Stable sets of primes in number fields

Alexander Ivanov

We define a new class of sets — stable sets — of primes in number fields. For example, Chebotarev sets $P_{M/K}(\sigma)$, with $M/K$ Galois and $\sigma \in G(M/K)$, are very often stable. These sets have positive (but arbitrarily small) Dirichlet density and they generalize sets with density one in the sense that arithmetic theorems such as certain Hasse principles, the Grunwald–Wang theorem, and Riemann’s existence theorem hold for them. Geometrically, this allows us to give examples of infinite sets $S$ with arbitrarily small positive density such that $\text{Spec} \mathcal{O}_{K,S}$ is a $K(\pi, 1)$ (simultaneously for all $p$).

1. Introduction

The main goal of this article is to define a new class of sets of primes of positive Dirichlet density in number fields — stable sets. These sets have a positive but arbitrarily small density and they generalize, in many aspects, sets of density one. In particular, most of the arithmetic theorems, such as certain Hasse principles, the Grunwald–Wang theorem, Riemann’s existence theorem, $K(\pi, 1)$-property, etc., which hold for sets of density one (see [NSW 2008, Chapters IX and X]), also hold for stable sets. Our goals are on the one hand to prove these arithmetic results, and on the other hand to give many examples of stable sets.

The idea is as follows: let $\lambda > 1$. A set $S$ of primes in a number field $K$ is $\lambda$-stable for the extension $\mathcal{L}/K$ if there is a subset $S_0 \subseteq S$, a finite subextension $\mathcal{L}/L_0/K$...
and some $a > 0$ such that we have $\delta_L(S_0) \in [a, \lambda a)$ for all finite subextensions $\mathcal{L}/L/L_0$, where $\delta_L$ is the Dirichlet density. We call the field $L_0$ a $\lambda$-stabilizing field for $S$ for $\mathcal{L}/K$. A more restrictive version is the notion of persistent sets: $S$ is persistent if the function $L \mapsto \delta_L(S_0)$ gets constant in the tower $\mathcal{L}/K$ beginning from some finite subextension $L_0/K$ (see Definition 2.4). In particular, for any $\lambda > 1$, a $\lambda$-stable set is persistent.

The main result in this article is the following theorem, which links stability to vanishing of certain Shafarevich–Tate groups. Let $\mathcal{X}_1$ denote the usual Shafarevich–Tate group, consisting of global cohomology classes which vanish locally in a given set of primes. If $A$ is a module over a finite group $G$, then $H^1(G, A)$ denotes the subgroup of $H^1(G, A)$ consisting of precisely those classes which vanish after restriction to all cyclic subgroups of $G$. Moreover, if $\mathcal{L}/L$ is a Galois extension of fields, then $G_{\mathcal{L}/L}$ denotes its Galois group, and if $A$ is a $G_{\mathcal{L}/L}$-module, then $L(A)/L$ denotes the trivializing extension of $A$.

**Theorem 4.1.** Let $K$ be a number field, $T$ a set of primes of $K$ and $\mathcal{L}/K$ a Galois extension. Let $A$ be a finite $G_{\mathcal{L}/K}$-module. Assume that $T$ is $p$-stable for $\mathcal{L}/K$, where $p$ is the smallest prime divisor of $|A|$. Let $L$ be a $p$-stabilizing field for $T$ for $\mathcal{L}/K$. Then

$$\mathcal{X}_1(\mathcal{L}/L, T; A) \subseteq H^1(L(A)/L, A).$$

In particular, if $H^1(L(A)/L, A) = 0$, then $\mathcal{X}_1(\mathcal{L}/L, T; A) = 0$.

This theorem has numerous applications to the structure of the Galois group $G_{K, S} := \text{Gal}(K_S/K)$, where $K$ is a number field and $S$ is stable. To explain our results, we need some notation. If $S, R$ are two sets of primes of a number field $K$, then we denote by $K^R_S$ the maximal extension of $K$, which is unramified outside $S$ and completely split in $R$. Moreover, we denote by $G^R_{K, S}$ the Galois group of $K^R_S/K$. Let $\mathcal{L}/K$ be any Galois extension. For a prime $p$ of $K$ we denote by $\mathcal{L}_p$ the completion of $\mathcal{L}$ at a (any) extension of $p$ to $\mathcal{L}$ (the isomorphism class of the completion $\mathcal{L}_p$ does not depend on the particular choice of the extension of $p$ to $\mathcal{L}$ as $\mathcal{L}/K$ is Galois, and we suppress this choice in our notation). Furthermore, $\mathcal{G}_p$ denotes the absolute Galois group of $K_p$, and $K_p(p)$ (resp. $K_p^{nr}(p)$) denotes the maximal (resp. maximal unramified) pro-$p$ extension of $K_p$. Moreover, for a profinite group $G$, we denote the pro-$p$ completion of $G$ by $G(p)$. For more notation, see also the end of this introduction.

**Theorem** (cf. Theorems 5.1 and 6.4). Let $K$ be a number field, $p$ a rational prime, $p$ a prime of $K$ and $T \supseteq S \supseteq R$ sets of primes of $K$ with $R$ finite. Assume that $S$ is $p$-stable for $K_S^R(\mu_p)/K$. Then:

---

1In fact a weaker condition would suffice; see Theorem 5.1.
Stable sets of primes in number fields

(A) (Local extensions)

\[ K_{S, p}^R \supseteq \begin{cases} K_p(p) & \text{if } p \in S \setminus R, \\ K_p^{nr}(p) & \text{if } p \notin S. \end{cases} \]

(B) (Riemann’s existence theorem) Let \( I'_p(p) \) denote the Galois group of the maximal pro-\( p \)-extension of \( K_{S, p}^R \) and let \( K_T'/K_S^R \) denote the maximal pro-\( p \) subextension of \( K_T/K_S^R \). The natural map

\[ \phi_{T, S}^R : \bigotimes_{p \in R(K_S^R)} I'_p(p) \rightarrow G_{K_T'(p)/K_S^R} \]

is an isomorphism (where \( \bigotimes \) is to be understood in the sense of [NSW 2008, Chapter IV]).

(C) (Cohomological dimension) Assume that either \( p \) is odd or \( K \) is totally imaginary. Then

\[ \text{cd}_p G_{K, S}^R = \text{scd}_p G_{K, S}^R = 2. \]

(D) (K(\( \pi, 1 \))-property) Assume additionally that \( R = \emptyset \), \( S \supseteq S_\infty \) and that either \( p \) is odd or \( K \) is totally imaginary. Then \( \text{Spec} \mathcal{O}_{K, S} \) is a K(\( \pi, 1 \)) for \( p \) (see Definition 6.1).

There are also corresponding results for the maximal pro-\( p \) quotient \( G_{K, S}^R(p) \) of \( G_{K, S}^R \). These results are essentially well-known (see [NSW 2008]) if \( \delta_K(S) = 1 \) and respectively if \( S \supseteq S_p \cup S_\infty \). Also, A. Schmidt showed recently that if \( T_0 \) is any fixed set with \( \delta_K(T_0) = 1 \) and \( S \) is an arbitrary finite set of primes, then there is a finite subset \( T_1 \subseteq T_0 \) (depending on \( S \)) such that the pro-\( p \) versions of the above results essentially (e.g., except the result on \( \text{scd}_p \)) hold if one replaces \( S \) by \( S \cup T_1 \) (see [Schmidt 2007; 2009; 2010]).

A further application of stable sets concerns a generalization of the Neukirch–Uchida theorem, which is a result of anabelian nature. More details on this can be found in [Ivanov 2013, Section 6]. Now we see many examples of stable (even persistent) sets:

**Corollary 3.4.** Let \( M/K \) be finite Galois and let \( \sigma \in G_{M/K} \). Let \( S \subseteq P_{M/K}(\sigma) \) (i.e., up to a density-zero subset, \( S \) is equal to \( P_{M/K}(\sigma) \)). Let \( \mathcal{L}/K \) be any extension. Then \( S \) is persistent — or, equivalently, stable (see Corollary 3.6) — for \( \mathcal{L}/K \) if and only if

\[ G_{M/M \cap \mathcal{L}} \cap C(\sigma; G_{M/K}) \neq \emptyset, \]

where \( C(\sigma; G_{M/K}) \) denotes the conjugacy class of \( \sigma \) in \( G_{M/K} \). In particular:

(i) If \( \sigma = 1 \), then \( S \subseteq P_{M/K}(1) = \text{cs}(M/K) \) is persistent for any extension \( \mathcal{L}/K \).

(ii) If \( M \cap \mathcal{L} = K \), then \( S \subseteq P_{M/K}(\sigma) \) is persistent for \( \mathcal{L}/K \).
**Outline.** In Section 2 we introduce stable, sharply $p$-stable, strongly $p$-stable and persistent sets. Section 3 is devoted to examples: in particular, we introduce almost Chebotarev sets, which provide us with a rich supply of persistent sets (Section 3B), and we show essentially that an almost Chebotarev set is sharply and strongly $p$-stable for almost all $p$ (Section 3C). In Section 4A we prove our main result which is a general Hasse principle. In Sections 4B–4D we discuss some further Hasse principles and uniform bounds on Shafarevich–Tate groups for stable sets. In Section 5 we deduce arithmetic applications such as the Grunwald–Wang theorem, realization of local extensions, Riemann’s existence theorem and cohomological dimension. In Section 6 we use results from Section 5 to deduce the $K(\pi,1)$-property at $p$ for Spec $\mathcal{O}_{K,S}$ with $S$ being sharply $p$-stable.

**Notation.** Our notation essentially coincides with the notation in [NSW 2008]. We collect some of the most important notation here. For a profinite group $G$ we denote by $G(p)$ its maximal pro-$p$ quotient. For a subgroup $H \subseteq G$, we denote by $N_G(H)$ its normalizer in $G$. If $\sigma \in G$, then we write $C(\sigma; G)$ for its conjugacy class. For two finite groups $H \subseteq G$, we write $m_H^G$ (or $m_H$, if $G$ is clear from the context) for the character of the induced representation $\text{Ind}^G_H 1_H$.

For a Galois extension $M/L$ of fields, $G_{M/L}$ denotes its Galois group and $L(p)$ denotes the maximal pro-$p$ extension of $L$ (in a fixed algebraic closure). By $K$ we always denote an algebraic number field, that is, a finite extension of $\mathbb{Q}$. If $p$ is a prime of $K$ and $L/K$ is a Galois extension, then $D_{p,L/K} \subseteq G_{L/K}$ denotes the decomposition subgroup of $p$. We write $\Sigma_K$ for the set of all primes of $K$ and $S, T, R, \ldots$ will usually denote subsets of $\Sigma_K$. If $L/K$ is an extension and $S$ a set of primes of $K$, then we denote the pull-back of $S$ to $L$ by $S_L$, $S(L)$ or $S$ (if no ambiguity can occur). We write $K^G_S/K$ for the maximal extension of $K$, which is unramified outside $S$ and completely split in $R$, and we write $G^G_S := G_{K^G_S}$ for its Galois group. We use the abbreviations $K_S := K^\varnothing_S$ and $G_S := G_{K^\varnothing_S}$. Further, for $p \leq \infty$ a (archimedean or nonarchimedean) prime of $\mathbb{Q}$, we let $S_p = S_p(K)$ denote the set of all primes of $K$ lying over $p$. Further, if $S \subseteq \Sigma_K$, we write $\mathbb{N}(S) := \mathbb{N} \cap \mathcal{O}_{K,S}^\times$, i.e., $p \in \mathbb{N}(S)$ if and only if $S_p \subseteq S$.

We write $\delta_K$ for the Dirichlet density on $\Sigma_K$. For $S, T$ subsets of $\Sigma_K$, we use (following [NSW 2008, Definition 9.1.2])

$$S \subseteq T :\iff \delta_K(S \setminus T) = 0, \quad S \supseteq T :\iff (S \subseteq T) \text{ and } (T \subseteq S).$$

Thus $S \subseteq T$ if $S$ is contained in $T$ up to a set of primes of density zero. For a finite Galois extension $M/K$ and $\sigma \in G_{M/K}$, we have the Chebotarev set

$$P_{M/K}(\sigma) = \{ p \in \Sigma_K : p \text{ is unramified in } M/K \text{ and } (p, M/K) = C(\sigma; G_{M/K}) \},$$

where $(p, M/K)$ denotes the conjugacy class of Frobenius elements corresponding to primes of $M$ lying over $p$. 

2. Stable and persistent sets

2A. Warm-up: preliminaries on Dirichlet density. Let $\mathcal{P}_K$ denote the set of all subsets of $\Sigma_K$. The Dirichlet density $\delta_K$ is not defined for all elements in $\mathcal{P}_K$. Moreover, there are examples of finite extensions $L/K$ and $S \in \mathcal{P}_K$ such that $S$ has a density but the pull-back $S_L$ of $S$ to $L$ has no density. To avoid dealing with such sets we make the following convention, which holds until the end of this article.

Convention 2.1. If $S \in \mathcal{P}_K$ is a set of primes of $K$, then we assume implicitly that, for all finite extensions $L/K$, all finite Galois extensions $M/L$ and all $\sigma \in G_{M/L}$, the set $S_L \cap P_{M/L}(\sigma)$ has a Dirichlet density.\footnote{The optimal way to omit sets having no density would be to find an appropriate sub-$\sigma$-algebra of $\mathcal{P}_K$ (for any $K$) such that the restriction of $\delta_K$ to it is a measure (and the pull-back maps $\mathcal{P}_K \to \mathcal{P}_L$ attached to finite extensions $L/K$ restrict to pull-back maps on these sub-$\sigma$-algebras). Unfortunately, there is no satisfactory way to find such a $\sigma$-algebra $\mathcal{P}_K$, at least not if one requires that if $S \in \mathcal{P}_K$, then $T \in \mathcal{P}_K$ for any $T \supseteq S$, or, if one requires the weaker condition that any finite set of primes of $K$ lies in $\mathcal{P}_K$. Indeed, countability of $\Sigma_K$ would imply $\mathcal{P}_K = \mathcal{P}_K$ in this case, but not all elements of $\mathcal{P}_K$ have a Dirichlet density.}

Convention 2.1 is satisfied for all sets lying in the rather large subset

$$\mathcal{A}_K := \left\{ S \subseteq \Sigma_K : S \simeq \bigcup_i P_{L_i/K_i}(\sigma_i)_{K_i} \text{ for some } K/K_i/\mathbb{Q} \right. \\
\left. \text{and } L_i/K_i \text{ finite Galois and } \sigma_i \in G_{L_i/K_i} \right\}$$

of $\mathcal{P}_K$, where the unions are disjoint and countable (or finite or empty). The set $\mathcal{A}_K$ cannot be closed simultaneously under (arbitrary) unions and complements: otherwise it would be a $\sigma$-algebra and hence would be equal to $\mathcal{P}_K$.

To compute the density of pull-backs of sets we use the following two lemmas. Let $L/K$ be a finite extension of degree $n$ (not necessarily Galois). For $0 \leq m \leq n$, define the sets

$$P_m(L/K) := \{ p \in \Sigma_K : p \text{ is unramified and has exactly } m \text{ degree-1 factors in } L \}.$$

In particular, $P_n(L/K) = cs(L/K)$, $P_{n-1}(L/K) = \emptyset$. Recall that if $H \subseteq G$ are finite groups, then $m_H$ denotes the character of the $G$-representation $\text{Ind}_H^G 1$. One has

$$m_H(\sigma) = |\{ gH : \langle \sigma \rangle^g \subseteq H \}| = |\{ \langle \sigma \rangle gH : \langle \sigma \rangle^g \subseteq H \}|,$$

where $\langle \sigma \rangle \subseteq G$ denotes the subgroup generated by $\sigma$ and where $\langle \sigma \rangle^g := g^{-1} \langle \sigma \rangle g$. The second equality follows immediately from the fact that if $\langle \sigma \rangle^g \subseteq H$, then $gH = \langle \sigma \rangle gH$.

Lemma 2.2. Let $L/K$ be a finite extension and $N/K$ a finite Galois extension containing $L$, with Galois group $G$, such that $L$ corresponds to a subgroup $H \subseteq G$. 

The optimal way to omit sets having no density would be to find an appropriate sub-$\sigma$-algebra of $\mathcal{P}_K$ (for any $K$) such that the restriction of $\delta_K$ to it is a measure (and the pull-back maps $\mathcal{P}_K \to \mathcal{P}_L$ attached to finite extensions $L/K$ restrict to pull-back maps on these sub-$\sigma$-algebras). Unfortunately, there is no satisfactory way to find such a $\sigma$-algebra $\mathcal{P}_K$, at least not if one requires that if $S \in \mathcal{P}_K$, then $T \in \mathcal{P}_K$ for any $T \supseteq S$, or, if one requires the weaker condition that any finite set of primes of $K$ lies in $\mathcal{P}_K$. Indeed, countability of $\Sigma_K$ would imply $\mathcal{P}_K = \mathcal{P}_K$ in this case, but not all elements of $\mathcal{P}_K$ have a Dirichlet density.
Then
\[ P_m(L/K) \cong \{ \mathfrak{p} \in P_m(L/K) : \mathfrak{p} \text{ is unramified in } N/K \} = \bigcup_{C(\sigma; G) \subseteq G, m_H(\sigma) = m} P_{N/K}(\sigma), \]
where the right-hand side is a disjoint union. In particular, \( P_m(L/K) \in \mathcal{A}_K \) and
\[ \delta_K(P_m(L/K)) = |G|^{-1} \sum_{C(\sigma; G) \subseteq G, m_H(\sigma) = m} |C(\sigma; G)|. \]

\textit{Proof.} The proof of the first statement is an elementary exercise in Galois theory: if \( \mathfrak{p} \) is a prime of \( K \) unramified in \( N \), then the primes of \( L \) lying over \( \mathfrak{p} \) are in one-to-one correspondence with double cosets \( \langle \sigma \rangle gH \), where \( \sigma \) is arbitrary in the Frobenius class of \( \mathfrak{p} \); the residue field extension of a prime belonging to the coset \( \langle \sigma \rangle gH \) over \( \mathfrak{p} \) has the Galois group \( \langle \sigma \rangle^g / \langle \sigma \rangle^g \cap H \). The second statement follows from the first and the Chebotarev density theorem. \( \square \)

**Lemma 2.3.** Let \( L/K \) be a finite extension of degree \( n \), let \( S \) be a set of primes of \( K \) and let \( N/K \) be a Galois extension containing \( L \) such that \( G := G_{N/K} \supseteq G_{N/L} =: H \). Then
\[ \delta_L(S) = \sum_{m=1}^n m\delta_K(S \cap P_m(L/K)) = \sum_{C(\sigma; G) \subseteq G} m_H(\sigma)\delta_K(S \cap P_{N/K}(\sigma)). \]
If, in particular, \( L/K \) is Galois, we get the well-known formula
\[ \delta_L(S) = [L : K]\delta_K(S \cap \text{cs}(L/K)). \]

\textit{Proof.} The first equation is an easy computation and the second follows from Lemma 2.2. \( \square \)

**2B. Key definitions.** Let \( K \) be a number field and \( S \) a set of primes. If \( \delta_K(S) = 0 \) (resp. \( = 1 \)), then \( \delta_L(S) = 0 \) (resp. \( = 1 \)) for all finite \( L/K \). Now, if \( 0 < \delta_K(S) < 1 \), then it can happen that there is some finite \( L/K \) with \( \delta_L(S) = 0 \) (take a finite Galois extension \( L/K \) and set \( S := \Sigma_K \setminus \text{cs}(L/K) \); then \( \delta_K(S) = 1 - [L : K]^{-1} \) and \( \delta_L(S_L) = 0 \)). For stable sets, defined below, this possibility is excluded.

**Definition 2.4.** Let \( S \) be a set of primes of \( K \), let \( L/K \) be any extension and let \( \lambda > 1 \). A finite subextension \( L/L_0/K \) is called \( \lambda \)-stabilizing for \( S \) for \( L/K \) if there exists a subset \( S_0 \subseteq S \) and some \( a \in (0, 1] \) such that \( \lambda a > \delta_L(S_0) \geq a > 0 \) for all finite subextensions \( L/L/L_0 \). Moreover, we call \( L_0 \) persisting for \( S \) for \( L/K \) if there exists a subset \( S_0 \subseteq S \) such that \( \delta_L(S_0) = \delta_{L_0}(S_0) > 0 \) for all finite subextensions \( L/L_0 \). Further:

(i) We call \( S \) \( \lambda \)-stable (resp. persistent) for \( L/K \) if it has a \( \lambda \)-stabilizing (resp. persisting) extension for \( L/K \).
(ii) We call $S$ stable for $\mathcal{L}/K$ if there is a $\lambda > 1$ such that $S$ is $\lambda$-stable for $\mathcal{L}/K$.
Assume that $\lambda = p$ is a rational prime.

(iii) We call $S$ sharply $p$-stable for $\mathcal{L}/K$ if $\mu_p \subseteq \mathcal{L}$ and if $S$ is $p$-stable for $\mathcal{L}/K$, or if $\mu_p \not\subseteq \mathcal{L}$ and if $S$ is stable for $\mathcal{L}(\mu_p)/K$.

In applications we will often use the case $\mathcal{L} = K_S$. Therefore, we provide the following definition:

**Definition 2.5.** Let $S$ be a set of primes of $K$ and let $\lambda > 1$.

(i) We call $S$ $\lambda$-stable (resp. stable, resp. persistent) if $S$ is $\lambda$-stable (resp. stable, resp. persistent) for $K_S/K$.
Assume that $\lambda = p$ is a rational prime.

(ii) We call $S$ sharply $p$-stable if $S$ is sharply $p$-stable for $K_S/K$. Moreover, we define the exceptional set $E_{\text{sharp}}(S)$ to be the set of all rational primes $p$ such that $S$ is not sharply $p$-stable.

(iii) We call $S$ strongly $p$-stable if $S$ is $p$-stable for $K_{S\cup S_p\cup S_\infty}/K$ with a $p$-stabilizing field contained in $K_S$. Further, we call $S$ strongly $\infty$-stable if $S$ is stable for $K_{S\cup S_\infty}/K$. Moreover, we define the exceptional set $E_{\text{strong}}(S)$ to be the set of all rational primes $p$ or $p = \infty$ such that $S$ is not strongly $p$-stable.

Clearly, a strongly $p$-stable set is also $p$-stable and sharply $p$-stable. In particular, we have $E_{\text{strong}}(S) \supseteq E_{\text{sharp}}(S)$. On the other side, in general, neither one of the properties ‘$p$-stable’ or ‘sharply $p$-stable’ implies the other.

**Lemma 2.6.** Let $\mathcal{L}/K$ be an extension and $S$ a set of primes of $K$.

(i) Let $\lambda \geq \mu > 1$. If $S$ is $\mu$-stable with $\mu$-stabilizing field $L_0$, then $S$ is $\lambda$-stable with $\lambda$-stabilizing field $L_0$.

(ii) If $L_0$ is a $\lambda$-stabilizing (resp. persisting) field for $S$ for $\mathcal{L}/K$, then any finite subextension $\mathcal{L}/L_1/L_0$ has the same property.

(iii) Let $S'$ be a further set of primes of $K$. If $S \subseteq S'$ and if $S$ is $\lambda$-stable (resp. persistent) for $\mathcal{L}/K$, then $S'$ also has this property. Any $\lambda$-stabilizing (resp. persisting) field for $S$ has the same property for $S'$.

(iv) Let $\mathcal{L}/N/M/K$ be subextensions. If $S$ is $\lambda$-stable (resp. persistent) for $\mathcal{L}/K$ with $\lambda$-stabilizing (resp. persisting) field $L_0 \subseteq N$, then $S_M$ is $\lambda$-stable (resp. persistent) for $N/M$.

**Lemma 2.7.** Let $\mathcal{L}/K$ be an extension and $S$ a set of primes of $K$. Assume that $S$ is sharply $p$-stable for $\mathcal{L}/K$. There is a finite subextension $\mathcal{L}/L_0/K$ such that, for any subextensions $\mathcal{L}/N/L/L_0$ (with $L/L_0$ finite), $S$ is sharply $p$-stable for $N/L$. 
The proofs of these lemmas are straightforward. The following proposition gives another characterization of stable sets and shows, in particular, that if $S$ is stable for $\mathcal{L}/K$, then any finite subfield $\mathcal{L}/L/K$ is $\lambda$-stabilizing for $S$ with a certain $\lambda > 1$ depending on $L$.

**Proposition 2.8.** Let $S$ be a set of primes of $K$ and $\mathcal{L}/K$ an extension. The following are equivalent:

(i) $S$ is stable for $\mathcal{L}/K$.

(ii) There exists some $\lambda > 1$ such that $S$ is $\lambda$-stable for $\mathcal{L}/K$ with $\lambda$-stabilizing field $K$.

(iii) There exists some $\epsilon > 0$ such that $\delta_L(S) > \epsilon$ for all finite $\mathcal{L}/L/K$.

**Proof.** The directions (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) are trivial. We prove (i) $\Rightarrow$ (iii). Let $\lambda > 1$ and let $S$ be $\lambda$-stable for $\mathcal{L}/K$ with $\lambda$-stabilizing field $L_0$. Then there is some $a > 0$ and a subset $S_0 \subseteq S$ such that $a \leq \delta_L(S_0) < \lambda a$ for all $\mathcal{L}/L/L_0$. Suppose there is no $\epsilon > 0$ such that $\delta_L(S_0) > \epsilon$ for all $\mathcal{L}/L/K$. This implies that there is a family $(M_i)_{i=1}^{\infty}$ of finite subextensions of $\mathcal{L}/K$ with $\delta_{M_i}(S_0) \to 0$ as $i \to \infty$. Then $d_i = [L_0 M_i : M_i] = [L_0 : L_0 \cap M_i]$ is bounded from above by $[L_0 : K]$ and hence, by Lemma 2.3,

$$\delta_{L_0 M_i}(S_0) = \sum_{m=1}^{d_i} m \delta_{M_i}(S_0 \cap P_m(L_0 M_i/M_i)) \leq [L_0 : K] \delta_{M_i}(S_0) \to 0$$

for $i \to \infty$. This contradicts the $\lambda$-stability of $S_0$ with respect to the $\lambda$-stabilizing field $L_0$. \qed

Here is a brief overview of the use of these conditions and some examples:

- Most examples of stable sets are given by (almost) Chebotarev sets, i.e., sets of the form $S \simeq P_{M/K}(\sigma)$, or sets containing them (see Section 3B).
- If $S$ is stable for $\mathcal{L}/K$, then $\delta_L(S) > 0$ for all finite $\mathcal{L}/L/K$. The converse is not true in general (see [Ivanov 2013, Section 3.5.4]), but it is true for almost Chebotarev sets (see Section 3B).
- If an almost Chebotarev set is stable for an extension, then it is also persistent for it (see Corollary 3.6). It is not clear whether there are examples of stable but not persistent sets (but see [Ivanov 2013, Section 3.5.4]).
- For a stable almost Chebotarev set $S$, the set $E^{\text{sharp}}(S)$ is finite and $E^{\text{strong}}(S)$ is either $\Sigma_\mathbb{Q}$ or finite (see Section 3C).
- Roughly speaking, $p$-stability (for $\mathcal{L}/K$) is enough to prove Hasse principles in dimension 1 for $p$-primary $(G_{\mathcal{L}/K})$-modules. See Section 4.
• To prove Hasse principles in dimension 2 and Grunwald–Wang-type results for $p$-primary $G_{K,S}$-modules, we need strong $p$-stability. We will give examples of persistent sets $S$ together with a finite set $T$ such that Grunwald–Wang (even stably) fails, i.e., $\text{coker}^1(K_{S\cup T}/L; \mathbb{Z}/p\mathbb{Z}) \neq 0$ for all finite subextensions $K_S/L/K$. But it is not clear whether one can find such an example with the additional requirement that $T \subseteq S$ (and necessarily $S$ being not strongly $p$-stable). See Section 5B.

• On the other side, for applications of Grunwald–Wang (i.e., to prove Riemann’s existence theorem, to realize local extensions by $K_S/K$, to compute (strict) cohomological dimension, etc.), it is enough to require that $S$ is sharply $p$-stable. See Sections 5A, 5C and 5D.

3. Examples

In this section we construct examples of stable sets. First, in Section 3A, we see to which extent ‘stable’ is more general than ‘of density 1’. Then, in Sections 3B and 3C, we introduce almost Chebotarev sets and determine when they are stable, strongly $p$-stable, and sharply $p$-stable. Finally, in Section 3D, we construct a stable almost Chebotarev set $S$ with $\mathbb{N}(S) = \{1\}$.

3A. Sets of density one. Stable and persistent sets generalize sets of density one. In particular, every set of primes of $K$ of density one is persistent for any extension $\mathcal{L}/K$ with persisting field $K$ and is strongly $p$-stable for each $p$. Nevertheless, sets of density one have some properties which stable and persistent sets do not have in general:

(i) The intersection of two sets of density one again has density one, which is not true for stable and persistent sets: the intersection of two sets persistent for $\mathcal{L}/K$ can be empty (see Corollary 3.4 and explicit examples below).

(ii) If $S \subseteq \Sigma_K$ has density one, then there are infinitely many primes $p \in \Sigma_Q$ such that $S_p \subseteq S$ (otherwise, for all primes $p \in \text{cs}(K/Q)$ one could choose a prime $p \in S_p \setminus S$ of $K$ and we would have $\delta_K(S) \leq 1 - [K^n : \mathbb{Q}]^{-1}$, where $K^n$ denotes the normal closure of $K$ over $\mathbb{Q}$). On the other side, it is easy to construct a persistent set $S \subseteq \Sigma_K$ with $\mathbb{N}(S) = \{1\}$, i.e., $S_\ell \not\subseteq S$ for all $\ell \in \Sigma_Q$ (see Section 3D for an example).

Observe that, for sets $S$ with $\mathbb{N}(S) = \{1\}$, mentioned above, none of the $\ell$-adic representations $\rho_{A,\ell} : G_K \to \text{GL}_d(\overline{\mathbb{Q}}_\ell)$ which come from an abelian variety $A/K$ factor through the quotient $G_K \to G_{K,S}$ (indeed, the Tate-pairing on $A$ shows that the determinant of $\rho_{A,\ell}$ is the $\ell$-part of the cyclotomic character of $K$ and, in particular, $\rho_{A,\ell}$ is highly ramified at all primes of $K$ lying over $\ell$. If $\rho_{A,\ell}$ factored over $G_{K,S}$, then we would have $S_\ell \subseteq S$). In particular, this makes it very hard, if
not impossible, to study the group $G_{K,S}$ via the Langlands program (for example, like in [Chenevier 2007] and [Chenevier and Clozel 2009], where a prime $\ell \in \mathbb{N}(S)$ is always necessary). If $S$ is additionally stable, then methods involving stability allow us to study $G_{K,S}$.

3B. Almost Chebotarev sets.

**Definition 3.1.** Let $K$ be a number field and $S$ a set of primes of $K$. Then $S$ is called a Chebotarev set (resp. an almost Chebotarev set) if $S = \mathcal{P}_{M/K}(\sigma)$ (resp. $S \supset \mathcal{P}_{M/K}(\sigma)$), where $M/K$ is a finite Galois extension and $\sigma \in G_{M/K}$.

**Remark 3.2.** $M$ and the conjugacy class of $\sigma$ are not unique, i.e., there are pairs $(M/K, \sigma)$, $(N/K, \tau)$ such that $M \neq N$ and $\mathcal{P}_{M/K}(\sigma) \subset \mathcal{P}_{N/K}(\tau)$ (or even $\mathcal{P}_{M/K}(\sigma) = \mathcal{P}_{N/K}(\tau)$). If one restricts attention to pairs $(M/K, \sigma)$ such that $\sigma$ is central in $G_{M/K}$, then $(M/K, \sigma)$ is indeed unique. See [Ivanov 2013, Remark 3.13].

**Proposition 3.3.** Let $M/K$ be a finite Galois extension and let $\sigma \in G_{M/K}$. Let $L/K$ be any finite extension and set $L_0 := L \cap M$. Then

$$\delta_L(\mathcal{P}_{M/K}(\sigma)_L) = \frac{|C(\sigma; G_{M/K}) \cap G_{M/L_0}|}{|G_{M/L_0}|}.$$ 

Thus $\delta_L(\mathcal{P}_{M/K}(\sigma)_L) \neq 0$ if and only if $C(\sigma; G_{M/K}) \cap G_{M/L_0} \neq \emptyset$. In particular, this is always the case if $L_0 = K$ or if $\sigma = 1$.

**Proof.** Let $N/K$ be a finite Galois extension with $N \supseteq ML$. Define $H := G_{N/L}$ and $\bar{H} := G_{M/L_0}$. We have a natural surjection $H \twoheadrightarrow \bar{H}$. Let $1_\sigma$ denote the class function on $G_{M/K}$, which has value 1 on $C(\sigma; G_{M/K})$ and 0 outside. Finally, let $m_H$ denote the character on $G := G_{N/K}$ of the induced representation $\text{Ind}_{H}^{G} 1_{H}$. Then we have

$$\delta_L(\mathcal{P}_{M/K}(\sigma)_L) = \sum_{C(g;G)\rightarrow C(\sigma;G_{M/K})} \delta_L(\mathcal{P}_{N/K}(g)_L) = \sum_{C(g;G)\rightarrow C(\sigma;G_{M/K})} m_H(g) \delta_K(\mathcal{P}_{N/K}(g))$$ 

$$= \sum_{C(g;G)\rightarrow C(\sigma;G_{M/K})} m_H(g) \frac{|C(g; G)|}{|G|} = \frac{1}{|G|} \sum_{g \in C(\sigma; G_{M/K})} m_H(g)$$ 

$$= \langle m_H, \text{inf}^G_{G_{M/K}} 1_{\sigma} \rangle_G = \langle 1_H, \text{inf}^H_{G_{M/K}} 1_{\sigma} \rangle_H = \langle 1_H, 1_{\sigma}, 1_{\bar{H}} \rangle_{\bar{H}}$$ 

$$= \frac{|C(\sigma; G_{M/K}) \cap \bar{H}|}{|\bar{H}|}.$$ 

The first equality follows from [Wingberg 2006, Proposition 2.1] and the second from Lemma 2.3. The third-to-last equality is Frobenius reciprocity, and the second-to-last equality follows from the easy fact that if $H \twoheadrightarrow \bar{H}$ is a surjection of finite groups and $\chi$, $\rho$ are two characters of $\bar{H}$, then $\langle \text{inf}^H_{\bar{H}} \chi, \text{inf}^H_{\bar{H}} \rho \rangle_H = \langle \chi, \rho \rangle_{\bar{H}}$. □
Corollary 3.4. Let $M/K$ be finite Galois and let $\sigma \in G_{M/K}$. Let $\mathcal{L}/K$ be any extension and set $L_0 := M \cap L$. Then a set $S \preceq P_{M/K}(\sigma)$ is persistent for $\mathcal{L}/K$ if and only if

$$C(\sigma; G_{M/K}) \cap G_{M/L_0} \neq \emptyset.$$ If this is the case, $L_0$ is a persistent field for $S$ for $\mathcal{L}/K$. In particular:

(i) Any set $S \preceq \text{cs}(M/K)$ is persistent for any extension $\mathcal{L}/K$.

(ii) Any set $S \preceq P_{M/K}(\sigma)$ is persistent for any extension $\mathcal{L}/K$ with $L_0 \cap M = K$.

Example 3.5 (a persistent set). Let $K$ be a number field, $M/K$ a finite Galois extension which is totally ramified in a prime $p$ of $K$. Let $\sigma \in G_{M/K}$ and let $S$ be a set of primes of $K$ such that $S \preceq P_{M/K}(\sigma)$ and $p \notin S$. Then $S$ is persistent with persisting field $K$. Indeed, we have $K_S \cap M = K$ by construction, and the claim follows from Corollary 3.4.

Corollary 3.6. Let $S$ be an almost Chebotarev set and $\mathcal{L}/K$ an extension. Then the following are equivalent:

(i) $S$ is stable for $\mathcal{L}/K$.

(ii) $S$ is persistent for $\mathcal{L}/K$.

(iii) $\delta_L(S) > 0$ for all finite $\mathcal{L}/L/K$.

Proof. Suppose $S \preceq P_{M/K}(\sigma)$, with $M/K$ a finite Galois extension and $\sigma \in G_{M/K}$. By Proposition 3.3, the density of $S$ is constant and equal to some $d \geq 0$ in the tower $\mathcal{L}/L_0$ with $L_0 = L \cap M$. There are two cases: either $d = 0$ or $d > 0$. If $d = 0$, then $S$ is not stable and hence also not persistent for $\mathcal{L}/K$ by Proposition 2.8, i.e., (i), (ii) and (iii) do not hold in this case. If $d > 0$, then $S$ is obviously persistent for $\mathcal{L}/K$ with persisting field $L_0$ and hence also stable, i.e., (i), (ii), (iii) hold. □

Remark 3.7. If $S$ is any stable set, then (ii) $\Rightarrow$ (i) $\Rightarrow$ (iii) still holds. But (iii) $\Rightarrow$ (i) fails in general (see [Ivanov 2013, Section 3.5.4]) and it is not clear whether (i) $\Rightarrow$ (ii) holds.

3C. Finiteness of $E^{\text{sharp}}(S)$, $E^{\text{strong}}(S)$ and examples.

Proposition 3.8. Let $S \preceq P_{M/K}(\sigma)$ with $\sigma \in G_{M/K}$.

(i) If $\infty \in E^{\text{strong}}(S)$, then $E^{\text{strong}}(S)$ contains all rational primes. If $\infty \notin E^{\text{strong}}(S)$, then $E^{\text{strong}}(S)$ is finite.

(ii) Assume $S$ is stable. If $\mu_p \subset K_S$ or if $M/K$ is unramified in $S_p \setminus S$, then $S$ is sharply $p$-stable. In particular, if $S$ is stable, then $E^{\text{sharp}}(S)$ is finite.

Proof. (i) If $\infty \in E^{\text{strong}}(S)$, then $S$ does not have a stabilizing field for $K_{S_{\infty}}/K$ which is contained in $K_S$. This is, by Proposition 2.8, equivalent to the fact that $S$ is not stable for $K_{S_{\infty}}/K$, which in turn is equivalent, by Corollary 3.6, to the...
fact that \( \delta_L(S) = 0 \) for all \( K_{S\cup S_\infty}/L/L_0 \), where \( L_0 \) is some finite subextension of \( K_{S\cup S_\infty}/K \). Thus \( p \in E^{\text{strong}}(S) \) for any \( p \).

Now assume \( \infty \notin E^{\text{strong}}(S) \). Set \( L_0 := M \cap K_{S\cup S_\infty} \) and \( L_p := M \cap K_{S\cup S_p\cup S_\infty} \). By Proposition 3.3, the density of \( S \) is constant in the towers \( K_{S\cup S_\infty}/L_0 \) and \( K_{S\cup S_p\cup S_\infty}/L_p \) and equal to some real numbers \( d_0 \) and \( d_p \), respectively. Since \( S \) is stable for \( K_{S\cup S_\infty}/K \), we have \( d_0 > 0 \).

We claim that for almost all \( p \)s we have \( L_p = L_0 \). More precisely, this is true for all \( p \)s such that the set

\[
\{ p \in (S_p \setminus S)L_0 : p \text{ is ramified in } M/L_0 \}
\]

is empty. In fact, if this set is empty for \( p \), then the extension \( L_p/L_0 \) is unramified in \( S_p \setminus S(L_0) \), since it is contained in \( M/L_0 \). But, being contained in \( K_{S\cup S_p\cup S_\infty} \) and unramified in \( S_p \setminus S(L_0) \), it is contained in \( K_{S\cup S_\infty} \), and hence also in \( M \cap K_{S\cup S_\infty} = L_0 \), which proves our claim.

Now let \( p \) be such that \( L_p = L_0 \). We claim that \( S \) is \( (\lceil L_0 : K \rceil d_0^{-1}) \)-stable for \( K_{S\cup S_p\cup S_\infty}/K \) with \( (\lceil L_0 : K \rceil d_0^{-1}) \)-stabilizing field \( K \). Indeed, as \( L_p = L_0 \), we have \( d_p = d_0 > 0 \). Let \( K_{S\cup S_p\cup S_\infty}/N/K \) be any finite subextension. We have

\[ d_0 = \delta_{L_0N}(S) = \lceil L_0N : N \rceil \delta_N(S \cap cs(L_0N/N)) \leq \lceil L_0 : K \rceil \delta_N(S), \]

i.e., \( \delta_N(S) \geq \lceil L_0 : K \rceil^{-1} d_0 \) for all \( N \), and our claim follows.

Finally, almost all primes satisfy \( p > \lceil L_0 : K \rceil^{-1} d_0 \) and \( L_p = L_0 \). For such primes, \( S \) is \( p \)-stable for \( K_{S\cup S_p\cup S_\infty}/K \) with stabilizing field \( K \).

(ii) The second assertion of (ii) follows from the first. If \( \mu_p \subseteq K_S \), then \( S \) is sharply \( p \)-stable by Corollary 3.6. Assume \( M/K \) is unramified in \( S_p \setminus S \). Set \( L_0 := M \cap K_S \), \( L'_0 := L_0(\mu_p) \cap K_S \) and \( L_p := M \cap K_S(\mu_p) \). From these definitions and from our assumption on \( M/K \) we have

\[
\begin{align*}
(1) \quad G_{K_S(\mu_p)/L'_0} & \cong G_{K_S/L_0} \times G_{L_0(\mu_p)/L'_0}, \\
(2) \quad L_p \cap K_S & = L_0, \text{ and} \\
(3) \quad L_p/L_0 & \text{ is unramified in } S_p \setminus S.
\end{align*}
\]

By item (3) the extension \( L_pL'_0/L'_0 \) is unramified in \( S_p \setminus S \), and by item (1) we get \( L_p \subseteq L_pL'_0 \subseteq K_S \). Hence, (2) gives \( L_p = L_0 \). Thus for all \( K_S(\mu_p)/L/L_0 \) we have, by Proposition 3.3, \( \delta_L(S) = \delta_{L_p}(S) = \delta_{L_0}(S) > 0 \), since \( S \) is stable.

\[ \Box \]

**Remark 3.9.** Suppose \( S \cong P_{M/K}(\sigma) \). We have the following equivalences:

\[ p \notin E^{\text{sharp}}(S) \iff S \text{ stable for } K_S(\mu_p)/K \iff C(\sigma; G_{M/K}) \cap G(M/M \cap K_S(\mu_p)) \neq \emptyset. \]

**Example 3.10** (persistent sets with \( E^{\text{strong}}(S) \) finite but nonempty). Let \( K \) be a totally imaginary number field and let \( M/K \) be a finite Galois extension such that
Then, by the same reasoning as in the preceding example, Example 3.11 and Example 3.12 and

- \( M/K \) is totally ramified in a prime \( p \in S_p(K) \),
- \( d := [M : K] > p. \)

Let \( \sigma \in G_{M/K} \) and let \( S \) be a set of primes of \( K \) such that

- \( S \supseteq P_{M/K}(\sigma) \),
- \( \text{Ram}(M/K) \setminus S = \{p\}. \)

Then \( S \) is persistent \( (\delta_L(S) = d^{-1} \) for all \( K_S/L/K \)) with persisting field \( K \). Further, \( S \) is not strongly \( p \)-stable, i.e., \( p \in E^\text{strong}(S) \) and \( \infty \notin E^\text{strong}(S) \), i.e., \( E^\text{strong}(S) \) is finite by Proposition 3.8. Indeed, \( M \subseteq K_{S\cup S_p\cup S_\infty} \) and there are two cases, \( \sigma = 1 \) and \( \sigma \neq 1 \). In the second case, the density of \( S \) in \( K_{S\cup S_p\cup S_\infty}/K \) is zero beginning from \( M \), hence \( S \) is nonstable for this extension, and \( S \) is not strongly \( p \)-stable. In the first case, we have \( \delta_L(S) = 1 \) for all \( K_{S\cup S_p\cup S_\infty}/L/M \). Assume there is a \( p \)-stabilizing field \( N \subseteq K_S \) for \( K_{S\cup S_p\cup S_\infty}/K \), i.e., there is some \( S_0 \subseteq S \) and some \( a \in (0, 1] \) with \( a \leq \delta_L(S_0) < pa \) for all \( K_{S\cup S_p\cup S_\infty}/L/N \). But this leads to a contradiction. Indeed,

\[
\delta_{MN}(S_0) = [MN : N]\delta_N(S_0 \cap \text{cs}(MN/N)) = [M : K]\delta_N(S_0) \geq p\delta_N(S_0),
\]

since \( N \cap M = K \) and \( S_0 \subseteq S \supseteq \text{cs}(M/K) \).

