b + 3q^{(n/2+1)N}\right)$$

since $q^{n/2+1} \geqslant 4$. Therefore, by (12.1),

$$\dim(Z) \leqslant b^2(1 + 3b^{-1}q^{(n/2+1)N})^2 \cdot \underbrace{\left(\frac{(2N+1)^2}{(2N+1)^2 + 2\sum_{i=1}^{2N} i^2}\right)}_{\leq \frac{3}{4N}}.$$

Choose N the largest integer with $q^{(n/2+1)N} < b/3$; we get

$$\dim(Z) \leqslant 3b^2/N.$$

<div align="right">□</div>

REFERENCES

[1] *Revêtements étales et groupe fondamental (SGA 1)*, volume 3 of *Documents Mathématiques (Paris) [Mathematical Documents (Paris)]*. Société Mathématique de France, Paris, 2003. Séminaire de géométrie algébrique du Bois Marie 1960–61. [Algebraic Geometry Seminar of Bois Marie 1960-61], Directed by A. Grothendieck, With two papers by M. Raynaud, Updated and annotated reprint of the 1971 original [Lecture Notes in Math., 224, Springer, Berlin; MR0354651 (50 #7129)].

[2] Benjamin Bakker and Jacob Tsimerman. The Ax–Schanuel conjecture for variations of Hodge structures.

[3] Michael Bate, Benjamin Martin, and Gerhard Röhrle. A geometric approach to complete reducibility. *Invent. Math.*, 161(1):177–218, 2005.

[4] Arnaud Beauville. Le groupe de monodromie des familles universelles d'hypersurfaces et d'intersections complètes. In *Complex analysis and algebraic geometry (Göttingen, 1985)*, volume 1194 of *Lecture Notes in Math.*, pages 8–18. Springer, Berlin, 1986.

[5] Pierre Berthelot. *Cohomologie cristalline des schémas de caractéristique $p > 0$*. Lecture Notes in Mathematics, Vol. 407. Springer-Verlag, Berlin-New York, 1974.

[6] Pierre Berthelot and Arthur Ogus. *Notes on crystalline cohomology*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1978.

[7] Armand Borel. *Linear algebraic groups*, volume 126 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[8] Olivier Brinon and Brian Conrad. CMI summer school notes on p-adic Hodge theory.

[9] Sourav Chatterjee and Persi Diaconis. A central limit theorem for a new statistic on permutations. *Indian J. Pure Appl. Math.*, 48(4):561–573, 2017.

[10] Theo de Jong and Gerhard Pfister. *Local analytic geometry*. Advanced Lectures in Mathematics. Friedr. Vieweg & Sohn, Braunschweig, 2000. Basic theory and applications.

[11] P. Deligne. Théorème de Lefschetz et critères de dégénérescence de suites spectrales. *Inst. Hautes Études Sci. Publ. Math.*, (35):259–278, 1968.

[12] Pierre Deligne, James S. Milne, Arthur Ogus, and Kuang-yen Shih. *Hodge cycles, motives, and Shimura varieties*, volume 900 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1982.

[13] Alexandre Grothendieck; Jean Dieudonné. éléments de géométrie algébrique: Iv. étude locale des schémas et des morphismes de schémas, troisième partie. *Pub. Math. IHES.*, 28:5–255, 1966.

[14] François Digne and Jean Michel. *Representations of finite groups of Lie type*, volume 21 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1991.

[15] Gerd Faltings. Endlichkeitss atze f ur abelsche variet aten uber zahlk orpern. *Invent. math.*, 73:349–366, 1983.

[16] Gerd Faltings. Crystalline cohomology and p-adic Galois-representations. In *Algebraic analysis, geometry, and number theory (Baltimore, MD, 1988)*, pages 25–80. Johns Hopkins Univ. Press, Baltimore, MD, 1989.

[17] Benson Farb and Dan Margalit. *A Primer on Mapping Class Groups*. Princeton University Press, 2012.

[18] Jean-Marc Fontaine. Périodes p-adiques. In *Astérisque*, volume 223. Société Mathématique de France, 1994.

[19] Alexander Lubotzky Fritz Grunewald, Michael Larsen and Justin Malestein. Arithmetic quotients of the mapping class group. *Geometric and Functional Analysis*, 25:1493–1542, 2015.

[20] A. Grothendieck. Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. II. *Inst. Hautes Études Sci. Publ. Math.*, (24):231, 1965.

[21] Oliver Johnson and Christina Goldschmidt. Preservation of log-concavity on summation. *ESAIM Probab. Stat.*, 10:206–215, 2006.