**Example 3.11** (persistent sets with \( E^\text{strong}(S) = \emptyset \)). Let \( M/K \) be a finite Galois extension of degree \( d \), with \( K \) totally imaginary, which is totally ramified in at least two primes \( p \) and \( l \) with different residue characteristics \( \ell_1 \) and \( \ell_2 \), respectively. Let \( S \supseteq P_{M/K}(\sigma) \) for some \( \sigma \in G_{M/K} \) such that \( p, l \notin S \). Then \( M \cap K_S = K \), hence \( S \) is persistent with persisting field \( K \). Let \( p \) be a rational prime. Then \( M \cap K_{S\cup S_p\cup S_\infty} = K \), since \( M/K \) is totally ramified over primes with different residue characteristics \( \ell_1 \) and \( \ell_2 \). Hence \( S \) is strongly \( p \)-stable for every prime \( p \) and \( K \) is a persisting field for \( S \) for \( K_{S\cup S_p\cup S_\infty}/K \).

**Example 3.12** (persistent sets with \( E^\text{strong}(S) = \emptyset \)). There is also another possibility, to construct sets \( S \) with \( E^\text{strong}(S) = \emptyset \), using the same idea as in the preceding example. Assume for simplicity that \( K \) is totally imaginary. Let \( M_1, M_2/K \) be two Galois extensions of \( K \), and let \( \sigma_1 \in G_{M_1/K} \), \( \sigma_2 \in G_{M_2/K} \). Assume \( M_i/K \) is totally ramified in a nonarchimedean prime \( p_i \) of \( K \) such that the residue characteristics of \( p_1, p_2 \) are unequal. Then let \( S \) be a set of primes of \( K \) such that

- \( S \supseteq P_{M_1/K}(\sigma_1) \cup P_{M_2/K}(\sigma_2) \),
- \( \{p_1, p_2\} \notin S \).

Then, by the same reasoning as in the preceding example, \( S \) is persistent with persisting field \( K \) and \( E^\text{strong}(S) = \emptyset \). Moreover, for each rational prime \( p \), the field \( K \) is persisting for \( S \) for \( K_{S\cup S_p\cup S_\infty}/K \).
3D. Stable sets with $\mathbb{N}(S) = \{1\}$. Let $M/K/K_0$ be two finite Galois extensions of a number field $K_0$. Then the natural map $G_{M/K_0} \to \text{Aut}(G_{M/K})$ induces an exterior action

$$G_{K/K_0} \to \text{Out}(G_{M/K}),$$

thus inducing a natural action of $G_{K/K_0}$ on the set of all conjugacy classes of $G_{M/K}$. For any $g \in G_{K/K_0}$ and $\sigma \in G_{M/K}$, we choose a representative of the conjugacy class $g \cdot C(\sigma; G_{M/K})$ and denote it by $g \cdot \sigma$. Further, $G_{K/K_0}$ acts naturally on $\Sigma_K$, and we have

$$g \cdot P_{M/K}(\sigma) = P_{M/K}(g \cdot \sigma).$$

Let $K_0 = \mathbb{Q}$ and let $\sigma \in G_{M/K}$ be an element such that $C(\sigma; G_{M/K})$ is not fixed by the action of $G_{K/Q}$. Then set

$$S := \text{cs}(K/\mathbb{Q})_K \cap P_{M/K}(\sigma).$$

If $p \in \Sigma_K \setminus \text{cs}(K/\mathbb{Q})$, then $S \cap S_p = \emptyset$. If $p \in \text{cs}(K/\mathbb{Q})$ such that $S_p \cap S \neq \emptyset$, then the action of $g \in G_{K/K_0}$, chosen such that $C(\sigma; G_{M/K}) \neq C(g \cdot \sigma; G_{M/K})$, defines an isomorphism between the disjoint sets $S_p \cap P_{M/K}(\sigma)$ and $S_p \cap P_{M/K}(g \cdot \sigma)$, hence the last of these two sets is nonempty. From this we obtain $S_p \not\subseteq S$. Thus $\mathbb{N}(S) = \{1\}$. Moreover, if we choose $\sigma$ such that the stabilizer of $C(\sigma; G_{M/K})$ in $G_{K/Q}$ is trivial, then for any $p$ the intersection $S_p \cap S$ is either empty or contains exactly one element.

Now we have to choose $M$ in such a way that $S$ is stable. This is easy: e.g, take $M/K$ such that it is totally ramified in a fixed prime which is (by definition of $S$) not contained in $S$. Then $K_S \cap M = K$, i.e., $S$ is stable for $K_S/K$ with stabilizing field $K$, as $\delta_K(\text{cs}(K/\mathbb{Q})_K) = 1$ and hence $S \subseteq P_{M/K}(\sigma)$.

4. Shafarevich–Tate groups of stable sets

In this section we generalize many Hasse principles to stable sets and additionally prove finiteness and uniform bounds of certain Shafarevich–Tate groups associated with stable sets. The main result is the Hasse principle in Theorem 4.1. Further, there are two variants of uniform bounds on the size of $\mathbb{I}^1$: on the one side we can vary the coefficients, and on the other side the base field. We study both variants, the first in Section 4C and the second in Section 4D. These results are used in later sections.

4A. Stable sets and $\mathbb{I}^1$: key result. Let $K$ be a number field and $\mathcal{L}/K$ a (possibly infinite) Galois extension. Let $A$ be a finite $G_{\mathcal{L}/K}$-module. Let now $T$ be a set of primes of $K$. Consider the $i$-th Shafarevich–Tate group with respect to $T$,

$$\mathbb{I}^i(\mathcal{L}/K, T; A) := \ker(\text{res}^i : H^i(\mathcal{L}/K, A) \to \prod_{p \in T} H^i(\mathcal{O}_P, A)),$$
Stable sets of primes in number fields

where \( \mathcal{G}_p = \text{Gal}_{K_p} \) is the local absolute Galois group (the map res is essentially independent of the choice of this separable closure and we suppress it in the notation). We also write \( \Pi^i(K_S/K; A) \) instead of \( \Pi^i(K_S/K, S; A) \). We denote by \( K(A) \) the \emph{trivializing extension} for \( A \), i.e., the smallest field between \( K \) and \( K_p \) such that the subgroup \( G_{\mathcal{L}/K(A)} \) of \( G_{\mathcal{L}/K} \) acts trivially on \( A \). It is a finite Galois extension of \( K \).

Let \( G \) be a finite group and \( A \) a \( G \)-module. Following Serre [1964, §2] and Jannsen [1982], let \( H^i_\ast(G, A) \) be defined by exactness of the sequence

\[
0 \to H^i_\ast(G, A) \to H^i(G, A) \to \prod_{H \subset G \text{ cyclic}} H^i(H, A). 
\]

Our key result is the following theorem. All subsequent results make use of this theorem in a crucial way.

\textbf{Theorem 4.1.} Let \( K \) be a number field, \( T \) a set of primes of \( K \) and \( \mathcal{L}/K \) a Galois extension. Let \( A \) be a finite \( \text{Gal}_{\mathcal{L}/K} \)-module. Assume that \( T \) is \( p \)-stable for \( \mathcal{L}/K \), where \( p \) is the smallest prime divisor of \( |A| \). Let \( L \) be a \( p \)-stabilizing field for \( T \). Then

\[
\Pi^1(\mathcal{L}/L, T; A) \subseteq H^1_\ast(L(A)/L, A). 
\]

In particular, if \( H^1_\ast(L(A)/L, A) = 0 \), then \( \Pi^1(\mathcal{L}/L, T; A) = 0 \).

\textbf{Lemma 4.2.} Let \( \mathcal{L}/L/K \) be two Galois extensions of \( K \) and \( T \) a set of primes of \( K \). Let \( A \) be a \( \text{Gal}_{\mathcal{L}/K} \)-module such that for any \( p \in T \) one has \( A^{\mathcal{G}_{\mathcal{L}/L}} = A^{\mathcal{G}_{\mathcal{L}/L}} \). Then there is an exact sequence

\[
0 \to \Pi^1(L/K, T; A^{\mathcal{G}_{\mathcal{L}/L}}) \to \Pi^1(\mathcal{L}/K, T; A) \to \Pi^1(\mathcal{L}/L, T_L; A). 
\]

\textit{Proof.} The proof is an easy and straightforward exercise. \( \square \)

\textbf{Lemma 4.3.} Let \( L/K \) be a finite Galois extension, \( T \) a set of primes of \( K \) and \( A \) a finite \( \text{Gal}_{L/K} \)-module. Let \( i > 0 \). Assume that \( T \) is \( p \)-stable for \( L/K \) with \( p \)-stabilizing field \( K \), where \( p \) is the smallest prime divisor of \( |A| \). Then

\[
\Pi^i(L/K, T; A) \subseteq H^i_\ast(L/K, A). 
\]

\textit{Proof.} Since any \( p \)-stable set is \( \ell \)-stable for all \( \ell > p \), we can assume that \( A \) is \( p \)-primary. We have to show that any cyclic \( p \)-subgroup of \( \text{Gal}_{L/K} \) is a decomposition subgroup of a prime in \( T \). This is the content of the next lemma. \( \square \)

\textbf{Lemma 4.4.} Let \( L/K \) be a finite Galois extension, \( T \) a set of primes of \( K \) and \( p \) a rational prime such that \( T \) is \( p \)-stable for \( L/K \) with \( p \)-stabilizing field \( K \). Then any cyclic \( p \)-subgroup of \( \text{Gal}_{L/K} \) is the decomposition group of a prime in \( T \).

\textbf{Remark 4.5.} (i) This lemma shows automatically that there are infinitely many primes in \( T \) for which the given cyclic group is a decomposition group.
(ii) In some sense this lemma ‘generalizes’ Chebotarev’s density theorem, which says, in particular, that if $S$ has density one and $L/K$ is finite Galois, then any element of $G_{L/K}$ is a Frobenius of a prime in $S$.

Proof. Assume that the cyclic $p$-subgroup $H \subseteq G_{L/K}$ is not a decomposition group of a prime in $T$. Let $pH \subseteq H$ be the subgroup of index $p$. Then one computes directly $m_{pH}(\sigma) = pm_{H}(\sigma)$ for any $\sigma \in pH$. Since $H$ is not a decomposition subgroup of a prime $p \in T$, no generator of $H$ is a Frobenius at $T$, i.e., $P_{L/K}^{T}(\sigma) \cap T = \emptyset$ for any $\sigma \in H \setminus pH$. By $p$-stability of $T$, there is a subset $T_0 \subseteq T$ and an $a > 0$ such that $pa > \delta_{L}^{T}(T_0) \geq a$ for all $L/L'/K$. Let $L_0 = L^H$ and $L_1 = L^{bH}$. Then, by Lemma 2.3,

$$
\delta_{L_0}(T_0) = \sum_{\sigma \in H} m_{H}(\sigma)\delta_{K}(P_{L/K}^{T}(\sigma) \cap T_0) = \sum_{\sigma \in pH} m_{H}(\sigma)\delta_{K}(P_{L/K}^{T}(\sigma) \cap T_0)
$$

$$
= p^{-1}\sum_{\sigma \in pH} m_{pH}(\sigma)\delta_{K}(P_{L/K}^{T}(\sigma) \cap T_0) = p^{-1}\delta_{L_1}(T_0).
$$

This contradicts our assumption on $T_0$. □

Proof of Theorem 4.1. We can assume that $L = K$. By applying Lemma 4.2 to $L/K(A)/K$ and using Lemma 4.3, we are reduced to showing that if $A$ is a trivial $G$-module, then $\Pi^1(L/K, T; A) = 0$. Let $T_0 \subseteq T$ and $a > 0$ be such that $pa > \delta_{L}^{T}(T_0) \geq a$ for all $L/L'/K$. Let $G_{L/K}^{T}$ be the quotient of $G_{L/K}$, corresponding to the maximal subextension of $L/K$, which is completely split in $T$. We have then

$$
\Pi^1(L/K, T; A) = \ker\left(\text{Hom}(G_{L/K}^{T}, A) \to \prod_{p \in T} \text{Hom}(\mathfrak{g}_p, A)\right) = \text{Hom}(G_{L/K}^{T}, A).
$$

If $0 \neq \phi \in \text{Hom}(G_{L/K}^{T}, A)$, then $M := L^{\text{ker}(\phi)} / K$ is a finite extension inside $L/K$ with Galois group $\text{im}(\phi) \neq 0$ and completely decomposed in $T$, and in particular in $T_0$. Thus

$$
pa > \delta_{M}(T_0) = [M : K]\delta_{K}(T_0 \cap M/K)) = |\text{im}(\phi)|\delta_{K}(T_0) \geq pa,
$$

since $\delta_{K}(T_0) \geq a$. This is a contradiction, and hence we obtain

$$
\Pi^1(L/K, T; A) = \text{Hom}(G_{L/K}^{T}, A) = 0.
$$

□

4B. Hasse principles. Let $K$, $S$, $T$ be a number field and two sets of primes of $K$. Various conditions on $S$, $T$, $A$ which imply the Hasse principles in cohomological dimensions 1 and 2 are considered in [NSW 2008, Chapter IX, §1]. We prove analogous results for stable sets. Before stating them, we refer the reader to [NSW 2008, Definitions 9.1.5 and 9.1.7] for definitions of the special cases.

Corollary 4.6. Let $K$ be a number field, let $T$ and $S$ be sets of primes of $K$, and let $A$ be a finite $G_{K,S}$-module. Assume that $T$ is $p$-stable for $K_S/K$, where $p$ is the
smallest prime divisor of $|A|$. If $L$ is a $p$-stabilizing field for $T$ for $K_S/K$ and if $H_1^T(L(A)/L, A) = 0$, then

$$\Pi^1(K_S/L, T; A) = 0.$$ 

In particular:

(i) Let $L_0$ be a $p$-stabilizing field for $T$ for $K_S/K$ which trivializes $A$. Then $\Pi^1(K_S/L, T; A) = 0$ for any finite $K_S/L_0$.

(ii) Assume $S \supseteq S_{\infty}$ and $n \in \mathbb{N}(S)$ with the smallest prime divisor equal $p$. If $L_0$ is a $p$-stabilizing field for $T$ for $K_S/K$, then $\Pi^1(K_S/L, T; \mu_n) = 0$ for any finite $K_S/L_0$ such that we are not in the special case $(L, n, T)$. In the special case $(L, n, T)$ we have $\Pi^1(K_S/L, T; \mu_n) = \mathbb{Z}/2\mathbb{Z}$.

The same also holds if one replaces $G_{K,S}$ by the quotient $G_{K,S}(c)$, where $c$ is a full class of finite groups in the sense of [NSW 2008, Definition 3.5.2].

Proof. (i) The first statement follows directly from Theorem 4.1. Since $L_0$ is a $p$-stabilizing field trivializing $A$, any finite subextension $L$ of $K_S/L_0$ has the same property. Hence (i) follows.

(ii) To prove (ii), we can assume $n = p^r$. If we are not in the special case $(L, p^r)$, [NSW 2008, Proposition 9.1.6] implies $H^1(L(\mu_{p^r})/L, \mu_{p^r}) = 0$, i.e., we are done by Theorem 4.1. Assume we are in the special case $(L, p^r)$. In particular, we have $p = 2$. Then $H^1(L(\mu_{2^r})/L, \mu_{2^r}) = \mathbb{Z}/2\mathbb{Z}$. Since by Theorem 4.1 we have

$$\Pi^1(K_S/L(\mu_{2^r}), T; \mu_{2^r}) = 0,$$

we see from Lemma 4.2 that

$$\Pi^1(K_S/L, T; \mu_{2^r}) = \Pi^1(L(\mu_{2^r})/L, T; \mu_{2^r}).$$

Now the same argument as in the proof of [NSW 2008, Theorem 9.1.9(ii)] finishes the proof.

We turn to $\Pi^2$. For a $G_{K,S}$-module $A$ such that $|A| \in \mathbb{N}(S)$, we denote by

$$A' := \text{Hom}(A, G_{K_S,S}^*)$$

the dual of $A$. As in [NSW 2008, Corollary 9.1.10], we obtain:

**Corollary 4.7.** Let $K$ be a number field, $S \supseteq S_{\infty}$ a set of primes of $K$, and $A$ a finite $G_{K,S}$-module with $|A| \in \mathbb{N}(S)$. Assume that $S$ is $p$-stable (i.e., $p$-stable for $K_S/K$), where $p$ is the smallest prime divisor of $|A|$. Let $L$ be a $p$-stabilizing field for $S$ for $K_S/K$ such that $H^1_*(L(A')/L, A') = 0$. Then

$$\Pi^2(K_S/L; A) = 0.$$
In particular:

(i) Let $L_0$ be a $p$-stabilizing field for $S$ for $K_S/K$ which trivializes $A'$. Then $\Pi^2(K_S/L; A) = 0$ for any finite $K_S/L/L_0$.

(ii) Let $n \in \mathbb{N}(S)$ with smallest prime divisor $p$. If $L$ is a $p$-stabilizing field for $S$ and we are not in the special case $(L, n, S)$, then $\Pi^2(K_S/L, \mathbb{Z}/n\mathbb{Z}) = 0$. In the special case, we have $\Pi^2(K_S/L; \mathbb{Z}/n\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}$.

**Remark 4.8.** The condition $|A| \in \mathbb{N}(S)$ is not necessary if $A$ is trivial: we postpone the proof of this until all necessary ingredients (in particular the Grunwald–Wang theorem, Riemann’s existence theorem and $\text{cd}_p G_{K,S} = 2$) are proven. See Proposition 5.13.

**Proof of Corollary 4.7.** By Poitou–Tate duality [NSW 2008, Theorem 8.6.7] (this is the reason why we need $S \supseteq S_\infty$ and $|A| \in \mathbb{N}(S)$) we have

$$\Pi^2(K_S/L, A) \cong \Pi^1(K_S/L, A')^\vee,$$

where $X^\vee := \text{Hom}(X, \mathbb{R}/\mathbb{Z})$ is the Pontrjagin dual. An application of Theorem 4.1 to $K_S/K$, the sets $S = T$ and the module $A'$ gives the desired result. Now (i) and (ii) follow from Corollary 4.6. \qed

**4C. Finiteness of the Shafarevich–Tate group with divisible coefficients.** As a version of Corollary 4.6(i), we have the following proposition.

**Proposition 4.9.** Let $K$ be a number field, $\mathcal{L}/K$ a Galois extension, $p^m$ some rational prime power ($m \geq 1$). Let $T$ be a set of primes of $K$ which is $p^m$-stable for $\mathcal{L}/K$, with $p^m$-stabilizing field $L_0$. Then

$$|\Pi^1(\mathcal{L}/L, T; \mathbb{Z}/p^r\mathbb{Z})| < p^m$$

for any $r > 0$ and any finite subextension $\mathcal{L}/L/L_0$.

**Proof.** Let $T_0 \subseteq T$ and $a > 0$ be such that $a \leq \delta_L(T_0) < p^m a$ for all finite $\mathcal{L}/L/L_0$. Let $\mathcal{L}/L/L_0$ be a finite extension. Assume that $|\Pi^1(\mathcal{L}/L, T; \mathbb{Z}/p^r\mathbb{Z})| \geq p^m$. Then

$$|\Pi^1(\mathcal{L}/L, T_0; \mathbb{Z}/p^r\mathbb{Z})| \geq p^m$$

and we have

$$\Pi^1(\mathcal{L}/L, T_0; \mathbb{Z}/p^r\mathbb{Z}) \cong \text{Hom}(G^{T_0}_{\mathcal{L}/L}(p), \mathbb{Z}/p^r\mathbb{Z}) = (G^{T_0}_{\mathcal{L}/L}(p)^{ab}/p^r)^\vee.$$

Thus, $|\Pi^1(\mathcal{L}/L, T_0; \mathbb{Z}/p^r\mathbb{Z})| \geq p^m$ implies $|G^{T_0}_{\mathcal{L}/L}(p)^{ab}/p^r| \geq p^m$. Now, if $M/L$ is the subextension of $\mathcal{L}/L$ corresponding to $G^{T_0}_{\mathcal{L}/L}(p)^{ab}/p^r$, then it has a finite subextension $M_1$ of degree at least $p^m$ which is completely split in $T_0$. Hence, we have $\delta_{M_1}(T_0) \geq p^m \delta_L(T_0)$, which is a contradiction to the $p^m$-stability of $T_0$. \qed
Corollary 4.10. Let \( K \) be a number field, \( \mathscr{L}/K \) a Galois extension, and \( T \) a set of primes of \( K \) stable for \( \mathscr{L}/K \). Then \( \Pi^1(\mathscr{L}/K, T; \mathbb{Q}_p/\mathbb{Z}_p) \) is finite for any \( p \). Moreover, \( \Pi^1(\mathscr{L}/K, T; \mathbb{Q}/\mathbb{Z}) \) is finite.

Proof. For the first statement it is enough to show that \( |\Pi^1(\mathscr{L}/K, T; \mathbb{Z}/p^r\mathbb{Z})| \) is uniformly bounded for \( r > 0 \). By Proposition 2.8, there is some \( m \geq 1 \) such that \( K \) is a \( p^m \)-stabilizing field for \( T \) for \( \mathscr{L}/K \). Then Proposition 4.9 implies \( |\Pi^1(\mathscr{L}/K, T; \mathbb{Z}/p^r\mathbb{Z})| < p^m \). For the last statement, we note the decomposition \( \Pi^1(\mathscr{L}/K, T; \mathbb{Q}/\mathbb{Z}) = \bigoplus_p \Pi^1(\mathscr{L}/K, T; \mathbb{Q}_p/\mathbb{Z}_p) \). The proven part shows that each of the summands is finite. Moreover, almost all are zero: there is some \( \lambda > 1 \) such that \( K \) is a \( \lambda \)-stabilizing field for \( T \) for \( \mathscr{L}/K \). Thus, for any \( p \geq \lambda \), the group \( \Pi^1(\mathscr{L}/K, T; \mathbb{Q}_p/\mathbb{Z}_p) \) vanishes. \( \square \)

4D. Uniform bound. For later needs (see Section 5C) we prove the following uniform bounds. The results of this section are not part of [Ivanov 2013].

Proposition 4.11. Let \( \mathcal{M}/\mathscr{L}/K \) be Galois extensions, let \( A \) be a finite \( G_{\mathcal{M}/K} \)-module and let \( S \) be stable for \( \mathcal{M}(A)/K \). Then there is some \( C > 0 \) such that

\[
|\Pi^1(\mathcal{M}/L, S; A)| < C
\]

for all finite subextensions \( \mathcal{M}/L/K \).

Proof. For each \( \mathcal{L}/L/K \), Lemma 4.2 applied to \( \mathcal{M}/L(A)/L \) gives an exact sequence

\[
0 \to \Pi^1(L(A)/L, S; A) \to \Pi^1(\mathcal{M}/L, S; A) \to \Pi^1(\mathcal{M}/L(A), S_{L(A)}; A). \quad (4-1)
\]

Now \( \Pi^1(L(A)/L, S; A) \subseteq H^1(L(A)/L, A) \) and we have that \( G_{L(A)/L} \) is a subgroup of the finite group \( G_{K(A)/K} \), thus for all \( \mathcal{L}/L/K \) we have

\[
|\Pi^1(L(A)/L, S; A)| < m := 1 + \max_{H \leq G_{K(A)/K}} H^1(H, A).
\]

As \( S \) is stable for \( \mathcal{L}(A)/K \), by Proposition 2.8 there is some \( \epsilon > 0 \) such that \( \delta_N(S) > \epsilon \) for all \( \mathcal{L}(A)/N/K(A) \). Suppose that \( |\Pi^1(\mathcal{M}/L(A), S, A)| \geq \epsilon^{-1} \) for some \( \mathcal{L}/L/K \). Then, exactly as in the proof of Proposition 4.9, there is an extension \( M/L(A) \) of degree \( \geq \epsilon^{-1} \) which is completely split in \( S \). We obtain

\[
\delta_M(S) = [M : L(A)]\delta_{L(A)}(S) > \epsilon^{-1}\epsilon = 1,
\]

which is a contradiction. Taking into account (4-1), we obtain the statement of the proposition with respect to \( C := m\epsilon^{-1} \). \( \square \)

Corollary 4.12. Let \( K \) be a number field, \( S \) and \( T \) sets of primes of \( K \), and \( n \) a natural number.
(i) Assume that $K_S/K$ is a subextension such that $S$ is stable for $L/K$ and $T$ has density 0. Then there is some real $C > 0$ such that for any $L/K$ one has

$$|\Sha^I(K_S \cup T, S \setminus T, \mathbb{Z}/n\mathbb{Z})| < C.$$  

(ii) Assume that $T \supseteq (S_\infty \setminus S)$ has density 0 and that $n \in \mathbb{Q}_K^{*}$. Let $K_S/L/K$ be a subextension such that $S$ is stable for $\mathcal{L}(\mu_n)/K$. There is some real $C > 0$ such that for any $L/K$ one has

$$|\Sha^I(K_S \cup T, S \setminus T, \mu_n)| < C.$$  

Remark 4.13. The case $S$ stable for $L/K$, but not stable for $\mathcal{L}(\mu_p)/K$, still remains mysterious: one can neither show such a uniform bound by the same methods nor find counterexamples. Moreover, the same kind of arguments do not even show that $\Sha(K_S \cup T, S \setminus T, \mu_p)$ must be finite.

5. Arithmetic applications

5A. Overview and results. In this section we will be interested in the applications of the Hasse principles proven in the preceding section for stable sets. In particular, we will show two versions of the Grunwald–Wang theorem for them, with varying assumptions: we will have a strong Grunwald–Wang result if we assume strong $p$-stability (Section 5B) and only a weaker lim-version (which is still enough for applications) after weakening the assumption to sharp $p$-stability (Section 5C). After this we will consider realization of local extensions, Riemann’s existence theorem and the cohomological dimension of $G_{K,S}$. For each of these three results there is a profinite and a pro-$p$ version respectively. We state them below and give proofs in Section 5D. Further, in Section 5E we prove a Hasse principle for $\Sha^2$ for constant $p$-primary coefficients without the assumption $\mu_p \in \mathbb{Q}_K^{*}$ (see Corollary 4.7 and Remark 4.8).

Theorem 5.1. Let $K$ be a number field, $p$ a rational prime and $T \supseteq S \supseteq R$ sets of primes of $K$ with $R$ finite.

(A) Assume $S$ is sharply $p$-stable for $K^R_S(p)/K$. Then

$$K^R_S(p)_p = \begin{cases} K_p(p) & \text{if } p \in S \setminus R, \\ K^\text{nr}_p(p) & \text{if } p \not\in S. \end{cases}$$

(A) Assume $S$ is sharply $p$-stable for $K^R_S/K$. Then

$$K^R_{S,p} \supseteq \begin{cases} K_p(p) & \text{if } p \in S \setminus R, \\ K^\text{nr}_p(p) & \text{if } p \not\in S. \end{cases}$$
(Bp) Assume $S$ is sharply $p$-stable for $K^R_S(p)/K$. Then the natural map
\[
\phi^{R}_{T,S}(p) : \bigstar_{p \in R(K^R_S(p))} \langle \mathcal{O}_p(p) \rangle \bigstar_{p \in (T\setminus S)(K^R_S(p))} I_p(p) \rightarrow G_{K_T(p)/K^R_S(p)}
\]
is an isomorphism, where $I_p(p) := G_{K_p(p)/K^p}(p) \subseteq \mathcal{O}_p(p) := G_{K_p(p)/K_p}$.

Let $K'_T(p)/K^R_S$ denote the maximal pro-$p$ subextension of $K_T/K^R_S$.

(B) Assume $S$ is sharply $p$-stable for $K^R_S/K$. Then the natural map
\[
\phi^{R}_{T,S}(p) : \bigstar_{p \in R(K^R_S)} \langle \mathcal{O}_p(p) \rangle \bigstar_{p \in (T\setminus S)(K^R_S)} I'_p(p) \rightarrow G_{K'_T(p)/K^R_S}
\]
is an isomorphism, where $I'_p(p)$ denotes the Galois group of the maximal pro-$p$ extension of $K^R_{S,p}$.

Assume $p$ is odd or $K$ is totally imaginary.

(Cp) Assume $S$ is sharply $p$-stable for $K^R_S(p)/K$. Then
\[
\text{cd} G^R_{K,S}(p) = \text{scd} G^R_{K,S}(p) = 2.
\]

(C) Assume $S$ is sharply $p$-stable for $K^R_S/K$. Then
\[
\text{cd}_p G^R_{K,S} = \text{scd}_p G^R_{K,S} = 2.
\]

5B. Grunwald–Wang theorem and strong $p$-stability. Consider the cokernel of the global-to-local restriction homomorphism
\[
coker^i(K_S/K, T; A) := \text{coker}(\text{res}^i : H^i(K_S/K, A) \rightarrow \prod_{p \in T} H^i(\mathcal{O}_p, A)),
\]
where $A$ is a finite $G_{K,S}$-module, $T$ is a subset of $S$ and $\prod'$ means that almost all classes are unramified. If $A$ is a trivial $G_{K,S}$-module, then the vanishing of this cokernel is equivalent to the existence of global extensions unramified outside $S$, which realize given local extensions at primes in $T$. If $S$ has density 1, the set $T$ is finite, $A$ is constant and we are not in a special case, this vanishing is essentially the statement of the Grunwald–Wang theorem. Certain conditions on $S$, $T$, $A$, under which this cokernel vanishes are considered in [NSW 2008, Chapter IX, §2]. All of them require $S$ to have certain minimal density. We prove analogous results for stable sets.

**Corollary 5.2.** Let $K$ be a number field, $T \subseteq S$ sets of primes of $K$ with $S_\infty \subseteq S$. Let $A$ be a finite $G_{K,S}$-module with $|A| \in \mathbb{N}(S)$. Assume that $T$ is finite and $S$ is $p$-stable, where $p$ is the smallest prime divisor of $|A|$. For any $p$-stabilizing field $L$ for $S$ for $K_S/K$ such that $H^1_s(L(A')/L, A') = 0$, we have
\[
coker^i(K_S/L, T; A) = 0.
\]
Proof. Since $T$ is finite and $S$ is $p$-stable for $K_S/K$, we have that $S \setminus T$ is also $p$-stable for $K_S/K$, and the $p$-stabilizing fields for $S$ and $S \setminus T$ are equal. Let $L$ be as in the corollary. By Theorem 4.1 applied to $K_S/L$, $S \setminus T$ and $A'$, we obtain $\Pi^1(K_S/L, S \setminus T; A') = 0$. Then [NSW 2008, Lemma 9.2.2] implies that $\coker^1(K_S/L, T; A) = 0$. \hfill \Box

Now we give a generalization of [NSW 2008, Theorem 9.2.7].

**Theorem 5.3.** Let $K$ be a number field, $S$ a set of primes of $K$. Let $T_0$, $T \subseteq S$ be two disjoint subsets such that $T_0$ is finite. Let $p$ be a rational prime and $r > 0$ an integer. Assume $S \setminus T$ is $p$-stable for $K_{S \cup S_p \cup S_\infty}/K$ with $p$-stabilizing field $L_0$, which is contained in $K_S$. Then, for any finite $K_S/L/L_0$ such that we are not in the special case $(L, p', S \setminus (T_0 \cup T))$, the canonical map

$$H^1(K_S/L, \mathbb{Z}/p^r\mathbb{Z}) \to \bigoplus_{p \in T_0(L)} H^1(\mathfrak{g}_p, \mathbb{Z}/p^r\mathbb{Z}) \oplus \bigoplus_{p \in T(L)} H^1(\mathfrak{g}_p, \mathbb{Z}/p^r\mathbb{Z})$$

is surjective, where $\mathfrak{g}_p \subseteq \mathfrak{g}_p = G_{K_p^{\text{sep}}/L_p}$ is the inertia subgroup. If we are in the special case $(L, p', S \setminus (T_0 \cup T))$, then $p = 2$ and the cokernel of this map is of order 1 or 2.

**Proof.** This follows from Corollary 4.6(ii) in exactly the same way as [NSW 2008, Theorem 9.2.7] follows from [NSW 2008, Theorem 9.2.3(ii)]. \hfill \Box

**Remarks 5.4.** (i) If $\delta_K(T) = 0$, the condition $\mathcal{S} \setminus T$ is $p$-stable for $K_{S \cup S_p \cup S_\infty}/K$ with a $p$-stabilizing field contained in $K_S$ is equivalent to $\mathcal{S}$ is strongly $p$-stable.

(ii) If $\delta_K(S) = 1$ and $\delta_K(T) = 0$, then $L_0 = K$ is a persisting field for $S \setminus T$ for any $\mathcal{L}/K$ and the condition in the theorem is automatically satisfied. Thus our result is a generalization of [NSW 2008, Theorem 9.2.7]. To show that it is a proper generalization, we give the following example. Let $N/M/K$ be finite Galois extensions of $K$ such that $N/K$ (and hence also $M/K$) is totally ramified in a nonarchimedean prime $\mathfrak{l}$ of $K$, lying over the rational prime $\ell$. Suppose $\sigma \in G_{M/K}$ and let $\tilde{\sigma} \in G_{N/K}$ be a preimage of $\sigma$. Let $S \supseteq T$ be such that

$$S \subseteq P_{M/K}(\sigma), \mathfrak{l} \not\in S \text{ and } T \subseteq P_{M/K}(\sigma) \setminus P_{N/K}(\tilde{\sigma}).$$

Then $S \setminus T \subseteq P_{N/K}(\tilde{\sigma})$ is persistent for $K_{S \cup S_p \cup S_\infty}/K$ for any $p \neq \ell$, and, moreover, $K$ is a persisting field (indeed, this follows from $K_{S \cup S_p \cup S_\infty} \cap N = K$). Hence the sets $S \supseteq T$ satisfy the conditions of the theorem with respect to each $p \neq \ell$. Observe that in this example $T$ is itself persistent for $K_{S \cup S_p \cup S_\infty}/K$ with persisting field $K$. In [NSW 2008, Theorem 9.2.7], the set $T$ must have density zero.

From this we obtain the following classical form of the Grunwald–Wang theorem. The proof is the same as in [NSW 2008, Theorem 9.2.8].
Corollary 5.5. Let $T \subseteq S$ be sets of primes of a number field $K$. Let $A$ be a finite abelian group. Assume that $T$ is finite and that, for any prime divisor $p$ of $|A|$, $S$ is $p$-stable for $K_{S \cup \mathfrak{p} \cup \infty}/K$ with stabilizing field $K$. For all $\mathfrak{p} \in T$, let $L_{\mathfrak{p}}/K_{\mathfrak{p}}$ be a finite abelian extension such that its Galois group can be embedded into $A$. Assume that we are not in the special case $(K, \exp(A), S \setminus T)$. Then there exists a global abelian extension $L/K$ with Galois group $A$, unramified outside $S$, such that $L$ has completion $L_{\mathfrak{p}}$ at $\mathfrak{p} \in T$.

Example 5.6 (a set with persistent subset for which Grunwald–Wang stably fails). Let $p$ be an odd prime and assume $\mu_p \subset K$ (in particular, $K$ is totally imaginary and we can ignore the infinite primes). Let $S$ be a set of primes of $K$. For $V = S_{\mathfrak{p}} \setminus S$, let $T \supseteq V$ be a finite set of primes of $K$. By [NSW 2008, Theorem 9.2.2] we have for all $K_S/L/K$ a short exact sequence (recall that $\mu_p \cong \mathbb{Z}/p\mathbb{Z}$ by assumption)

$$0 \to \text{III}^1(K_{SUT}/L, S \cup T; \mathbb{Z}/p\mathbb{Z}) \to \text{III}^1(K_{SUT}/L, S \setminus T; \mathbb{Z}/p\mathbb{Z}) \to \text{coker}^1(K_{SUT}/L, T; \mathbb{Z}/p\mathbb{Z})^\vee \to 0.$$ 

Assume now that $S$ is $p$-stable with $p$-stabilizing field $K$. Then

$$\text{III}^1(K_{SUT}/L, S \cup T; \mathbb{Z}/p\mathbb{Z}) \subseteq \text{III}^1(K_{S}/L, S; \mathbb{Z}/p\mathbb{Z}) = 0$$

and hence we have

$$\text{coker}^1(K_{SUT}/L, T; \mathbb{Z}/p\mathbb{Z}) \cong \text{III}^1(K_{SUT}/L, S \setminus T; \mathbb{Z}/p\mathbb{Z})^\vee.$$ 

We can find such a set $S$ for which additionally $\text{III}^1(K_{SUT}/L, S \setminus T; \mathbb{Z}/p\mathbb{Z}) \neq 0$ for each $K_S/L/K$. For an explicit example, assume $K = \mathbb{Q}(\mu_p)$ and let $T \supseteq S_{\mathfrak{p}}(K)$ be a finite set of primes of $K$ ($S_{\mathfrak{p}}(K)$ consists of exactly one prime). Let $M/K$ be a Galois extension of degree $p$ with $\emptyset \neq \text{Ram}(M/K) \subseteq T$ (e.g., $M = \mathbb{Q}(\mu_{p^2})$). Define $S := \text{cs}(M/K)$. Then $M \cap K_S = K$ and hence $ML \cap K_S = L$ for each $K_S/L/K$. Thus $S$ is persistent with persisting field $K$. Further, $ML/L$ is a Galois extension of degree $p$ which is completely split in $S \setminus T$ and unramified outside $S \cup T$, hence the subgroup $G_{K_{SUT}/ML} \subsetneq G_{K_{SUT}/L}$ is the kernel of a nontrivial homomorphism $0 \neq \phi_M \in \text{III}^1(K_{SUT}/L, S \setminus T; \mathbb{Z}/p\mathbb{Z})$. Hence this group is nontrivial.

Thus, $S$ is persistent but not strongly $p$-stable—in particular, no $p$-stabilizing field for $S \cup T$ for $K_{S \cup \mathfrak{p} \cup \infty}/K$ is contained in $K_S$—and Grunwald–Wang does not hold for $S \cup T \supseteq T$ (i.e., the cokernel in Theorem 5.3 is nonzero). It is still unclear whether there is an example of sets $\tilde{S} \supseteq \tilde{T}$ such that $\tilde{S}$ is persistent but not strongly $p$-stable and Grunwald–Wang fails for $\tilde{S} \supseteq \tilde{T}$.

Finally, we have two corollaries generalizing [NSW 2008, Theorems 9.2.4 and 9.2.9] to stable sets.

Corollary 5.7. Let $K$ be a number field, $T \subseteq S$ sets of primes of $K$ with $T$ finite. Let $K_S/L/K$ be a finite Galois subextension with Galois group $G$. Let $p$ be a
prime and \( A = \mathbb{F}_p[G]^n \) a \( G_{K,S} \)-module. Assume \( S \) is \( p \)-stable for \( K_{S \cup S_p \cup S_\infty} / K \) with \( p \)-stabilizing field \( L \). Then the restriction map

\[
H^1(K_S / K, \ A) \to \bigoplus_{p \in T} H^1(\mathfrak{g}_p, \ A)
\]

is surjective.

**Proof.** (See [NSW 2008, Corollary 9.2.4]) We have the following commutative diagram, in which the vertical maps are Shapiro-isomorphisms:

\[
\begin{array}{ccc}
H^1(K_S / K, \ A) & \to & \bigoplus_{p \in T} H^1(\mathfrak{g}_p, \ A) \\
\downarrow \sim & & \downarrow \sim \\
H^1(K_S / L, \mathbb{F}_p^m) & \to & \bigoplus_{\mathfrak{q} \in T(L)} H^1(\mathfrak{g}_{\mathfrak{q} p}, \mathbb{F}_p^m)
\end{array}
\]

The lower map is surjective by Theorem 5.3, and so is the upper. \( \square \)

**Corollary 5.8.** Let \( K \) be number field, \( S \) a set of primes of \( K \). Let \( K_S / L / K \) be a finite Galois subextension with Galois group \( G \). Let \( p \) be a prime and \( A = \mathbb{F}_p[G]^n \) a \( G_{K,S} \)-module. Assume that \( S \) is \( p \)-stable for \( K_{S \cup S_p \cup S_\infty} / L \) with \( p \)-stabilizing field \( L \). Then the embedding problem

\[
\begin{array}{cccccc}
1 & \to & A & \to & E & \to & G & \to & 1
\end{array}
\]

is properly solvable.

**Proof.** It follows from Corollary 5.7 in the same way as [NSW 2008, Proposition 9.2.9] follows from [NSW 2008, Corollary 9.2.4]. \( \square \)

**5C. Grunwald–Wang cokernel in the limit and sharp \( p \)-stability.** If one is interested (motivated by Theorem 5.1, we are) in the vanishing of the direct limit over \( K_S / L / K \) of the Grunwald–Wang cokernel, rather than in the vanishing of the cokernel for each \( L \), one can use sharp \( p \)-stability instead of strong \( p \)-stability, which is considerably weaker.

**Theorem 5.9.** Let \( K \) be a number field, \( S \) a set of primes of \( K \) and \( \mathcal{L} \subseteq K_S \) a subextension normal over \( K \) such that \( S \) is sharply \( p \)-stable for \( \mathcal{L} / K \). Let \( T \) be a finite set of primes of \( K \) containing \((S_p \cup S_\infty) \setminus S\). If \( p^\infty | [\mathcal{L} : K] \), then

\[
\lim_{\mathcal{L} / L / K, \text{res}} \coker^1(K_{S \cup T} / L, T, \mathbb{Z} / p \mathbb{Z}) = 0.
\]
Proof. For any finite subextension \( \mathcal{L}/L/K \) we have the short exact sequence
\[
0 \to \mathbb{III}^1(K_{SUT}/L, S \cup T; \mu_p) \to \mathbb{III}^1(K_{SUT}/L, S \setminus T; \mu_p) \\
\quad \to \text{coker}^1(K_{SUT}/L, T; \mathbb{Z}/p\mathbb{Z})^\vee \to 0.
\]

Dualizing, we see that it is enough to show that
\[
\lim_{\mathcal{L}/L/K, \text{corv}} \mathbb{III}^1(K_{SUT}/L, S \setminus T; \mu_p)^{\vee} = 0.
\]

For any two finite subextensions \( \mathcal{L}/L'/L/K \) we have the maps
\[
\text{res}_{L}^{L'} : \mathbb{III}^1(K_{SUT}/L, S \setminus T; \mu_p) \to \mathbb{III}^1(K_{SUT}/L', S \setminus T; \mu_p) : \text{cor}_{L}^{L'}.
\]

Lemma 5.10. There is a finite subextension \( \mathcal{L}/L_1/K \) such that, for all \( \mathcal{L}/L'/L/L_1 \), the map \( \text{res}_{L}^{L'} \) is an isomorphism.

Proof. First we claim that \( \text{res}_{L}^{L'} \) is injective if \( L \) is big enough. Assume first that \( \mu_p \subseteq \mathcal{L} \) and that \( S \) is \( p \)-stable for \( \mathcal{L}/K \). Let \( \mathcal{L}/L_0/K \) be a finite subextension which \( p \)-stabilizes \( S \) and contains \( \mu_p \). Then any finite subextension \( \mathcal{L}/L/K_0 \) satisfies the same. Assume \( \text{res}_{L}^{L'} \) is not injective, i.e., there is some nonzero \( \phi \) in \( \mathbb{III}^1(K_{SUT}/L, S \setminus T; \mathbb{Z}/p\mathbb{Z}) \) with \( \text{res}_{L}^{L'}(\phi) = 0 \) (we have chosen some trivialization of \( \mu_p \)). This \( \phi \) can be seen as a homomorphism \( \phi : G_{K_{SUT}} \to \mathbb{Z}/p\mathbb{Z} \) which is trivial on all decomposition subgroups of primes in \( S \setminus T \). Define \( M := (K_{SUT})_{\text{ker} \phi} \). This is a finite Galois extension of \( L \) with Galois group \( \mathbb{Z}/p\mathbb{Z} \) and \( \text{cs}(M/L) \subseteq S \setminus T \).

But then
\[
\delta_M(S) = [M : L]\delta_L(S \cap \text{cs}(M/L)) = p \delta_L(S),
\]
since \( T \) is finite. Now \( \text{res}_{L}^{L'}(\phi) = 0 \) implies \( M \subseteq L' \subseteq \mathcal{L} \) and hence we get a contradiction to the \( p \)-stability of \( S \).

Now assume that \( \mu_p \not\subseteq K_S \). Then \( \text{res}_{L}^{L'} \) is always injective. Indeed, suppose there is a nonzero \( x \) in
\[
\mathbb{III}^1(K_{SUT}/L, S \setminus T; \mu_p)
\]
with \( \text{res}_{L}^{L'}(x) = 0 \). This implies \( x \in L'^{*,p} \). Let \( y^p = x \) with \( y \in L' \). Then \( L(y) \subseteq L' \subseteq \mathcal{L} \). Since the polynomial \( T^p - x \) is irreducible over \( L \) (since \( x \not\in L^{*,p} \)), the conjugates of \( y \) over \( L \) are precisely the roots of this polynomial, which are clearly \( \{\zeta_i^j y\}_{i=0}^{p-1} \) for \( \zeta \in \mu_p(K) \setminus \{1\} \). Since \( \mathcal{L} \) is normal over \( L \), these conjugates lie in \( \mathcal{L} \). In particular, we deduce that \( \zeta \in \mathcal{L} \), which contradicts \( \mu_p \not\subseteq \mathcal{L} \). This finishes the proof of the injectivity claim.

By Corollary 4.12(ii), there is a constant \( C > 0 \) such that
\[
|\mathbb{III}^1(K_{SUT}/L, S \setminus T, \mu_p)| < C
\]
for all $\mathcal{L}/L/K$. Together with the injectivity shown above, this shows that there is a finite subextension $\mathcal{L}/L_1/K$ such that, for all $\mathcal{L}/L_1'/L/L_1$, the map $\operatorname{res}_{L_1'}$ is bijective.

Now we can finish the proof of Theorem 5.9. Assume $L_1$ is as in Lemma 5.10. Let $\mathcal{L}/L_1/L$. Since $p^\infty[[\mathcal{L}:K]]$, there is a further extension $\mathcal{L}/L_1'/L$ such that $p$ divides $[L_1'/L]$. In the situation of (5-1) we have $\operatorname{cor} \circ \operatorname{res} = [L_1':L] = 0$ since $\mu_p$ is $p$-torsion. Dualizing gives $\operatorname{res}^\vee \circ \operatorname{cor}^\vee = (\operatorname{cor} \circ \operatorname{res})^\vee = 0$. But, along with $\operatorname{res}$, $\operatorname{res}^\vee$ is also an isomorphism, hence we obtain $\operatorname{cor}^\vee = 0$. This shows that

$$\lim_{\mathcal{L}/L/K.\operatorname{res}} \operatorname{III}^1(K_{S\cup T}/L, S \setminus T; \mu_p)^\vee = 0.$$  

We have the same arguments for $\operatorname{III}^2$.

**Proposition 5.11.** Let $K$ be a number field, $S$ a set of primes of $K$ and $\mathcal{L} \subseteq K_S$ a subextension normal over $K$ such that $S$ is sharply $p$-stable for $\mathcal{L}/K$. Let $T \supseteq S \cup S_p \cup S_\infty$ be a further set of primes. If $p^\infty[[\mathcal{L}:K]]$, then

$$\lim_{\mathcal{L}/L/K.\operatorname{res}} \operatorname{III}^2(K_T/L, T; \mathbb{Z}/p\mathbb{Z}) = 0.$$  

**Proof.** By Poitou–Tate duality this is equivalent to

$$\lim_{\mathcal{L}/L/K.\operatorname{cor}^\vee} \operatorname{III}^1(K_T/L, T; \mu_p)^\vee = 0.$$  

This follows in the same way as in the proof of Theorem 5.9.  

**5D. Consequences.** Here we prove Theorem 5.1.

**Lemma 5.12.** Let $S \supseteq R$ be sets of primes of $K$. Assume that $R$ is finite and that $S \cap \text{cs}(K(\mu_p)/K)$ is infinite. Then $p^\infty[[K^R_S(p):K]].$

**Proof.** By [NSW 2008, Corollary 10.7.7], for any $C > 0$ there is some finite subset $S_C \subseteq S \cap \text{cs}(K(\mu_p)/K)$ such that $R \subseteq S_C$ and

$$\dim_{p} H^1(G^R_{K,S_C}(p), \mathbb{Z}/p\mathbb{Z}) > C.$$  

Since each group $G^R_{K,S_C}(p)$ is a quotient of $G^R_{K,S}(p)$, the lemma follows.  

**Proof of Theorem 5.1.** (A$_p$), (A): Let $p$ be a prime of $K$ which is not contained in $R$. Since the local group $\mathfrak{g}_p(p)$ is solvable and the assumptions carry over to extensions of $K$ in $K^R_S(p)$, it is enough to show that any class $\alpha_p \in H^1(\mathfrak{g}_p(p), \mathbb{Z}/p\mathbb{Z})$ (which has to be unramified if $p \not\in S$) is realized by a global class after a finite extension. Define $T := \{p\} \cup R \cup S_p \cup S_\infty$ and let $(\alpha_q) \in \prod_{q \in T} H^1(\mathfrak{g}_p(p), \mathbb{Z}/p\mathbb{Z})$ such that $\alpha_q$ is unramified if $q \not\in S$ and 0 if $p \in R$. By Theorem 5.9, there is some finite extension $K^R_S(p)/L/K$ such that $(\alpha_q)$ comes from a global class $\alpha \in H^1(G^R_{L,S\cup T}(p), \mathbb{Z}/p\mathbb{Z})$.  

The \( \mathbb{Z}/p\mathbb{Z} \)-extension of \( L \) corresponding to \( \alpha \) is unramified outside \( S \), completely split in \( R \) and hence contained in \( K_S^R(p) \). (A) has analogous proof.

(B\( _p \)): The proof of this part essentially coincides with the proofs of [NSW 2008, Theorem 10.5.8] and [Ivanov 2013, Theorem 4.26]. As done there, we can restrict ourselves to the case \( T \supseteq S_p \cup S_\infty \). All cohomology groups in the proof have \( \mathbb{Z}/p\mathbb{Z} \)-coefficients and we omit them from the notation. After computing the cohomology on the left side, by [NSW 2008, Proposition 1.6.15] we have to show that the map

\[
H^i(\phi^R_{T,S}(p)) : H^i(K(p)/K^R_S(p)) \to \bigoplus_{p \in R(K^R_S(p))} H^i(\mathfrak{g}_p(p)) \oplus \bigoplus_{p \in (T \setminus S)(K^R_S(p))} H^i(I_p(p))
\]

induced by \( \phi^R_{T,S}(p) \) in the cohomology is bijective for \( i = 1 \) and injective for \( i = 2 \). (Here \( \bigoplus \) means the restricted direct sum in the sense of [NSW 2008, Definition 4.3.13].) Now \( H^1(\phi^R_{T,S}(p)) \) is injective since \( \phi^R_{T,S}(p) \) is clearly injective. To show surjectivity for \( i = 1 \), consider, for any finite subset \( T_1 \subset T \setminus S \) which contains \((S_p \cup S_\infty) \setminus S \) and any finite \( K^R_S(p)/L/K \), the composed maps

\[
H^1(K_{S \cup T_1}(p)/L) \twoheadrightarrow \bigoplus_{p \in R(U_{T_1}(L))} H^1(\mathfrak{g}_p \leq T_1(p)/L) \oplus \bigoplus_{p \in T_1(L)} H^1(I_p(p)),
\]

where \( \mathfrak{g}_p = I_{K_p}/L_p \subseteq G_{K_p}/L_p = \mathfrak{g}_p \) is the inertia subgroup. Passing to the direct limit over \( K^R_S(p)/L/K \), we obtain by Theorem 5.9 the surjection

\[
H^1(K_{S \cup T_1}(p)/K^R_S(p)) \twoheadrightarrow \bigoplus_{p \in R(K^R_S(p))} H^1(\mathfrak{g}_p(p)) \oplus \bigoplus_{p \in T_1(K^R_S(p))} H^1(I_{K_p}/K_p)^{G_{K_p}/K^R_S(p)}(p),
\]

which is, after passing to the direct limit over all finite \( T_1 \subset T \setminus S \), exactly \( H^1(\phi^R_{T,S}(p)) \), since by (A\( _p \)) we have \( K^R_S(p) = K^R_{\infty}(p) \) for \( p \in T \setminus S \) and hence

\[
H^1(I_{K_p}/K_p)^{G_{K_p}/K^R_S(p)}(p) = H^1(I_p(p))
\]

(see the proofs of [NSW 2008, Theorem 10.5.8] and [Ivanov 2013, Theorem 4.26]). Finally, the injectivity of \( H^2(\phi^R_{T,S}(p)) \) follows by passing to the limit and using Proposition 5.11.

(B): By Lemma 2.7, there is some \( K^R_S/L_0/K \) such that, for all \( K^R_S/L/L_0 \), the set \( S \) is sharply \( p \)-stable for \( L^S_S(p)/L \). Thus (B) follows from (B\( _p \)) as we have

\[
I'_{p}(p) = \lim_{\rightarrow} I_{L_0(p)/L_p} \quad \text{and} \quad G_{K^r_T(p)/K^R_S} = \lim_{\rightarrow} G_{L_T(p)/L^S_S(p)}.
\]

(C\( _p \), (C)): The proof essentially coincides with the proofs of [NSW 2008, Theorem 10.5.10 and Corollary 10.5.11] and [Ivanov 2013, Theorem 4.31, Corollary 4.33]. To avoid many repetitions, we only recall the argument for \( \text{cd} G_{K^r_S(p)} \leq 2 \) in the case \( R = \emptyset \) (which differs in one aspect from the cited proofs). Therefore, set
$V = (S_p \cup S_\infty) \setminus S$ and consider the Hochschild–Serre spectral sequence $(E^{ij}_n, \delta^{ij}_n)$ for the Galois groups of the global extensions $K_{S \cup V}(p)/K_S(p)/K$. By [NSW 2008, Proposition 8.3.18 and Corollary 10.4.8], we have

$$\cd G_{K, S \cup V}(p) \leq \cd_p G_{K, S \cup V} \leq 2.$$ 

By Riemann’s existence theorem, $(B_p)$, the group $G_{K_{S \cup V}(p)/K_S(p)}$ is a free pro-$p$ group. Hence $E^{ij}_n$ degenerates in the second tableau and, in particular, we have (omitting $\mathbb{Z}/p\mathbb{Z}$-coefficients from the notation)

$$\coker(\delta^{11}_2) = E^{30}_3 = E^{30}_\infty \subseteq H^3(G_{K, S \cup V}(p)) = 0,$$

i.e., $\delta^{11}_2$ is surjective. Again by Riemann’s existence theorem we have

$$H^1(K_{S \cup V}(p)/K_S(p)) \cong \bigoplus_{p \in V} \text{Ind}_{D_{p, K_S(p)/K}}^G(H^1(I_p(p))).$$

This and Shapiro’s lemma imply

$$E^{11}_2 = \bigoplus_{p \in V} H^2(K_p(p)/K_p). \quad (5-2)$$

Further, we have the following commutative diagram with exact rows and columns:

$$\begin{array}{ccc}
\bigoplus_{p \in S} H^2(K_p(p)/K_p) & \rightarrow & H^0(K_{S \cup V}/K, \mu_p)^\vee \\
\downarrow & & \downarrow = \\
H^2(K_{S \cup V}(p)/K) & \rightarrow & \bigoplus_{p \in S \cup V} H^2(K_p(p)/K_p) \rightarrow H^0(K_{S \cup V}/K, \mu_p)^\vee \\
\downarrow & & \downarrow & & \downarrow \\
H^1(K_S(p)/K, H^1(K_{S \cup V}(p)/K_S(p))) & \rightarrow & \bigoplus_{p \in V} H^2(K_p(p)/K_p) & \rightarrow & 0 \\
\downarrow & & \downarrow \delta^{11}_2 & & \downarrow \\
H^3(K_S(p)/K)
\end{array}$$

The second row comes from the Poitou–Tate long exact sequence. The first map in the third row is the isomorphism (5-2). The map in the first row is surjective since its dual map $\mu_p(K) \rightarrow \bigoplus_{p \in S} \mu_p(K_p)$ is injective. Now (in contrast to proofs cited from [NSW 2008] and [Ivanov 2013]) the first map in the second row is not necessarily injective, but one can simply replace the first entry in the second row by $H^2(K_{S \cup V}(p)/K)/\Pi^2(K_{S \cup V}/K, S \cup V; \mathbb{Z}/p\mathbb{Z})$, as both maps in the diagram which start at this entry factor through this quotient. Now apply the snake lemma to the second and third row and obtain $H^3(K_S(p)/K) = 0$ and hence also $\cd G_{K, S}(p) \leq 2$ by [NSW 2008, Proposition 3.3.2].
Vanishing of $\Pi^2(G_S; \mathbb{Z}/p\mathbb{Z})$ without $p \in \mathcal{O}_{K,S}^*$. We generalize Corollary 4.7 for the constant module. The proof makes use of Theorem 5.1 parts (A), (B), (C) along with the result of Neumann showing the vanishing of certain cohomology groups. Its special case $\delta_K(S) = 1$ is not contained in [NSW 2008].

**Proposition 5.13.** Let $K$ be a number field, $S$ a set of primes of $K$. Let $p$ be a rational prime, $r > 0$ an integer. Assume that either $p$ is odd or $K_S$ is totally imaginary. Then the following hold:

(i) [Ivanov 2013, Proposition 4.34] Assume $S$ is strongly $p$-stable and $L_0$ is a $p$-stabilizing field for $S$ for $K_{S∪S_p∪S_∞}/K$. Assume $p$ is odd or $L_0$ is totally imaginary. Then

$$\Pi^2(K_S/L; \mathbb{Z}/p^r\mathbb{Z}) = 0$$

for any finite $K_S/L/L_0$ such that we are not in the special case $(L, p^r, S)$.

(ii) Let $K_S/L/K$ be a normal subextension. Assume that $S$ is sharply $p$-stable for $L/K$ and $p^∞|[L : K]$. Then

$$\lim_{\mathcal{L}/L/K} \Pi^2(K_S/L; \mathbb{Z}/p^r\mathbb{Z}) = 0.$$  

**Proof.** Define $V := (S_p ∪ S_∞) \setminus S$. We write $H^*(\cdot)$ instead of $H^*(\cdot, \mathbb{Z}/p^r\mathbb{Z})$ and $\Pi^*(\cdot, \cdot)$ instead of $\Pi^*(\cdot, \cdot; \mathbb{Z}/p^r\mathbb{Z})$. Let $K'_{S∪V}(p)$ be the maximal pro-$p$ subextension of $K_{S∪V}/K_S$. Let $K_S/L/K$ be a finite subextension and consider the tower of extensions

\[
\begin{tikzcd}
& K_{S∪V} \arrow{dr}{H} \arrow{d}{N} \\
K_{S∪V}'(p) \arrow{dr}{H} \arrow{d}{G_{L,S∪V}} & K_S \arrow{dr}{H} \arrow{d}{G_{L,S}} \\
L & G_{L,S∪V}'(p) \arrow{d}{G_{L,S∪V}}
\end{tikzcd}
\]

with $N := G_{K_{S∪V}/K_{S∪V}'(p)}$, $H := G_{K_{S∪V}'(p)/K_S}$ and $G_{L,S∪V}'(p) := G_{K_{S∪V}'(p)/L}$. We claim that for any such $L$ we have under the assumptions of (i) the natural isomorphisms

$$\Pi^2(K_{S∪V}'(p)/L, S∪V) = \Pi^2(K_{S∪V}/L, S∪V) \quad \text{for any } K_S/L/K, $$

$$\Pi^2(K_{S}/L, S) = \Pi^2(K_{S∪V}'(p)/L, S∪V) \quad \text{for any } K_S/L/L_0, $$  

and under (ii) the natural isomorphism

$$\lim_{\mathcal{L}/L/K} \Pi^2(K_{S}/L, S) = \lim_{\mathcal{L}/L/K} \Pi^2(K_{S∪V}'(p)/L, S∪V).$$
Once this claim is shown, (i) follows immediately from Corollary 4.7 and (ii) follows from Proposition 5.11. Thus it is enough to prove the above claim. The first isomorphism in (5-3) follows immediately from the definition of $\text{III}^2$, once we know that the inflation map $H^2(G'_{L,SUV}(p)) \to H^2(G_{L,SUV})$ is an isomorphism. To show this last assertion, consider the Hochschild–Serre spectral sequence

$$E_2^{ij} = H^i(G'_{L,SUV}(p), H^j(N)) \Rightarrow H^{i+j}(G_{L,SUV}).$$

A result of Neumann [NSW 2008, Theorem 10.4.2] applied to $K_{SUV}/K'_{SUV}(p)$ (the upper field is $p$-(S $\cup$ V)-closed, the lower is $p$-$(S_p \cup S_\infty)$-closed) implies $E_2^{ij} = 0$ for $j > 0$. Hence the sequence degenerates in the second tableau and

$$H^i(G'_{SUV}(p)) = H^i(G_{SUV}),$$

for $i \geq 0$, proving our claim. Thus we are reduced to showing that the second map in (5-3) and the map in (5-4) are isomorphisms. For $p \in V$, let $K'_p(p)$ denote the maximal pro-$p$ extension of $K_{S,p}$. Define

$$I'_p(p) := G_{K'_p(p)/K_{S,p}}.$$  

(Observe that if $p \in S_\infty$, then $I'_p(p) = 1$. Indeed, if $p > 2$, this is always the case, and if $p = 2$, then $K_{S,p} = \mathbb{C}$ using the assumption that $K_S$ is totally imaginary.) By [Ivanov 2013, Lemma 4.23] (which was only shown there under strong $p$-stability assumption on $S$, but due to Theorem 5.1(A) it also holds under sharp $p$-stability assumption with exactly the same proof), we have $I'_p(p) = D_{p,K'_{SUV}(p)/K_{S}}$. Next, by Riemann’s existence theorem, Theorem 5.1(B), applied to $K'_{SUV}(p)/K_{S}/K$, we have

$$H \cong \bigoplus_{p \in V(K_S)} I'_p(p).$$

By [Ivanov 2013, Corollary 4.24], the groups $I'_p(p)$ are free pro-$p$ groups, and hence $H$ is a free pro-$p$ group. Thus $\text{cd}_p H \leq 1$. Consider the exact sequence

$$1 \to H \to G'_{L,SUV}(p) \to G_{L,S} \to 1$$

and the corresponding Hochschild–Serre spectral sequence

$$E_2^{ij} = H^i(G_{L,S}, H^j(H)) \Rightarrow H^{i+j}(G'_{L,SUV}(p)).$$

Since by Theorem 5.1(C) we know that $\text{cd}_p G_{L,S} = 2$, we have $E_2^{ij} = 0$ if $i > 2$ or $j > 1$. Moreover, we have

$$H^1(H) = \bigoplus_{V(K_S)} H^1(I'_p(p)) = \bigoplus_{V(L)} \text{Ind}^{G_{L,S}}_{D_{p,K_S/L}} H^1(I'_p(p))$$

as $G_{L,S}$-modules, where $D_{p,K_S/L} \subseteq G_{L,S}$ is the decomposition group at $p$, which is in particular procyclic and has an infinite $p$-Sylow subgroup (by Theorem 5.1(A)). Using this, an easy computation involving Frobenius reciprocity, Shapiro’s lemma
and [Ivanov 2013, Lemma 4.24] allows us to compute the terms \( E_2^{01} \) and \( E_2^{11} \). We obtain the exact sequence

\[
0 \longrightarrow H^1(G_{L,S}) \longrightarrow H^1(G_{L,S \cup V}(p)) \longrightarrow \bigoplus_{V(L)} H^1(I_p'(p))^{D_{p,K_S/L}} \longrightarrow \]

\[
\delta \longrightarrow H^2(G_{L,S}) \longrightarrow H^2(G_{L,S \cup V}(p)) \longrightarrow d \bigoplus_{V(L)} H^2(\mathcal{O}_p) \longrightarrow 0,
\]

where \( \delta := \delta_2^{01} : E_2^{01} \rightarrow E_2^{20} \) denotes the differential in the second tableau. Assume first that we are in the situation of (i) and let \( L \) be as introduced there. Then we have the surjections

\[
H^1(G_{L,S \cup V}(p)) \longrightarrow \bigoplus_{p \in V(L)} H^1(\mathcal{O}_p) = \bigoplus_{p \in V(L)} H^1(D_{p,K_{S \cup V}(p)/L}) \longrightarrow \bigoplus_{V(L)} H^1(I_p'(p))^{D_{p,K_S/L}}.
\]

The first map is surjective by Grunwald–Wang (Theorem 5.3), and the second and the third maps follow from [Ivanov 2013, Lemma 4.24]. Hence the map preceding \( \delta \) is surjective and hence \( \delta = 0 \). Thus the lower row of the above 6-term exact sequence gives the short exact sequence

\[
0 \longrightarrow \text{III}^2(K_S/L, S) \longrightarrow \text{III}^2(K_{S \cup V}'(p)/L, S) \longrightarrow \bigoplus_{V(L)} H^2(\mathcal{O}_p).
\]

On the other side, by definition of \( \text{III}^2 \), we have that the kernel of \( d \) is precisely \( \text{III}^2(K_{S \cup V}'(p)/K, S \cup V) \), which shows the second equality in (5-3). The equality in (5-4) follows from the assumptions in (ii) by the same arguments after taking \( \lim \) over \( \mathcal{L}/L/K \) (and using Theorem 5.9 instead of Theorem 5.3).

\[\square\]

### 6. \( K(\pi, 1) \)-property

Assume that either \( p \) is odd or \( K \) is totally imaginary, and let \( X = \text{Spec} \mathcal{O}_K,S \). While it is well known that \( X \) is a \( K(\pi, 1) \) for \( p \) if either \( S \supseteq S_p \cup S_\infty \) (‘wild case’) or \( \delta_K(S) = 1 \), it is a challenging problem to determine whether \( X \) is a \( K(\pi, 1) \) if \( S \) is finite and does not necessarily contain \( S_p \cup S_\infty \). Until recently there were no nontrivial examples of \( (K, S) \) such that \( X \) is a \( K(\pi, 1) \) for \( p \) or a pro-\( p \) \( K(\pi, 1) \) and, say, \( S \cap S_p = \emptyset \). Recent results of Schmidt [2007; 2009; 2010] show that any point of \( \text{Spec} \mathcal{O}_K \) has a basis for Zariski-topology consisting of pro-\( p \) \( K(\pi, 1) \)-schemes. More precisely, given \( K \), a finite set \( S \) of primes of \( K \), a rational prime \( p \) and any set \( T \) of primes of \( K \) of density 1, Schmidt showed that one can find a finite subset \( T_1 \subseteq T \) such that \( X \setminus T_1 \) is pro-\( p \) \( K(\pi, 1) \). The main ingredient in the proof is the theory of mild pro-\( p \) groups, developed by Labute. We conjecture that one can replace the condition \( \delta_K(T) = 1 \) in Schmidt’s work by the weaker condition that \( T \) is strongly \( p \)-stable (or even that \( T \) is sharply \( p \)-stable for \( K_T(p)/K \)).
In the present section we enlarge the set of the examples of such pairs \((K, S)\) for which \(X\) is a \(K(\pi, 1)\) for \(p\) and prove essentially that if \(S\) is sharply \(p\)-stable, then \(X\) is a \(K(\pi, 1)\) for \(p\). In particular, if \(S\) is a stable almost Chebotarev set with \(S_\infty \subseteq S\), then \(X\) is a \(K(\pi, 1)\) for almost all primes \(p\) (see Proposition 3.8 and Example 3.10), and if \(E^{\text{sharp}}(S) = \emptyset\) and \(K\) is totally imaginary, then \(X\) is a \(K(\pi, 1)\).

6A. Generalities on the \(K(\pi, 1)\)-property. There are many equivalent ways to characterize the \(K(\pi, 1)\)-property of schemes (see [Stix 2002, Appendix A], where they are discussed in detail). Without repeating all of it, we want to introduce a small refinement of terminology which is better adapted to formulating our results.

Let \(X\) be a connected scheme, \(X_{\text{et}}\) the étale site on \(X\). Fix a geometric point \(\bar{x} \in X\) and let \(\pi := \pi_1(X, \bar{x})\) be the étale fundamental group of \(X\). Let \(\mathcal{B}\pi\) denote the site of continuous \(\pi\)-sets endowed with the canonical topology. Further, let \(p\) be a rational prime and let \(\mathcal{B}\pi^p\) denote the site of continuous \(\pi^{(p)}\)-sets, where \(\pi^{(p)}\) is the pro-\(p\) completion of \(\pi\). As in [Stix 2002, Appendix A.1], we have natural continuous maps of sites:

\[
\begin{array}{ccc}
X_{\text{et}} & \xrightarrow{\gamma} & \mathcal{B}\pi \\
\downarrow \gamma_p & & \downarrow \gamma_p \\
\mathcal{B}\pi^p & & \\
\end{array}
\]

For a site \(Y\), let \(\mathcal{S}(Y)\) denote the category of sheaves of abelian groups on \(Y\), let \(\mathcal{S}(Y)_f\) be the subcategory of locally constant torsion sheaves, and \(\mathcal{S}(Y)_p\) the subcategory of locally constant \(p\)-primary torsion sheaves. Let \(A \in \mathcal{S}(\mathcal{B}\pi)_f\) and \(B \in \mathcal{S}(\mathcal{B}\pi^p)_p\). Then we have the natural transformations of functors \(\text{id} \rightarrow \mathbb{R}y_*y^*\) and \(\text{id} \rightarrow \mathbb{R}_{Y_p,*}y^*_p\), which induce maps in the cohomology:

\[c^i_A : H^i(\pi, A) \rightarrow H^i(X_{\text{et}}, y^*A) \quad \text{and} \quad c^i_{p,B} : H^i(\pi^{(p)}, B) \rightarrow H^i(X_{\text{et}}, y^*_pB)\]

Let \(\tilde{X}\) (resp. \(\tilde{X}(p)\)) denote the universal (resp. universal pro-\(p\)) covering of \(X\). Since

\[H^1(\tilde{X}_{\text{et}}, A) = H^1(\tilde{X}(p)_{\text{et}}, B) = 0\]

for each \(A, B\), the maps \(c^i_A, c^i_{p,B}\) are isomorphisms for \(i = 0, 1\) and injective for \(i = 2\).

Definition 6.1. Let \(X\) be a connected scheme.

(i) \(X\) is a \(K(\pi, 1)\) if \(c^i_A\) is an isomorphism for all \(A \in \mathcal{S}(\mathcal{B}\pi)_f\) and \(i \geq 0\).

(ii) \(X\) is a \(K(\pi, 1)\) for \(p\) if \(c^i_A\) is an isomorphism for all \(A \in \mathcal{S}(\mathcal{B}\pi)_p\) and \(i \geq 0\).

(iii) \(X\) is a pro-\(p\) \(K(\pi, 1)\) if \(c^i_{p,B}\) is an isomorphism for all \(B \in \mathcal{S}(\mathcal{B}\pi^p)_p\) and \(i \geq 0\).

Note that we use a shift in definitions compared with [Schmidt 2007] or [Wingberg 2007]: what there is called a \(K(\pi, 1)\) for \(p\), we call here a pro-\(p\) \(K(\pi, 1)\). Parts (i) and (iii) of our definition coincide with the definition of a \(K(\pi, 1)\) in [Stix 2002, Definition A.1.2]. By decomposing any sheaf into \(p\)-primary components we obtain:
Lemma 6.2. $X$ is a $K(\pi, 1)$ if and only if it is a $K(\pi, 1)$ for all $p$.

Now we have a criterion for being $K(\pi, 1)$. For a scheme $X$, let $\text{Fet}_X$ (resp. $\text{Fet}_X^{(p)}$) denote the category of all finite étale coverings (resp. finite étale $p$-coverings) of $X$. For a number field $K$, let

$$\delta_K = \begin{cases} 1 & \text{if } \mu_p \subseteq K, \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 6.3. Let $K$ be a number field, $S \supseteq S_\infty$ a set of primes of $K$ such that either $\delta_K = 0$ or $S_f \neq \emptyset$. Assume that either $p$ is odd or $K$ is totally imaginary. Let $X = \text{Spec} \mathcal{C}_{K,S}$. The following are equivalent:

(i) $X$ is a $K(\pi, 1)$ for $p$.
(ii) $\lim_{Y \in \text{Fet}_X} H^2(Y_{\text{ét}}, \mathbb{Z}/p\mathbb{Z}) = 0$.

The same also holds if one replaces ‘$K(\pi, 1)$ for $p$’ by ‘pro-$p$ $K(\pi, 1)$’ and ‘$\text{Fet}_X$’ by ‘$\text{Fet}_X^{(p)}$’, respectively.

Proof. For the full proof, see [Ivanov 2013, Proposition 5.5]. For convenience, we sketch here the main steps. (i) $\Rightarrow$ (ii) holds for any connected scheme and follows from [Stix 2002, Proposition A.3.1] and (ii) $\Rightarrow$ (i) follows from the well-known criterion [Stix 2002, Proposition A.3.1] and the fact that, for every $q > 0$ and every locally constant $p$-primary torsion sheaf $A$ on $X_{\text{ét}}$, we have

$$\lim_{Y \in \text{Fet}_X} H^q(Y_{\text{ét}}, A|_Y) = 0.$$ 

Since $A$ is trivialized on some $Y \in \text{Fet}_X$, we can assume that $A$ is constant. By dévissage we are reduced to the case $A = \mathbb{Z}/p\mathbb{Z}$. The elements of $H^1(Y_{\text{ét}}, \mathbb{Z}/p\mathbb{Z})$ can be interpreted as torsors, which kill themselves, i.e., the case $q = 1$ follows. Further by [Artin et al. 1973, Exposé X, Proposition 6.1], $H^q(Y_{\text{ét}}, \mathbb{Z}/p\mathbb{Z}) = 0$ for $q > 3$. The case $q = 3$ follows from Artin–Verdier duality. Finally, (ii) implies the case $q = 2$. The pro-$p$ case has a similar proof. \qed

6B. $K(\pi, 1)$ and sharp $p$-stability.

Theorem 6.4. Let $K$ be a number field, $S \supseteq S_\infty$ a set of primes of $K$ and $p$ a rational prime. Assume that either $p$ is odd or $K$ is totally imaginary. Then:

(i) If $S$ is sharply $p$-stable for $K_S(p)/K$, then $\text{Spec} \mathcal{C}_{K,S}$ is a pro-$p$ $K(\pi, 1)$.
(ii) If $S$ is sharply $p$-stable, then $\text{Spec} \mathcal{C}_{K,S}$ is a $K(\pi, 1)$ for $p$.

Remark 6.5. If $K$ is totally imaginary or in the pro-$p$ case, the assumption $S_\infty \subseteq S$ is superfluous as $G_S(p) = G_{S \cup S_\infty}(p)$: if $p > 2$, then this is true in general and if $p = 2$, then this is true since we have assumed that $K$ is totally imaginary.
**Corollary 6.6.** Let $K$ be a number field, $S \supseteq S_\infty$ a stable set of primes of $K$ such that $E^{\text{sharp}}(S)$ is finite (in particular, $S$ can be any stable almost Chebotarev set with $S \supseteq S_\infty$). Then $	ext{Spec} \mathcal{O}_{K,S}$ is a $K(\pi, 1)$ for almost all primes $p$. If $E^{\text{sharp}}(S) = \emptyset$ and $K$ is totally imaginary, then $	ext{Spec} \mathcal{O}_{K,S}$ is a $K(\pi, 1)$.

**Example 6.7.** Let $K$ be totally imaginary. Define $\tilde{K} := \bigcup_p K(\mu_p)$. Let $M/K$ be finite Galois with $M \cap \tilde{K} = K$ and let $\sigma \in G_{M/K}$. Assume that $S \supseteq P_{M/K}(\sigma)$ is stable. Then $	ext{Spec} \mathcal{O}_{K,S}$ is a $K(\pi, 1)$.

**Proof of Theorem 6.4.** The proof essentially coincides with that of [Ivanov 2013, Theorem 5.12]. We only prove (ii) (the pro-$p$ case (i) has a similar proof). Define $X := \text{Spec} \mathcal{O}_{K,S}$. As $L$ goes through finite subextensions of $K_\mathfrak{S}/K$, the normalization $Y$ of $X$ in $L$ goes through all finite étale connected coverings of $X$. Define $V := S_p \setminus S$. For any such $Y$ we have a decomposition

$$Y \setminus V \hookrightarrow Y \overset{i}{\hookrightarrow} V$$

in an open and a closed part. Now we see that $Y \setminus V$ is a $K(\pi, 1)$ for $p$ and that $\pi_1(Y \setminus V) = G_{L,S \cup V}$. Hence

$$c_A^i : H^i(G_{L,S \cup V}) \hookrightarrow H^i((Y \setminus V)_{\text{ét}}, A)$$

is an isomorphism for any $i \geq 0$ and any $p$-primary $G_{L,S \cup V}$-module $A$. We have the Lary spectral sequence for $j$:

$$E_2^{mn} = H^m(Y, R^nj_*\mathbb{Z}/p\mathbb{Z}) \Rightarrow H^{m+n}(Y \setminus V, \mathbb{Z}/p\mathbb{Z}).$$

Let us compute the terms in this spectral sequence. First of all we have

$$R^nj_*\mathbb{Z}/p\mathbb{Z} = \begin{cases} \mathbb{Z}/p\mathbb{Z} & \text{if } n = 0, \\ \bigoplus_{p \in V} H^1(\mathfrak{g}_p, \mathbb{Z}/p\mathbb{Z}) & \text{if } n = 1, \\ 0 & \text{if } n > 1, \end{cases}$$

where $\mathfrak{g}_p \subseteq \mathfrak{G}_p$ denotes the inertia subgroup of the full local Galois group at $p$. Thus

$$E_2^{01} = \bigoplus_{p \in V} H^1(\mathfrak{g}_p, \mathbb{Z}/p\mathbb{Z})^{\mathfrak{g}_p^\text{nr}},$$

$$E_2^{11} = H^1(Y_{\text{ét}}, \bigoplus_{p \in V} H^1(\mathfrak{g}_p, \mathbb{Z}/p\mathbb{Z})) = \bigoplus_{p \in V} H^2(\mathfrak{g}_p, \mathbb{Z}/p\mathbb{Z}),$$

and $E_2^{mn} = 0$ if $n > 1$ or if $n = 1$ and $m > 1$ (as $\text{cd}_p(\mathfrak{g}_p^\text{nr}) = 1$). Further, $E_2^{m0} = 0$ for $m > 3$, as $\text{cd}_p Y \leq 3$ and $E_2^{30} = H^3(Y, \mathbb{Z}/p\mathbb{Z}) = 0$ by [Ivanov 2013, Lemma 5.9], and

$$E_2^{10} = H^1(Y_{\text{ét}}, \mathbb{Z}/p\mathbb{Z}) = H^1(G_{L,S}, \mathbb{Z}/p\mathbb{Z}).$$
Thus we have the following nonzero entries in the second tableau:

\[
\begin{array}{ccc}
\bigoplus_{p \in V} H^1(\mathcal{O}_p, \mathbb{Z}/p\mathbb{Z}) & \bigoplus_{p \in V} H^2(\mathcal{O}_p, \mathbb{Z}/p\mathbb{Z}) & 0 \\
\otimes \mathbb{Z}/p\mathbb{Z} & H^1(G_{L,S}, \mathbb{Z}/p\mathbb{Z}) & H^2(Y_{\text{et}}, \mathbb{Z}/p\mathbb{Z})
\end{array}
\]

From this and the isomorphism (6-1) we obtain the following exact sequence (from now on, we omit the \(\mathbb{Z}/p\mathbb{Z}\)-coefficients):

\[
\begin{array}{c}
0 \longrightarrow H^1(G_{L,S}) \longrightarrow H^1(G_{L,S\cup V}) \longrightarrow \bigoplus_{p \in V} H^1(\mathcal{O}_p, \mathbb{Z}/p\mathbb{Z}) \xrightarrow{\delta^0_1} \\
\longrightarrow H^2(Y_{\text{et}}) \longrightarrow H^2(G_{L,S\cup V}) \longrightarrow \bigoplus_{p \in V} H^2(\mathcal{O}_p) \longrightarrow 0
\end{array}
\]

By Proposition 6.3 it is enough to show that \(\lim_{Y \in \text{Fet}_X} H^2(Y_{\text{et}}) = 0\). Taking the limit over all \(Y \in \text{Fet}_X\) of this sequence, we see by Theorem 5.9 that the direct limit of the maps preceding \(\delta^0_1\) is surjective, hence we obtain

\[
\lim_{Y \in \text{Fet}_X} H^2(Y_{\text{et}}) \cong \lim_{Y \in \text{Fet}_X} \Pi^2(K_{S\cup V}/L, V; \mathbb{Z}/p\mathbb{Z}).
\]

To finish the proof consider the following commutative diagram with exact rows:

\[
\begin{array}{c}
0 \longrightarrow \bigoplus_{p \in V} H^2(\mathcal{O}_p) \longrightarrow H^2(G_{L,S\cup V}) \xrightarrow{\mu_p(L)^\vee} 0 \\
0 \longrightarrow \bigoplus_{p \in V} H^2(\mathcal{O}_p) \longrightarrow \bigoplus_{p \in V} H^2(\mathcal{O}_p) \longrightarrow 0 \longrightarrow 0
\end{array}
\]

Here the first map in the upper row becomes injective after taking the limit by Proposition 5.11. The snake lemma shows that

\[
\begin{array}{c}
\lim_{Y \in \text{Fet}_X} H^2(Y_{\text{et}}) \cong \lim_{Y \in \text{Fet}_X} \Pi^2(K_{S\cup V}/L, V; \mathbb{Z}/p\mathbb{Z}) \subseteq \lim_{Y \in \text{Fet}_X} \bigoplus_{p \in S} H^2(\mathcal{O}_p),
\end{array}
\]

and the last limit vanishes as \(p^{\infty} || K_{S:p} : K_p\) for all \(p \in S\) by Theorem 5.1(A). This finishes the proof of (ii).

\[\square\]

Acknowledgements

Some of the results in this article coincide with the author’s Ph.D. thesis [Ivanov 2013], written under the supervision of Jakob Stix at the University of Heidelberg. The author is grateful to him for the very good supervision, and to Kay Wingberg,
Johannes Schmidt and many other people for helpful remarks and interesting discussions. The work on the author’s thesis was partially supported by the Mathematics Center Heidelberg and the Mathematical Institute Heidelberg. The author thanks both institutions for their hospitality and the excellent working conditions.

References


Communicated by Bjorn Poonen
Received 2014-06-23 Revised 2015-09-07 Accepted 2015-10-23

ivanov@ma.tum.de Zentrum Mathematik, Technischen Universität, Boltzmannstraße 3, D-85747 Garching bei München, Germany
Hopf–Galois structures arising from groups with unique subgroup of order $p$

Timothy Kohl

For $\Gamma$ a group of order $mp$, where $p$ is a prime with $\gcd(p, m) = 1$, we consider the regular subgroups $N \leq \text{Perm}(\Gamma)$ that are normalized by $\lambda(\Gamma)$, the left regular representation of $\Gamma$. These subgroups are in one-to-one correspondence with the Hopf–Galois structures on separable field extensions $L/K$ with $\Gamma = \text{Gal}(L/K)$. Elsewhere we showed that if $p > m$ then all such $N$ lie within the normalizer of the Sylow $p$-subgroup of $\lambda(\Gamma)$. Here we show that one only need assume that all groups of a given order $mp$ have a unique Sylow $p$-subgroup, and that $p$ not be a divisor of the order of the automorphism groups of any group of order $m$. We thus extend the applicability of the program for computing these regular subgroups $N$ and concordantly the corresponding Hopf–Galois structures on separable extensions of degree $mp$.

Introduction

The motivation and antecedents for this work lie in the subject of Hopf–Galois theory for separable field extensions. Specifically, we extend the results of [Kohl 2013] on Hopf–Galois structures on Galois extensions of degree $mp$ for $p$ a prime where $p > m$. We will not delve into all the particulars of Hopf–Galois theory, since this discussion focuses on the group-theoretic underpinnings of this class of examples. For the general theory, one may consult [Chase and Sweedler 1969] for basic definitions and initial examples and [Greither and Pareigis 1987] for the theory as applied to separable extensions, which is the category in which our earlier paper and others fall. In brief, let $L/K$ be a finite Galois extension with $\Gamma = \text{Gal}(L/K)$. Such an extension is canonically Hopf–Galois for the $K$-Hopf algebra $H = K[\Gamma]$, but also for potentially many other $K$-Hopf algebras. Their enumeration is determined by the following paraphrased variant of the main theorem in [Greither and Pareigis 1987]. (Recall that a regular subgroup of the group of permutations of a set is one whose action on the set is transitive and free.)

MSC2010: primary 20B35; secondary 20D20, 20D45, 16T05.
Keywords: Hopf–Galois extension, regular subgroup.
Theorem [Greither and Pareigis 1987]. If $L/K$ is a finite Galois extension with $\Gamma = \text{Gal}(L/K)$ then the Hopf algebras which act to make the extension Hopf–Galois correspond in a one-to-one fashion with regular subgroups $N \leq B = \text{Perm}(\Gamma)$ such that $\lambda(\Gamma) \leq \text{Norm}_B(N)$, where $\lambda(\Gamma)$ is the image of the left regular representation of $\Gamma$ in $B$.

Each such $N$ gives rise to the Hopf algebra $H = L[N]^\Gamma$, the fixed ring of the group ring $L[N]$ under the action of $\Gamma$ simultaneously on the coefficients, by virtue of $\Gamma = \text{Gal}(L/K)$, and the group elements by virtue of $\lambda(\Gamma)$ normalizing $N$. The problem of enumerating such $N$ for different classes of extensions has been the subject of much recent work by Byott (e.g., [2004]), Childs (e.g., [2003]), the author, and others. This discussion will be strictly focused on the enumeration of these groups, for a particular set of cases, keyed to the order of $\Gamma$ and consequently of any such regular subgroup $N$ which satisfies the conditions of the above theorem. Again, no discussion of Hopf–Galois structures is required from here on since we are looking at the purely group-theoretic translation arising from the above theorem.

1. Preliminaries

As it is so essential, let’s briefly review regularity in the context of finite groups.

Definition 1.1. Let $X$ be a finite set and let $N \leq \text{Perm}(X)$ be a group of permutations of $X$. We say that $N$ is semiregular if it acts freely, that is, if each element of $N$ apart from the identity acts without fixed points. If $N$ acts transitively on $X$ and $|N| = |X|$ then $N$ is semiregular; and if $N$ is semiregular its action is transitive if and only if $|N| = |X|$. Thus any two of these conditions imply the third. A group satisfying these conditions is called regular.