[22] Nicholas M. Katz and William Messing. Some consequences of the Riemann hypothesis for varieties over finite fields. *Invent. Math.*, 23:73–77, 1974.

[23] Nicholas M. Katz and Tadao Oda. On the differentiation of de Rham cohomology classes with respect to parameters. *J. Math. Kyoto Univ.*, 8:199–213, 1968.

[24] Minhyong Kim. The motivic fundamental group of $\mathbf{P}^1 \backslash \{0, 1, \infty\}$ and the theorem of Siegel. *Invent. Math.*, 161(3):629–656, 2005.

[25] Minhyong Kim. The unipotent Albanese map and Selmer varieties for curves. *Publ. Res. Inst. Math. Sci.*, 45(1):89–133, 2009.

[26] Minhyong Kim. Tangential localization for Selmer varieties. *Duke Math. J.*, 161(2):173–199, 2012.

[27] Brian Lawrence. Two results on period maps. *Stanford University PhD Thesis*, 2017.

[28] Eduard Looijenga. Prym representations of mapping class groups. *Geom. Dedicata*, 64(1):69–83, 1997.

[29] Barry Mazur. Frobenius and the Hodge filtration (estimates). *Annals of Mathematics*, 98(1), 1973.

[30] Shinichi Mochizuki. The geometry of the compactification of the hurwitz scheme. *Publ. Res. Inst. Math. Sci*, 1995.

[31] A. N. Paršin. Algebraic curves over function fields. I. *Izv. Akad. Nauk SSSR Ser. Mat.*, 32:1191–1219, 1968.

[32] Jonathan Pila and Jacob Tsimerman. Ax?schanuel for the j-function. *Duke Math. J.*, 165(13):2587–2605, 09 2016.

[33] Richard Pink. l-adic algebraic monodromy groups, cocharacters, and the Mumford-Tate conjecture. *J. Reine Angew. Math.*, 495:187–237, 1998.

[34] Kenneth A. Ribet. Galois action on division points of Abelian varieties with real multiplications. *Amer. J. Math.*, 98(3):751–804, 1976.

[35] R. W. Richardson, Jr. Conjugacy classes in Lie algebras and algebraic groups. *Ann. of Math. (2)*, 86:1–15, 1967.

[36] Neantro Saavedra Rivano. *Catégories Tannakiennes*. Lecture Notes in Mathematics, Vol. 265. Springer-Verlag, Berlin-New York, 1972.

[37] Nick Salter and Bena Tshishiku. Arithmeticity of the monodromy of some Kodaira fibrations. *https://arxiv.org/abs/1805.06789*.

[38] Hans Rudolf Schneebeli. Group extensions whose profinite completion is exact. *Arch. Math. (Basel)*, 31(3):244–253, 1978/79.

[39] Shankar Sen. Lie algebras of Galois groups arising from Hodge-Tate modules. *Ann. of Math. (2)*, 97:160–170, 1973.

[40] Jean-Pierre Serre. Géométrie algébrique et géométrie analytique. *Ann. Inst. Fourier, Grenoble*, 6:1–42, 1955–1956.

[41] Jean-Pierre Serre. *Abelian l-adic representations and elliptic curves*. Advanced Book Classics. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, second edition, 1989. With the collaboration of Willem Kuyk and John Labute.

[42] Jean-Pierre Serre. *Galois cohomology*. Springer-Verlag, Berlin, 1997. Translated from the French by Patrick Ion and revised by the author.

[43] Jean-Pierre Serre. Complète réductibilité. *Astérisque*, (299):Exp. No. 932, viii, 195–217, 2005. Séminaire Bourbaki. Vol. 2003/2004.

[44] The Stacks Project Authors. *Stacks Project*. http://stacks.math.columbia.edu, 2018.

[45] André Weil. On a certain type of characters of the idèle-class group of an algebraic number-field. In *Proceedings of the international symposium on algebraic number theory, Tokyo & Nikko, 1955*, pages 1–7. Science Council of Japan, Tokyo, 1956.

[46] Jean-Pierre Wintenberger. Un scindage de la filtration de Hodge pour certaines variétés algébriques sur les corps locaux. *Ann. of Math. (2)*, 119(3):511–548, 1984.

[47] G. Wüstholz. The finiteness theorems of Faltings. In *Rational points (Bonn, 1983/1984)*, Aspects Math., E6, pages 154–202. Friedr. Vieweg, Braunschweig, 1984.