In view of the cardinality condition, in order to organize the enumeration of the regular $N \leq \text{Perm}(\Gamma)$ that may arise for a given Galois group $\Gamma$, we consider, for $[M]$ the isomorphism class of a given group of the same order as $\Gamma$, the set

$$R(\Gamma, [M]) = \{N \leq B \mid N \text{ regular, } N \cong M, \lambda(\Gamma) \leq \text{Norm}_B(N)\}$$

and let $R(\Gamma)$ be the union of the $R(\Gamma, [M])$ over all isomorphism classes of groups of the same order as $\Gamma$.

Since they are important in the enumeration of $R(\Gamma, [M])$, we recall some facts from [Kohl 2013], henceforth referred to as [K].

Lemma 1.2 [K, Corollary 3.3 and Proposition 3.4]. Let $N$ be a regular subgroup of $B$ and define the opposite of $N$ as $N^{\text{opp}} = \text{Cent}_B(N)$. Then:

- $N^{\text{opp}}$ is also a regular subgroup of $B$.
- $N \cap N^{\text{opp}} = Z(N)$ (so $N = N^{\text{opp}}$ if and only if $N$ is abelian).
- \((N_{opp})_{opp} = N\).
- \(N \in R(\Gamma, [M])\) if and only if \(N_{opp} \in R(\Gamma, [M])\).

Again, in the cases considered in [K] it was assumed that \(|\Gamma| = mp\) for \(p\) prime and \(p > m\). Our goal is to extend those results to groups of order \(mp\) where \(gcd(p, m) = 1\), but where one need not assume that \(p > m\).

We begin by briefly reviewing the setup in [K], where we considered groups \(\Gamma\) of order \(mp\) with a unique and therefore characteristic Sylow \(p\)-subgroup due to the assumption that \(p > m\). Since \(p > m\) obviously implies \(gcd(p, m) = 1\), by the Schur–Zassenhaus lemma \(\Gamma\) may be written as \(PQ\) for \(P\) and \(Q\) subgroups of \(\Gamma\) where \(|P| = p\) and \(|Q| = m\). More specifically, there is a split exact sequence \(P \to \Gamma \to Q\) whereby \(\Gamma = P \rtimes Q\), with \(\tau : Q \to Aut(P)\) induced by conjugation within \(\Gamma\) by the complementary subgroup \(Q\). Using \(Q\) for the quotient of \(\Gamma\) by \(P\) and the image of the section in \(\Gamma\) is admittedly a slight abuse of notation. The condition \(p > m\) is sufficient, of course, to make the Sylow \(p\)-subgroup unique and have order \(p\).

Going forward, we wish to drop the assumption that \(p > m\) and consider groups of order \(mp\) for \(p\) prime, with \(gcd(p, m) = 1\) and where congruence conditions force any group of order \(mp\), including \(\Gamma\) and any \(N \in R(\Gamma)\), to have a unique Sylow \(p\)-subgroup.

With \(\Gamma\) as above, if \(\lambda : \Gamma \to Perm(\Gamma) = B\) is the left regular representation then we define \(\mathcal{P} = P(\lambda(\Gamma))\) to be the Sylow \(p\)-subgroup of \(\lambda(\Gamma)\) and \(Q\) to be the complementary subgroup to \(\mathcal{P}\) in \(\lambda(\Gamma)\). We wish to prove the following strengthening of [K, Theorem 3.5] which will allow us to apply the program developed in Sections 1–3 of [K] (and applied in subsequent sections therein) to a much larger class of groups.

**Theorem 1.3.** Let \(\Gamma\) have order \(mp\) where \(gcd(p, m) = 1\) and \(p \nmid |Aut(Q)|\) for any group \(Q\) of order \(m\), and assume any group of order \(mp\) has a unique Sylow \(p\)-subgroup. If \(N \in R(\Gamma)\) then \(N\) is a subgroup of \(Norm_B(\mathcal{P})\).

To prove this, we need to modify certain key results from [K], starting with Lemma 1.1 regarding the \(p\)-torsion of \(Aut(\Gamma)\).

**Lemma 1.4.** Let \(\Gamma\) have order \(mp\), where \(gcd(p, m) = 1\), and assume \(\Gamma\) has a unique Sylow \(p\)-subgroup, so that \(\Gamma \cong P \rtimes Q\) as above. Assume also that \(p\) does not divide \(|Aut(Q)|\).

(a) If \(\tau\) is trivial (whence \(\Gamma \cong P \times Q\)) then \(p\) does not divide \(|Aut(\Gamma)|\).

(b) If \(\tau\) is nontrivial then \(Aut(\Gamma)\) has a unique Sylow \(p\)-subgroup, consisting of the inner automorphisms induced by conjugation by elements of \(P\).

**Proof.** In (a), if \(\Gamma\) is such a direct product then \(Aut(\Gamma) = Aut(P) \times Aut(Q)\) and so if \(p \nmid |Aut(Q)|\) then \(p \nmid |Aut(\Gamma)|\). The proof is basically the same as in [K]. For (b),
since $\Gamma / P \cong Q$, any $\psi \in \text{Aut}(\Gamma)$ induces $\bar{\psi} \in \text{Aut}(\Gamma / P) \cong \text{Aut}(Q)$ and if $\psi$ has $p$-power order then $\bar{\psi}$ is trivial since $p$ does not divide $|\text{Aut}(Q)|$. And, as also observed in [K], when $\Gamma$ is not a direct product, $|P \cap Z(\Gamma)| = 1$ and conjugation in $\Gamma$ by elements of $P$ yields the order-$p$ subgroup of $\text{Aut}(\Gamma)$. \hfill \square

The condition that $p \mid |\text{Aut}(Q)|$ was automatic when $p > m$, but it often holds true even when $p < m$. For example, if $p = 5$ and $m = 8$ then Sylow theory easily shows that any group of order 40 will have a unique Sylow $p$-subgroup. One could also consider the groups of order eight — $C_8$, $C_4 \times C_2$, $C_2 \times C_2 \times C_2$, $D_4$ and $Q_2$ — whose automorphism groups have orders 4, 8, 168, 8, 24, respectively, none of which is divisible by 5.

The cycle structure of a regular permutation group’s elements is greatly circumscribed by the condition that all nontrivial elements of the group act freely. Any such element, because it and all its nontrivial powers lack fixed points, must be a product of cycles of the same length, and the sum of the lengths must equal $|X|$. For example, if $X = \{1, 2, 3, 4, 5, 6\}$ then $(1, 2)(3, 4)(5, 6)$ and $(1, 2, 3)(4, 5, 6)$ satisfy this property. In contrast, $\mu = (1, 2, 3, 4)(5, 6)$ cannot belong to a regular subgroup of $\text{Perm}(X)$ even though it does not have fixed points, because $\mu^2 = (1, 3)(2, 4)$.

Since $\mathcal{P}$ is a nontrivial subgroup of the canonically regular permutation group $\lambda(\Gamma)$, we must have

$$\mathcal{P} = \langle \pi \rangle = \langle \pi_1 \pi_2 \cdots \pi_m \rangle,$$

where the $\pi_i$ are disjoint $p$-cycles. In a similar fashion, if $N$ is any regular subgroup of $B$ then its Sylow $p$-subgroup $P(N)$ is also cyclic of order $p$ and therefore of the form

$$P(N) = \langle \theta \rangle = \langle \theta_1 \theta_2 \cdots \theta_m \rangle,$$

where the $\theta_i$ are also disjoint $p$-cycles. For $N \in R(\Gamma)$ we are looking at those regular $N$ which are normalized by $\lambda(\Gamma)$. Now, $P(N)$ is characteristic in $N$, so $\lambda(\Gamma)$ normalizing $N$ implies that $\lambda(\Gamma)$ (and therefore $\mathcal{P}$) normalizes $P(N)$. In [K] the assumption $p > m$ was used to show that $\mathcal{P}$ and $P(N)$ must, in fact, centralize each other. In particular Proposition 1.2 there showed that if $p > m$ then (after renumbering the $\theta_i$ if necessary) one has $\theta_i = \pi_i^{a_i}$ for $a_i \in \mathbb{F}_p^\times$. The reason for this was that for $p > m$ the group $S_m$ contains no elements of order $p$, and so $\theta$ is a product of the same $\pi_i$ that comprise the generator of $\mathcal{P}$.

As it turns out, this is not automatically true if we just assume that $\gcd(p, m) = 1$. For example, if $p = 5$ and $m = 8$ then in $S_{40}$ let

$$\pi_i = (1 + (i - 1)5, 2 + (i - 1)5, 3 + (i - 1)5, 4 + (i - 1)5, 5 + (i - 1)5)$$

for $i = 1, \ldots, 8$ and let $\theta_j = (j, j + 5, j + 10, j + 15, j + 20)$ for $j = 1, \ldots, 5$ and $\theta_6 = \pi_6$, $\theta_7 = \pi_7$, $\theta_8 = \pi_8$. One may verify that $\pi = \pi_1 \cdots \pi_8$ is centralized by $\theta = \theta_1 \cdots \theta_8$ but $\theta_j$, for $j = 1, \ldots, 5$, is not a power of any $\pi_i$. 
This example shows that the $P(N) \leq N$ being normalized, and thus centralized, by $P$ is insufficient to guarantee that $P(N) \leq \langle \pi_1, \pi_2, \ldots, \pi_m \rangle$. This does not nullify the possibility of the program in [K] being generalized. This example merely shows that $\text{Cent}_B(P)$ contains many semiregular subgroups of order $p$ that are not subgroups of $\langle \pi_1, \pi_2, \ldots, \pi_m \rangle$. However, it turns out that for those $N$ normalized by $\lambda(\Gamma)$, since $P(N)$ is characteristic in $N$ and therefore normalized by $\lambda(\Gamma)$, the possible $P(N)$ that can arise are still contained in $\langle \pi_1, \pi_2, \ldots, \pi_m \rangle$. To arrive at this conclusion, we need to recall some facts about the structure of $\text{Norm}_B(P)$ and $\text{Cent}_B(P)$.

With $P = \langle \pi \rangle = \langle \pi_1 \cdots \pi_m \rangle$, where the $\pi_i$ are disjoint $p$-cycles, we can define $V = \langle \pi_1, \pi_2, \ldots, \pi_m \rangle$, the elementary abelian subgroup of $B$ generated by the $\pi_i$. Also, we can choose $\gamma_i \in \Gamma$ for $i = 1, \ldots, m$ such that $\pi_i = (\gamma_i, \pi(\gamma_i), \ldots, \pi^{p-1}(\gamma_i))$ and if we let $\Pi_i$ be the support of $\pi_i$, then the $\Pi_i$ are, of course, disjoint and their union is $\Gamma$ as a set. Define $S \leq B$ to be those permutations $\alpha$ such that for each $i \in \{1, \ldots, m\}$ there exists a single $j \in \{1, \ldots, m\}$ satisfying $\alpha(\pi^i(\gamma_i)) = \pi^j(\gamma_j)$ for each $t \in \mathbb{Z}_p$. Equivalently, $\alpha$ operates on the blocks $\Pi_i$ as follows:

$$\alpha(\langle \gamma_i, \pi(\gamma_i), \ldots, \pi^{p-1}(\gamma_i) \rangle) = \langle \gamma_j, \pi(\gamma_j), \ldots, \pi^{p-1}(\gamma_j) \rangle.$$ 

It is clear that $S$ is isomorphic to $S_m$ viewed as $\text{Perm}([\Pi_1, \ldots, \Pi_m])$, where $\alpha \in S$ corresponds to a permutation $x \in S_m$ which permutes the $m$ blocks $\Pi_i$ amongst each other. In a similar fashion, we may define another subgroup $U \leq B$ keyed to $\pi$ and the $\pi_i$. For a cyclic group $C = \langle x \rangle$ of order $p$, the automorphisms are given by $x \mapsto x^c$ for $c \in U_p = \mathbb{F}_p^\times = \langle u \rangle$. Within $B$, therefore, since $P$ is cyclic of order $p$, there exists a product $u_1 \cdots u_m$ of $m$ disjoint $(p-1)$-cycles with the property that $u_i \pi_i u_i^{-1} = \pi_i^u$, and $u_i \pi_j u_i^{-1} = \pi_j$ for $j \neq i$. (Note that the support of each $u_i$ is $\Pi_i \sim \{\text{one point}\}$.) Therefore, $(u_1 \cdots u_m) \pi (u_1 \cdots u_m)^{-1} = \pi^u$ and we define $U = \langle u_1 \cdots u_m \rangle$. With these three subgroups of $B$ so defined, we can easily describe $\text{Cent}_B(P)$ and $\text{Norm}_B(P)$ as in [K] by

$$\text{Cent}_B(P) = \forall S \equiv C_p \ltimes S_m \cong C_p^m \rtimes S_m,$$

$$\text{Norm}_B(P) = \forall U \equiv C_p^m \rtimes (\text{Aut}(C_p) \times S_m),$$

where $C_p$ denotes the cyclic group of order $p$ and $S_m$ is the $m$-th symmetric group.

The semidirect product formulation is useful and may be closely connected to the intrinsic (as a subgroup of $B$) description. We may view $V = \langle \pi_1, \ldots, \pi_m \rangle$ naturally as $C_p^m$ but also, more fruitfully, as $\mathbb{F}_p^m$, the dimension-$m$ vector space over $\mathbb{F}_p$, so that we may equate $\pi_1^{a_1} \cdots \pi_m^{a_m}$ with $[a_1, \ldots, a_m]$, a vector in $\mathbb{F}_p^m$. As the group $S$ permutes the $\pi_i$ amongst themselves, we may identify it with permutations $\alpha \in S_m$ acting by coordinate shifts on the vectors $\hat{a} = [a_1, \ldots, a_m]$ and where $u \in U_p$ acts by scalar multiplication. Thus we may represent a typical element of $\text{Norm}_B(P)$ by a triple $(\hat{a}, u^t, \alpha)$, with $\hat{a} \in \mathbb{F}_p^m$, $u \in U_p$ and $\alpha \in S_m$, where (as permutations)
(\hat{a}, u^r, \alpha)(\pi_i^k(\gamma_i)) = \pi_{\alpha(i)}^{ku^r+a_{\alpha(i)}}(\gamma_{\alpha(i)}) \text{ and where the multiplication (and resulting conjugation operations) are defined by}

\begin{align*}
(\hat{a}, u^r, \alpha)(\hat{b}, u^s, \beta) &= (\hat{a} + u^r \alpha(\hat{b}), u^{r+s}, \alpha \beta), \quad (1) \\
(\hat{b}, u^s, \beta)(\hat{a}, u^r, \alpha)(\hat{b}, u^s, \beta)^{-1} &= (\hat{b} + u^s \beta(\hat{a}) - u^r (\beta \alpha \beta^{-1})(\hat{b}), u^r, \beta \alpha \beta^{-1}), \quad (2) \\
(\hat{a}, u^r, \alpha)^n &= \left(\sum_{i=0}^{n-1} u^{ri} \alpha^i(\hat{a}), u^{rn}, \alpha^n\right). \quad (3)
\end{align*}

In this setting the elements of \(\mathcal{V}\) correspond to tuples of the form \((\hat{v}, 1, I)\), where \(I\) is the identity of \(S_m\); in particular \(\pi = \pi_1 \pi_2 \cdots \pi_m = ([1, 1, \ldots, 1], 1, I)\). The elements of \(\text{Cent}_B(\mathcal{P})\) correspond to those tuples where \(r = 0\) (i.e., the middle coordinate is 1), which leads us back to the discussion of \(P(N)\) for \(N\) a regular subgroup of \(B\) normalized by \(\lambda(\Gamma)\). In this situation we have \(P(N) = \{\theta\}\), where \(\theta = (\hat{a}, 1, \alpha)\) has order \(p\) and no fixed points. If \(P(N) \not\subset \mathcal{V}\) then \(\alpha \neq I\), implying (since \(\alpha \in S_m\) with \(m\) coprime to \(p\)) that \(\alpha\) has fixed points in \([1, \ldots, m]\). If \(\alpha(i) = i\) then

\[\theta(\pi_i^k(\gamma_i)) = \pi_{\alpha(i)}^{ku^r+a_{\alpha(i)}}(\gamma_{\alpha(i)}) = \pi_i^{k+a_i}(\gamma_i),\]

which means that \(a_i \neq 0\) and more importantly that \(\theta\) restricted to \(\Pi_i\) equals \(\pi_i^{a_i}\). And for those \(j\) not fixed by \(\alpha\), the restriction of \(\theta\) to \(\Pi_j\) is not a power of \(\pi_j\). That is, \(\theta = \theta_1 \theta_2 \cdots \theta_m\), where \(\theta_i = \pi_i^{a_i}\) for at least one \(i\), and \(\theta_j \not\in \mathcal{V}\) for at least one \(j\). The example given above for \(S_{40}\) is an instance of this; in particular, the fixed-point-free element of order 5 is \(([1, 1, 1, 1, 1, 1, 1, 1], 1, (1, 2, 3, 4, 5))\), which is in \(\text{Cent}_{S_{40}}(\langle \pi_1 \pi_2 \cdots \pi_8 \rangle)\).

The above example motivates the following.

**Definition 1.5.** For \(\theta \in B\) and \(\pi_i\) as above, we say \(\pi_i\) divides \(\theta\) (denoted \(\pi_i \mid \theta\)) if the cycle structure of \(\theta\) contains some nontrivial power of \(\pi_i\). Similarly we write \(\pi_i \nmid \theta\) if no power of \(\pi_i\) is a factor in the cycle structure of \(\theta\).

Observe that \(\pi_i \mid \theta\) if and only if \(\pi_i \mid \theta^e\) for all \(e \in U_p\).

The requirement that \(N\) be normalized by \(\lambda(\Gamma)\), together with the fact that \(P(N)\) is characteristic, means that \(P(N)\) is normalized by \(\lambda(\Gamma)\). The upshot of this is the following recapitulation of [K, Proposition 1.2], which (together with Lemma 1.4) will allow us to deduce our main result, namely that \(N \in R(\Gamma)\) implies \(N \leq \text{Norm}_B(\mathcal{P})\) under weaker hypotheses:

\[p \text{ and } m \text{ are coprime, } p \text{ does not divide } |\text{Aut}(Q)| \text{ for any group } Q \text{ of order } m, \]
\[\text{and any group of order } mp \text{ has a unique Sylow } p\text{-subgroup.} \quad (4)\]

This is the core result, since it guarantees that \(P(N) \leq \mathcal{V} \leq \text{Norm}_B(\mathcal{P})\) for all \(N \in R(\Gamma)\).
Theorem 1.6. If $N$ is a regular subgroup of $B$ normalized by $\lambda(\Gamma)$ and $P(N)$ is its Sylow $p$-subgroup, then $P(N)$ is a semiregular subgroup of $\mathcal{V} = \langle \pi_1, \pi_2, \ldots, \pi_m \rangle$. That is, $P(N) = \langle \pi_1^{a_1} \cdots \pi_m^{a_m} \rangle$, where $a_i \in U_p = \mathbb{F}_p^\times$ for $i = 1, \ldots, m$.

Proof. If $P(N) = \langle \theta \rangle$ is not a subgroup of $\mathcal{V}$ then, as shown above, $\theta = \theta_1 \theta_2 \cdots \theta_m$, where $\pi_i | \theta$ for some $i$ and $\pi_j | \theta$ for some $j \neq i$. In [K, Proposition 3.8] it is shown that if $(\hat{b}, u^s, \beta) \in \text{Norm}_B(P)$ has order coprime to $p$ then the permutation coordinate $\beta$ acts without fixed points. It is important to note that the proof of this is not dependent on whether $p < m$ or $p > m$. Applying this to $\lambda(\Gamma)$, which normalizes $\mathcal{P} = \langle \pi_1 \cdots \pi_m \rangle$, has no fixed points, and contains elements $(\hat{b}, u^s, \beta)$ of order coprime to $p$, we conclude that $\beta \in S_m$ is fixed-point-free. In fact, if $Q$ is the complementary subgroup of $P$ in $\lambda(\Gamma)$ and if $t$ maps $(\hat{b}, u^s, \beta)$ to $\beta$ then $t(Q)$ must be a regular subgroup of $S_m$. The reason is that the elements of $Q$ have order relatively prime to $p$, so $t(Q)$ is a semiregular subgroup of $S_m$; but if $|t(Q)| < m$ there would exist $(\hat{a}, u^r, \alpha)$ and $(\hat{a}', u'^r, \alpha)$ in $Q$, which would imply that $(\hat{a}' - u'^r \hat{a}, u'^r - r, I) \in Q$, which, again by [K, Proposition 3.8], must have order divisible by $p$. But since this element is in $Q$ it must be the identity, whence $r = r'$ and so $\hat{a} = \hat{a}'$. This means that $t(Q)$ is regular; by transitivity we pick an element $g = (\hat{b}, u^s, \beta)$ in $Q$ with $\beta(i) = j$, and then $g([1, 1, \ldots, 1], 1, I)g^{-1} = (u^s \beta([1, 1, \ldots, 1]), 1, I) = (u^s[1, 1, \ldots, 1], 1, I)$, where, in particular,

$$g \pi_1 \pi_2 \cdots \pi_m g^{-1} = \pi_{t(1)}^{u^s} \pi_{t(2)}^{u^s} \cdots \pi_{t(m)}^{u^s}.$$ 

And since $g(\theta_1 \theta_2 \cdots \theta_m)g^{-1} = (g \theta_1 g^{-1})(g \theta_2 g^{-1}) \cdots (g \theta_m g^{-1})$, we have $g \theta_i g^{-1} = g \pi_i a_i g^{-1} = \pi_{t(i)}^{u^s a_i} = \pi_{t(j)}^{u^s a_j}$; therefore $\pi_j | g \theta g^{-1}$. The problem now is that $g \theta g^{-1} = \theta^e$ for some $e \in U_p$ implies that $\pi_j | \theta$, contrary to the assumption that $\pi_j | \theta$. We conclude that any such $\theta$ must, in fact, be a fixed-point-free subgroup of $\mathcal{V}$ and therefore of the form asserted in the statement of the theorem. \[\square\]

With this in place, we can review the remaining foundational elements in [K], which, without serious modifications, imply, under the weaker hypotheses (4), that $N \leq \text{Norm}_B(P)$ for all $N \in R(\Gamma)$. We do this also in order to provide some applications for classes of groups where these weaker hypotheses hold.

We have just shown that $N \in R(\Gamma)$ implies $P(N) \leq \text{Norm}_B(P)$ and that $P(N) = \langle \pi_1^{a_1} \cdots \pi_m^{a_m} \rangle$. Define $Q(N)$ to be the complementary subgroup to $P(N)$ inside $N$. Then, since $Q(N)$ normalizes $P(N)$, we have $q \pi_i^{a_i} q^{-1} = \pi_j^{b_j}$ for $q \in Q(N)$, where the mapping of $i \mapsto j = q(i)$ for $i, j \in \{1, \ldots, m\}$ makes $Q(N)$ (abstractly) a regular subgroup of $S_m$. Let $Q = Q(\lambda(\Gamma))$ be the complementary subgroup to $P$ inside $\lambda(\Gamma)$. If $N \in R(\Gamma)$, then the elements of $\lambda(\Gamma)$ that act nontrivially on $P(N)$ are those in $Q$. If we define $\hat{u}_i = [0, \ldots, 1, \ldots, 0] = \pi_i$, then we have the following result (whose proof does not require that $p < m$) about the possibilities for $P(N)$ for any $N \in R(\Gamma)$. 

Hopf–Galois structures arising from groups with unique subgroup of order $p$
Theorem 1.7 [K, Theorem 2.1]. Any semiregular subgroup of $\mathcal{V}$ of order $p$ that is normalized by $Q$ is generated by

$$\hat{\rho}_\chi = \sum_{\gamma \in Q} \chi(\gamma)\hat{v}_\gamma(1),$$

where $\chi : Q \to U_p = \mathbb{F}_p^*$ is a linear character of $Q$.

This theorem allows one to compute “potential” $P(N)$, one for each such linear character. For example, $\mathcal{P}$ is generated by $[1, 1, \ldots, 1] = \hat{\rho}$, where $i : Q \to \mathbb{F}_p^*$ is the trivial character. All of these order-$p$ elements have the form $(\hat{\rho}_\chi, 1, I)$ in the semidirect product formulation of $\text{Norm}_B(\mathcal{P})$. The remaining component is to show that not only does $N \in R(\Gamma)$ imply $P(N) \leq \text{Norm}_B(\mathcal{P})$ but also that $N$ itself is contained in $\text{Norm}_B(\mathcal{P})$. The assumption $p > m$ in [K] was not actually used in the proof of this main result, but rather in the following lemma:

Lemma 1.8 [K, Lemma 2.2]. Let $\chi_1, \chi_2$ be distinct linear characters of $Q$ in $\mathbb{F}_p^*$. Then $\langle \hat{\rho}_{\chi_1}, \hat{\rho}_{\chi_2} \rangle$ cannot contain $\hat{\rho}$.

A careful reading of the proof of this in [K] shows it is not necessary to assume that $p > m$, but merely gcd$(p, m) = 1$. With this completed, Theorem 1.3, saying that $N \in R(\Gamma)$ implies $N \leq \text{Norm}_B(\mathcal{P})$, follows in the exact same fashion as in [K], since the proof does not hinge on the relationship between $p$ and $m$ beyond the fact that they are relatively prime. It does require that $p$ not divide the order of $\text{Aut}(Q)$, as in Lemma 1.4. This assumption on $|\text{Aut}(Q)|$ is needed to control the size and structure of the Sylow $p$-subgroup of $\text{Norm}_B(N)$. Specifically, it is either cyclic of order $p$ if $P(N)$ is central in $N$, or elementary abelian of order $p^2$ if $P(N)$ is noncentral. Again, in [K] this was automatic from assuming that $p > m$.

The application of this theorem, which is the actual program in [K] (demonstrated in Sections 4 and 5 therein), is based on the observation (in [K, Proposition 3.11]) that any two regular subgroups of $B$ that are isomorphic as abstract groups are, in fact, conjugate subgroups of $B$. That being said, to enumerate $R(\Gamma)$, one can avoid the complications of working with left regular representations and instead:

1. replace $B = \text{Perm}(\Gamma)$ by $S_{mp} = \text{Perm}([1, \ldots, mp])$,
2. choose $\mathcal{P} = \langle \pi_1 \cdots \pi_m \rangle$ where $\pi_i = (1 + (i - 1)p, \ldots, pi)$,
3. determine $Q$ corresponding to each such $\Gamma$ where $\Gamma = \mathcal{P} Q$,
4. embed the $\Gamma$ as subgroups of the semidirect product formulation of $\text{Norm}_{S_{mp}}(\mathcal{P})$,
5. enumerate the characters $\chi : Q \to \mathbb{F}_p^*$ and concordantly the potential $P(N)$ as $\langle \hat{\rho}_\chi \rangle$ also embedded in $\text{Norm}_{S_{mp}}(\mathcal{P})$,
6. compute the possible $N \in R(\Gamma)$ that may arise, also as subgroups of $\text{Norm}_{S_{mp}}(\mathcal{P})$. 

If one is more interested in the sizes of the different $R(\Gamma, [M])$ one may use the fact (in [K, Theorem 3.5]) that, for each $N \in R(\Gamma, [M])$, if $P(N)$ is central in $N$ then $P(N) = P(N^{\text{opp}})$, and otherwise either $P(N) = P$ or $P(N^{\text{opp}}) = P$. The point is that one can enumerate those $N \in R(\Gamma, [M])$ for which $P(N) = P$ and, depending on whether $P(N)$ is central, use the above fact to infer the count of those for $N$ for which $P(N) \neq P$ (if any). The virtue of this is that one need not calculate the characters of $Q$ in $\mathbb{F}_p^*$, nor the resulting potential $P(N)$.

We shall demonstrate applications of this program, where we now have a wider class of examples to choose from, based upon the conditions on $p$, $m$ and $|\text{Aut}(Q)|$ as discussed above.

2. Groups of order $p_1p_2p_3$

To be slightly formal, if $n_p$ denotes the number of Sylow $p$-subgroups of a group, we define the following subsets of $\mathbb{N} \times \mathbb{N}$:

$$F_0 = \{ (p, m) \mid p \text{ prime}, \gcd(p, m) = 1, p^{|\text{Aut}(Q)|} \text{ for all groups } Q \text{ of order } m \},$$

$$F_3 = \{ (p, m) \mid p \text{ prime}, \gcd(p, m) = 1, n_p = 1 \text{ for all groups of order } mp \}.$$

The program in [K] for enumerating Hopf–Galois structures on Galois extensions of order $mp$ may be used for those $(p, m) \in F_0 \cap F_3$. As in [K], $(p, m) \in F_0 \cap F_3$ for $p$ prime when $p > m$, but we want to now consider other $p$ and $m$. The case of $p = 5$ and $m = 8$ as indicated already is one such example. In lieu of working out the enumeration of all the $14^2$ possible pairings $R(\Gamma, [M])$ for order 40, we shall instead conclude with an overview of some (classes of) choices for $|\Gamma| = |N| = n = pm$ which force $(p, m) \in F_0 \cap F_3$. Such forcing conditions have appeared in group theory literature including recent examples such as [Pakianathan and Shankar 2000].

Our example will be somewhat more narrow, but is in this same spirit.

If $p_1 < p_2 < p_3$ are primes, then by Sylow theory $n_{p_1} \equiv 1 \pmod{p_1}$ and $n_{p_3} \mid p_1p_2$. However, $n_{p_1} \neq 1$ and $n_{p_3} \neq p_2$ since $p_3$ is larger than $p_1$ and $p_2$. If $n_{p_3} = p_1p_2$ then one must have $p_1p_2(p_3 - 1)$ elements of order $p_3$ and so, by necessity, $n_{p_2} = 1$. Thus $\Gamma$ has a normal subgroup $P_2$ of order $p_2$ and so $\Gamma/P_2$ (having order $p_1p_3$) has a normal subgroup of order $p_3$, which gives rise to a normal abelian subgroup $\Delta \leq \Gamma$ of order $p_2p_3$. We have that $P_2 \leq \Delta$ but also that $\Delta$ must contain a normal (in particular characteristic) subgroup $P_3$ of order $p_3$, which means that $P_3 \triangleleft \Gamma$, so that, in fact, $n_{p_3} = 1$. Hence $(p_3, p_1p_2)$ is guaranteed to be in $F_3$. Moreover, the complementary subgroup $Q$ is either a cyclic or metacyclic group of order $p_1p_2$. The question then is whether $(p_3, p_1p_2) \in F_0$ as well. However, this is easy since

$$|\text{Aut}(Q)| = \begin{cases} (p_1 - 1)(p_2 - 1) & \text{if } Q \text{ is abelian,} \\ p_2(p_2 - 1) & \text{if } Q \text{ is nonabelian,} \end{cases}$$

and so, if $p_1 < p_2 < p_3$ one has $(p_3, p_1p_2) \in F_3 \cap F_Q$. 

Hence $(p_3, p_1p_2) \notin F_3$. This allows us to use the program to count all the results for $R(\Gamma, [M])$ of order $p_1p_2p_3$.
If $|\Gamma| = p_1p_2p_3$ then $\Gamma = PQ$ where $P$ is cyclic of order $p_3$, of course, and normalized by $Q$ which has order $p_1p_2$, which is itself either a direct or semidirect product of cyclic groups. That is, $\Gamma \cong C_{p_3} \rtimes g(C_{p_2} \rtimes f C_{p_1})$ where $f : C_{p_1} \to \text{Aut}(C_{p_2})$ and $g : C_{p_2} \rtimes f C_{p_1} \to \text{Aut}(C_{p_1})$. There are at most two groups $Q$ of order $p_1p_2$ depending on whether $f$ is nontrivial. The group $\Gamma$ being an iterated (semi)direct product, the role of $g$ must also be factored into the enumeration of the distinct groups of order $p_1p_2p_3$. In particular, we must consider whether the $C_{p_2}$ and/or the $C_{p_1}$ components of $Q$ act nontrivially on $C_{p_1}$. These possibilities for $f$ and $g$ are keyed to congruence conditions on the $p_i$, in particular, whether $p_1 \mid (p_2 - 1)$ and/or $p_1 \mid (p_3 - 1)$ and/or $p_2 \mid (p_3 - 1)$. Alonso [1976] (while exploring an explicit formula due to Hölder [1895] for the number of groups of square-free order) works through the enumeration of groups having order equal to the product of three distinct primes. In particular we give Table 1 therein of the number of groups of order $p_1p_2p_3$ (with our notation for the three primes):

<table>
<thead>
<tr>
<th>$p_2 \mid (p_3 - 1)$</th>
<th>$p_1 \mid (p_3 - 1)$</th>
<th>$p_1 \mid (p_2 - 1)$</th>
<th># groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>no</td>
<td>no</td>
<td>no</td>
<td>1</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>2</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>2</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>$p_1 + 2$</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>2</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>3</td>
</tr>
<tr>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>4</td>
</tr>
<tr>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>$p_1 + 4$</td>
</tr>
</tbody>
</table>

For two of the eight cases, the number of groups varies linearly with $p_1$ (specifically when $C_{p_1}$ acts nontrivially on both $C_{p_2}$ and $C_{p_3}$) but for the others the size is constant. For the case of $p_2 \mid (p_3 - 1)$, $p_1 \mid (p_3 - 1)$, and $p_1 \mid (p_2 - 1)$, it follows that $p_3 > p_1p_2 = m$ which falls into the category of cases dealt with in [K]. Indeed, therein we enumerated $R(\Gamma, [M])$ for all groups of order $mp$ where $p = 2q + 1$ for $q$ a prime (making $p$ a safe-prime) and $m = \phi(p) = 2q$ so that $mp = 2 \times q \times (2q + 1)$, the product of three distinct primes! Therefore we will instead consider the case where $p_3 \equiv 1 \pmod{p_2}$ but where $p_1$ divides both $p_2 - 1$ and $p_3 - 1$, for this includes cases where $p_3 < m = p_1p_2$, for example, $(p_1, p_2, p_3) = (3, 7, 13)$. The $p_1 + 2$ cases can be presented explicitly and again we refer to [Alonso 1976] for the particulars with a slight modification of his notation.

Given $(p_1, p_2, p_3)$, the groups of order $p_1p_2p_3$ are iterated semidirect products, the number of which, as mentioned above, are keyed to elements of order $p_1$ in $U_{p_2}$ and $U_{p_3}$. Specifically, if $U_{p_3} = \langle u_3 \rangle$ and $U_{p_2} = \langle u_2 \rangle$, then the conditions

$p_3 \equiv 1 \pmod{p_1}$, \hspace{1em} $p_2 \equiv 1 \pmod{p_1}$, \hspace{1em} $v_3 = u_3^{(p_3 - 1)/p_1}$, \hspace{1em} $v_2 = u_2^{(p_2 - 1)/p_1}$
together imply that then \(|v_3| = p_1\) and \(|v_2| = p_1\). We have the following presentations for groups of order \(p_1p_2p_3\) which are the cases listed as (10)–(13) in [Alonso 1976, p. 634]. (Note that Alonso adopts the ordering \(p_3 < p_2 < p_1\), the reverse of our convention.)

**Proposition 2.1.** If \(p_1, p_2\) and \(p_3\) are distinct odd primes, where \(p_1 < p_2 < p_3\) and where \(p_3 \equiv 1 \pmod{p_1}\), \(p_2 \equiv 1 \pmod{p_1}\), but \(p_3 \not\equiv 1 \pmod{p_2}\), then the groups of order \(p_1p_2p_3\) are

\[
C_{p_1p_2p_1} = \langle x, y, z \mid x^p, y^p, z^p, yxy^{-1} = x, zxz^{-1} = x, zyz^{-1} = y \rangle,
\]

\[
C_{p_2} \times (C_{p_3} \rtimes C_{p_1}) = \langle x, y, z \mid x^p, y^p, z^p, yxy^{-1} = x, zxz^{-1} = x^v, zyz^{-1} = y \rangle,
\]

\[
C_{p_3} \times (C_{p_2} \rtimes C_{p_1}) = \langle x, y, z \mid x^p, y^p, z^p, yxy^{-1} = x, zxz^{-1} = x, zyz^{-1} = y^v \rangle,
\]

\[
C_{p_3} \rtimes C_{p_2} \rtimes C_{p_1} = \langle x, y, z \mid x^p, y^p, z^p, yxy^{-1} = x, zxz^{-1} = x^v, zyz^{-1} = y^{v_2} \rangle,
\]

where \(i = 1, \ldots, p_1 - 1\).

Our goal is to examine \(R(\Gamma, [M])\) for all groups of this order, as presented above. Following the program as laid out at the end of Section 1, we shall work within the ambient symmetric group \(B = S_{mp}\) where, in this case, \(p = p_3\) and \(m = p_1p_2\). Also, we will choose representative regular subgroups of \(\text{Norm}_B(P)\) where \(P\) is generated by the product of \(m\) disjoint \(p\)-cycles. The elements of \(\text{Norm}_B(P)\) shall be tuples \((\hat{x}, u, \xi)\) where \(\hat{x}\) is a vector in \(\mathbb{F}_p^m\), and where \(u \in U_p\), \(\xi \in S_m\). We note that \(P\) is embedded in \(\text{Norm}_B(P)\) as \(\langle ([1, 1, \ldots, 1], 1, I) \rangle\). The representation of each \(\Gamma\) (as a regular subgroup of \(\text{Norm}_B(P)\)) from among the \(p_1 + 2\) different isomorphism classes is somewhat arbitrary but will be selected for computational convenience. Also, all will be chosen to have their Sylow \(p\)-subgroup be \(P\). The differences will lie in the representation of the complementary subgroups of order \(m\), of which there are two possibilities, up to isomorphism, given that \(p_2 \equiv 1 \pmod{p_1}\).

**Lemma 2.2.** If we define \(\tilde{v}_2 = v_2^{-1}\) in \(U_{p_2}\) and

\[
\sigma = \prod_{k=1}^{p_1} (k, k + p_1, k + 2p_1, \ldots, k + (p_2 - 1)p_1),
\]

\[
\tilde{\sigma} = \prod_{k=1}^{p_1} (k, k + p_1, k + 2p_1, \ldots, k + (p_2 - 1)p_1) \tilde{v}_2^k,
\]

\[
\tau = \prod_{i=0}^{p_2-1} (1 + ip_1, 2 + ip_1, \ldots, p_1 + ip_1),
\]

\[
\delta = \left( \prod_{i=0}^{p_2-1} (1 + ip_1, 2 + \tilde{v}_2ip_1, \ldots, p_1 + \tilde{v}_2^{p_1-1}ip_1) \right)^{-1}.
\]
then \( \langle \sigma, \tau \rangle \cong C_{p^2 p_3} \) and \( \langle \sigma, \delta \rangle \cong C_{p^2} \rtimes C_{p_1} \), and both are regular subgroups of \( S_m \) where \( m = p_1 p_2 \). Moreover, \( \langle \hat{\sigma}, \tau \rangle = (\langle \sigma, \delta \rangle)^{opp} = \text{Cent}_{S_n}(\langle \sigma, \delta \rangle) \).

**Proof.** For the metacyclic group \( C_{p^2} \rtimes C_{p_1} \), there exist generators of orders \( p_2 \) and \( p_1 \) where conjugating the order-\( p_2 \) generator by the order-\( p_1 \) generator raises the order-\( p_2 \) generator to the power \( v_2 \), where \( v_2 \) has order \( p_1 \) in \( U_{p_2} \). This is possible given that \( p_2 \equiv 1 \pmod{p_1} \). For \( C_{p^2 p_1} \), the two generators must, of course, centralize each other. If one writes \( \sigma \) above as \( \sigma_1 \cdots \sigma_{p_1} \) then one may verify that \( \tau \sigma_i \tau^{-1} = \sigma_{i+1} \) and that \( \delta \sigma_i \delta^{-1} = \sigma_{i-1}^{v_2} \). As to regularity, one recalls that if \( N \) is a semiregular subgroup of \( S_n \) of order \( n \) then \( N \) is regular, which is certainly the case for the groups \( \langle \sigma, \tau \rangle \) and \( \langle \sigma, \delta \rangle \). The last assertion is a matter of verifying that the respective generators centralize each other, for example, that \( \sigma \delta \sigma^{-1} = \tilde{\sigma} \).

For the groups of order \( p_1 p_2 p_3 \), the generators of order \( p_2 \) centralize the order-\( p_3 \) generator, but the order-\( p_1 \) generator may or may not centralize the generators of orders \( p_2 \) and \( p_3 \) as presented in Proposition 2.1. Using these presentations, we define the following \( \Gamma \) of each isomorphism class embedded in \( \text{Norm}_B(\mathcal{P}) \), just as in [K, p. 2230].

**Proposition 2.3.** The regular subgroups of \( \text{Norm}_B(\mathcal{P}) \) from each of the isomorphism classes of groups of order \( mp = (p_1 p_2) p_3 \) given in Proposition 2.1 are

\[
\Gamma = C_{p^2 p_2 p_3} = \langle (\hat{1}, 1, I), (\hat{0}, 1, \sigma), (\hat{0}, 1, \tau) \rangle,
\]

\[
\Gamma = C_{p_2} \times (C_{p_3} \rtimes C_{p_1}) = \langle (\hat{1}, 1, I), (\hat{0}, 1, \sigma), (\hat{0}, v_3, \tau) \rangle,
\]

\[
\Gamma = C_{p_3} \times (C_{p_2} \rtimes C_{p_1}) = \langle (\hat{1}, 1, I), (\hat{0}, 1, \tilde{\sigma}), (\hat{0}, 1, \tilde{\tau}) \rangle,
\]

\[
\Gamma = C_{p_3 p_2} \rtimes_j C_{p_1} = \langle (\hat{1}, 1, I), (\hat{0}, 1, \tilde{\sigma}), (\hat{0}, v_3, \tau^j) \rangle,
\]

where \( j = 1, \ldots, p_1 - 1 \).

**Proof.** We note that \( \mathcal{P} \leq \text{Norm}_B(\mathcal{P}) \) is generated by \( (\hat{1}, 1, I) \). We can prove that these groups have the asserted structure by using (2) above. First, we note that \( \langle \hat{0}, 1, \beta \rangle \) centralizes \( (\hat{1}, 1, I) \) for any \( \beta \in S_m \), and that \( \langle \hat{0}, 1, \sigma \rangle \) is centralized by \( \langle \hat{0}, 1, \tau \rangle \). Next we have

\[
(\hat{0}, v_3, \tau)(\hat{1}, 1, I)(\hat{0}, v_3, \tau)^{-1} = (\hat{0} + v_3 \tau(\hat{1}) - \hat{0}, v_3 \cdot 1 \cdot v_3^{-1}, \tau I \tau^{-1})
\]

\[
= (v_3 \tau(\hat{1}), 1, I) = (v_3 \hat{1}, 1, I) = (\hat{1}, 1, I)^{v_3}
\]

and similarly \( (\hat{0}, v_3, \tau^j)(\hat{1}, 1, I)(\hat{0}, v_3, \tau^j)^{-1} = (\hat{1}, 1, I)^{v_3} \). We also have

\[
(\hat{0}, 1, \tau)(\hat{0}, 1, \tilde{\sigma})(\hat{0}, 1, \tau)^{-1} = (\hat{0}, 1, \tau \tilde{\sigma} \tau^{-1}) = (\hat{0}, 1, \tilde{\sigma} v_2) = (\hat{0}, 1, \tilde{\sigma})^{v_2}
\]

and \( (\hat{0}, v_3, \tau^j)(\hat{0}, 1, \tilde{\sigma})(\hat{0}, v_3, \tau^j)^{-1} = (\hat{0}, 1, \tilde{\sigma})^{v_3} \). The proof is finished by recalling [K, Proposition 3.8], which states that if \( \langle \hat{a}, u, \alpha \rangle \) in \( \text{Norm}_B(\mathcal{P}) \) has order coprime to \( p \) then it is fixed-point-free if and only if \( \alpha \in S_m \) is fixed-point-free. In each of
the groups listed here, the order-$p$ subgroup $\mathcal{P}$ is unique and acts freely, and the elements outside of $\mathcal{P}$ have fixed-point-free permutation coordinates. Thus each group $\Gamma$ given in the statement of the proposition is semiregular of order $mp$ and therefore regular.

**Theorem 2.4.** The cardinality of $R(\Gamma, [M])$ for the $p_1 + 2$ classes of groups of order $p_1 p_2 p_3$ given in Proposition 2.1 is as follows, where the rows correspond to different $\Gamma$ and the columns are the classes $[M]$:

<table>
<thead>
<tr>
<th>$\Gamma$</th>
<th>$M \rightarrow C_{p_3 p_2 p_1}$</th>
<th>$C_{p_3} \times (C_{p_2} \times C_{p_1})$</th>
<th>$C_{p_2} \times (C_{p_3} \times C_{p_1})$</th>
<th>$C_{p_3 p_2} \ltimes_i C_{p_1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{p_3 p_2 p_1}$</td>
<td>1</td>
<td>$2(p_1-1)$</td>
<td>$2(p_1-1)$</td>
<td>$4(p_1-1)$</td>
</tr>
<tr>
<td>$C_{p_1} \times (C_{p_2} \times C_{p_1})$</td>
<td>$p_2$</td>
<td>$2(1+p_2(p_1-2))$</td>
<td>$2p_2(p_1-1)$</td>
<td>$4(1+p_2(p_1-2))$</td>
</tr>
<tr>
<td>$C_{p_2} \times (C_{p_3} \times C_{p_1})$</td>
<td>$p_3$</td>
<td>$2p_3(p_1-1)$</td>
<td>$2(1+p_3(p_1-2))$</td>
<td>$4(1+p_3(p_1-2))$</td>
</tr>
<tr>
<td>$C_{p_3 p_2} \ltimes_j C_{p_1}$</td>
<td>$p_3 p_2$</td>
<td>$2p_3(1+p_2(p_1-2))$</td>
<td>$2p_2(1+p_3(p_1-2))$</td>
<td>—</td>
</tr>
</tbody>
</table>

The cardinality of $R(\Gamma, [C_{p_3 p_2} \ltimes_i C_{p_1}])$ is independent of $i \in U_{p_1}$ for the $\Gamma$ listed on the first three rows, and for the last it depends on the relationship between $i$ and $j$ in $U_{p_1}$ thus:

<table>
<thead>
<tr>
<th>$i$, $j$</th>
<th>$R(C_{p_3 p_2} \ltimes_j C_{p_1}, [C_{p_3 p_2} \ltimes_i C_{p_1}])$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j = i, -i$</td>
<td>$2(1 + p_3 + p_2 + (2p_1 - 5)p_2p_3)$</td>
</tr>
<tr>
<td>$j \neq i, -i$</td>
<td>$2(2p_3 + 2p_2 + (2p_1 - 6)p_2p_3)$</td>
</tr>
</tbody>
</table>

As there are $p_1 + 2$ classes of groups of order $(p_1 p_2) p_3$ and therefore $(p_1 + 2)^2$ different possible $R(\Gamma, [M])$, segmented into different classes depending on the different possibilities for $P(N)$, fully detailing the enumeration of all these would tax the patience of the reader. Moreover, given that many pairings give rise to very similar calculations, we shall instead give a sampling of the computations of $R(\Gamma, [M])$. In particular, we shall focus on the enumeration of $R(C_{p_3 p_2} \ltimes_j C_{p_1}, [C_{p_3 p_2} \ltimes_i C_{p_1}])$ since these can effectively be captured in one single computational framework. We will enumerate those $N$ where $P(N) = \mathcal{P}$ and double the resulting figure to account for the corresponding $N^{opp}$ which arise, and therefore have the full count.

**Proof.** We have $\Gamma \cong C_{p_3 p_2} \ltimes_j C_{p_1}$ which is embedded in $\text{Norm}_B(\mathcal{P})$ as

$$\langle (\hat{1}, 1, I), (\hat{0}, 1, \tilde{\sigma}), (\hat{0}, \nu_3, \tau^j) \rangle$$

and we are looking at those regular $N \leq \text{Norm}_B(\mathcal{P})$ isomorphic to $C_{p_3 p_2} \ltimes_i C_{p_1}$ and normalized by this $\Gamma$, where $i, j \in U_{p_1}$. Moreover, as we will be focusing on those $N$ such that $P(N) = \mathcal{P}$, we have

$$N = \langle (\hat{1}, 1, I), (\hat{a}, \nu_3^s, \alpha), (\hat{b}, \nu_3^s, \beta) \rangle,$$
where \((\hat{a}, u_3^\alpha, \alpha)\) has order \(p_2\) and centralizes \((\hat{1}, 1, I)\), and \((\hat{b}, u_3^\beta, \beta)\) has order \(p_1\) and conjugates \((\hat{1}, 1, I)\) to \((\hat{1}, 1, I)^\nu_3 = (v_3 \hat{1}, 1, I)\) and \((\hat{a}, u_3^\alpha, \alpha)\) to \((\hat{a}, u_3^\alpha, \alpha)^v_2\), in accordance with the presentations of the abstract groups as in Proposition 2.1. We will consider those conditions on the components of these 3-tuples which govern order and (semi)regularity and guarantee that \(\Gamma\) normalizes \(N\). The computations themselves will require the basic operational facts about \(\text{Norm}_B(\mathcal{P})\) as given in (1), (2), and (3) together with the fact, mentioned earlier, that if \((\hat{v}, u, \zeta)\) has order coprime \(p = p_3 = |\mathcal{P}|\) and acts without fixed points, then \(\zeta \in S_m\) (\(m = p_2p_3\)) must act without fixed points on the \(m\) coordinates of \(\hat{v}\). We shall proceed \textit{ad hoc}, playing off the different requirements against each other in order to limit the choices for the components.

To begin with, we have (by virtue of the isomorphism class of \(N\))
\[
(\hat{a}, u_3^\alpha)(\hat{1}, 1, I)(\hat{a}, u_3^\alpha)^{-1} = (u_3^\alpha\hat{1}, 1, \alpha I \alpha^{-1}) = (u_3^\alpha\hat{1}, 1, I) = (\hat{1}, 1, I)^u_2, 
\]
which means that \(u_3^\alpha = 1\) since the order-\(p_2\) generator of \(N\) must centralize \(\mathcal{P}\), so that \((\hat{a}, u_3^\alpha, \alpha) = (\hat{a}, 1, \alpha)\). And since \(|(\hat{a}, 1, \alpha)| = p_2\), we have
\[
\left(\sum_{t=0}^{p_2-1} \alpha^t(\hat{a}), 1, \alpha^{p_2}\right) = (\hat{0}, 1, I), 
\]
which, since \((\hat{a}, 1, \alpha)\) is fixed-point-free of order coprime to \(p_3\), means that \(\alpha\) equals \(\alpha_1\alpha_2 \cdots \alpha_{p_1}\), a product of \(p_1\) disjoint \(p_2\)-cycles.

Also, in \(N\) we have
\[
(\hat{b}, u_3^\beta)(\hat{1}, 1, I)(\hat{b}, u_3^\beta)^{-1} = (u_3^\beta\hat{1}, 1, I) = (\hat{1}, 1, I)^u_3, 
\]
which means that \(u_3^\beta = v_3\), so that \((\hat{b}, u_3^\beta, \beta) = (\hat{b}, v_3, \beta)\). And since this must have order \(p_1\), we have
\[
\left(\sum_{t=0}^{p_1-1} v_3^\beta\alpha^t(\hat{b}), v_3^{p_1}, \beta^{p_1}\right) = (\hat{0}, 1, I), 
\]
which, again by the fixed-point-freeness condition on this element of order coprime to \(p_3\), means that \(\beta\) is a fixed-point-free element of order \(p_1\) in \(S_m\).

Again in \(N\), we must have, by virtue of how the order-\(p_1\) generator must act on the order-\(p_2\) generator, that
\[
(\hat{b}, v_3, \beta)(\hat{a}, 1, \alpha)(\hat{b}, v_3, \beta)^{-1} = \hat{b} + v_3\beta(\hat{a}) - (\beta\alpha\beta^{-1})(\hat{b}), 1, \beta\alpha\beta^{-1}, 
\]
where the right-hand side must equal
\[
(\hat{a}, 1, \alpha)^{v_2} = \left(\sum_{t=0}^{v_2-1} \alpha^t(\hat{a}), 1, \alpha^{v_2}\right). 
\]
In particular, we have $\beta \alpha \beta^{-1} = \alpha^{1/2}$ and, more broadly, $\beta \in \text{Norm}_{S_m}(\langle \alpha \rangle)$ where $\alpha$ is a product of $p_1$ disjoint $p_2$-cycles in $S_m$ where $m = p_1 p_2$. The implications of this are that $\text{Norm}_{S_m}(\langle \alpha \rangle)$ is itself another twisted wreath product just like $\text{Norm}_\beta(\mathcal{P})$ and so we shall use the same sort of 3-tuple representation for understanding the relationship between $\alpha$ and $\beta$ within this smaller normalizer. We shall return to the analysis of this relationship shortly.

Looking outward, we now start imposing restrictions on the generators of $N$ imposed by $N$ being normalized by $\Gamma$. To start with, we observe that in $C_{p_2p_3} \rtimes C_{p_1}$ the order-$p_2$ and order-$p_3$ subgroups are characteristic since they are unique of those orders. Thus

$$(\hat{0}, v_3, \tau^j)(\hat{a}, 1, \alpha)(\hat{0}, v_3, \tau^j)^{-1} = (v_3 \tau^j(\hat{a}), 1, \tau^j \alpha \tau^{-j}),$$

where the right-hand side must be an element of $\langle (\hat{a}, 1, \alpha) \rangle$; hence $\tau^j$ lies in $\text{Norm}_{S_m}(\langle \alpha \rangle)$, and thus $\tau \in \text{Norm}_{S_m}(\langle \alpha \rangle)$. We also have

$$(\hat{0}, 1, \tilde{\sigma})(\hat{a}, 1, \alpha)(\hat{0}, 1, \tilde{\sigma})^{-1} = (\tilde{\sigma}(\hat{a}), 1, \tilde{\sigma} \alpha \tilde{\sigma}^{-1}),$$

where the right-hand side must equal $(\hat{a}, 1, \alpha)$, since $\text{Aut}(C_{p_2})$ has no $p_2$-torsion. This implies that $\tilde{\sigma}(\hat{a}) = \hat{a}$ and $\tilde{\sigma} \in \text{Cent}_{S_m}(\alpha)$ which means that $\alpha \in \text{Cent}_{S_m}(\tilde{\sigma})$. The latter observation implies that $\alpha \in \langle \sigma_1, \sigma_2, \ldots, \sigma_{p_1} \rangle$, where $\tilde{\sigma} = \sigma_1^{q_2} \cdots \sigma_{p_1}^{q_1}$. The reason for this is that $\text{Cent}_{S_m}(\tilde{\sigma})$ is isomorphic to the wreath product $C_{p_2} \wr S_{p_1} \cong C_{p_2}^{p_1} \rtimes S_{p_1}$, where the base group of the wreath product, $C_{p_2}^{p_1}$, corresponds to $\langle \sigma_1, \sigma_2^{q_2}, \ldots, \sigma_{p_1}^{q_{p_1}} \rangle = \langle \sigma_1, \ldots, \sigma_{p_1} \rangle$. Therefore, in $C_{p_2}^{p_1} \rtimes S_{p_1}$ the only $p_2$-torsion is in this base group since $S_{p_1}$ cannot have any $p_2$-torsion. As to $\hat{a} = [a_1, \ldots, a_m]$, the condition $\tilde{\sigma}(\hat{a}) = \hat{a}$ means that, for any $k \in \{1, \ldots, m\}$,

$$a_k = a_{\tilde{\sigma}(k)} = a_{\tilde{\sigma}^2(k)} = \cdots = a_{\tilde{\sigma}^{p_2-1}(k)}.$$

But now, since $|(\hat{a}, 1, \alpha)| = p_2$, we have $\sum_{t=0}^{p_2-1} \alpha^t(\hat{a}) = \hat{0}$, and since $\alpha$ is in $\langle \sigma_1, \sigma_2, \ldots, \sigma_{p_1} \rangle$, the orbit of $k \in \{1, \ldots, m\}$ under $\alpha$ is the same as under $\tilde{\sigma}$. Thus $p_2 a_k$ vanishes for any $a_k$ in $\hat{a}$. Since $\hat{a} \in \mathbb{F}_{p_2}^m$, this means $a_k = 0$ since $p_2 \neq 0$ in $\mathbb{F}_{p_2}$. The end result is that $(\hat{a}, 1, \alpha) = (\hat{0}, 1, \alpha)$.

As we saw above, $\alpha$ belongs to $\langle \sigma_1, \sigma_2, \ldots, \sigma_{p_1} \rangle$, and since $\alpha$ acts freely, we must have $\alpha = \sigma_1^{q_1} \sigma_2^{q_2} \cdots \sigma_{p_1}^{q_{p_1}}$, where $q_k \in U_{p_2}$ for $k = 1, \ldots, p_1$. Since $\tau \sigma_1 \tau^{-1} = \sigma_1^{p_1+1}$, if $\tau \alpha \tau^{-1} = \alpha^w$ for some $w \in U_{p_2}$ then, given that $|\tau| = p_1$, we have $w^{p_1} \equiv 1 \pmod{p_2}$ and so $w = v_2^f$ for some $f \in \mathbb{Z}_{p_2}$. And since

$$\tau(\sigma_1^{q_1} \sigma_2^{q_2} \cdots \sigma_{p_1}^{q_{p_1}}) \tau^{-1} = \sigma_1^{q_1} \sigma_2^{q_2} \cdots \sigma_{p_1}^{q_{p_1}},$$

we have $wq_1 = q_{p_1}$, $wq_2 = q_1$, $\ldots$, $wq_{p_1} = q_{p_1-1}$, which means that

$$\hat{a} = [q_1, q_2, \ldots, q_{p_1}] = q_1[1, w^{p_1-1}, w^{p_1-2}, \ldots, w^2, w],$$

Hopf–Galois structures arising from groups with unique subgroup of order $p$ 51
where now \( q_1w = q_1w' \) if and only if \( w = w' \), and if \( q_1 \neq q_1' \) then \( \hat{q} \neq \hat{q}' \), which gives, ostensibly, \( \phi(p_2)p_1 \) choices for \( \alpha \). However, recalling that \((\hat{0}, 1, \alpha)^k = (\hat{0}, 1, \alpha^k)\), we may divide this count by \( \phi(p_2) \) (i.e., assume \( q_1 = 1 \)) to get \( p_1 \) unique choices for \( \langle \alpha \rangle \) and thus \( p_1 \) unique \((\hat{0}, 1, \alpha)\), one of each \( w = v_1' \) for \( f \in \mathbb{Z}_{p_1} \).

Now that we have the enumeration of \((\hat{0}, 1, \alpha)\) given above, how many \((\hat{b}, v_3, \beta)\) are there? What first must be observed is that \( C_{p_2p_3} \times_i C_{p_1} \) has \( p_3p_2\phi(p_1) \) elements of order \( p_1 \). In the presentation

\[
\langle x, y, z \mid x^{p_3}, y^{p_2}, z, yxy^{-1} = x, zxz^{-1} = x^{v_3}, zyz^{-1} = y^{v_3^s} \rangle
\]

the elements of order \( p_1 \) are of the form \(x^ry^sz^t\) for \( r \in \mathbb{Z}_{p_3}, s \in \mathbb{Z}_{p_2}, t \in U_{p_1} \), so the Sylow \( p_1 \)-subgroups are far from characteristic. Thus, when a generator of \( \Gamma \) acts on \((\hat{b}, v_3, \beta)\) by conjugation, the result is one of these \( p_3p_2\phi(p_1) \) other elements of order \( p_1 \). Consider the action of conjugation by \((\hat{0}, 1, \tilde{\sigma})\)

\[\hat{0}, 1, \tilde{\sigma})(\hat{b}, v_3, \beta)(\hat{0}, 1, \tilde{\sigma})^{-1} = (\tilde{\sigma}(\hat{b}), v_3, \tilde{\sigma}\beta\tilde{\sigma}^{-1}) = (\hat{1}, 1, I)^r(\hat{0}, 1, \alpha)^s(\hat{b}, v_3, \beta)^t,\]

which implies that \( t \) must equal 1 and so

\[\hat{0}, 1, \tilde{\sigma})(\hat{b}, v_3, \beta)(\hat{0}, 1, \tilde{\sigma})^{-1} = (r\hat{1} + \alpha^s(\hat{b}), v_3, \alpha^s\beta).\]

For the action of \((\hat{0}, v_3, \tau^j)\) we get

\[\hat{0}, v_3, \tau^j)(\hat{b}, v_3, \beta)(\hat{0}, v_3, \tau^j)^{-1} = (v_3\tau^j(\hat{b}), v_3, \tau^j\beta\tau^{-j}) = (\hat{1}, 1, I)^r(\hat{0}, 1, \alpha)^s(\hat{b}, v_3, \beta)^t,\]

which implies that \( t' \) must equal 1 and so

\[\hat{1}, 1, I)^r(\hat{0}, 1, \alpha)^s(\hat{b}, v_3, \beta)^t' = (r'\hat{1} + \alpha^{s'}(\hat{b}), v_3, \alpha^{s'}\beta).\]

This leads to four “normalization conditions” which must be satisfied:

\[
\tilde{\sigma}(\hat{b}) - \alpha^s(\hat{b}) \in \langle \hat{1} \rangle, \quad (n1)
\]

\[
\tilde{\sigma}\beta\tilde{\sigma}^{-1} = \alpha^s\beta, \quad (n2)
\]

\[
v_3\tau^j(\hat{b}) - \alpha^{s'}(\hat{b}) \in \langle \hat{1} \rangle, \quad (n3)
\]

\[
\tau^j(\hat{b}) - \alpha^{s'} \beta = \alpha^s\beta. \quad (n4)
\]

We can deal with \( (n1) \) immediately by looking once more at how the generators of \( N \) interact. We have

\[\hat{b}, v_3, \beta)(\hat{0}, 1, \alpha)(\hat{b}, v_3, \beta)^{-1} = (\hat{b} - (\beta\alpha\beta^{-1})(\hat{b}), 1, \beta\alpha\beta^{-1}),\]

which must equal \((\hat{0}, 1, \alpha)^{v_3^s} = (\hat{0}, 1, \alpha^{v_3^s})\), implying that \( \hat{b} = (\beta\alpha\beta^{-1})(\hat{b}) = \alpha^{v_3^s}(\hat{b}) \).

By a similar argument to that above, the components of \( \hat{b} \in \mathbb{F}_{p_3}^m \) are also constant along the supports of \( \sigma_1, \ldots, \sigma_{p_1} \), so in fact we may simply observe that \( \alpha(\hat{b}) = \hat{b} \).
Thus $\tilde{\alpha}(\tilde{b}) - \alpha^s(\tilde{b}) = \tilde{b} - \tilde{b} = 0$, which lies in $\langle \hat{1} \rangle$ automatically. This allows us, by the way, to rewrite (n3) as $v_3\tau^i(\tilde{b}) - \tilde{b} \in \langle \hat{1} \rangle$. For (n2) and (n4) we must use that $\beta \in \text{Norm}_{S_m}(\langle \alpha \rangle)$ and represent $\beta$ as an element of this normalizer. Since $\alpha = \sigma_1\sigma_2^{w_1^{-1}}\cdots\sigma_{p_1}^w$ (a product of $p_1$ disjoint $p_2$-cycles), given that $m = p_1p_2$ we have

$$\text{Norm}_{S_m}(\langle \alpha \rangle) \cong \langle \sigma_1, \sigma_2^{w_1^{-1}}, \ldots, \sigma_{p_1}^w \rangle \rtimes (U_{p_2} \times S_{p_1}) \cong \mathbb{F}_{p_2}^5 \rtimes (U_{p_2} \times S_{p_1}).$$

Thus we may write $\alpha$ as the 3-tuple $(\hat{1}, 1, I)$; moreover $\beta = (\hat{c}, v_2^i, \mu)$ for some $\hat{c}$ and $\mu$, since $\beta$ normalizes $\langle \alpha \rangle$ and in view of the following calculation, which uses (2) and the fact that $\mu(\hat{1}) = \hat{1}$ for all $\mu$:

$$(\hat{c}, v_2^i, \mu)(\hat{1}, 1, I)(\hat{c}, v_2^i, \mu)^{-1} = (v_2^i\hat{1}, 1, I) = \alpha^{v_2^i}.$$

At first glance, this permits a fairly large number of possible $\beta$ for a given $\alpha$, but the requirement that $\Gamma$ normalizes $N$ imposes quite a number of restrictions. We begin by observing that

$$\alpha = (\hat{1}, 1, I),$$

$$\tilde{\sigma} = ([1, w\bar{v}_2, w^2\bar{v}_2^2, \ldots, w^{p_1-1}\bar{v}_2^{p_1-1}], 1, I) = (\hat{d}, 1, I),$$

$$\tau = (\hat{0}, w, (1, 2, \ldots, p_1)),$$

and, with this in mind, we see that (n4) translates into conditions on the components of these 3-tuples in $\text{Norm}_{S_m}(\langle \alpha \rangle)$. In particular, with respect to $\beta = (\hat{c}, v_2^i, \mu)$,

$$\tau^j \beta^{-j} = (\hat{0}, w, (1, \ldots, p_1)^j)(\hat{c}, v_2^i, \mu)(\hat{0}, w, (1, \ldots, p_1))^{-j} = (w^j(1, \ldots, p_1)^j(\hat{c}), v_2^i, (1, \ldots, p_1)^j\mu(1, \ldots, p_1)^{-j}),$$

$$\alpha^s \beta = (s\hat{1}, 1, I)(\hat{c}, v_2^i, \mu) = (s\hat{1} + \hat{c}, v_2^i, \mu).$$

Since $\beta = (\hat{c}, v_2^i, \mu)$ has order $p_1$ (and is therefore coprime to $|\alpha|$ in $\text{Norm}_{S_m}(\langle \alpha \rangle)$), $\mu$ must be fixed-point-free of order $p_1$ in $S_{p_1}$ and thus a $p_1$-cycle. But now (n4) implies that $w^j(1, \ldots, p_1)^j\mu(1, \ldots, p_1)^{-j} = \mu$, which means that $\mu = (1, \ldots, p_1)^e$ for some $e \in U_{p_1}$. Furthermore, (n4) also implies that $w^j(1, \ldots, p_1)^j(\hat{c}) - \hat{c} \in \langle \hat{1} \rangle$, a condition which we shall get back to shortly. Since

$$\tilde{\sigma} \beta \tilde{\sigma} = (\hat{d}, 1, I)(\hat{c}, v_2^i, \mu)(\hat{d}, 1, I)^{-1} = (\hat{d} + \hat{c} - v_2^i\mu(\hat{d}), v_2^i, \mu),$$

$$\alpha^s \beta = (s\hat{1} + \hat{c}, v_2^i, \mu),$$

condition (n2) implies that $\hat{d} - v_2^i\mu(\hat{d}) \in \langle \hat{1} \rangle$. (All instances of $\hat{1}$ here refer to the vector $[1, \ldots, 1] \in \mathbb{F}_{p_2}^5$ which is the base group for the twisted wreath product $\text{Norm}_{S_m}(\langle \alpha \rangle)$.)

Recalling that $\hat{d} = [1, w\bar{v}_2, w^2\bar{v}_2^2, \ldots, w^{p_1-1}\bar{v}_2^{p_1-1}]$ and that $\mu = (1, \ldots, p_1)^e$, the condition $\hat{d} - v_2^i\mu(\hat{d}) \in \langle \hat{1} \rangle$ can be analyzed by looking at the components and observing that, in $\langle \hat{1} \rangle$, all the components of a given vector are equal. That is,
\[ \hat{d} = [1, w \bar{v}_2, w^2 \bar{v}_2^2, \ldots, w^{p_1-1} \bar{v}_2^{p_1-1}], \]
\[ v^j_2 \mu(\hat{d}) = [(w \bar{v}_2)^{p_1-e} v^j_2, (w \bar{v}_2)^{p_1-e+1} v^j_2, \ldots, (w \bar{v}_2)^{p_1-e+(p_1-1)} v^j_2]. \]

In particular, the difference of the first components of \( \hat{d} \) and \( v^j_2 \mu(\hat{d}) \) equals the difference of their second components, and so
\[ [1 - w^{p_1-e} \bar{v}_2^{p_1-e-i}] = w \bar{v}_2 [1 - w^{p_1-e} \bar{v}_2^{p_1-e-i}] \]
in \( \mathbb{F}_{p_2} \). If we let \( x = 1 - w^{p_1-e} \bar{v}_2^{p_1-e-i} \) then the above implies that \( x = w \bar{v}_2 x \), so either \( x = 0 \), or \( w \bar{v}_2 = 1 \) regardless of \( x \). If \( w \bar{v}_2 = 1 \), then, since (as determined above) \( w = v^f_2 \) for some \( f \in \mathbb{Z}_{p_1} \), it must be that \( f = 1 \). Otherwise, if \( x = 0 \) then \( w^{p_1-e} = v^f_2^{p_1-e-i} \) and, since (as determined above) \( w = v^f_2 \), this means that \( fe \equiv e + i \) (mod \( p_1 \)), which, by the way, is impossible if \( f = 1 \) since \( i \neq 0 \).

Thus, if \( f = 1 \) then there are no restrictions on \( e \in U_{p_1} \), and if \( f \neq 1 \) then \( fe = e + i \) which implies that \( e = (f-1)^{-1} \) (mod \( p_1 \)).

Now if we go back to (n4), we have the condition \( w^j(1, \ldots, p_1)^{i}(\hat{c}) - \hat{c} \in \{ 1 \} \)
which means that
\[ [c_1, \ldots, c_{p_1}] = [l, l, \ldots, l] + w^j [c_{-j+1}, c_{-j+2}, \ldots, c_{-j+p_1}] \]
for some \( l \in \mathbb{F}_{p_2} \). Looking at the coordinates of \( \hat{c} \) we get
\[ c_{1+j} = l + w^j c_1, \]
\[ c_{1+2j} = l + w^j c_{1+j} = l(1 + w^j) + w^{2j} c_1, \]
\[ \vdots \]
\[ c_{1+kj} = l(1 + w^j + (w^j)^2 + \cdots + (w^j)^{k-1}) + w^{kj} c_1, \]
\[ = \begin{cases} 
& l k + c_1 \quad \text{if } w = 1 \text{ (i.e., } f = 0), \\
& l (\frac{1-(w^j)^k}{1-w^j}) c_1 \quad \text{if } w \neq 1 \text{ (i.e., } f \neq 0). 
\end{cases} \]

This gives a partial parametrization of the possible \( \hat{c} \), but we must also include the conditions imposed by the fact that \( |\beta| = |(\hat{c}, v^j_2, \mu)| = p_1 \), that is,
\[ \sum_{t=0}^{p_1-1} v^j_2 \mu^t(\hat{c}) = \hat{0}, \tag{*} \]
which, since \( \mu^t = (1, \ldots, p_1)^{et} \), means that \( \sum_{t=0}^{p_1-1} v^j_2 (1, \ldots, p_1)^{et}(\hat{c}) = \hat{0} \). So if we let \( \hat{c} = [c_1, \ldots, c_{p_1}] \) then \( (1, \ldots, p_1)^{e+1}(\hat{c}) = [c_{-et+1}, c_{-et+2}, \ldots, c_{-et+p_1}] \), which translates into the (single) condition
\[ \sum_{i=0}^{p_1-1} v^j_2 c_{-et+1} = 0 \tag{**} \]
since the vector equation \( (*) \) consists of a system of equations, all of which are
equivalent to (**). To utilize this information, together with the above parametrization of $\hat{c}$ (in terms of $c_1$ and ‘$l$’ above), we first observe that $1 + kj = -et + 1$ implies $k = j^{-1}(-et)$ and so

$$c_{-et+1} = \begin{cases} l(j^{-1}(-et)) + c_1 & \text{if } w = 1 \text{ i.e., } f = 0, \\ l(\frac{1}{1-w^{-et}}) + w^{-et}c_1 & \text{if } w \neq 1 \text{ i.e., } f \neq 0. \end{cases}$$

For the case $w = 1$ ($f = 0$), we have $e = i(-1)^{-1} = -i$ and so

$$\sum_{t=0}^{p_1-1} v^{it}_2 c_{-et+1} = \sum_{t=0}^{p_1-1} v^{it}_2 (l(j^{-1}(-et)) + c_1) = \sum_{t=0}^{p_1-1} v^{it}_2 (l(j^{-1}(it)) + c_1)$$

$$= \sum_{t=0}^{p_1-1} v^{it}_2 (lj^{-1}it + c_1) = c_1 \sum_{t=0}^{p_1-1} v^{it}_2 + l \sum_{t=0}^{p_1-1} j^{-1}it v^{it}_2$$

$$= lj^{-1}i \sum_{t=0}^{p_1-1} tv^{it}_2 = lj^{-1}i \frac{p_1}{v^{it}_2 - 1}. $$

The last two lines of the above calculation are justified as follows. Since $v^{p_1}_2 = 1$ in $\mathbb{F}_{p_2}$, we have

$$\sum_{t=0}^{p_1-1} v^{it}_2 = 0, \quad \sum_{t=0}^{p_1-1} tx^t = x \left( \frac{p_1 x^{p_1-1}}{x - 1} - \frac{x^{p_1} - 1}{(x - 1)^2} \right),$$

and substituting in $x = v^{it}_2$ we get

$$\sum_{t=0}^{p_1-1} tv^{it}_2 = \frac{p_1 v^{ip_1}_2}{v^{it}_2 - 1} = \frac{p_1}{v^{it}_2 - 1}.$$

This being the case, $\sum_{t=0}^{p_1-1} v^{it}_2 c_{-et+1} = 0$ if and only if $l = 0$, which means that $c_{1+kj} = c_1$ for all $k$, and therefore $\hat{c} \in \langle 1 \rangle$.

For the case where $w \neq 1$ (i.e., $f \neq 0$), we have

$$\sum_{t=0}^{p_1-1} v^{it}_2 c_{-et+1} = \sum_{t=0}^{p_1-1} v^{it}_2 \left( l \left( \frac{1 - (wi)^j}{1 - w^j} \right) + w^{-et}c_1 \right)$$

$$= \frac{l}{1-w^j} \sum_{t=0}^{p_1-1} v^{it}_2 (1-w^{-et}) \sum_{t=0}^{p_1-1} v^{it}_2 w^{-et} c_1$$

$$= \frac{l}{1-w^j} \sum_{t=0}^{p_1-1} v^{it}_2 - \frac{l}{1-w^j} \sum_{t=0}^{p_1-1} v^{it}_2 w^{-et} + c_1 \sum_{t=0}^{p_1-1} v^{it}_2 w^{-et}$$
\[
\left( c_1 - \frac{l}{1 - w} \right) \sum_{i=0}^{p_1-1} v_2^{(i-fe)t}.
\]

If \( i \neq fe \) then the last sum above equals 0 for all \( c_1, l \in \mathbb{F}_{p_2} \), yielding \( p_2^2 \) choices. If \( i = fe \) then the last sum is 0 only when \( c_1 = l/(1-w^j) \), which means only \( p_2 \) choices.

The requirements of condition (n3), that \( v_3^j \tau^j(\hat{b}) = \hat{b} \) and equivalently \( \alpha(\hat{b}) = \hat{b} \). We also must factor in order considerations, just as in the above enumeration of \( \hat{c} \), namely that \( \sum_{i=0}^{p_1-1} v_3^{j} \tau^j(\hat{b}) = \hat{b} \). Since the components of \( \hat{b} \) (a vector in \( \mathbb{F}_{p_3}^{p_1} \)) are equal on the supports of the cycles that make up \( \hat{\sigma} \) (a product of \( p_1 \) disjoint \( p_2 \)-cycles) and since \( \tau \sigma_i \tau^{-1} = \sigma_{i+1} \), by identifying together these identical components, we can proceed, for the moment, as if \( \hat{b} \) were a vector in \( \mathbb{F}_{p_3}^{p_1} \).

With this identification, \( \tau \) acts on this \( \hat{b} \) as \((1, \ldots, p_1) \) and therefore \( \tau^j \) acts like \((1, \ldots, p_1)^j \). Similarly, since \( \beta = (\hat{c}, v_3^2, (1, \ldots, p_1)^\epsilon) \), it acts on \( \hat{b} \) as \((1, \ldots, p_1)^\epsilon \). Consequently, if we set \( \hat{b} = [b_1, b_2, \ldots, b_{p_1}] \) then (n3) implies that \([b_1, \ldots, b_{p_1}] = [l, l, \ldots, l] + v_3[b_{e+1}, b_{e+2}, \ldots, b_3] \) for some \( l \in \mathbb{F}_{p_3} \) so that, in a similar fashion to the computation of \( \hat{c} \) a few pages back, we have

\[
b_{1+kj} = l(1 + v_3 + v_3^2 + \cdots + v_3^{k-1}) + v_3^k b_1 = l(1 - v_3^k) + v_3^k b_1.
\]

Since \( 1 + kj = 1 - et \) implies \( k = -j^{-1}et \), we get

\[
b_{1-et} = l\left(1 - v_3^{-j^{-1}et}\right) + v_3^{-j^{-1}et} b_1
\]

for the parametrization of \( \hat{b} \). The question is: how many ‘degrees of freedom’ do we have since, ostensibly, we can choose \( l, b_1 \in \mathbb{F}_{p_3} \) ? The order requirement becomes \( \sum_{i=0}^{p_1-1} v_3^j(1, \ldots, p_1)^{et}(\hat{b}) = \hat{b} \) which reduces to \( \sum_{i=0}^{p_1-1} v_3^i b_{1-et} = 0 \). In a similar fashion to the calculations above, we get that \(|(\hat{b}, v_3, \beta)| = p_1 \) if and only if

\[
(b_1 - \frac{l}{1 - v_3}) \sum_{i=0}^{p_1-1} v_3^{(1-j^{-1}et)} = 0,
\]

which comes down to two possibilities given that \( v_3^{p_1} \equiv 1 \pmod{p_3} \). If \( e \neq j \) then we may choose \( b_1 \) and \( l \) in \( \mathbb{F}_{p_3} \) arbitrarily (i.e., \( p_2^2 \) choices); otherwise, one must choose \( b_1 \) and \( l \) such that \( b_1 = l/(1-v_3) \), which yields \( p_2 \) choices.
The enumeration of the possible \((b, v_3, \beta)\) comes down to the interaction between the parameters \(f, e, i, \) and \(j\) as determined by order conditions on \((\hat{c}, v_3, \beta)\), where \(\beta = (\hat{c}, v_2, (1, \ldots, p_1)^{e'})\), and the normalization conditions (n1)–(n4). The parameters \(i\) and \(j\) are chosen at the outset, but the core parameter is \(f \in \mathbb{Z}_{p_1}\) which determines the possible \(\alpha\). Subsequently, \(e\) is determined by \(\beta\) since \((b, v_3, \beta)\) normalizes \(\langle \hat{0}, 1, \alpha \rangle\). The different possibilities are summarized as follows:

\[
\begin{align*}
  f = 0 & \text{ implies } e = i(f - 1)^{-1} = -i, \\
  f = 1 & \text{ allows } e = 1, \ldots, p_1 - 1, \\
  f = 2, \ldots, p_1 - 1 & \text{ implies } e = i(f - 1)^{-1}, \\
  f = 0 & \text{ implies } p_2 \text{ choices for } \hat{c}, \\
  f > 0 & \text{ implies } \begin{cases} p_2 \text{ choices for } \hat{c} \text{ when } i = fe, \\
p_2^2 \text{ choices for } \hat{c} \text{ when } i \neq fe. \end{cases} \\
  j = e & \text{ implies } p_3 \text{ choices of } \hat{b}, \\
  j \neq e & \text{ implies } p_3^2 \text{ choices of } \hat{b}.
\end{align*}
\]

If we denote by \(s_0, s_1, s_{>1}\) the number of \((\hat{b}, v_3, \beta)\) for the different choices of \(f\), then we want to know \(s_0 + s_1 + s_{>1}\).

For \(f = 0\) we have \(e = -i\) and so there are \(p_2\) different \(\beta = (\hat{c}, v_3, (1, \ldots, p_1)^{e'})\).

If \(j = -i = e\) there are \(p_3\) choices for \(\hat{b}\), and if \(j \neq -i\) there are \(p_3^2\). Hence

\[
s_0 = \begin{cases} p_2 p_3 & j = -i, \\
p_2 p_3^2 & j \neq -i. \end{cases}
\]

For \(f = 1\) we have \(e = 1, \ldots, p_1 - 1\) and \(fe = f\), so \(fe = i\) for exactly one \(e\) and \(j = e\) also exactly once. Depending on whether \(i = j\) or not, for potentially the same \(e\), this results in different possibilities for the number of \(\hat{c}\) and \(\hat{b}\). We have

\[
s_1 = \begin{cases} p_2 p_3 + (p_1 - 2) p_2^2 p_3^2 & j = i, \\
p_2 p_3^2 + p_2^2 p_3 + (p_1 - 3) p_2^2 p_3^2 & j \neq i. \end{cases}
\]

For \(f > 1\) we have \(e = i(f - 1)^{-1}\) and so \(fe = f\) which is never equal to \(i\); thus there are \(p_2^2\) choices for \(\hat{c}\). If \(j = i(f - 1)^{-1} = e\) then \(f - 1 = ij^{-1}\) which is impossible if \(f - 1 = -1\), that is, \(j = -i\); thus there are \(p_3^2\) different \(\hat{b}\) for each \(f > 1\). We have then the count for \(f > 1\):

\[
s_{>1} = \begin{cases} (p_1 - 2) p_2^2 p_3^2 & j = -i, \\
(p_1 - 3) p_2 p_3^2 + p_2^2 p_3 & j \neq -i. \end{cases}
\]

Now, for \(i, j \in U_{p_1}\), we have that \(j = i\) implies \(j \neq -i\) since \(p_1 > 2\) and, similarly, if \(j = -i\) then \(j \neq i\). So we have

\[
s_0 + s_1 + s_{>1} = \begin{cases} p_2 p_3 + p_2^2 p_3 + p_2^2 p_3 + (2p_1 - 5)p_2^2 p_3 & j = i, -i, \\
2p_2^2 p_3 + 2p_2^2 p_3 + (2p_1 - 6)p_2^2 p_3 & j \neq i, -i. \end{cases}
\]
What $s_0 + s_1 + s_{>1}$ represents is the number of those

$$\{(0, 1, \alpha), (\hat{b}, v_2, \beta)\}$$

which generate $Q(N)$, where $P(N) = \mathcal{P}$ and where, for distinct $f$, the resulting $\langle \alpha \rangle$ and thus $\langle (0, 1, \alpha) \rangle$ are distinct. Thus, to remove duplicate $Q(N)$s (arising from $(\hat{b}, v_3, \beta)$ which generate the same $Q(N)$ with $(0, 1, \alpha)$) we must divide by $p_2p_3$. The reason for this is that, as we mentioned above, in the abstract groups $C_{p_2p_3} \times_i C_{p_1}$, if one multiplies a given element of order $p_1$ by an element of order $p_2$ or $p_3$ (or both) one gets another element of order $p_1$.

We have now completely enumerated those $N \in R(C_{p_2p_3} \times_j C_{p_1}, [C_{p_2p_3} \times_i C_{p_1}])$ where $P(N) = \mathcal{P}$. Since the order-$p_3$ subgroup is not a direct factor, we now double this figure since the groups in $R(C_{p_2p_3} \times_j C_{p_1}, [C_{p_2p_3} \times_i C_{p_1}])$ are evenly distributed between those classes where $P(N) = \mathcal{P}$ versus those for which $P(N) \neq \mathcal{P}$. The count given in the statement of the theorem for $|R(C_{p_2p_3} \times_j C_{p_1}, [C_{p_2p_3} \times_i C_{p_1}])|$ is now verified.

\[ \square \]

3. Square-free groups where $p < m$

There are many prime triples $(p_1, p_2, p_3)$, where $p_3 \equiv 1 \pmod{p_1}$, $p_2 \equiv 1 \pmod{p_1}$, and $p_3 \not\equiv 1 \pmod{p_2}$ (which give rise to groups of the type studied in Theorem 2.4) but where $p = p_3 < p_1p_2 = m$. Indeed, if one takes prime triples from $\{2, \ldots, 113\}$ then, of these, 474 have the property implying that groups of order $p_1p_2p_3$ are in the category studied in Theorem 2.4, and, of these, 246 have the property $p < m$. If we look beyond to groups of order $p_1p_2p_3p_4$, where $p_1 < p_2 < p_3 < p_4$, which are also explored in [Alonso 1976], then the analog of Theorem 2.4 is the case where $\{p_4, p_3, p_2\}$ are all equivalent to 1 (mod $p_1$) but none of $\{p_4, p_3, p_2\}$ are equivalent to 1 mod each other. In this case, the number of groups of order $p_1p_2p_3p_4$ is $p_1^2 + p_1 + 2$. If one looks at the 4-tuples of distinct primes chosen in $\{2, \ldots, 113\}$ then, of these, 3173 satisfy the congruence conditions of this class, and if $m = p_1p_2p_3$ and $p = p_4$ then, of these, 3151 have the property that $p < m$.

Acknowledgement

The author wishes to thank the referee for the careful reading given to this paper and for the helpful comments and suggestions.

References


Hopf–Galois structures arising from groups with unique subgroup of order \( p \)


Communicated by Susan Montgomery

Received 2014-08-05 Revised 2015-10-01 Accepted 2015-11-27

tkohl@math.bu.edu Department of Mathematics and Statistics, Boston University, 111 Cummington Mall, Boston, MA 02215, United States
On tensor factorizations of Hopf algebras

Marc Keilberg and Peter Schauenburg

We prove a variety of results on tensor product factorizations of finite dimensional Hopf algebras (more generally Hopf algebras satisfying chain conditions in suitable braided categories). The results are analogs of well-known results on direct product factorizations of finite groups (or groups with chain conditions) such as Fitting’s lemma and the uniqueness of the Krull–Remak–Schmidt factorization. We analyze the notion of normal (and conormal) Hopf algebra endomorphisms, and the structure of endomorphisms and automorphisms of tensor products. The results are then applied to compute the automorphism group of the Drinfeld double of a finite group in the case where the group contains an abelian factor. (If it doesn’t, the group can be calculated by results of the first author.)

Introduction

The larger part of this paper is concerned with general results on Hopf algebras in braided categories generalizing well-known results from the theory of finite groups (or groups with chain conditions), such as Fitting’s lemma, the Krull–Remak–Schmidt decomposition, and a description of endomorphisms and automorphisms of products of Hopf algebras. The last section deals with the description of the automorphism group of the Drinfeld double \( D(G) \) of a finite group \( G \). This last problem was the starting point of our work.

In the case that \( G \) has no nontrivial abelian direct factors, a complete description of the automorphisms was given in [Keilberg 2015]. The case when \( G \) has such an abelian factor was left open. We will write such a group as \( G = C \times H \), where \( H \) has no nontrivial abelian direct factors and \( C \) is abelian. In this case we naturally have that \( D(G) \cong D(C) \otimes D(H) \) is a tensor product of Hopf algebras.

Thus, we are naturally led to analyze endomorphisms and automorphisms of a tensor product of two Hopf algebras. In [Bidwell et al. 2006; Bidwell 2008] an analysis of the automorphisms of direct products of groups was provided. The basic idea is to describe such automorphisms by a matrix of morphisms between the factors. The machinery of normal group endomorphisms and Fitting’s lemma then allows one to

---

Research partially supported through a FABER Grant by the Conseil régional de Bourgogne.

MSC2010: primary 16T05; secondary 18D35, 18D10, 20D99, 81R05.

Keywords: Hopf algebra, factorization, Fitting’s lemma, Krull–Remak–Schmidt.
deduce conditions on the various morphisms from conditions on the factors. For example, when the two factors have no common direct factors, then the diagonal terms of the matrix have to be automorphisms. In Section 8 we derive suitably analogous results for tensor product Hopf algebras. Before this can be done, however, we have to carry over to our Hopf algebraic setting some basic notions and classical results from group theory. In Section 2 we develop the terminology of commuting morphisms (for groups these are just morphisms whose images commute) and dually of cocommuting morphisms, and in Section 3 the notions of normal and conormal Hopf endomorphisms. The analog of Fitting’s lemma which will produce tensor product decompositions from binormal endomorphisms and thus, under suitable circumstances, common tensor factors from certain endomorphisms of tensor products, will be proved in Section 5. An important application of Fitting’s lemma in group theory is the uniqueness of the Krull–Remak–Schmidt decomposition, which we prove in Section 6. Extensions of the Krull–Remak–Schmidt decomposition were studied previously in [Burciu 2011] for decompositions of semisimple Hopf algebras into simple semisimple tensor factors. By contrast our techniques make no use of semisimplicity but only of chain conditions. It is also worth noting that the Krull–Remak–Schmidt result shows that our results are specific to Hopf algebras and cannot be readily generalized to finite or even fusion tensor categories: Müger [2003] gives an example where the factors in the decomposition of a fusion category into prime factors are not unique.

In fact the above results on the structure theory of finite dimensional Hopf algebras over a field \( k \) will be developed in greater generality for Hopf algebras in braided abelian tensor categories that fulfill chain conditions on Hopf subalgebras and quotient Hopf algebras. Apart from the fact that the results will thus immediately apply to objects like super-Hopf algebras, for some purposes the categorical setting is simply very natural, since it allows treating mutually dual notions like normality and conormality or ascending and descending chain conditions on the same footing. If the braiding of the base category is not a symmetry, then some of our basic objects of study may be hard to come by: it is well known that the tensor product of two Hopf algebras in a braided monoidal category can only be formed if the two factors are “unbraided”, that is, if the braiding between them behaves like a symmetry. On the other hand, some of our results imply that tensor product decompositions have to exist in certain situations. Thus these results also imply that the braiding has to be “partially trivial”. For example, if nonnilpotent normal endomorphisms of a Hopf algebra exist, they have to be isomorphisms by Fitting’s lemma unless the braiding is partially trivial. An automorphism of a tensor product of nonisomorphic Hopf algebras (necessarily “unbraided” between each other) has to induce automorphisms on the factors, unless the braiding is partially trivial on one of the factors.

Section 4 deals with some technical issues raised by our categorical framework. In preparation for Fitting’s lemma we decompose a Hopf algebra with chain conditions,
for which a Hopf algebra endomorphism is given, into a Radford biproduct (in the generalized braided version due to Bespalov and Drabant [1998]). A technical result on (co)invariants under Hopf algebra endomorphisms has some bearing on the notions of epimorphisms and monomorphisms studied notably for infinite dimensional Hopf algebras in [Chirvăsitu 2010].

In Section 9 we present the application of the general results on the structure of finite Hopf algebras and their automorphisms to the study of automorphisms of Drinfeld doubles of groups. Letting \( G = C \times H \) as before, taking the field to be the complex numbers, and defining \( \hat{H} \) to be the group of linear characters of \( H \), then under the isomorphisms \( \mathcal{D}(C) \cong \mathbb{C}(\hat{C} \times C) \) and \( \hat{C} \times C \cong C^2 \) the result can be stated as

\[
\text{Aut}(\mathcal{D}(C \times H)) \cong \begin{pmatrix}
\text{Aut}(C^2) & \text{Hom}^c(\mathcal{D}(H), \mathbb{C}C^2) \\
\text{Hom}(C^2, \hat{H} \times Z(H)) & \text{Aut}(\mathcal{D}(H))
\end{pmatrix}.
\]

The only term not explicitly determined by [Keilberg 2015] or standard methods for finite abelian groups is \( \text{Hom}^c(\mathcal{D}(H), \mathbb{C}C^2) \), which we define in Section 8. In this case the morphisms can be described entirely in terms of group homomorphisms and central subgroups of \( G \) satisfying certain relations [Agore et al. 2014; Keilberg 2015], so the description is not a significant problem. In Example 9.10 we completely describe \( \text{Aut}(\mathcal{D}(D_{2n})) \) where \( D_{2n} \) is the dihedral group of order \( 2n \), for the case \( n \equiv 2 \mod 4 \) and \( n > 2 \). This is precisely when there is an isomorphism \( D_{2n} \cong \mathbb{Z}_2 \times D_n \). From this we can easily provide a formula for the order of \( \text{Aut}(\mathcal{D}(D_{2n})) \). In particular we find that \( \text{Aut}(\mathcal{D}(D_{12})) \) has order \( 1152 = 2^73^2 \).

1. Preliminaries and notation

Throughout the paper, \( \mathcal{B} \) is an abelian braided tensor category with braiding \( \tau \); we will assume that \( \mathcal{B} \) is strict, backed up by the well-known coherence theorems. Algebras, coalgebras, bialgebras, Hopf algebras are in \( \mathcal{B} \). All undecorated Homs, Ends, etc., will be for morphisms of Hopf algebras or groups, as appropriate. We will use the following graphical notations to do computations in \( \mathcal{B} \): the braiding is

\[
\tau_{VW} = \begin{array}{c}
\begin{array}{c}
W \\
V \\
\end{array}
\end{array}
\quad \text{and} \quad \tau_{VW}^{-1} = \begin{array}{c}
\begin{array}{c}
W \\
V \\
\end{array}
\end{array}.
\]

We shall say that the objects \( V \) and \( W \) are unbraided if \( \tau_{VW} = \tau_{WV}^{-1} \).

Multiplication and unit of an algebra \( A \), and comultiplication and counit of a coalgebra \( C \) are

\[
\nabla_A = \begin{array}{c}
\begin{array}{c}
A \\
A \\
\end{array}
\end{array}, \quad \eta_A = \begin{array}{c}
\begin{array}{c}
A \\
\end{array}
\end{array}, \quad \Delta_C = \begin{array}{c}
\begin{array}{c}
C \\
C \\
\end{array}
\end{array}, \quad \varepsilon_C = \begin{array}{c}
\begin{array}{c}
C \\
\end{array}
\end{array}.
\]
The antipode of a Hopf algebra and, if it exists, its inverse are

\[
S = \begin{array}{c}
H \\
H
\end{array}
\quad \text{and} \quad
S^{-1} = \begin{array}{c}
H \\
H
\end{array}.
\]

In order to have a straightforward notion of Hopf subalgebra and quotient Hopf algebra of a given Hopf algebra, we shall assume that tensor products in \( \mathcal{B} \) are exact in each argument.

An object in \( \mathcal{B} \) satisfies the ascending chain condition on subobjects if and only if it satisfies the descending chain condition on quotient objects, by which we understand the descending chain condition on subobjects in the opposite category. For Hopf algebras we will use the descending chain conditions on Hopf subalgebras and on quotient Hopf algebras. When a Hopf algebra satisfies the descending chain conditions on both Hopf subalgebras and quotient Hopf algebras, we simply say that it satisfies both chain conditions.

If \( f : H \to G \) is a Hopf algebra morphism, we define the right and left \( f \)-coinvariant subobjects of \( H \) as being the equalizers

\[
0 \longrightarrow H^{co f} \longrightarrow H \xrightarrow{\begin{array}{c}(H \otimes f) \Delta \\
H \otimes \eta
\end{array}} H \otimes G
\]

And dually, the left and right invariant quotients by coequalizers

\[
G \otimes H \xrightarrow{\begin{array}{c}\nabla (f \otimes G) \\
\varepsilon \otimes G
\end{array}} G \longrightarrow H \setminus G \longrightarrow 0
\]

We note that the coinvariant subobjects are subalgebras of \( H \), and the invariant quotients are quotient coalgebras of \( G \).

We will say a Hopf algebra is abelian if it is both commutative and cocommutative. When working over a field, abelian semisimple Hopf algebras are precisely the duals of abelian group algebras, up to a separable field extension [Montgomery 1993, Theorem 2.3.1]. We will say a Hopf algebra is nonabelian when it is not abelian.

2. Commuting and cocommuting morphisms

In this section we formulate an obvious commutation condition for morphisms to an algebra (for ordinary algebras it just means that elements in the respective images
commute) and its dual, and we collect equally obvious consequences that will be useful in later calculations. We note that for each and every fact on Hopf algebras in a braided category there is a dual fact. We will not always state, but still freely use the duals of our statements.

Let $A$ be an algebra, $V, W \in \mathcal{B}$, and $f : V \to A$, $g : W \to A$ morphisms in $\mathcal{B}$. We say that $f$ and $g$ multiplication commute and write $f \nabla g$ if $\nabla(g \otimes f) = \nabla(f \otimes g) \tau$ ($= \nabla \tau(g \otimes f)$), or graphically

$$f \nabla g = g \nabla f.$$ 

Dually, we say two morphisms $f : C \to V$ and $g : C \to W$ from a coalgebra $C$ in $\mathcal{B}$ comultiplication commute, or cocommute for short, and write $f \nabla c g$ if

$$g \nabla f = f \nabla g.$$ 

We say that $f, g$ bicommutate if both $f \nabla g$ and $f \nabla h$.

If $A$ and $B$ are algebras in $\mathcal{B}$, then the natural maps $f : A \to A \otimes B$ and $g : B \to A \otimes B$ satisfy $f \nabla g$, but they only satisfy $g \nabla f$ if $A$ and $B$ are unbraided. In fact:

$$f \nabla g A B = g A B \nabla f A B \quad \text{and} \quad A B \nabla f A B = g A B \nabla f A B.$$

**Lemma 2.1.** Let $A$ be an algebra, $C$ a coalgebra, and $U, V, W, X, Y$ objects in $\mathcal{B}$.

(i) Let $f : U \to A$, $g : V \to A$ and $h : W \to A$.

(a) If $f \nabla g$ and $f \nabla h$, then $f \nabla (g \otimes h)$.

(b) If $f \nabla h$ and $g \nabla h$, then $(\nabla(f \otimes g)) \nabla h$.

(c) If $f \nabla g$, then $f a \nabla g b$ for any $a : X \to U$ and $b : Y \to V$.

(ii) Let $f, g, h : C \to A$.

(a) If $f \nabla g$ and $f \nabla h$ then $f \nabla (g * h)$.

(b) If $f \nabla h$ and $g \nabla h$ then $(f * g) \nabla h$.

(c) If $f \nabla g$ and $g \nabla f$, then $f * g = g * f$.

(iii) Let $f, g : C \to A$.

(a) If $C$ is a bialgebra, $f, g$ are algebra morphisms, and $f \nabla g$, then $f * g$ is an algebra morphism.

(b) If $A, C$ are bialgebras, $f, g$ are bialgebra morphisms, $f \nabla g$ and $f \nabla h$, then $f * g$ is a bialgebra morphism.

(c) If $A$ is a bialgebra, $C$ a Hopf algebra, and $f, g$ are unital coalgebra morphisms, then $f \nabla g \iff f * g$ is a coalgebra morphism.
Note that $f \rightsquigarrow g$ is not necessarily equivalent to $g \rightsquigarrow f$ in the braided setting. The first part of the following result says, however, that the two properties are equivalent for Hopf algebras with sufficiently well-behaved antipodes. On the other hand, the second part says that if both properties are fulfilled then either the braiding is close to being a symmetry, or the morphisms are close to being trivial.

**Proposition 2.2.** Let $H$, $K$, and $A$ be Hopf algebras, and $f : H \to A$, $g : K \to A$ Hopf algebra morphisms.

(i) If $f \rightsquigarrow g$, and if the antipode of $A$ is a monomorphism or the antipodes of $H$ and $K$ are epimorphisms, then $g \rightsquigarrow f$.

(ii) If $f \rightsquigarrow g$ and $g \rightsquigarrow f$, then

$$H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g}$$

which implies $g \rightsquigarrow f$ if the antipodes of $H$ and $K$ are epimorphisms. A similar argument shows the same if the antipode of $A$ is a monomorphism.

Proof. For the first claim, we calculate

which implies $g \rightsquigarrow f$ if the antipodes of $H$ and $K$ are epimorphisms. A similar argument shows the same if the antipode of $A$ is a monomorphism.

We now turn to the second claim. We have

$$H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g} H K \xrightarrow{f \otimes g}$$
In other words,

\[ H \otimes X \otimes A = X \otimes A \otimes H \]

does not depend on the choice of \( X \in \{ \left\langle \begin{array}{c} \emptyset \\ \emptyset \end{array} \right\rangle \} \). But since \( f \circ S \) and \( g \circ S \) are convolution inverses to \( f \) and \( g \), respectively, we have

\[ f \otimes S = g \otimes \] \[ H \otimes K \otimes A = H \otimes K \otimes A \]

That the latter expression does not depend on the choice of \( X \) is the claim.  

As special cases one recovers two known facts that show how badly usual Hopf algebra constructions behave in a “truly braided” tensor category: a Hopf algebra cannot be commutative (or cocommutative) as a (co)algebra in \( B \) unless the braiding on the Hopf algebra is an involution [Schauenburg 1998], and the tensor product of two Hopf algebras cannot be a Hopf algebra unless the two factors are unbraided.

3. Normal endomorphisms

Recall that the left adjoint action and the left coadjoint coaction of a Hopf algebra \( H \) on itself are

\[ \text{ad} = \] \[ H \otimes H \otimes H \]

and

\[ \text{coad} = \] \[ H \otimes H \otimes H \]

We note that the adjoint action is characterized by a twisted commutativity condition:

\[ \text{ad} = \] \[ H \otimes H \]

\[ \text{(3-1)} \]
Definition 3.1. Let \( f : H \to H \) be a morphism in \( \mathcal{B} \), with \( H \) a Hopf algebra.

(i) \( f \) is normal if it is left \( H \)-linear with respect to the adjoint action.
(ii) \( f \) is conormal if it is left \( H \)-colinear with respect to the coadjoint coaction.
(iii) \( f \) is binormal if it is both normal and conormal.

For group algebras considered in the category of \( \mathbb{C} \)-vector spaces, the definition of a normal morphism agrees with the one used in group theory [Rotman 1995]. Since group algebras are cocommutative, every group endomorphism is trivially conormal. We will be primarily concerned with normal, conormal, and binormal endomorphisms of Hopf algebras.

Lemma 3.2. Let \( f : H \to H \) be an endomorphism of the Hopf algebra \( H \).

(i) The following are equivalent:
   (a) \( f \) is normal.
   (b) \( f \bowtie ((f S) \ast \text{id}_H) \).
   (c) \( (f S) \ast \text{id}_H \) is an algebra morphism.

(ii) The following are equivalent:
   (a) \( f \) is binormal.
   (b) \( f \bowtie ((f S) \ast \text{id}_H) \) and \( f \bowtie ((f S) \ast \text{id}_H) \).
   (c) \( (f S) \ast \text{id}_H \) is a bialgebra morphism.

Proof. We only show part (i). For the equivalence of (b) and (c) we apply the bijection

\[
\mathcal{B}(H \otimes H, H) \ni T \mapsto \begin{array}{c}
\begin{array}{c}
\text{H}
\end{array}
\end{array} \in \mathcal{B}(H \otimes H, H)
\]

to the two sides of the equation expressing multiplicativity of \( g := f S \ast \text{id} \). We get

and
which are the two sides of (b), up to composition with the isomorphism $H \otimes S$. For the equivalence of (a) and (b), we apply the bijection

$$B(H \otimes H, H) \ni T \mapsto T \in B(H \otimes H, H)$$

to the two sides of (a) and once again get two sides of (b):

and

□
We recall Radford’s theorem [1985] on Hopf algebras with a projection, which was
generalized to a categorical setting even more general than the one in the present
paper by Bespalov and Drabant [1998]:

**Theorem 4.1.** Let $H$ be a Hopf algebra, and $\pi$ an idempotent Hopf algebra endomorphism of $H$. Then $H \cong \text{Im}(\pi) \otimes \text{Im}(p)$, where $p = (\pi \circ S) \ast \text{id}_H$ is an idempotent endomorphism of the object $H$ in $B$ (but not necessarily a Hopf endomorphism). $B := \text{Im}(p)$ is a subalgebra and a quotient coalgebra of $H$. The algebra structure of $\text{Im}(\pi) \otimes B$ is a semidirect product with respect to a certain action of $K = \text{Im}(\pi)$ on $B$, and the coalgebra structure is the cosemidirect product with respect to a certain coaction.

Moreover $\text{Im}(p) \cong \text{co}\pi H \cong \pi \setminus H$.

**Proof.** Only the last statement is not in [Bespalov and Drabant 1998], who avoid using coinvariant subobjects altogether to generalize [Radford 1985] to categories that might not have equalizers. We check the first isomorphism. We find

and if some morphism $t : T \to H$ satisfies $(\pi \otimes \text{id}_H) \Delta t = \eta \otimes t$, then

**Proposition 4.2.** Let $H$ be a Hopf algebra, and $f$ a Hopf algebra endomorphism of $H$. Assume that $H$ satisfies both chain conditions. Then there is $n \in \mathbb{N}$ such that $H \cong \text{Im}(f^n) \otimes \text{co}\, f^n H$ is a Radford biproduct.

**Proof.** Consider the epi-mono factorization $f = (H \xrightarrow{\varepsilon} B \xrightarrow{\mu} H)$, where we identify $B = \text{Im}(f) = \text{Coim}(f)$. Then the endomorphism $t = em$ of $B$ satisfies $mt = fm$ and $te = ef$. The chain conditions on $H$ imply that the ascending chain of the
kernels of \( f^n \) and the descending chain of the images, hence the ordered chain of quotient objects formed by the cokernels of \( f^n \) stabilize. Then, replacing \( f \) by a suitable power \( f^n \), we can assume that \( t \) is an isomorphism. Then \( \pi = mt^{-1}e \) is an idempotent endomorphism of \( H \), since \( \pi^2 = mt^{-1}emt^{-1}e = mt^{-1}t^{-1}e = mt^{-1}e = \pi \).

Thus \( H \cong \text{Im}(\pi) \otimes \text{co} \pi H \) is a Radford biproduct. Moreover, \( \text{Im}(\pi) = \text{Im}(f) \), and \( \text{co} \pi H = \text{co} f H \).

**Proposition 4.3.** Let \( H \) be a Hopf algebra in \( \mathcal{B} \) that satisfies both chain conditions, and \( f \) a Hopf algebra endomorphism of \( H \).

(i) If the left or right \( f \)-coinvariants of \( H \) are trivial, then \( f \) is a monomorphism in \( \mathcal{B} \).

(ii) If the left or right \( f \)-invariant quotient of \( H \) is trivial, then \( f \) is an epimorphism in \( \mathcal{B} \).

**Proof.** We prove the first part. By Proposition 4.2, \( H \cong \text{Im}(f^n) \otimes \text{co} f^n H \) is a Radford biproduct for some \( n \). If \( \text{co} f^n H \) were trivial without \( f^n \) being monic, it would follow that \( H \) is isomorphic to a proper quotient of itself, contradicting the chain conditions. Now assume for some \( m > 1 \) that \( \text{co} f^m H \) is nontrivial. Let \( j : \text{co} f^m H \to H \) be the inclusion. By exactness of tensor products in \( \mathcal{B} \), we have an equalizer

\[
0 \longrightarrow \text{co} f^m H \otimes H \longrightarrow H \xrightarrow{(f \otimes H) \Delta \otimes H} H \otimes H \otimes H
\]

and by the calculation

\[
((f^{m-1} \otimes H) \Delta \otimes H) (f \otimes H) \Delta j = (f^m \otimes f \otimes H) (\Delta \otimes H) \Delta j
\]

\[
= (f^m \otimes (f \otimes H) \Delta) \Delta j = \eta \otimes (f \otimes H) \Delta j
\]

we see that \( (f \otimes H) \Delta j \) factors through this equalizer. We conclude that if \( (f \otimes H) \Delta j \) were not trivial, then it would follow that \( \text{co} f^{m-1} H \otimes H \) is not \( I \otimes H \), which implies that \( \text{co} f^{m-1} H \) is nontrivial. We can conclude by induction that \( \text{co} f^n H \) is nontrivial after all. \( \square \)

**Remarks 4.4.** Let \( f : H \to G \) be a Hopf algebra homomorphism in \( \mathcal{B} \).

(i) Clearly, if \( f \) is a monomorphism in \( \mathcal{B} \), then it is a monomorphism in \( \text{HopfAlg}(\mathcal{B}) \).

(ii) If \( f \) has trivial left or right coinvariants, then \( f \) is a monomorphism in \( \text{Coalg}(\mathcal{B}) \).

(iii) If \( f \) is normal, and a monomorphism in \( \text{HopfAlg}(\mathcal{B}) \), then \( f \) has trivial left and right coinvariants.

Thus the preceding result shows that normal endomorphisms of a Hopf algebra in \( \mathcal{B} \) satisfying both chain conditions are monic (epic) if and only if they are so considered as morphisms in \( \mathcal{B} \).
Proof. If $C$ is a coalgebra and $g, h : C \to H$ are coalgebra morphisms with $fg = fh$, then

\[
\begin{align*}
C & \xrightarrow{g} H \otimes G = C \xrightarrow{h} H \otimes G = C \xrightarrow{g} H \otimes G = C \xrightarrow{h} H \otimes G \xrightarrow{g} H \otimes G = C \xrightarrow{h} H \otimes G = C \xrightarrow{g} H \otimes G
\end{align*}
\]

and thus, if $H^{co f}$ is trivial, $g \ast Sh = \eta \varepsilon$, whence $g = h$. If $f$ is normal, then the coinvariants are a Hopf subalgebra. \hfill \Box

Remark 4.5. In general it is false that monic is equivalent to trivial coinvariants, or that epic is equivalent to trivial invariants. In finite dimensions these concepts agree by the Nichols–Zoeller theorem [1989]; see also [Scharfschwerdt 2001]. In infinite dimensions, however, counterexamples are known [Chirvăsitu 2010].

Lemma 4.6. Let $H$ be a Hopf algebra in $\mathcal{B}$ that satisfies both chain conditions. Assume further that the braiding $\tau_{HH}$ has finite order. Then the antipode of $H$ is an automorphism in $\mathcal{B}$.

Proof. Depict the iterates of the antipode by

\[
S^m = \underbrace{m \, m \, m \, \ldots}_{m \text{ times}}.
\]

One has

\[
\underbrace{m \, m \, m \, \ldots \, m}_{m \text{ times}} \xrightarrow{\tau^m} m.
\]

Using this, we can show inductively that the coinvariants of $H$ under an iterate of the antipode are trivial as follows. Let $t : T \to H$ be a morphism factoring through $^{coS^{2n}}H$, i.e., $(S^{2n} \otimes H) \Delta t = \eta \otimes t$. We will show that $(S^m \otimes H) \Delta t = \eta \otimes t$ for any $m$, whence (taking $m = 0$) $t = \eta \varepsilon t$. 
Assume \((S^{m+1} \otimes H) \Delta t = \eta \otimes t\), or pictorially

\[
\begin{array}{c}
\bullet & \overset{T}{\bullet} \\
\otimes m & \quad H \\
\Delta H & \otimes H \quad H
\end{array}
\]

Then

\[
\begin{array}{c}
\bullet & \overset{T}{\bullet} \\
\otimes m & \quad H \\
\Delta H & \otimes H \quad H
\end{array}
\]

Since the braiding on \(H\) has finite order by assumption, some even power of the antipode is a Hopf algebra endomorphism of \(H\). Therefore that even power of the antipode is a monomorphism in \(B\). By the dual reasoning it is also an epimorphism, and therefore \(S\) itself is an automorphism in \(B\).  

5. Fitting’s lemma

**Proposition 5.1** (Fitting’s lemma). Let \(H\) be a Hopf algebra, and \(f\) a Hopf algebra endomorphism of \(H\). Assume that \(H\) satisfies both chain conditions, so that there is an \(n \in \mathbb{N}\) such that \(H \cong \text{Im}(f^n) \otimes \text{co}f^n H\) is a Radford biproduct.

If \(f\) is normal, the action of \(\text{Im}(f^n)\) on \(\text{co}f^n H\) is trivial, so that, as an algebra, \(H\) is the tensor product of \(\text{Im}(f^n)\) and \(\text{co}f^n H\) in \(B\). Similarly if \(f\) is conormal, then the coalgebra \(H\) is a tensor product of coalgebras in \(B\). In particular, if \(f\) is binormal then \(\text{Im}(f^n)\) and \(\text{co}f^n H\) are unbraided Hopf algebras in \(B\), and \(H\) is isomorphic to their tensor product.

**Proof.** We continue the proof of Proposition 4.2, assuming that \(\text{Ker}(f^2) = \text{Ker}(f)\) and \(\text{Coker}(f^2) = \text{Coker}(f)\) after replacing \(f\) by a power of \(f\). We now add the observation that normality of \(f\) implies that \(p = (f \ast (Sf \ast \text{id}_H))p = (Sf \ast \text{id}_H)p\). Therefore \(f \triangleright p\), and dually \(f \triangleleft p\) if \(f\) is conormal. This in turn implies that the Radford biproduct is just an ordinary tensor product algebra or tensor product coalgebra, as appropriate.

**Definition 5.2.** Let \(H\) be a Hopf algebra. If \(H \cong A \otimes B\) for two Hopf algebras \(A\) and \(B\), we say that \(A\) is a tensor factor of \(H\). (This implies that \(A\) and \(B\) are unbraided.) We say that \(H\) is tensor indecomposable if it does not have a nontrivial tensor factor. An endomorphism \(f\) of \(H\) is nilpotent if there is an \(n \in \mathbb{N}\) such that \(f^n = \eta \varepsilon\).
Corollary 5.3. If $H$ is a tensor indecomposable Hopf algebra satisfying both chain conditions, then every binormal endomorphism of $H$ is nilpotent or an automorphism.


Of course, a Hopf algebra satisfying both chain conditions can be (inductively) decomposed as a tensor product of indecomposable Hopf subalgebras. We shall now show that the Hopf algebraic analog of the Krull–Remak–Schmidt theorem asserting the uniqueness of such a decomposition also holds. A version of this for completely reducible semisimple Hopf algebras was established in [Burciu 2011]. In general, it cannot be hoped that this result has a categorical version. In [Müger 2003] it was shown that a nondegenerate fusion category factorizes into a product of prime ones, but that this was generally not unique. Therefore, such decompositions are rather specific to Hopf algebras.

Lemma 6.1. Let $f, g$ be bicommuting, binormal endomorphisms of a tensor indecomposable Hopf algebra $H$.

If $f$ and $g$ are nilpotent, then so is $f \ast g$.

Proof. Otherwise $f \ast g$ is a normal automorphism, and after composing $f$ and $g$ with its inverse, we can assume that $f \ast g = \text{id}_H$. In particular $f$ composition commutes with $g$. Then one can show by induction that $\text{id}_H = (f \ast g)^n$ is a convolution product of terms of the form $f^k g^{n-k}$ for $0 \leq k \leq n$ (in fact this is a binomial formula with binomial coefficients, but writing it is cumbersome because addition is replaced with convolution products). If $f^m = \eta \varepsilon = g^m$, this implies $(f \ast g)^{2m} = \eta \varepsilon$, since each term contains an $m$-th power of either $f$ or $g$. □

Remark 6.2. Let $H$ and $H_1, \ldots, H_k$ be Hopf algebras in $B$. Decomposing $H$ as a tensor product Hopf algebra

\[ H \cong H_1 \otimes H_2 \otimes \cdots \otimes H_k \]

amounts to specifying a system of injections $\iota_i : H_i \to H$ and projections $\pi_i : H \to H_i$, all of them Hopf algebra morphisms, which commute and cocommutate pairwise and satisfy $\pi_i \iota_i = \text{id}_{H_i}, \pi_i \iota_j = \eta \varepsilon$ if $i \neq j$, and $\iota_1 \pi_1 \ast \iota_2 \pi_2 \ast \cdots \ast \iota_k \pi_k = \text{id}_H$. The isomorphisms between $H$ and the tensor product are then given by

\[ H \xrightarrow{\Delta^{(k-1)}} H \otimes^k \xrightarrow{\pi_1 \otimes \cdots \otimes \pi_k} H_1 \otimes \cdots \otimes H_k \]

and

\[ H_1 \otimes \cdots \otimes H_k \xrightarrow{\iota_1 \otimes \cdots \otimes \iota_k} H \otimes^k \xrightarrow{\nabla^{(k-1)}} H. \]

Note that the $H_i$ need to be pairwise unbraided for the tensor product Hopf algebra to make sense.
\textbf{Theorem 6.3.} Let $H$ be a Hopf algebra in $\mathcal{B}$, and let
\[ H = H_1 \otimes H_2 \otimes \cdots \otimes H_k \]
\[ = G_1 \otimes G_2 \otimes \cdots \otimes G_\ell \]
be two tensor decompositions of $H$ in tensor indecomposable factors.

Then $k = \ell$, and $H_i \cong G_i$ after a suitable permutation of the indices.
Moreover, letting
\[ H_i \xleftrightarrow{\pi_i} H \xleftrightarrow{q_j} G_j \]
denote the systems of injections and projections associated to the decompositions into tensor factors, then the factors can be so numbered that for any $1 \leq m \leq k$
\[ H \xrightarrow{\triangle^{(k-1)}} H^\otimes k \xrightarrow{\pi_1 \otimes \cdots \otimes \pi_m \otimes p_{m+1} \otimes \cdots \otimes p_k} H_1 \otimes \cdots \otimes H_m \otimes G_{m+1} \otimes \cdots \otimes G_k \quad (6-1) \]
and
\[ H_1 \otimes \cdots \otimes H_m \otimes G_{m+1} \otimes \cdots \otimes G_k \xrightarrow{\ell_1 \otimes \cdots \otimes \ell_m \otimes q_{m+1} \otimes \cdots \otimes q_k} H^\otimes k \xrightarrow{V^{(k-1)}} H \quad (6-2) \]
are isomorphisms.

\textbf{Proof.} There is nothing to show if one of the decompositions consists of only one factor. Otherwise we consider
\[ \text{id}_{H_1} = \pi_1 \ell_1 = \pi_1 (q_1 p_1 \ast \cdots \ast q_\ell p_\ell) \ell_1 = \pi_1 q_1 p_1 \ell_1 \ast \cdots \ast \pi_1 q_\ell p_\ell \ell_1. \]
Since $H_1$ is indecomposable, and the terms in the last convolution product are bicommuting binormal endomorphisms, we know that one of $\pi_1 q_j p_j \ell_1$ is an isomorphism. Without loss of generality we assume this happens for $j = 1$, and that $\pi_1 q_1 p_1 \ell_1 = \text{id}_{H_1}$. It follows that $\pi_1 q_1$ and $p_1 \ell_1$ are mutually inverse isomorphisms between $H_1$ and $G_1$. Now put $f = q_2 p_2 \ast \cdots \ast q_\ell p_\ell$ and $t = \ell_1 \pi_1 q_1 p_1 \ast f$. Since $p_1 t = p_1 \ell_1 \pi_1 q_1 p_1 = p_1$, we have $H^{\text{co}t} \subset H^{\text{co} p_1 t} = H^{\text{co} p_1}$. Thus, for $j : H^{\text{co} t} \to H$ the inclusion, we find
\[ \Delta j = (H \otimes q_1 p_1 \ast \cdots \ast q_\ell p_\ell) \Delta j = (H \otimes f) \Delta j = (H \otimes t) \Delta j = (H \otimes \eta) j, \]
and therefore $H^{\text{co}t}$ is trivial. We conclude that $t$ is an automorphism of $H$.

Write $\tilde{\pi} : H \to H_2 \otimes \cdots \otimes H_k =: \tilde{H}$ and $\tilde{\iota} : \tilde{H} \to H$ for the natural projection and injection morphisms, and similarly for $\tilde{p} : H \to \tilde{G}$, $\tilde{q} : \tilde{G} \to H$. Since $t q_1 = \ell_1 \pi_1 q_1$, we have $\tilde{\pi} t q_1 = \eta \varepsilon$, and thus $\tilde{\pi} t = \tilde{\pi} t \tilde{\eta} \tilde{p}$ and $\tilde{\pi} t \tilde{\eta} \tilde{p} t^{-1} \tilde{\iota} = \tilde{\pi} \tilde{\iota} = \text{id}_{\tilde{H}}$. It follows that $\tilde{\pi} \tilde{\eta}$ and $\tilde{p} \tilde{t}^{-1} \tilde{\iota}$ are mutually inverse isomorphism between $\tilde{G}$ and $\tilde{H}$.

Thus, by an inductive argument we have $k = \ell$, and we can rearrange the indices to get $H_i \cong G_i$ for all $i$. 


Note further that the automorphism $t$ above is the composition of the isomorphism

$$
H \xrightarrow{\Delta^{(k-1)}} H \otimes H \xrightarrow{p_1 \otimes \cdots \otimes p_k} G_1 \otimes \cdots \otimes G_k \xrightarrow{\pi_1 \otimes G_2 \otimes \cdots \otimes G_k} H \otimes G_2 \otimes \cdots \otimes G_k \xrightarrow{\eta \otimes q_2 \otimes \cdots \otimes q_k} H \otimes k \xrightarrow{\nabla^{(k-1)}} H,
$$

with the morphism

$$
H_1 \otimes G_2 \otimes \cdots \otimes G_k \xrightarrow{\eta \otimes q_2 \otimes \cdots \otimes q_k} H \otimes k \xrightarrow{\nabla^{(k-1)}} H,
$$

whence the latter is an isomorphism. Again by an inductive argument, we get that (6-2) is an isomorphism; the reasoning for (6-1) is similar. □

7. Endomorphisms of tensor products

Let $H$ and $K$ be two Hopf algebras in $B$, unbraided so that one can form the tensor product bialgebra $H \otimes K$. Let $A$ be an algebra in $B$. It is well known that there is a bijection

$$\text{Alg}(H \otimes K, A) \cong \{(a, b) \in \text{Alg}(H, A) \times \text{Alg}(K, A) | a \gamma b \}.$$

In fact, a pair $(a, b)$ of multiplication commuting algebra morphisms induces $f = \nabla_A(a \otimes b)$, and

$$
\begin{array}{c}
H \otimes K \\
\downarrow f \\
A
\end{array}
= 
\begin{array}{c}
H \\
\downarrow a \\
A
\end{array}
\begin{array}{c}
K \\
\downarrow b \\
A
\end{array}
= 
\begin{array}{c}
H \\
\downarrow a \\
A
\end{array}
\begin{array}{c}
K \\
\downarrow b \\
A
\end{array}
= 
\begin{array}{c}
H \\
\downarrow a \\
A
\end{array}
\begin{array}{c}
K \\
\downarrow b \\
A
\end{array}
\begin{array}{c}
K \\
\downarrow a \otimes b \\
A
\end{array}

shows that $f$ is multiplicative. Conversely, given $f : H \otimes K \to A$, define $a = f(H \otimes \eta)$ and $b = f(\eta \otimes K)$. Then, with $T := H \otimes K$:

$$
\begin{array}{c}
H \\
\downarrow a \otimes b \\
A
\end{array}
= 
\begin{array}{c}
H \\
\downarrow f \\
A
\end{array}
\begin{array}{c}
K \\
\downarrow f \\
A
\end{array}
= 
\begin{array}{c}
H \\
\downarrow f \\
A
\end{array}
= f
$$

and

$$
\begin{array}{c}
K \\
\downarrow b \otimes a \\
A
\end{array}
= 
\begin{array}{c}
K \\
\downarrow f \\
A
\end{array}
\begin{array}{c}
H \\
\downarrow f \\
A
\end{array}
= 
\begin{array}{c}
K \\
\downarrow f \\
A
\end{array}
\begin{array}{c}
H \\
\downarrow f \\
A
\end{array}
= 
\begin{array}{c}
K \\
\downarrow f \\
A
\end{array}
\begin{array}{c}
H \\
\downarrow f \\
A
\end{array}
= \begin{array}{c}
K \otimes K \\
\downarrow a \otimes b \\
A
\end{array}.
$$

Assume that $A$ is a bialgebra in $B$, and $a$, $b$, $f$ are as above. Then $f$ is a bialgebra homomorphism if and only if $a$ and $b$ are.
Dually, for a coalgebra $C$ in $B$, a bijection
\[
\{(a, b) \in \mathrm{Coalg}(C, H) \times \mathrm{Coalg}(C, H) \mid a \Join b\} \longrightarrow \mathrm{Coalg}(C, H \otimes K)
\]
is given by $(a, b) \mapsto (a \otimes b)\Delta$, and it induces bijections on the subsets containing (pairs of) bialgebra maps.

Putting the above together, one obtains a bijection between $\mathrm{End}(H \otimes K)$ and
\[
\{(a, b, c, d) \mid a \in \mathrm{End}(H), b \in \mathrm{Hom}(K, H), c \in \mathrm{Hom}(H, K), d \in \mathrm{End}(K),
\]

with the endomorphism of $H \otimes K$ corresponding to a quadruple of Hopf algebra map “components” given by
\[
\begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{pmatrix} H & K \\ \downarrow & \downarrow \\ \downarrow & \downarrow \\ H & K \end{pmatrix}.
\]

Consider a second endomorphism $g$ of $H \otimes K$ dissected analogously into a matrix $\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$ of Hopf algebra endomorphisms. Then it is straightforward to check that $gf$ corresponds to $\begin{pmatrix} a' & a \ast b' & c' & d' \\ c' & a & b' & d' \end{pmatrix}$.

**Proposition 7.1.** Let $H$ and $K$ be as above, and $f \in \mathrm{End}(H \otimes K)$ described by a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Assume that the antipodes of $H$ and $K$ are automorphisms in $B$.

Then $f$ is normal if and only if $a$ and $c$ are normal, $b \Join \mathrm{id}_H$, and $d \Join \mathrm{id}_K$; a similar characterization holds for conormal endomorphisms.

**Proof.** We fix projections and injections for the tensor product $P := H \otimes K$:
\[
H \underset{\pi_H}{\overset{i_H}{\rightleftarrows}} P \underset{\pi_K}{\overset{i_K}{\rightleftarrows}} K
\]

First assume that $f$ is normal. Since $f \Join (fS \ast \mathrm{id}_P)$, $a = \pi_H f i_H$ commutes with $\pi_H(fS \ast \mathrm{id}_P)i_H = aS \ast \mathrm{id}_H$. Similarly $c$ is normal. Using (3-1) we have

so that $b \Join \mathrm{id}_H$. Similarly $d \Join \mathrm{id}_K$. 
Now suppose that the stated normality and commutation conditions on \(a, b, c, d\) hold. Writing \(\hat{a} = \iota_H a \pi_H, \hat{b} = \iota_H b \pi_K\) etc. we can write \(f = \hat{a} \ast \hat{b} \ast \hat{c} \ast \hat{d}\) as a convolution product of four commuting and cocommuting endomorphisms of \(P\).

We are claiming that this product commutes with \(f \ast \text{id}_P = \hat{a}_S \ast \hat{b}_S \ast \hat{c}_S \ast \hat{d}_S \ast \iota_K \pi_K \ast \iota_H \pi_H\).

(The last equality using that \(\hat{a}\) bicommutes with \(\hat{b}, \hat{c}, \hat{d}\), and \(\iota_K\).) Now \(\hat{a}\) commutes with \(\hat{a}_S \ast \iota_H \pi_H\) since \(a\) is normal, with \(\hat{b}_S\) since \(a \overset{\gamma}{\ast} b\), and with \(\iota_K \pi_K\) and \(\hat{c}_S\) since \(\iota_H \overset{\gamma}{\ast} \iota_K\). The next factor \(\hat{b}\) commutes with \(\hat{a}_S, \hat{b}_S,\) and \(\iota_H \pi_H\) since \(b \overset{\gamma}{\ast} \text{id}_H\), and it commutes with \(\hat{c}_S, \hat{d}_S,\) and \(\iota_K \pi_K\), since \(\iota_K \overset{\gamma}{\ast} \iota_H\). Similar arguments deal with the convolution factors \(\hat{c}\) and \(\hat{d}\).

\(\square\)

**Remark 7.2.** Similarly, an endomorphism \(f\) of a tensor product of several pairwise unbraided Hopf algebras \(H_1, \ldots, H_k\) can be described by a matrix \((v_{ij})\) of Hopf algebra homomorphisms between the factors. By inductive arguments one can show that \(f\) is normal if and only if all the diagonal terms are normal, and the off-diagonal terms commute with the identities on their targets.

An interesting case arises when there are no nontrivial homomorphisms \(H_i \to H_j\) commuting with \(\text{id}_{H_j}\). In this case any normal endomorphism preserves the decomposition into tensor factors. One can deduce from this that the Krull–Remak–Schmidt decomposition is unique in a stronger sense than up to permutation and isomorphism; in the original case of decompositions of groups the uniqueness of the subgroups in a direct decomposition into directly indecomposable factors follows as stated in Remak’s thesis [1911].

### 8. Automorphisms of tensor products

We consider now the automorphisms of tensor products of Hopf algebras. These are the natural extensions of the corresponding results in group theory [Bidwell et al. 2006; Bidwell 2008].

Throughout this section we let \(H\) and \(K\) be unbraided Hopf algebras, so that we can form the tensor product \(H \otimes K\), and we assume that the antipodes of \(H\) and \(K\) are automorphisms in \(B\).

Identify endomorphisms of \(H \otimes K\) with matrices of Hopf algebra homomorphisms as in Section 7. Let \(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{End}(H \otimes K)\). If \(a\) is an automorphism, then by (c) of Lemma 2.1 the condition \(a \overset{\gamma}{\ast} b\) implies \(\text{id}_H \overset{\gamma}{\ast} b\) (and \(x \overset{\gamma}{\ast} b\) for any \(x : X \to H\)). Similarly \(\text{id}_K \overset{\gamma}{\ast} c\), and, if \(d\) is also an automorphism, then \(b \overset{\gamma}{\ast} \text{id}_H\) and \(c \overset{\gamma}{\ast} \text{id}_K\).

Define

\[
\mathcal{A} = \begin{pmatrix}
\text{Aut}(H) & \text{Hom}^c(K, H) \\
\text{Hom}^c(H, K) & \text{Aut}(K)
\end{pmatrix},
\]

where \(c\) is cohomology.
where $\text{Hom}^c(K, H) := \{ b \in \text{Hom}(K, H) \mid b \triangleright \text{id}_H \text{ and } b \triangleleft \text{id}_H \}$. This is easily seen to be an abelian group under convolution product. Indeed, the image of any such morphism determines an abelian Hopf subalgebra of $H$. Note that $b \triangleright \text{id}_H \iff b \triangleright \alpha$ for some/all $\alpha \in \text{Aut}(H)$, and similarly $b \triangleleft \text{id} \iff b \triangleleft \alpha$ for some/all $\alpha \in \text{Aut}(H)$.

Consider an automorphism $f$ of $H \otimes K$, and its decomposition as a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of Hopf algebra homomorphisms as in Section 7. Let $f^{-1}$ correspond in the same way to a matrix $\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$. Then we have $\text{id}_K = (\varepsilon \otimes K)f^{-1}f(\eta \otimes K) = c'b * d'd$. Since $c' \triangleright d'$ and $b \triangleleft d$, we have that $c'b \triangleright d'd = (c'bS) * \text{id}_K$ and $c'b \triangleleft (c'bS) * \text{id}_K$. In other words, $c'b$ is a binormal endomorphism of $K$. In the same way $bc'$ is a binormal endomorphism of $H$. If we assume both chain conditions on $H$ and $K$, then for $n$ sufficiently large $b$ and $c'$ induce mutually inverse isomorphisms between the images of $(c'b)^n$ and $(bc')^n$. Thus, using Fitting’s lemma, the image of $(c'b)^n$ is a common tensor factor of $H$ and $K$.

This gives part of the following result.

**Theorem 8.1.** Suppose that $H$ and $K$ satisfy both chain conditions. Then $A \subseteq \text{Aut}(H \otimes K)$ if and only if $H$ and $K$ have no nontrivial common abelian direct tensor factors. On the other hand, $\text{Aut}(H \otimes K) = A$ if and only if $H$ and $K$ have no nontrivial common direct tensor factors.

**Proof.** If $H$ and $K$ have a common nontrivial direct tensor factor, then permutations of this factor in $H \otimes K$ are automorphisms of $H \otimes K$ not contained in $A$.

By the preceding remarks, to show $\text{Aut}(H \otimes K) \subseteq A$ it remains to prove that the common tensor factor in $H \otimes K$ that we found is necessarily nontrivial if $d$ is not an automorphism. A similar argument will apply to show that $a$ is an automorphism, and the commutation and cocommutation conditions for the components of an endomorphism will be equivalent to the off-diagonal terms (co)commuting with the identity instead of the automorphisms on the diagonal.

Thus suppose that $d$ is not an automorphism. Then we can assume without loss of generality that the right $d$-coinvariant subobject $D$ of $H$ is nontrivial. If $\iota: D \rightarrow H$ is the inclusion, then $c'bi = \triangledown (c'b \otimes \eta)\iota = \triangledown (c'b \otimes d'd)\Delta \iota = (c'b * d'd)\iota = \iota$, hence $(c'b)^n\iota = \iota$ for all $n$, and the image of $(c'b)^n$ is nontrivial as desired.

The desired equality in the second part will then hold once we have proven the first equivalence.

To this end we first consider the forward direction by contrapositive. Suppose that $H$ and $K$ have a common abelian direct tensor factor $L$, and write $H = H' \otimes L$ and $K = K' \otimes L$. Since $L$ is abelian, its antipode $S_L$ is a Hopf endomorphism of $L$. Taking $a = \text{id}_H$, $b = \text{id}_K$, $b = \eta_K, \varepsilon_K, S_L$, and $c = \eta_H, \varepsilon_H, \otimes S_L$ we find that $\psi = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in A$. However, $L$ is a subobject of the right $\psi$-coinvariant subobject, whence $\psi \not\in \text{Aut}(H \otimes K)$.
For the remaining direction, assume that \( f = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) belongs to \( \mathcal{A} \); in particular \( f \) is a Hopf algebra endomorphism of \( H \otimes K \). After multiplying with the obvious automorphism of \( H \otimes K \), namely \( \left( a^{-1} \eta e, d^{-1} \eta e \right) \), we may assume that \( a = \text{id}_H \) and \( d = \text{id}_K \). Now consider \( g = \left( cS_H \eta e, bS_K \eta e \right) \), another Hopf endomorphism of \( H \otimes K \). One computes \( gf = \left( cS_H \eta e, bS_K \eta e \right) = \left( dS_K \eta e, cS_H \eta e \right) \). By the chain conditions on \( H \) and \( K \), for \( n \) sufficiently large \( b \) and \( c \) induce mutually inverse isomorphisms between the images of \( (bc)^n \) and \( (cb)^n \). Fitting’s lemma implies that these isomorphic images are an abelian common tensor factor of \( H \) and \( K \). It can only be trivial if \( bc \) and \( cb \) are nilpotent, in which case \( \text{id} \ast bcS \) and \( cbS \ast \text{id} \) are automorphisms. In the latter case, \( f \) was an automorphism. \( \square \)

These results have obvious extensions to more than two factors by induction, which we leave to the reader. The results, however, do not cover the case of a (sub)category where the braiding is a symmetry.

**Theorem 8.2.** Let \( H \) be a tensor indecomposable nonabelian Hopf algebra satisfying both chain conditions in \( B \), and suppose the braiding of \( B \) is a symmetry. Fix \( n \in \mathbb{N} \), and let \( \mathcal{A}_n \) denote those \( \left( \alpha_{ij} \right) \in \text{End}(H^{\otimes n}) \) such that \( \alpha_{ii} \in \text{Aut}(H) \) and \( \alpha_{ij} \in \text{End}^c(H) \) for all \( i \) and \( j \neq i \). Then

\[
\text{Aut}(H^{\otimes n}) \cong \mathcal{A}_n \rtimes S_n.
\]

**Proof.** By assumptions on \( H \), \( \mathcal{A}_n \subseteq \text{Aut}(H^{\otimes n}) \). The group \( S_n \) acts on \( H^{\otimes n} \) by permuting factors, and so acts on \( \text{Aut}(H^{\otimes n}) \) by permuting columns. Conjugating by this action sends \( \mathcal{A}_n \) to itself. We need only show that every automorphism is a column permutation of an element of \( \mathcal{A}_n \).

So let \( \left( \alpha_{ij} \right) \in \text{Aut}(H^{\otimes n}) \), with inverse \( \left( \alpha'_{ij} \right) \). Then \( \alpha_{i1} \alpha'_{1i} \ast \cdots \ast \alpha_{in} \alpha'_{ni} = \text{id} \) holds for all \( i \). Since the \( \alpha_{ik} \alpha'_{ki} \) are all binormal endomorphisms the notation is unambiguous, and the terms of the convolution product can be arbitrarily reordered. Moreover, since \( H \) is indecomposable we may conclude that one of the \( \alpha_{ik} \alpha'_{ki} \) is an automorphism. In particular for all \( i \) there is a \( k \) such that \( \alpha_{ik} \) is an epimorphism and \( \alpha'_{ki} \) is a monomorphism. By the chain conditions it follows that \( \alpha_{ik} \) and \( \alpha'_{ki} \) are both automorphisms. Since \( H \) is nonabelian there is at most one such \( k \) for any given \( i \). This completes the proof. \( \square \)

9. Application to doubles of groups

For this section we work in the category of vector spaces over a field \( \mathbb{k} \). Throughout this section \( G, H, K, C \) will all be finite groups. For any group \( G \) let \( \hat{G} \) be the group of group-like elements of \( \mathbb{k}^G \), the dual of the group algebra \( \mathbb{k}G \). Note that \( \hat{G} \) is precisely the \( \mathbb{k} \)-linear characters of \( G \). We also define \( \Gamma_G = \hat{G} \times G \). We denote
The morphism is defined by $$k$$ whenever $$D$$ are all subject to the following compatibility relations:

(i) $$u$$ is an abelian normal subgroup of $$G$$, and the determination of $$Aut(D(G))$$ whenever $$G$$ has no nontrivial abelian direct factors. Such a group is said to be purely nonabelian. When $$G$$ is abelian we have $$\mathcal{D}(G) = kG$$, an abelian Hopf algebra, and the determination of $$Aut(D(G))$$ is then straightforward. Indeed, under mild assumptions on $$k$$ we have $$\mathcal{D}(G) \cong k(G \times G)$$. Subsequently in this case $$Aut(D(G))$$ can be computed by classical methods in group theory [Shoda 1928]. We note that the structure of such an automorphism group has been of more recent interest [Bidwell and Curran 2010; Hillar and Rhea 2007]. It is the goal of this section to complete the description of $$Aut(D(G))$$ when $$G$$ has an abelian direct factor.

So suppose that $$G = C \times H$$ with $$C$$ abelian. Then $$\mathcal{D}(G) \cong \mathcal{D}(C) \otimes \mathcal{D}(H)$$. Since $$\mathcal{D}(C)$$ is an abelian Hopf algebra the results of the previous section can be applied whenever $$\mathcal{D}(H)$$ has no abelian direct tensor factors. We will proceed to show this happens precisely when $$H$$ is purely nonabelian.

We have the following description of $$\text{Hom}(\mathcal{D}(G), \mathcal{D}(K))$$.

**Theorem 9.1** [Keilberg 2015; Agore et al. 2014]. The elements of the set $$\text{Hom}(\mathcal{D}(G), \mathcal{D}(K))$$ are in bijective correspondence with matrices $$(u, r, v)$$ where $$u : kG \to kK$$ is a morphism of unitary coalgebras, $$p : kG \to kK$$ is a morphism of Hopf algebras, and $$r : G \to K$$ and $$v : G \to K$$ are group homomorphisms. These are all subject to the following compatibility relations, for all $$a, b \in kG$$ and $$g \in G$$:

(i) $$u(g \mapsto a) = v(g) \mapsto u(a)$$, from which it follows that $$u^*v$$ is normal;

(ii) $$u \wr p$$;

(iii) $$u(ab) = u(a(1))(p(a(2)) \mapsto u(b))$$;

(iv) $$p(g \mapsto a) = v(g) \mapsto p(a)$$.

The morphism is defined by

$$a \# g \mapsto u(a(1))r(g) \# p(a(2))v(g).$$

Composition of such morphisms is given by matrix multiplication, as in Section 7.

The morphism $$p$$ is uniquely determined by an isomorphism $$kA \cong kB$$, where $$A$$ is an abelian normal subgroup of $$G$$ and $$B$$ is an abelian subgroup of $$K$$. In particular we must have $$kA \cong \hat{kA}$$. For the remainder of this section any use of $$A, B$$ refers to these subgroups. We note that the last relation says $$p \wr v$$ if and only if $$A \leq Z(G)$$, or equivalently $$p$$ is cocentral: $$p \wr \text{id}.$$

the conjugation action of $$G$$ on $$kG$$ and $$kG$$ both by $$\mapsto$$; for instance: $$g \mapsto x = gxg^{-1}$$ for all $$g, x \in G$$. We will be concerned with $$\mathcal{D}(G)$$, the Drinfeld double of a finite group $$G$$. As a coalgebra $$\mathcal{D}(G) = \mathcal{k}^{Gco} \otimes kG$$, and the algebra structure is given by having $$G$$ act on $$\mathcal{k}^{Gco}$$ by the conjugation action. We note that $$\Gamma_G$$ is the group of group-like elements of $$\mathcal{D}(G)$$. See [Dijkgraaf et al. 1991; Montgomery 1993] for further details on the construction and properties of this Hopf algebra.

In [Keilberg 2015] the first author gave a complete description of $$Aut(D(G))$$ whenever $$G$$ has no nontrivial abelian direct factors. Such a group is said to be purely nonabelian. When $$G$$ is abelian we have $$\mathcal{D}(G) = kG$$, an abelian Hopf algebra, and the determination of $$Aut(D(G))$$ is then straightforward. Indeed, under mild assumptions on $$k$$ we have $$\mathcal{D}(G) \cong k(G \times G)$$. Subsequently in this case $$Aut(D(G))$$ can be computed by classical methods in group theory [Shoda 1928].
By convention we implicitly identify any element of $\text{Hom}(D(G), D(K))$ or $\text{Hom}(k^G \otimes kG, k^K \otimes kK)$ with its quadruple of components $(u, r, p, v)$, or equivalently as a matrix \(egin{pmatrix} u & r \\ p & v \end{pmatrix}\).

The following is then immediate.

**Lemma 9.2.** A morphism $\psi \in \text{Hom}(D(G), D(K))$ is canonically an element of $\text{Hom}(k^G \otimes kG, k^K \otimes kK)$ precisely when $p \psi v$ and $u$ is a morphism of Hopf algebras. On the other hand, $\phi \in \text{Hom}(k^G \otimes kG, k^K \otimes kK)$ is canonically an element of $\text{Hom}(D(G), D(K))$ precisely when $u^* v$ is normal and $A \leq Z(G)$.

In the first case we call such a morphism untwistable, and in the second we call it twistable. Clearly any untwistable morphism is also twistable, and vice versa. The distinction is simply in the algebra structures we start with.

Now since $k^G \otimes kG$ and $k^K \otimes kK$ are canonically self-dual, any morphism $\psi \in \text{Hom}(k^G \otimes kG, k^K \otimes kK)$ yields a dual morphism $\psi^* \in \text{Hom}(k^K \otimes kK, k^G \otimes kG)$ with components $(v^*, r^*, p^*, u^*)$. The following is then clear.

**Corollary 9.3.** Both $\psi \in \text{Hom}(k^G \otimes kG, k^K \otimes kK)$ and $\psi^*$ are twistable if and only if the following all hold:

(i) $u^* v$ is normal;
(ii) $v u^*$ is normal;
(iii) $A \leq Z(G)$;
(iv) $B \leq Z(K)$.

In this case we may canonically view $\psi$ as an element of $\text{Hom}(D(G), D(K))$ and $\psi^*$ as an element of $\text{Hom}(D(K), D(G))$.

In [Keilberg 2015] a morphism $\psi = (u, r, p, v) \in \text{Hom}(D(G), D(K))$ was said to be flippable if also $(v^*, r^*, p^*, u^*) \in \text{Hom}(D(K), D(G))$. This is equivalent to saying that $\psi$ is untwistable and the corresponding dual $\psi^*$ is twistable. In particular the corollary gives a complete description of the flippable elements of $\text{Hom}(D(G), D(K))$, and shows that “flipping” an element of $\text{Hom}(D(G), D(K))$ can naturally be described as dualizing the morphism.

**Corollary 9.4.** For any group $G$, the group $\text{Aut}(D(G))$ is canonically a subgroup of $\text{Aut}(k^G \otimes kG)$ which is closed under dualization.

**Proof.** Follows from the preceding corollary, Section 8, and the properties of $\text{Aut}(D(G))$ established in [Keilberg 2015].

We now show that the act of untwisting a morphism is fairly well behaved whenever the image is commutative.

**Proposition 9.5.** Let $\psi \in \text{Hom}(D(G), D(K))$ be untwistable. For convenience, let $\psi' = \psi \in \text{Hom}(k^G \otimes kG, k^K \otimes kK)$. Then the following all hold.
(i) **If** \( G = H \), **then** \( \psi \) **is conormal if and only if** \( \psi' \) **is conormal.**

(ii) **If** \( \psi \) **has a commutative image then** \( \psi' \) **has commutative image.**

(iii) **If** \( \psi \) **has commutative image and** \( G = H \), **then** \( \psi' \) **is normal if and only if** \( v \) **is normal and** \( B \leq Z(G) \).

(iv) **If** \( \psi \) **has commutative image and** \( G = H \), **then** \( \psi' \) **is normal if and only if** \( \psi' \) **is normal and** \( G \) **acts trivially on** \( \text{Img}(u) \).

(v) **If** \( \psi \) **has commutative image and** \( G = H \), **then** \( \psi \) **is conormal.**

**Proof.** We first prove (i), as it is the only one that does not suppose that \( \psi \) has commutative image. To this end we compute

\[
a_{(3)} \# g \cdot S(a_{(1)} \# g) \otimes a_{(2)} \# g = a_{(3)} \# g \cdot (g^{-1} \rightarrow S(a_{(1)}) \# g^{-1}) \otimes a_{(2)} \# g = a_{(3)}S(a_{(1)}) \# 1 \otimes a_{(2)} \# g; \tag{9-1}
\]

\[
a_{(3)} \otimes g \cdot S(a_{(1)} \otimes g) \otimes a_{(2)} \otimes g = a_{(3)}S(a_{(1)}) \otimes 1 \otimes a_{(2)} \otimes g. \tag{9-2}
\]

The claim then follows.

For the remainder of the proof, suppose that \( \psi \) is untwistable and has commutative image. By checking the commutativity condition we can easily determine the following facts: \( p \not\supseteq v \); \( v \) has abelian image, or equivalently \( v \not\supseteq v; u(a(g \rightarrow b)) = u((h \rightarrow a)b) = u(ab) \) for all \( g, h \in G \) and \( a, b \in \mathbb{k}^G \), which implies \( u \not\supseteq \text{id} \). In particular, \( u(g \rightarrow a) = u(a) \) and \( p(g \rightarrow a) = p(a) \) for all such \( a, g \).

The first part is then immediate, as we have \( \psi(a \# g \cdot b \# h) = \psi'(a \otimes g \cdot b \otimes h) \). Another way of saying this is that when \( \psi \) has commutative image we may compute products in either the double or the tensor product without affecting the result. Furthermore \( \psi((g \rightarrow a) \# g) = \psi(a \# g) \) for all appropriate \( a, g \), and so

\[
\psi(S(a \# g)) = \psi(g^{-1} \rightarrow S(a) \# g^{-1}) = \psi(S(a) \# g^{-1}) = \psi'(S(a \otimes g)).
\]

Thus we may perform all computations with \( \psi \) in either \( \mathcal{D}(G) \) or \( \mathbb{k}^G \otimes \mathbb{k}G \) as we desire.

The last part of the result follows from Equations (9-1) and (9-2) and \( u \not\supseteq \text{id} \). We need only prove the parts concerning normality of \( \psi, \psi' \).

To determine when \( \psi' \) is normal, we first note that by commutativity we have

\[
\psi'(a_{(2)} \otimes g \cdot b \otimes h \cdot S(a_{(1)}) \otimes g^{-1}) = a(1)\psi'(b \otimes h).
\]

On the other hand,

\[
a_{(2)} \otimes g \cdot \psi'(b \otimes h) \cdot S(a_{(1)}) \otimes g^{-1} = a(1)r(h)u(b_{(1)}) \otimes gp(b_{(1)})v(h)g^{-1}.
\]

The map \( \psi' \) is normal precisely when these two expressions are the same, and we easily find this is equivalent to \( B \leq Z(H) \) and \( v \) normal.
Finally, we determine when $\psi$ is normal. By previous remarks, we have
\[
\psi((a_2 \# g \cdot b \# h \cdot S(a_1 \# g)) = a(1)\psi(b \# ghg^{-1})
\]
\[
= a(1)r(h)u(b_{(1)}) \# p(b_{(2)})v(ghg^{-1}). \tag{9-3}
\]
On the other hand, we have
\[
a_{(2)} \# g \cdot \psi(b \# h) \cdot (g^{-1} \to S(a_{(1)} \# g^{-1})
\]
\[
= r(h)a_{(2)}(g \to u(b_{(1)}))(gp(b_{(2)})v(h)g^{-1} \to S(a_{(1)})) \# gp(b_{(3)})v(h)g^{-1}.
\]
Applying $\varepsilon \# \text{id}$ to both expressions we get
\[
a(1)p(b)v(ghg^{-1})
\]
for the first and
\[
a(1)gp(b)v(h)g^{-1}
\]
for the second. These are equal for all $a, b, g, h$ if and only if $B \leq Z(H)$ and $v$ is normal; equivalently, $\psi'$ is normal. Note that if $v$ is normal and has abelian image, then its image is in fact central. Therefore $gp(b)v(h)g^{-1} \to S(a) = S(a)$ precisely when $\psi'$ is normal. Subsequently the previous equation simplifies to
\[
a(1)r(h)(g \to u(b_{(1)})) \# p(b_{(2)})v(ghg^{-1}).
\]
Comparing with (9-3) completes the proof. \hfill \qed

**Lemma 9.6.** Any commutative direct tensor factor of $\mathcal{D}(G)$ is also a commutative direct tensor factor of $\mathbb{k}^G \otimes \mathbb{k}G$.

**Proof.** Suppose $L$ is a commutative Hopf subalgebra of $\mathcal{D}(G)$ such that $\mathcal{D}(G) = M \otimes L$ for some Hopf subalgebra $M$. We then have a projection $\pi : \mathcal{D}(G) \to L$ with associated right inverse the imbedding $i : L \hookrightarrow \mathcal{D}(G)$.

The morphism $i\pi$ is an endomorphism of $\mathcal{D}(G)$. Since the image is central in $\mathcal{D}(G)$ it is easily seen to be untwistable and binormal. Therefore $i\pi$ is canonically a twistable, binormal, idempotent endomorphism of $\mathbb{k}^G \otimes \mathbb{k}G$ with image $L$. By Fitting’s lemma we conclude that $L$ is also a direct tensor factor of $\mathbb{k}^G \otimes \mathbb{k}G$. \hfill \qed

**Remark 9.7.** Since $\mathbb{k}^G$ is commutative we see that the converse will only hold when $G$ is abelian. Indeed since $\mathcal{D}(G)$ is quasitriangular any commutative direct tensor factor of $\mathcal{D}(G)$ is necessarily abelian.

The lemma gives one part of the following.
Theorem 9.8. Let $G$ be a finite group. Then the following are equivalent.

(i) $G$ is purely nonabelian.
(ii) $\mathbb{k}G$ is purely nonabelian.
(iii) $\mathbb{k}^G$ is purely nonabelian.
(iv) $\mathbb{k}^G \otimes \mathbb{k}G$ is purely nonabelian.
(v) $\mathcal{D}(G)$ is purely nonabelian.

Indeed, $\mathbb{k}^G \otimes \mathbb{k}G$ and $\mathcal{D}(G)$ have the same abelian direct tensor factors.

Proof. Since the dual of an abelian Hopf algebra is again abelian, the equivalence of the second and third is immediate. By Krull–Remak–Schmidt, any abelian indecomposable factor of $\mathbb{k}^G \otimes \mathbb{k}G$ is isomorphic to an abelian indecomposable factor of either $\mathbb{k}G$ or $\mathbb{k}^G$. Thus the fourth is equivalent to the second and third. Since any Hopf subalgebra of $\mathbb{k}G$ is a subgroup algebra, the first and second are equivalent. By the lemma the fourth implies the fifth. To prove the fifth implies the fourth, we need only show that any abelian factor of $\mathbb{k}^G \otimes \mathbb{k}G$ yields an abelian factor of $\mathcal{D}(G)$.

So let $L$ be an abelian tensor factor of $\mathbb{k}^G \otimes \mathbb{k}G$ with associated projection $\pi$ and inclusion $i$. We wish to show that $i \pi$ is canonically a binormal endomorphism of $\mathcal{D}(G)$. Writing $i \pi = \left( \begin{array}{cc} u & p \\ v & v \end{array} \right)$, the properties of $\text{End}(\mathcal{D}(G))$ and commutativity of the image easily imply the following: $p \gamma v$, $v \gamma v$, $u \cdot p$, $p \gamma \text{id}$, $v \gamma \text{id}$. In particular, $v$ and $p$ have central image, and $v$ is a (bi)normal group homomorphism. Since $(i \pi)^*$ is also an idempotent endomorphism with abelian image we similarly conclude that $p^*$ and $u^*$ have central image, and that $u^*$ is a (bi)normal group homomorphism. Centrality of the image of $u^*$ (indeed, that $u^*$ has abelian image and is thus a class function) implies that $G$ acts trivially on the image of $u$. Applying the proposition we conclude that $i \pi$ is canonically a binormal endomorphism of $\mathcal{D}(G)$ with image $L$. Fitting’s lemma then implies that $L$ is a direct tensor factor of $\mathcal{D}(G)$, as desired. This completes the proof. □

Thus for $G = C \times H$ with $C$ abelian and $H$ purely nonabelian we conclude that $\mathcal{D}(C)$ and $\mathcal{D}(H)$ have no common direct tensor factors. Therefore we may apply the results of the previous section to obtain the following.

Theorem 9.9. Let $G = C \times H$, where $C$, $H$ are finite groups with $C$ abelian and $H$ purely nonabelian. Then

$$\text{Aut}(\mathcal{D}(G)) = \begin{pmatrix} \text{Aut}(\mathcal{D}(C)) & \text{Hom}^c(\mathcal{D}(H), \mathcal{D}(C)) \\ \text{Hom}^c(\mathcal{D}(C), \mathcal{D}(H)) & \text{Aut}(\mathcal{D}(H)) \end{pmatrix}.$$
Whenever the field is such that $D$ is a field

Consider a field $D$


Example 9.10. Consider a field $k$ of characteristic not 2. For $n \geq 3$, let $G = D_{2n}$ be the dihedral group of order $2n$, and suppose that $n \equiv 2 \pmod{4}$. The group $G$ has an abelian direct factor precisely under this assumption on $n$, in which case $G \cong \mathbb{Z}_2 \times D_n$. So we take $C = \mathbb{Z}_2$ and $H = D_n$, and note $\Gamma_C \cong \mathbb{Z}_2$ and $\Gamma_H \cong \mathbb{Z}_2 \times D_n$. It is also well known that $\operatorname{Aut}(\Gamma_C) \cong S_3$. By [Keilberg 2015] we have $\operatorname{Aut}(D(D_n)) \cong \mathbb{Z}_2 \times \operatorname{Aut}(D_n) \cong \mathbb{Z}_2 \times \operatorname{Hol}(\mathbb{Z}_n^2)$. Here $\operatorname{Hol}(\mathbb{Z}_n) = \mathbb{Z}_n \rtimes \operatorname{Aut}(\mathbb{Z}_n)$ is the holomorph of $\mathbb{Z}_n$, a group of order $n\phi(n)$, where $\phi$ is the Euler totient function.

We have $Z(\Gamma_H) \cong \mathbb{Z}_2$, from which it follows that $\operatorname{Hom}(\Gamma_C, Z(\Gamma_H)) \cong \mathbb{Z}_2^2$ as groups. We claim that

$$\operatorname{Hom}^c(\mathcal{D}(H), \mathcal{D}(C)) \cong \mathbb{Z}_2^2$$

as well. Let $(u, r, p, v) \in \operatorname{Hom}(\mathcal{D}(H), \mathcal{D}(C))$. The abelian normal subgroups of $D_n$ all have odd order, so $p$ is necessarily trivial. By normality of $u^*v$, we have $u^*(b^x) = u^*(b) = u^*(b)x^{-1}$ for all $x \in D_n$ and $b \in \mathbb{Z}_2$. Since no order 2 subgroup of $D_n$ is normal we conclude that $u^*$ is trivial. From the preceding remarks it follows that $\operatorname{Hom}(\mathcal{D}(H), \mathcal{D}(C)) = \operatorname{Hom}^c(\mathcal{D}(H), \mathcal{D}(C))$. Since there are two possible homomorphisms $v : D_n \to \mathbb{Z}_2$, and two possible homomorphisms $r : D_n \to \mathbb{Z}_2$, all of which satisfy the necessary compatibilities, it quickly follows that $\operatorname{Hom}^c(\mathcal{D}(H), \mathcal{D}(C)) \cong \mathbb{Z}_2^2$ as desired.

As a consequence, $|\operatorname{Aut}(\mathcal{D}(D_{2n}))| = 2^5 \cdot 3 \cdot n \cdot \phi(n/2)$ whenever $n \equiv 2 \pmod{4}$. For $n = 6$ the order is 1152 = $2^7 \cdot 3^2$. The description and order of $\operatorname{Aut}(\mathcal{D}(D_{2n}))$ for $n \not\equiv 2 \pmod{4}$ is given in [Keilberg 2015].

References


Tensor factorizations of Hopf algebras


Communicated by Susan Montgomery
Received 2014-10-22 Accepted 2015-09-09

keilberg@usc.edu Institut de Mathématiques de Bourgogne, UMR5584 CNRS, Université Bourgogne Franche-Comté, F-21000 Dijon, France

peter.schauenburg@u-bourgogne.fr Institut de Mathématiques de Bourgogne, UMR5584 CNRS, Université Bourgogne Franche-Comté, F-21000 Dijon, France
Extension theorems for reductive group schemes

Adrian Vasiu

We prove several basic extension theorems for reductive group schemes via extending Lie algebras and via taking schematic closures. We also prove that, for each scheme $Y$, the category in groupoids of adjoint group schemes over $Y$ whose Lie algebra $\mathcal{O}_Y$-modules have perfect Killing forms is isomorphic, via the differential functor, to the category in groupoids of Lie algebra $\mathcal{O}_Y$-modules which have perfect Killing forms and which, as $\mathcal{O}_Y$-modules, are coherent and locally free.

1. Introduction

A group scheme $H$ over a scheme $S$ is called reductive if the morphism $H \to S$ has the following two properties: (i) it is smooth and affine (and therefore of finite presentation), and (ii) its geometric fibers are reductive groups over spectra of fields and therefore are connected (see [SGA 3 III 1970, Exposé XIX, Sections 2.7, 2.1, and 2.9]). If, moreover, the center of $H$ is trivial, then $H$ is called an adjoint group scheme over $S$. Let $\mathcal{O}_S$ be the structure ring sheaf of $S$. Let $\text{Lie}(H)$ be the Lie algebra $\mathcal{O}_S$-module of $H$. As an $\mathcal{O}_S$-module, $\text{Lie}(H)$ is coherent and locally free.

The main goal of the paper is to prove Theorems 1.2 and 1.4 below (see Sections 3 and 4) and to apply them and their proofs to obtain new extension theorems for homomorphisms between reductive group schemes (see Section 5). We begin by introducing two groupoids on sets (i.e., two categories whose morphisms are all isomorphisms).

1.1. Two groupoids on sets. Let $Y$ be an arbitrary scheme. Let $\text{Adj-perf}_Y$ be the category whose objects are adjoint group schemes over $Y$ with the property that their Lie algebra $\mathcal{O}_Y$-modules have perfect Killing forms (i.e., the Killing forms induce naturally $\mathcal{O}_Y$-linear isomorphisms from them into their duals) and whose morphisms are isomorphisms of group schemes. Let $\text{Lie-perf}_Y$ be the category whose objects are Lie algebra $\mathcal{O}_Y$-modules which have perfect Killing forms and

**MSC2010:** primary 14L15; secondary 11G18, 14F30, 14G35, 14K10, 14L17, 17B45.

**Keywords:** reductive group schemes, purity, regular rings, Lie algebras.
which as $\mathcal{O}_Y$-modules are coherent and locally free and whose morphisms are isomorphisms of Lie algebra $\mathcal{O}_Y$-modules.

**Theorem 1.2.** Let $\mathcal{L}_Y : \text{Adj-perf}_Y \to \text{Lie-perf}_Y$ be the functor which associates to a morphism $f : G \to H$ of $\text{Adj-perf}_Y$ the morphism $df : \text{Lie}(G) \to \text{Lie}(H)$ of $\text{Lie-perf}_Y$, which is the differential of $f$. Then the functor $\mathcal{L}_Y$ is an equivalence of categories.

We have a variant of this theorem for simply connected semisimple group schemes instead of adjoint group schemes; see Corollary 3.7. This theorem implies the classification of Lie algebras over fields of characteristic at least 3 that have nondegenerate Killing forms obtained previously by Curtis [1957], Seligman [1967], Mills [1957], Mills and Seligman [1957], Block–Zassenhaus [1964], and Brown [1969] (see Remark 3.6(a)). The functor $\mathcal{L}_Y$ is an equivalence of nonempty categories if and only if $Y$ is a $\text{Spec} \mathbb{Z}[\frac{1}{2}]$-scheme; see Corollary 3.8. Directly from Theorem 1.2 we get our first extension result:

**Corollary 1.3.** We assume that $Y = \text{Spec} A$ is an affine scheme. Let $K$ be the ring of fractions of $A$. Let $G_K$ be an adjoint group scheme over $\text{Spec} K$ such that the symmetric bilinear Killing form on the Lie algebra $\text{Lie}(G_K)$ of $G_K$ is perfect (i.e., it induces naturally a $K$-linear isomorphism $\text{Lie}(G_K) \cong \text{Hom}_K(\text{Lie}(G_K), K)$). We assume that there exists a Lie algebra $g$ over $A$ such that the following two properties hold:

(i) we have an identity $\text{Lie}(G_K) = g \otimes_A K$ and the $A$-module $g$ is projective and finitely generated;

(ii) the symmetric bilinear Killing form on $g$ is perfect.

Then there exists a unique adjoint group scheme $G$ over $Y$ that extends $G_K$, with an identity $\text{Lie}(G) = g$ that extends the identity of property (i).

Let $U$ be an open, Zariski-dense subscheme of $Y$. We call the pair $(Y, Y \setminus U)$ *quasipure* if each finite étale cover of $U$ extends uniquely to a finite étale cover of $Y$ (to be compared with [SGA 2 1968, Exposé X, Definition 3.1]).

**Theorem 1.4.** We assume that $Y$ is a normal, noetherian scheme and the codimension of $Y \setminus U$ in $Y$ is at least 2. Then the following two properties hold:

(a) Let $G_U$ be an adjoint group scheme over $U$. We assume that the Lie algebra $\mathcal{O}_U$-module $\text{Lie}(G_U)$ of $G_U$ extends to a Lie algebra $\mathcal{O}_Y$-module that is a locally free $\mathcal{O}_Y$-module. Then $G_U$ extends uniquely to an adjoint group scheme $G$ over $Y$.

(b) Let $H_U$ be a reductive group scheme over $U$. We assume that the pair $(Y, Y \setminus U)$ is quasipure and that the Lie algebra $\mathcal{O}_U$-module $\text{Lie}(G_U)$ of the adjoint group scheme $G_U$ of $H_U$ extends to a Lie algebra $\mathcal{O}_Y$-module that is a locally free $\mathcal{O}_Y$-module. Then $H_U$ extends uniquely to a reductive group scheme $H$ over $Y$. 
The proof of Theorem 1.2 that we include combines the cohomology theory of Lie algebras with a simplified variant of [Vasiu 1999, Claim 2, p. 464] (see Theorem 3.3 and Section 3.4). The proof of Theorem 1.4(a) is an application of [Colliot-Thélène and Sansuc 1979, Corollary 6.12] (see Section 4.1). The classical purity theorem of Nagata and Zariski (see [SGA 2 1968, Exposé X, Theorem 3.4(i)]) says that the pair \((Y, Y \setminus U)\) is quasipure, provided \(Y\) is regular and \(U\) contains all points of \(Y\) of codimension 1 in \(Y\). In such a case, a slightly weaker form of Theorem 1.4(b) was obtained in [Colliot-Thélène and Sansuc 1979, Theorem 6.13]. In general, the hypotheses of Theorem 1.4 are needed (see Remark 4.3). See [Moret-Bailly 1985] (resp. [Faltings and Chai 1990; Vasiu 1999; 2004; Vasiu and Zink 2010]) for different analogues of Theorem 1.4 for Jacobian (resp. abelian) schemes. For instance, in [Vasiu and Zink 2010, Corollary 1.5] it is proved that if \(Y\) is a regular, formally smooth scheme over the spectrum of a discrete valuation ring of mixed characteristic \((0, p)\) and index of ramification at most \(p - 1\) and if \(U\) contains all points of \(Y\) that are of either characteristic 0 or codimension 1 in \(Y\), then each abelian scheme over \(U\) extends uniquely to an abelian scheme over \(Y\).

Notation and basic results are presented in Section 2. In Section 3 we prove Theorem 1.2. In Section 4 we prove Theorem 1.4.

Section 5 contains two results on extending homomorphisms between reductive group schemes. Proposition 5.1 is an application of Theorem 1.4(b) and pertains to extensions of homomorphisms in codimension at least 2 over normal bases. Proposition 5.2 pertains to extensions of homomorphisms via schematic closures and refines [Vasiu 1999, Lemma 3.1.6]; its role is to achieve natural reductions such as the reduction to the case of either a torus or a semisimple group scheme.

Our main motivation for Theorems 1.2 and 1.4 stems from the meaningful applications to crystalline cohomology one gets by combining them with either Faltings’ results [1999, Section 4] (see [Vasiu 1999; 2008]) or de Jong’s extension theorem [1998, Theorem 1.1] (see [Vasiu 2012a; 2012b]). The manuscripts [Vasiu 2012a; 2012b] apply the results of the current paper to extend our prior work on integral canonical models of Shimura varieties of Hodge type in unramified mixed characteristic \((0, p)\) with \(p \geq 5\) (see [Vasiu 1999]) to unramified mixed characteristics \((0, 2)\) and \((0, 3)\). In addition, this paper can be used to get relevant simplifications of certain parts of the mentioned prior work (see [Vasiu 2008]).

2. Preliminaries

Our notation is gathered in Section 2.1, and then we include four basic results that are often used in Sections 3 to 5.

2.1. Notation and conventions. Let \(\bar{K}\) be an algebraic closure of a field \(K\). Let \(H\) be a reductive group scheme over a scheme \(S\). Let \(Z(H), H^{\text{der}}, H^{\text{ad}},\) and \(H^{\text{ab}}\)
denote the center, the derived group scheme, the adjoint group scheme, and the abelianization of $H$ (respectively). We have $H^{\text{ab}} = H/H^{\text{der}}$ and $H^{\text{ad}} = H/Z(H)$. The center $Z(H)$ is a group scheme of multiplicative type; see [SGA 3_III 1970, Exposé XXII, Corollary 4.1.7]. Let $Z^0(H)$ be the maximal torus of $Z(H)$; the quotient group scheme $Z(H)/Z^0(H)$ is a finite, flat group scheme over $S$ of multiplicative type. Let $H^{\text{sc}}$ be the simply connected semisimple group scheme cover of the derived group scheme $H^{\text{der}}$.

See [SGA 3_III 1970, Exposé XXII, Corollary 4.3.2] for the quotient group scheme $H/F$ of $H$ by a flat, closed, multiplicative type subgroup scheme $F$ of $Z(H)$. If $X$ or $X_S$ is an $S$-scheme, let $X_{A_1}$ (resp. $X_{S_1}$) be its pullback via a morphism $\text{Spec } A_1 \to S$ (resp. $S_1 \to S$).

If $S$ is either affine or integral, let $K_S$ be the ring of fractions of $S$. If $S$ is a normal, noetherian, integral scheme, let $\mathcal{O}(S)$ be the set of local rings of $S$ that are discrete valuation rings.

Let $\mathbb{G}_{m,S}$ be the rank-1 split torus over $S$; similarly, the group schemes $\mathbb{G}_{a,S}$, $\text{GL}_{d,S}$ with $d \in \mathbb{N}^*$, etc., will be understood to be over $S$. Let $\text{Lie}(H)$ be the Lie algebra $\mathbb{G}_S$-module of $H$. If $S = \text{Spec } A$ is affine, then let $\mathbb{G}_{m,A} := \mathbb{G}_{m,S}$, etc., and let $\text{Lie}(F)$ be the Lie algebra over $A$ of a closed subgroup scheme $F$ of $H$. For $A$-modules, we have $\text{Lie}(F) = \text{Ker}(F(A[x]/x^2) \to F(A))$, where the $A$-epimorphism $A[x]/(x^2) \to A$ takes $x$ to 0. The Lie bracket on $\text{Lie}(F)$ is defined by taking the (total) differential of the commutator morphism $[,] : F \times SF \to F$ at identity sections. If $S = \text{Spec } A$ is affine, then $\text{Lie}(H) = \text{Lie}(H)(S)$ is the Lie algebra over $A$ of global sections of $\text{Lie}(H)$ and it is a projective, finitely generated $A$-module.

If $N$ is a projective, finitely generated $A$-module, let $N^* := \text{Hom}_A(N, A)$, let $\text{GL}_N$ be the reductive group scheme over $\text{Spec } A$ of linear automorphisms of $N$, and let $\mathfrak{gl}_N := \text{Lie}(\text{GL}_N)$. Thus $\mathfrak{gl}_N$ is the Lie algebra associated to the $A$-algebra $\text{End}_A(N)$. A bilinear form $b_N : N \times N \to A$ on $N$ is called perfect if it induces an $A$-linear map $N \to N^*$ that is an isomorphism. If $b_N$ is symmetric, then its kernel is the $A$-submodule

$$\text{Ker}(b_N) := \{ a \in N \mid b_N(a, b) = 0 \text{ for all } b \in N \}$$

of $N$. For a Lie algebra $\mathfrak{g}$ over $A$ that is a projective, finitely generated $A$-module, let $\text{ad} : \mathfrak{g} \to \mathfrak{gl}_\mathfrak{g}$ be the adjoint representation of $\mathfrak{g}$ and let $\mathfrak{K}_\mathfrak{g} : \mathfrak{g} \times \mathfrak{g} \to A$ be the Killing form on $\mathfrak{g}$. For $a, b \in \mathfrak{g}$ we have $\text{ad}(a)(b) = [a, b]$, and $\mathfrak{K}_\mathfrak{g}(a, b)$ is the trace of the endomorphism $\text{ad}(a) \circ \text{ad}(b)$ of $\mathfrak{g}$. The kernel $\text{Ker}(\mathfrak{K}_\mathfrak{g})$ is an ideal of $\mathfrak{g}$.

We denote by $k$ an arbitrary field. Let $n \in \mathbb{N}^*$. See [Bourbaki 2002, Chapter VI, Section 4] and [Humphreys 1972, Chapter III, Section 11] for the classification of connected Dynkin diagrams. For

$$b \in \{A_n, B_n, C_n \mid n \in \mathbb{N}^*\} \cup \{D_n \mid n \geq 3\} \cup \{E_6, E_7, E_8, F_4, G_2\},$$
we say that $H$ is of isotypic $\mathbb{b}$ Dynkin type if the connected Dynkin diagram of each simple factor of an arbitrary geometric fiber of $H^{\text{ad}}$ is $\mathbb{b}$; if $H^{\text{ad}}$ is absolutely simple, we drop the word ‘isotypic’. We recall that $A_1 = B_1 = C_1$, $B_2 = C_2$, and $A_3 = D_3$.

**Proposition 2.2.** Let $Y$ be a normal, noetherian, integral scheme. Let $K := K_Y$.

(a) If $Y = \text{Spec } A$ is affine, then inside the field $K$ we have $A = \bigcap_{V \in \mathcal{V}(Y)} V$.

(b) Let $U$ be an open subscheme of $Y$ such that $Y \setminus U$ has codimension in $Y$ at least 2. Let $W$ be an affine $Y$-scheme of finite type. Then the natural restriction map $\text{Hom}_Y(Y, W) \to \text{Hom}_Y(U, W)$ is a bijection. If, moreover, $W$ is integral, normal and such that we have $\mathcal{D}(W) = \mathcal{D}(W_U)$, then $W$ is determined (up to unique isomorphism) by $W_U$.

(c) Suppose that $Y = \text{Spec } A$ is local, regular, and has dimension 2. Let $y$ be the closed point of $Y$ and let $U := Y \setminus \{y\}$. Then each locally free $\mathcal{O}_U$-module of finite rank extends uniquely to a free $\mathcal{O}_Y$-module.

**Proof.** See [Matsumura 1980, Theorem 38] for (a). To check (b), we can assume $Y = \text{Spec } A$ is affine. We write $W = \text{Spec } B$. The $A$-algebra of global functions of $U$ is $A$; see (a). We have $\text{Hom}_Y(U, W) = \text{Hom}_A(B, A) = \text{Hom}_Y(Y, W)$. If, moreover, $B$ is a normal ring and we have $\mathcal{D}(W) = \mathcal{D}(W_U)$, then $B$ is uniquely determined by $\mathcal{D}(W_U)$ (see (a)) and therefore by $W_U$. From this (b) follows. See [SGA 2 1968, Exposé X, Lemma 3.5] for (c). \hfill \Box

**Proposition 2.3.** Let $G$ be a reductive group scheme over a scheme $Y$. Then the functor on the category of $Y$-schemes that parametrizes maximal tori of $G$ is representable by a smooth, separated $Y$-scheme of finite type. Thus $G$ has split, maximal tori, locally in the étale topology of $Y$.

**Proof.** See [SGA 3 II 1970, Exposé XII, Corollary 1.10] for the first part. The second part follows easily from the first part (see also [SGA 3 III 1970, Exposé XIX, Proposition 6.1]). \hfill \Box

**Lemma 2.3.1.** Let $Y$ be a reduced scheme. Let $G$ be a reductive group scheme over $Y$. Let $K := K_Y$. Let $f_K : G'_K \to G_K$ be a central isogeny of reductive group schemes over Spec $K$. We assume that either $G$ is split or $Y$ is normal. We have:

(a) There exists (up to a canonical identification) at most one central isogeny $f : G' \to G$ that extends $f_K : G'_K \to G_K$. If $Y$ is integral (i.e., $K$ is a field), then there exists a unique central isogeny $f : G' \to G$ that extends $f_K : G'_K \to G_K$.

(b) If $Y$ is normal and integral, then $G'$ is the normalization of $G$ in (the field of fractions of) $G'_K$.

**Proof.** We first prove (a) in the case when $G$ is split. Let $T$ be a split, maximal torus of $G$. We first prove the existence part; thus $K$ is a field. As $f_K$ is a central isogeny, the inverse image $T'_K$ of $T_K$ in $G'_K$ is a split torus. Thus $G'_K$ is split. Let $\mathcal{R}' \to \mathcal{R}$ be
the 1-morphism of root data in the sense of [SGA 3_III 1970, Exposé XXI, Definition 6.8.1] which is associated to the central isogeny \( f_K : G'_K \rightarrow G_K \) that extends the isogeny \( T'_K \rightarrow T_K \). Let \( \tilde{f} : \tilde{G}' \rightarrow G \) be a central isogeny of split, reductive group schemes over \( Y \) which extends an isogeny of split tori \( \tilde{T}' \rightarrow T \) and for which the 1-morphism of root data associated to it and to the isogeny \( \tilde{T}' \rightarrow T \) is \( \mathcal{R}' \rightarrow \mathcal{R} \) (see [SGA 3_III 1970, Exposé XXV, Theorem 1.1]). From loc. cit. we also get that there exists an isomorphism \( i_K : \tilde{G}'_K \cong G'_K \) such that we have \( \tilde{f}_K = f_K \circ i_K \). Obviously, \( i_K \) is unique. Let \( G' \) be the unique group scheme over \( Y \) such that \( i_K \) extends (uniquely) to an isomorphism \( i : \tilde{G}' \cong G' \). Let \( f := \tilde{f} \circ i^{-1} \) : \( G' \rightarrow G \) be a central isogeny.

To check the uniqueness part, we consider two central isogenies \( G' \rightarrow G \) and \( G'_i \rightarrow G \) that extend a central isogeny \( f_K : G'_K \rightarrow G_K \) (thus \( G''_K = G'_{1,K} \)). Let \( G'_2 \) be the schematic closure of \( G'_K \) embedded diagonally into the product \( G' \times_Y G'_1 \). We are left to check that the two projections \( \pi_1 : G'_2 \rightarrow G' \) and \( \pi_2 : G'_2 \rightarrow G'_1 \) are isomorphisms, as in such a case the composite isomorphism \( \pi_2 \circ \pi_1^{-1} : G' \cong G'_1 \) is an isomorphism that extends the identity \( G'_K = G'_{1,K} \). This statement is local for the étale topology of \( Y \), and therefore we can assume, based on Proposition 2.3, that the inverse images of \( \tilde{T} \) to \( G' \) and \( G'_i \) are split tori. From this and [SGA 3_III 1970, Exposé XXIII, Theorem 4.1] we get that there exists a unique isomorphism \( \theta : G_1 \cong G'_1 \) which extends the identity \( G'_K = G_K \). This implies that \( G'_2 \) is the graph of \( \theta \), and therefore the two projections \( \pi_1 \) and \( \pi_2 \) are isomorphisms. We conclude that (a) holds if \( G \) is split.

We now prove simultaneously (a) and (b) in the case when \( Y \) is normal. If a \( G' \) as in (a) exists, then it is a smooth scheme over the normal scheme \( Y \) and thus it is a normal scheme; from this and the fact that \( f : G' \rightarrow G \) is a finite morphism, we get that \( G' \) is the normalization of \( G \) in \( G'_K \) and, in particular, that it is unique.

Thus to finish the proof of the lemma, it suffices to show that the normalization \( G' \) of \( G \) in \( G'_K \) is a reductive group scheme equipped with a central isogeny \( f : G' \rightarrow G \), locally in the étale topology of \( Y \). As each connected, étale scheme over \( Y \) is a normal scheme, based on Proposition 2.3 we can assume that \( G \) has a split, maximal torus \( T \). Thus the fact that \( G' \) is a reductive group scheme equipped with a central isogeny \( f : G' \rightarrow G \) follows from the previous three paragraphs.

\[ \square \]

**Lemma 2.3.2.** Let \( Y = \text{Spec } A \) be an affine scheme. Let \( K := K_Y \). Let \( T \) be a torus over \( Y \) equipped with a homomorphism \( \rho : T \rightarrow G \), where \( G \) is a reductive group scheme over \( Y \). Then the following three properties hold:

(a) the kernel \( \text{Ker}(\rho) \) is a group scheme over \( Y \) of multiplicative type;

(b) the kernel \( \text{Ker}(\rho) \) is trivial (resp. finite) if and only if the kernel \( \text{Ker}(\rho_K) \) is trivial (resp. finite);

(c) the quotient group scheme \( T/\text{Ker}(\rho) \) is a torus and we have a closed embedding homomorphism \( T/\text{Ker}(\rho) \hookrightarrow G \).
Proof. The statements of the lemma are local for the étale topology of \( Y \). Thus we can assume that \( Y \) is local and (see Proposition 2.3) that \( T \) and \( G \) are split. As \( Y \) is connected, the split reductive group scheme \( G \) has constant Lie type. Thus \( G \) is the pullback to \( Y \) of a reductive group scheme \( G \) over \( \text{Spec} \mathbb{Z} \); see [SGA 3ï¿½ 1970, Exposé XXV, Corollary 1.2]. As \( G \) can be embedded into a general linear group scheme over \( \text{Spec} \mathbb{Z} \) (for instance, see [SGA 3ï¿½ 1970, Exposé VI_B, Remark 11.11.1]), there exists a closed embedding homomorphism \( G \hookrightarrow \text{GL}_M \), with \( M \) a free \( \mathbb{A} \)-module of rank \( d \in \mathbb{N}^* \). By replacing \( \rho \) with its composite with this closed embedding homomorphism \( G \hookrightarrow \text{GL}_M \), we can assume that \( G = \text{GL}_M \) is a general linear group scheme over \( Y \). The representation of \( T \) on \( M \) is a finite direct sum of representations of \( T \) of rank 1; see [Jantzen 2003, Part I, Section 2.11]. Thus \( \rho \) factors as the composite of a homomorphism \( \rho_1 : T \rightarrow \mathbb{G}_{m,A}^m \) with a closed embedding homomorphism \( \mathbb{G}_{m,A}^m \hookrightarrow \text{GL}_M \). The kernel \( \ker(\rho_1) \) is a group scheme over \( Y \) of multiplicative type; see [SGA 3ï¿½ 1970, Exposé IX, Proposition 2.7(i)].

As \( \ker(\rho) = \ker(\rho_1) \), we get that (a) holds. As (a) holds, \( \ker(\rho) \) is flat over \( Y \) as well as the extension of a finite, flat group scheme \( T_1 \) by a torus \( T_0 \). But \( T_1 \) (resp. \( T_0 \)) is a trivial group scheme if and only if \( T_1,K \) (resp. \( T_0,K \)) is trivial. From this (b) follows. The quotient group scheme \( T/\ker(\rho) \) exists and is a closed subgroup scheme of \( \mathbb{G}_{m,A}^m \) that is of multiplicative type; see [SGA 3ï¿½ 1970, Exposé IX, Proposition 2.7(i)] and Corollary 2.5]. As the fibers of \( T/\ker(\rho) \) are tori, we get that \( T/\ker(\rho) \) is a torus. Thus (c) holds.

The following lemma is only a variant of [Vasiu 2005a, Lemma 2.1].

Lemma 2.4. Suppose that \( k = \overline{k} \). Let \( H \) be a reductive group over \( \text{Spec} k \). Let \( n \) be a nonzero ideal of \( \text{Lie}(H) \) which is a simple left \( H \)-module. We assume that there exists a maximal torus \( T \) of \( H \) such that we have \( \text{Lie}(T) \cap n = 0 \). Then \( \text{char}(k) = 2 \) and \( H^{\text{der}} \) has a normal, subgroup \( F \) which is isomorphic to \( \text{SO}_{2n+1,k} \) for some \( n \in \mathbb{N}^* \) and for which we have an inclusion \( n \subseteq \text{Lie}(F) \).

Remark 2.4.1. If \( n \) is assumed to be a restricted Lie subalgebra of \( \text{Lie}(H) \) (for instance, this holds if \( n \) is the Lie algebra of a subgroup of \( H \)), then there exists a purely inseparable isogeny \( H \rightarrow H/n \) (see [Borel 1991, Chapter V, Proposition 17.4]) and in this case Lemma 2.4 can be also deduced easily from [Prasad and Yu 2006, Lemma 2.2] applied to such isogenies with \( H^{\text{ad}} \) absolutely simple. In this paper, Lemma 2.4 will be applied only in such situations in which \( n \) is the Lie algebra of a subgroup of \( H \).

Theorem 2.5. Let \( f : G_1 \rightarrow G_2 \) be a homomorphism between group schemes over a scheme \( Y \). We assume that \( G_1 \) is reductive, that \( G_2 \) is separated and of finite presentation, and that all fibers of \( f \) are closed embeddings. Then \( f \) is a closed embedding.
Proof. As $G_1$ is of finite presentation over $Y$, the homomorphism $f$ is locally of finite type. As the fibers of $f$ are closed embeddings and thus monomorphisms, $f$ itself is a monomorphism (see [SGA 3, 1970, Exposé VI B, Corollary 2.11]). Thus the theorem follows from [SGA 3 II, 1970, Exposé XVI, Corollary 1.5(a)]. □

Lemma 2.5.1. Let $G$ be an adjoint group scheme over an affine scheme $Y = \text{Spec } A$. Let $\text{Aut}(G)$ be the group scheme over $Y$ of automorphisms of $G$. Then the natural adjoint representation $\text{Ad}: \text{Aut}(G) \to \text{GL}_{\text{Lie}(G)}$ is a closed embedding.

Proof. To prove the lemma, we can work locally in the étale topology of $Y$ and therefore (see Proposition 2.3) we can assume that $G$ is split and that $Y$ is connected. We have a short exact sequence $1 \to G \to \text{Aut}(G) \to C \to 1$ that splits (see [SGA 3 III, 1970, Exposé XXIV, Theorem 1.3]), where $C$ is a finite, étale, constant group scheme over $Y$. Thus $G$ is the identity component of $\text{Aut}(G)$, and $\text{Aut}(G)$ is a finite disjoint union of right translates of $G$ via certain $Y$-valued points of $\text{Aut}(G)$. If the fibers of $\text{Ad}$ are closed embeddings, then the restriction of $\text{Ad}$ to $G$ is a closed embedding (see Theorem 2.5), and thus also the restriction of $\text{Ad}$ to any right translate of $G$ via a $Y$-valued point of $\text{Aut}(G)$ is a closed embedding. The last two sentences imply that $\text{Ad}$ is a closed embedding. Thus, to finish the proof, we are left to check that the fibers of $\text{Ad}$ are closed embeddings. For this, we can assume that $A$ is an algebraically closed field.

As $G$ is adjoint and $A$ is a field, the restriction of $\text{Ad}$ to $G$ is a closed embedding. Thus the representation $\text{Ad}$ is a closed embedding if and only if each element $g \in \text{Aut}(G)(A)$ that acts trivially on $\text{Lie}(G)$ is trivial. We show that the assumption that there exists a nontrivial element $g$ leads to a contradiction. For this, we can assume that $G$ is absolutely simple and that $g$ is a nontrivial outer automorphism of $G$. Let $T$ be a maximal torus of a Borel subgroup $B$ of $G$ and let $n$ be the dimension of $T$.

For $t \in \text{Lie}(T)$, let $C_G(t)$ be its centralizer in $G$; it is a subgroup of $G$ that contains $T$. In this paragraph we check that, as $G$ is adjoint, we can choose $t$ such that $C_G(t)^0 = T$. We consider the root decomposition $\text{Lie}(G) = \text{Lie}(T) \bigoplus_{\alpha \in \Phi} g_{\alpha}$ with respect to $T$, where $\Phi$ is the root system of $G$ and where each $g_{\alpha}$ is a one-dimensional $A$-vector space normalized by $T$. Let $\Delta$ be the basis for $\Phi$ such that we have $\text{Lie}(B) = \text{Lie}(T) \bigoplus_{\alpha \in \Delta} g_{\alpha}$. As $G$ is adjoint, $\Delta$ is a basis for the dual $A$-vector space $\text{Lie}(T)^*$ (to be compared with [SGA 3 III, 1970, Exposé XXI, Definition 6.2.6 and Exposé XXII, Definition 4.3.3]). Thus for each root $\alpha \in \Delta$, $\text{Ker}(\alpha)$ is an $A$-vector subspace of $\text{Lie}(T)$ of dimension $n - 1$. As each $\alpha \in \Phi$ is conjugate under the Weyl group of $\Phi$ (equivalently of $G$) to an element of $\Delta$ (see [Humphreys 1972, Chapter III, Section 10, Theorem]), we get that for each $\alpha \in \Delta$ its kernel $\text{Ker}(\alpha)$ is an $A$-vector subspace of $\text{Lie}(T)$ of dimension $n - 1$. We choose $t \in \text{Lie}(T) \setminus \bigcup_{\alpha \in \Phi} \text{Ker}(\alpha)$. This implies that $\text{Lie}(C_G(t)) = \text{Lie}(T)$. From
this and the fact that \( T \) is a subgroup of \( C_G(t) \), we get that \( C_G(t) \) is a smooth group of dimension \( n \) and therefore that \( C_G(t)^0 = T \).

As \( g \) fixes \( t \) and \( \text{Lie}(B) \), it normalizes both \( C_G(t)^0 = T \) and \( B \). But it is well known that a nontrivial outer automorphism \( g \) of \( G \) that normalizes both \( T \) and \( B \) cannot fix \( \text{Lie}(B) \). Contradiction. Thus \( \text{Ad} \) is a closed embedding. \( \Box \)

We follow the ideas of [Vasiu 1999, Proposition 3.1.2.1(c) and Remark 3.1.2.2(3)] in order to prove the next proposition.

**Proposition 2.5.2.** Let \( V \) be a discrete valuation ring whose residue field is \( k \). Let \( Y = \text{Spec} \, V \) and let \( K := K_Y \). Let \( f : H_1 \rightarrow H_2 \) be a homomorphism between flat, finite type, affine group schemes over \( Y \) such that \( H_1 \) is a reductive group scheme and the generic fiber \( f_K : H_{1,K} \rightarrow H_{2,K} \) of \( f \) is a closed embedding. We have:

(a) The subgroup scheme \( \text{Ker}(f_K : H_{1,K} \rightarrow H_{2,K}) \) of \( H_{1,k} \) has a trivial intersection with each torus \( T_{1,k} \) of \( H_{1,k} \). In particular, \( \text{Lie}(\text{Ker}(f_k)) \cap \text{Lie}(T_{1,k}) = 0 \).

(b) The homomorphism \( f \) is finite.

(c) If \( \text{char}(k) = 2 \), we assume that \( H_{1,k} \) has no normal subgroup that is adjoint of isotypic \( B_{n} \) Dynkin type for some \( n \in \mathbb{N}^* \). Then \( f \) is a closed embedding.

**Proof.** Let \( \rho : H_2 \hookrightarrow \text{GL}_M \) be a closed embedding homomorphism, with \( M \) a free \( V \)-module of finite rank (see [SGA 3-I, Exposé VIb, Remark 11.11.1]). To prove the proposition we can assume that \( V \) is complete, that \( k = \overline{k} \), and that \( f_k : H_{1,k} \rightarrow H_{2,k} \) is an isomorphism. Let \( H_{0,k} := \text{Ker}(f_k) \). We now show that the group scheme \( H_{0,k} \cap T_{1,k} \) is trivial by adapting arguments from [Vasiu 1999, Remark 3.1.2.2(3) and proof of Lemma 3.1.6]. As \( V \) is strictly henselian, the maximal torus \( T_{1,k} \) of \( H_{1,k} \) is split and (see Proposition 2.3) it lifts to a maximal torus \( T_1 \) of \( H_1 \). The restriction of \( \rho \circ f \) to \( T_1 \) has a trivial kernel (as its fiber over \( \text{Spec} \, K \) is trivial; see Lemma 2.3.2(b)) and therefore it is a closed embedding (see Lemma 2.3.2(c)). Thus the restriction of \( f \) to \( T_1 \) is a closed embedding homomorphism \( T_1 \hookrightarrow H_2 \). Therefore, the intersection \( H_{0,k} \cap T_{1,k} \) is a trivial group scheme. Thus (a) holds.

We check (b). The identity component of the reduced scheme of \( \text{Ker}(f_k) \) is a reductive group that has 0 rank (see (a)) and therefore it is a trivial group. Thus \( f \) is a quasifinite, birational morphism. From Zariski’s main theorem (see [Grothendieck 1966, Theorem 8.12.6]) we get that \( H_1 \) is an open subscheme of the normalization \( H_2^n \) of \( H_2 \). Let \( H_3 \) be the smooth locus of \( H_2^n \) over \( \text{Spec} \, V \); it is an open subscheme of \( H_2^n \) that contains \( H_1 \). As \( H_3 \) is an open subscheme of the affine scheme \( H_2^n \), it is a quasaffine scheme.

As \( H_3 \) is smooth over \( \text{Spec} \, V \), the products \( H_3 \times_{\text{Spec} \, V} H_2^n \) and \( H_2^n \times_{\text{Spec} \, V} H_3 \) are smooth over \( H_2^n \) and thus are normal schemes. The product \( H_2^n \times_{\text{Spec} \, V} H_2^n \) is a flat scheme over \( \text{Spec} \, V \) whose generic fiber is smooth over \( \text{Spec} \, K \). Its
normalization \((H^n_2 \times \text{Spec} \mathcal{V} H^n_2)^n\) contains both \(H_3 \times \text{Spec} \mathcal{V} H^n_2\) and \(H^n_2 \times \text{Spec} \mathcal{V} H_3\) as open subschemes and is equipped with a finite surjective morphism given by \((H^n_2 \times \text{Spec} \mathcal{V} H^n_2)^n \to H^n_2 \times \text{Spec} \mathcal{V} H^n_2\) whose generic fiber is an isomorphism. The product morphism \(H_2 \times \text{Spec} \mathcal{V} H_2 \to H_2\) induces a natural product-type morphism \(\Theta : (H^n_2 \times \text{Spec} \mathcal{V}_2)^n \to H^n_2\). Its restrictions to \(H_3 \times \text{Spec} \mathcal{V} H^n_2\) and \(H^n_2 \times \text{Spec} \mathcal{V} H_3\) induce product-type morphisms \(H_3 \times \text{Spec} \mathcal{V} H^n_2 \to H^n_3\) and \(H^n_3 \times \text{Spec} \mathcal{V} H_3 \to H^n_3\). This implies that for each valued point \(z \in H^n_2(V)\) it makes sense to speak about the left \(zH_3\) and the right \(H_3z\) translations of \(H_3\) through \(z\); they are smooth open subschemes of \(H^n_2\) and thus of \(H_3\). This implies that \(H_3(V) = H^n_2(V)\) and that \(\Theta\) restricts to a product morphism \(H_3 \times \text{Spec} \mathcal{V} H_3 \to H_3\). The inverse automorphisms of the (\(\text{Spec} \mathcal{V}\))-schemes \(H_1\) and \(H_2\) induce an inverse automorphism of the (\(\text{Spec} \mathcal{V}\))-scheme \(H^n_2\) which restricts to an inverse automorphism of the (\(\text{Spec} \mathcal{V}\))-scheme \(H_3\). With respect to its product morphism, its inverse automorphism, and its identity section inherited from \(H_1\), the subscheme \(H_3\) gets the structure of a (quasiaffine) group scheme over \(\text{Spec} \mathcal{V}\) that is of finite type.

As \(V\) is complete, it is also excellent (see [Matsumura 1980, Section 34]). Thus the morphism \(H^n_2 \to H_2\) is finite. The homomorphism \(f\) is finite if and only if \(H_1 = H^n_2\) and thus if and only if the set \(H^n_2(k) \setminus H_1(k)\) is empty. We show that the assumption \(H_1 \neq H^n_2\) leads to a contradiction. Let \(x \in H^n_2(k) \setminus H_1(k)\). From [Vasiu 2012c, Lemma 4.1.5] applied to the completion of the local ring of \(x\) in \(H^n_2\), we get that there exists a finite, flat discrete valuation ring extension \(V'\) of \(V\) for which we have a valued point \(z' \in H^n_2(V')\) that lifts \(x\) (we recall that loc. cit. is only a local version of the global result [Grothendieck 1967, Corollary 17.16.2]). The flat (\(\text{Spec} \mathcal{V}'\))-scheme \(H^n_{2,V'}\) might not be normal but we have \(H_1 \neq H^n_2\) if and only if \(H_{1,V'} \neq H^n_{2,V'}\). Thus to reach a contradiction we can replace \(V\) by \(V'\), and therefore we can assume that there exists a valued point \(z \in H^n_2(V) = H_3(V)\) which lifts \(x\). As \(x \in H^n_2(k) \setminus H_1(k)\), we have \(z \in H_3(V) \setminus H_1(V)\). As \(H_1\) is a subgroup scheme of \(H_3\), all fibers of the homomorphism \(H_1 \to H_3\) are closed. From this and Theorem 2.5 we get that \(H_1\) is a closed subscheme of \(H_3\). Thus, as \(H_3\) is an integral scheme and as \(H_{3,K} = H_{1,K}\), we get that \(H_1 = H_3\). This contradicts the fact that \(z \in H_3(V) \setminus H_1(V)\). Thus (b) holds.

We check (c). We show that the assumption \(\text{Lie}(H_{0,k}) \neq 0\) leads to a contradiction. From Lemma 2.4 applied to \(H_{1,k}\) and to any simple \(H_{1,k}\)-submodule of the left \(H_{1,k}\)-module \(\text{Lie}(H_{0,k})\), we get that \(\text{char}(k) = 2\) and that \(H_{1,k}\) has a normal subgroup \(H_{4,k}\) isomorphic to \(SO_{2n+1,k}\) for some \(n \in \mathbb{N}^*\). As \(H_{4,k}\) is adjoint, we have a product decomposition \(H_{1,k} = H_{4,k} \times \text{Spec} \mathcal{V} H_{5,k}\) of reductive groups. It lifts (see [SGA 3\text{m} 1970, Exposé XXIV, Proposition 1.21]) to a product decomposition \(H_1 = H_4 \times \text{Spec} \mathcal{V} H_5\), where \(H_4\) is isomorphic to \(SO_{2n+1,V}\) and where \(H_5\) is a reductive group scheme over \(\text{Spec} \mathcal{V}\). This contradicts the extra hypothesis of (c). Thus we have \(\text{Lie}(H_{0,k}) = 0\). Therefore, \(H_{0,k}\) is a finite, étale, normal subgroup
of $H_{1,k}$. But $H_{1,k}$ is connected and thus its action on $H_{0,k}$ via inner conjugation is trivial. Therefore, we have $H_{0,k} \leq Z(H_{1}) \leq T_{1,k}$. Thus $H_{0,k} = H_{0,k} \cap T_{1,k}$ is the trivial group; see (a). In other words, the homomorphism $f_{k} : H_{1,k} \to H_{2,k}$ is a closed embedding. Thus $f : H_{1} \to H_{2}$ is a closed embedding homomorphism; see Theorem 2.5.

**Remark 2.5.3.** See [Vasiu 2005b, Theorem 1.2(b)] and [Prasad and Yu 2006, Theorem 1.2] for two other proofs of Proposition 2.5.2(c).

### 3. Lie algebras with perfect Killing forms

Let $A$ be a commutative $\mathbb{Z}$-algebra. Let $\mathfrak{g}$ be a Lie algebra over $A$ which as an $A$-module is projective and finitely generated. In this section we will assume that the Killing form $\mathcal{K}_{\mathfrak{g}}$ on $\mathfrak{g}$ is perfect. Let $U_{\mathfrak{g}}$ be the enveloping algebra of $\mathfrak{g}$, i.e., the quotient of the tensor algebra $T_{\mathfrak{g}}$ of $\mathfrak{g}$ by the two-sided ideal of $T_{\mathfrak{g}}$ generated by the subset $\{x \otimes y - y \otimes x - [x, y] \mid x, y \in \mathfrak{g}\}$ of $T_{\mathfrak{g}}$. Let $Z(U_{\mathfrak{g}})$ be the center of $U_{\mathfrak{g}}$. The categories of left $\mathfrak{g}$-modules and of left $U_{\mathfrak{g}}$-modules are canonically identified. We view $\mathfrak{g}$ as a left $\mathfrak{g}$-module via the adjoint representation $\text{ad} : \mathfrak{g} \to \mathfrak{gl}_{\mathfrak{g}}$; let $\text{ad} : U_{\mathfrak{g}} \to \text{End}(\mathfrak{g})$ be the $A$-homomorphism corresponding to the left $\mathfrak{g}$-module $\mathfrak{g}$. We refer to [Cartan and Eilenberg 1956, Chapter XIII] for the cohomology groups $H^{i}(\mathfrak{g}, \mathfrak{v})$ of a left $\mathfrak{g}$-module $\mathfrak{v}$ (here $i \in \mathbb{N}$). We denote also by $\mathcal{K}_{\mathfrak{g}} : \mathfrak{g} \otimes_{A} \mathfrak{g} \to A$ the $A$-linear map defined by $\mathcal{K}_{\mathfrak{g}} : \mathfrak{g} \times \mathfrak{g} \to A$. Thus we have $\mathcal{K}_{\mathfrak{g}} \in (\mathfrak{g} \otimes_{A} \mathfrak{g})^{*} = \mathfrak{g}^{*} \otimes_{A} \mathfrak{g}^{*}$.

Let $\phi : \mathfrak{g} \to \mathfrak{g}^{*}$ be the $A$-linear isomorphism defined naturally by $\mathcal{K}_{\mathfrak{g}}$. It induces an $A$-linear isomorphism $\phi^{-1} \otimes \phi^{-1} : \mathfrak{g}^{*} \otimes_{A} \mathfrak{g}^{*} \to \mathfrak{g} \otimes_{A} \mathfrak{g}$. The image $\Omega$ of $\phi^{-1} \otimes \phi^{-1}(\mathcal{K}_{\mathfrak{g}}) \in \mathfrak{g} \otimes_{A} \mathfrak{g} \subseteq T_{\mathfrak{g}}$ in $U_{\mathfrak{g}}$ is called the Casimir element of the adjoint representation $\text{ad} : \mathfrak{g} \to \mathfrak{gl}_{\mathfrak{g}}$.

**Lemma 3.1.** For the Casimir element $\Omega \in U_{\mathfrak{g}}$ the following four properties hold:

(a) if the $A$-module $\mathfrak{g}$ is free and if $\{x_{1}, \ldots, x_{m}\}$ and $\{y_{1}, \ldots, y_{m}\}$ are two $A$-bases for $\mathfrak{g}$ such that for all $i, j \in \{1, \ldots, m\}$ we have $\mathcal{K}_{\mathfrak{g}}(x_{i} \otimes y_{j}) = \delta_{ij}$, then $\Omega$ is the image of the element $\sum_{i=1}^{m} x_{i} \otimes y_{i}$ of $T_{\mathfrak{g}}$ in $U_{\mathfrak{g}}$;

(b) we have $\Omega \in Z(U_{\mathfrak{g}})$;

(c) the Casimir element $\Omega$ is fixed by the group of Lie automorphisms of $\mathfrak{g}$ (i.e., if $\sigma : U_{\mathfrak{g}} \to U_{\mathfrak{g}}$ is the $A$-algebra automorphism induced by a Lie algebra automorphism $\sigma : \mathfrak{g} \to \mathfrak{g}$, then we have $\sigma(\Omega) = \Omega$);

(d) the Casimir element $\Omega$ acts identically on $\mathfrak{g}$ (i.e., $\text{ad}(\Omega) = 1_{\mathfrak{g}}$).

**Proof.** Parts (a) and (b) are proved in [Bourbaki 1989, Chapter I, Section 3.7, Proposition 11]. Strictly speaking, loc. cit. is stated over a field but its proof applies over any commutative $\mathbb{Z}$-algebra. This is so, as the essence of the proof is [Bourbaki 1989, Chapter I, Section 3.5, Example 2] which is worked out over...
any commutative \(Z\)-algebra. In particular, [Bourbaki 1989, Chapter I, Section 3.5, Example 3] can be easily stated over a commutative \(Z\)-algebra (by involving a perfect invariant bilinear form over a commutative \(Z\)-algebra instead of a nondegenerate invariant bilinear form over a field). We recall here that \(\mathcal{H}_g\) is \(g\)-invariant, i.e., for all \(a, b, c \in g\) we have an identity \(\mathcal{H}_g(\text{ad}(a)(b), c) + \mathcal{H}_g(b, \text{ad}(a)(c)) = 0\) (see [Bourbaki 1989, Chapter I, Section 3.6, (13) and Proposition 8]), and this is the very essence of (b).

To check (c) and (d), we can assume that the \(A\)-module \(g\) is free. Let \(\{x_1, \ldots, x_m\}\) and \(\{y_1, \ldots, y_m\}\) be two \(A\)-bases for \(g\) as in (a). Thus \(\Omega\) is the image of \(\sum_{i=1}^m x_i \otimes y_i\) in \(U_g\). Therefore, \(\sigma(\Omega)\) is the image of \(\sum_{i=1}^m \sigma(x_i) \otimes \sigma(y_i)\) in \(U_g\). As we have \(\mathcal{H}_g(\sigma(x_i), \sigma(y_j)) = \delta_{ij}\) for \(i, j \in \{1, \ldots, m\}\), from (a) we get that the image of \(\sum_{i=1}^m \sigma(x_i) \otimes \sigma(y_i) \in T_g\) in \(U_g\) is \(\Omega\). Thus \(\sigma(\Omega) = \Omega\).

We check (d). Let \(z, w \in g\). We write \(\text{ad}(z) \circ \text{ad}(w)(x_i) = \sum_{j=1}^m a_{ji} x_j\), with the \(a_{ji}\)s in \(A\). Using the \(g\)-invariance of \(\mathcal{H}_g\) we compute

\[
\mathcal{H}_g(\text{ad}(\Omega)(z), w) = \mathcal{H}_g(\sum_{i=1}^m \text{ad}(x_i) \circ \text{ad}(y_i)(z), w) = -\sum_{i=1}^m \mathcal{H}_g(\text{ad}(y_i)(z), \text{ad}(x_i)(w))
\]

\[
= \sum_{i=1}^m \mathcal{H}_g(\text{ad}(z)(y_i), \text{ad}(x_i)(w)) = -\sum_{i=1}^m \mathcal{H}_g(y_i, \text{ad}(z) \circ \text{ad}(x_i)(w))
\]

\[
= \sum_{i=1}^m \mathcal{H}_g(y_i, \text{ad}(z) \circ \text{ad}(w)(x_i)) = \sum_{i=1}^m a_{ji} \delta_{ji} = \sum_{i=1}^m a_{ii} = \mathcal{H}_g(z, w),
\]

where the last equality is due to the very definition of \(\mathcal{H}_g\). This implies that, for each \(z \in g\), we have \(\text{ad}(\Omega)(z) - z \in \text{Ker}(\mathcal{H}_g) = 0\). Thus \(\text{ad}(\Omega)(z) = z\), i.e., (d) holds.

**Fact 3.2.** Let \(i \in \mathbb{N}\). Let \(v\) be a left \(g\)-module on which \(\Omega\) acts identically. Then the cohomology group \(H^i(g, v)\) is trivial.

**Proof.** We have an identity \(H^i(g, v) = \text{Ext}^i_{U_g}(A, v)\) of \(Z(U_g)\)-modules; see [Cartan and Eilenberg 1956, Chapter XIII, Sections 2 and 8]. As \(\Omega \in Z(U_g)\) acts trivially on \(A\) and identically on \(v\), the group \(\Omega \text{Ext}^i_{U_g}(A, v)\) is on one hand trivial and on the other hand equal to \(\text{Ext}^i_{U_g}(A, v)\). Therefore, \(\text{Ext}^i_{U_g}(A, v) = 0\), and \(H^i(g, v) = 0\).

**Theorem 3.3.** We recall that the Killing form \(\mathcal{H}_g\) on \(g\) is perfect. Then the group scheme \(\text{Aut}(g)\) over \(\text{Spec} A\) of Lie automorphisms of \(g\) is smooth and locally of finite presentation.

**Proof.** To check this, we can assume that the \(A\)-module \(g\) is free. The group scheme \(\text{Aut}(g)\) is a closed subgroup scheme of \(\text{GL}_g\) defined by a finitely generated ideal of the ring of functions of \(\text{GL}_g\). Thus \(\text{Aut}(g)\) is of finite presentation. Therefore, to show that \(\text{Aut}(g)\) is smooth over \(\text{Spec} A\), it suffices to show that, for each
We check that \( \theta \) we have to show that the restriction map \( \text{Aut}(g)(B) \to \text{Aut}(g)(B/j) \) is onto (see [Bosch et al. 1990, Chapter 2, Section 2.2, Proposition 6]). So as not to introduce extra notation by repeatedly tensoring with \( B \) over \( A \), we will assume that \( B = A \). Thus \( j \) is an ideal of \( A \) and we have to show that the restriction map \( \text{Aut}(g)(A) \to \text{Aut}(g)(A/j) \) is onto.

Let \( \bar{\sigma} : g/jg \to g/jg \) be a Lie automorphism. Let \( \sigma_0 : g \to g \) be an \( A \)-linear automorphism that lifts \( \bar{\sigma} \). Let \( jg_{\bar{\sigma}} \) be the left \( g \)-module which as an \( A \)-module is \( jg \) and whose left \( g \)-module structure is defined as follows: if \( x \in g \), then \( x \) acts on \( jg_{\bar{\sigma}} \) in the same way as \( \text{ad}(\bar{\sigma}(x)) \) (equivalently, as \( \text{ad}(\sigma_0(x)) \)) acts on the \( A \)-module \( jg = jg_{\bar{\sigma}} \); this makes sense as \( j^2 = 0 \). Let \( \theta : g \times g \to jg_{\bar{\sigma}} \) be the alternating map defined by the rule

\[
\theta(x, y) := [\sigma_0(x), \sigma_0(y)] - \sigma_0([x, y]) \quad \text{for all } x, y \in g. \tag{1}
\]

We check that \( \theta \) is a 2-cocycle, i.e., for all \( x, y, z \in g \) we have an identity

\[
d\theta(x, y, z) := x(\theta(y, z)) - y(\theta(x, z)) + z(\theta(x, y)) - \theta([x, y], z) + \theta([x, z], y) - \theta([y, z], x) = 0.
\]

Substituting (1) in the definition of \( d\theta \), we get that the expression \( d\theta(x, y, z) \) is a sum of 12 terms which can be divided into three groups as follows. The first group contains the three terms

\[-\sigma_0([[x, y], z]), \quad \sigma_0([[x, z], y]), \quad -\sigma_0([[y, z], x]);
\]

their sum is 0 due to the Jacobi identity and the fact that \( \sigma_0 \) is an \( A \)-linear map. The second group contains the six terms

\[[\sigma_0(x), \sigma_0(y, z)], \quad -[\sigma_0(x), \sigma_0(y, z)], \quad [\sigma_0(y), \sigma_0(x, z)], \quad -[\sigma_0(y), \sigma_0(x, z)], \quad [\sigma_0(z), \sigma_0(x, y)], \quad -[\sigma_0(z), \sigma_0(x, y)];
\]

obviously their sum is 0. The third group contains the three terms

\[[\sigma_0(x), [\sigma_0(y), \sigma_0(z)]], \quad -[\sigma_0(y), [\sigma_0(x), \sigma_0(z)]], \quad [\sigma_0(z), [\sigma_0(x), \sigma_0(y)]];\]

their sum is 0 due to the Jacobi identity. Thus, indeed, \( d\theta = 0 \).

As \( \Omega \) (i.e., \( \text{ad}(\Omega) \)) acts identically on \( g \) (see Lemma 3.1(d)), it also acts identically on \( jg \). But \( \Omega \) modulo \( j \) is fixed by the Lie automorphism \( \bar{\sigma} \) of \( g/jg \); see Lemma 3.1(c). Thus \( \Omega \) also acts identically on the left \( g \)-module \( jg_{\bar{\sigma}} \). From this and Fact 3.2 we get that \( H^2(g, jg_{\bar{\sigma}}) = 0 \). Thus \( \theta \) is the coboundary of a 1-cochain \( \delta : g \to jg_{\bar{\sigma}} \), i.e., we have

\[
\theta(x, y) = x(\delta(y)) - y(\delta(x)) - \delta([x, y]) \quad \text{for all } x, y \in g. \tag{2}
\]
Let \( \sigma : \mathfrak{g} \sim \rightarrow \mathfrak{g} \) be the \( A \)-linear isomorphism defined by the rule \( \sigma(x) := \sigma_0(x) - \delta(x) \); here \( \delta(x) \) is an element of the \( A \)-module \( j\mathfrak{g} = j\mathfrak{g}_\bar{\sigma} \). Due to formulas (1) and (2), we compute
\[
\sigma([x, y]) = \sigma_0([x, y]) - \delta([x, y]) \\
= [\sigma_0(x), \sigma_0(y)] - \theta(x, y) - \delta([x, y]) \\
= [\sigma_0(x), \sigma_0(y)] - x(\delta(y)) + y(\delta(x)) \\
= [\sigma_0(x), \sigma_0(y)] - \text{ad}(\bar{\sigma}(x))(\delta(y)) + \text{ad}(\bar{\sigma}(y))(\delta(x)) \\
= [\sigma_0(x), \sigma_0(y)] - \text{ad}(\sigma_0(x))(\delta(y)) + \text{ad}(\sigma_0)(y)(\delta(x)) \\
= [\sigma_0(x), \sigma_0(y)] - [\sigma_0(x), \delta(y)] + [\sigma_0(y), \delta(x)] \\
= [\sigma_0(x) - \delta(x), \sigma_0(y) - \delta(y)] - [\delta(x), \delta(y)] \\
= [\sigma(x), \sigma(y)] - [\delta(x), \delta(y)] \\
= [\sigma(x), \sigma(y)],
\]
where the last identity follows from \( j^2 = 0 \). Therefore, \( \sigma \) is a Lie automorphism of \( \mathfrak{g} \) that lifts the Lie automorphism \( \bar{\sigma} \) of \( \mathfrak{g} / j\mathfrak{g} \). Thus the restriction map \( \text{Aut}(\mathfrak{g})(A) \rightarrow \text{Aut}(\mathfrak{g})(A/j) \) is onto. \( \square \)

3.4. **Proof of Theorem 1.2.** The functor \( \mathcal{L}_Y \) is faithful; see Lemma 2.5.1. Thus to prove Theorem 1.2 it suffices to show that \( \mathcal{L}_Y \) is surjective on objects and that \( \mathcal{L}_Y \) is fully faithful. To check this, as \( \text{Adj-perf}_Y \) and \( \text{Lie-perf}_Y \) are groupoids on sets and as \( \mathcal{L}_Y \) is faithful, we can assume that \( Y = \text{Spec} \ A \) is affine. Thus to finish the proof we can check the following three properties:

(i) if \( \mathfrak{g} \) is an object of \( \text{Lie-perf}_Y \) (identified with a Lie algebra over \( A \)), then there exists a unique open subgroup scheme \( \text{Aut}(\mathfrak{g})^0 \) of \( \text{Aut}(\mathfrak{g}) \) which is an adjoint group scheme over \( Y \) and whose Lie algebra is the Lie subalgebra \( \text{ad}(\mathfrak{g}) \) of \( g|_\mathfrak{g} \) (therefore \( \mathfrak{g} = \text{ad}(\mathfrak{g}) \) is the image through \( \mathcal{L}_Y \) of the object \( \text{Aut}(\mathfrak{g})^0 \) of \( \text{Adj-perf}_Y \));

(ii) the group scheme \( \text{Aut}(\text{Aut}(\mathfrak{g})^0) \) of automorphisms of \( \text{Aut}(\mathfrak{g})^0 \) is \( \text{Aut}(\mathfrak{g}) \) acting on \( \text{Aut}(\mathfrak{g})^0 \) via inner conjugation (therefore \( \text{Aut}(\mathfrak{g})(A) = \text{Aut}(\text{Aut}(\mathfrak{g})^0)(A) \));

(iii) if \( G \) and \( H \) are two objects of \( \text{Adj-perf}_Y \) such that \( \text{Lie}(G) = \text{Lie}(H) \), then \( G \) and \( H \) are isomorphic.

To check the first two properties, we can assume that the \( A \)-module \( \mathfrak{g} \) is free and of rank \( m \in \mathbb{N}^* \). Let \( k \) be the residue field of an arbitrary point \( y \in Y \). It is well known that the Lie algebra \( \text{Lie}(\text{Aut}(\mathfrak{g})_k) \) is the Lie algebra of derivations of \( \mathfrak{g}_k := \mathfrak{g} \otimes_A k \). As this fact plays a key role in this paper, we include a proof of it. The tangent space of \( \text{Aut}(\mathfrak{g})_k \) at the identity element is identified with the set of
automorphisms $a$ of $g_k \otimes_k k[\varepsilon]/(\varepsilon^2)$, which modulo $\bar{\varepsilon} = \varepsilon + (\varepsilon^2)$ are the identity automorphism of $g_k$. We can write each such automorphism as
\[ a = 1_{g_k} \otimes_k k[\varepsilon]/(\varepsilon^2) + D_a \otimes \bar{\varepsilon}, \]
where $D_a$ is a $k$-linear endomorphism of $g_k$. The condition that $a$ respects the Lie bracket (i.e., $a([u, v] \otimes 1) = [a(u \otimes 1), a(v \otimes 1)]$ for all $u, v \in g_k$) is equivalent to the condition that $D_a$ is a derivation of $g_k$. The association $a \mapsto D_a$ identifies the tangent space of $\text{Aut}(g)_k$ at the identity element with the $k$-vector space of derivations of $g_k$. Under this identification, the Lie bracket of $a$ with an automorphism $b$ of $g_k \otimes_k k[\varepsilon_1]/(\varepsilon_1^2)$, which modulo $\bar{\varepsilon}_1 = \varepsilon_1 + (\varepsilon_1^2)$ is the identity automorphism of $g_k$, is the derivation of $g_k$ which corresponds to the automorphism
\[ aba^{-1}b^{-1} = 1_{g_k} \otimes k[\varepsilon_1]/(\varepsilon_1^2\varepsilon_1^2) + [D_a, D_b] \bar{\varepsilon}_1 \bar{\varepsilon}_1 \]
of $g_k \otimes_k k[\varepsilon_1]/(\varepsilon_1^2)$ and thus is the Lie bracket $[D_a, D_b] (\varepsilon_1$ is used here instead of $\varepsilon$ so that this last part makes sense). Therefore, $\text{Lie}(\text{Aut}(g)_k)$ is the Lie algebra of derivations of $g_k$.

As the Killing form $\mathfrak{H}_{g_k}$ is perfect, one argues as in [Humphreys 1972, Chapter II, Section 5.3, Theorem] that each derivation of $g_k$ is an inner derivation. Thus we have $\text{Lie}(\text{Aut}(g)_k) = \text{ad}(g) \otimes_A k$. As the group scheme $\text{Aut}(g)$ over $Y$ is smooth and locally of finite presentation (see Theorem 3.3), from [SGA 3, 1970, Exposé VIb, Corollary 4.4] we get that there exists a unique open subgroup scheme $\text{Aut}(g)_0$ of $\text{Aut}(g)$ whose fibers are connected. The fibers of $\text{Aut}(g)_0$ are open-closed subgroups of the fibers of $\text{Aut}(g)$ and thus are affine.

Let $N_k$ be a smooth, connected, unipotent, normal subgroup of $\text{Aut}(g)_0^l$. The Lie algebra $\text{Lie}(N_k)$ is a nilpotent ideal of $\text{Lie}(\text{Aut}(g)_0) = \text{ad}(g) \otimes_A k$. Thus
\[ \text{Lie}(N_k) \subseteq \text{Ker}(\mathfrak{H}_{\text{Lie}(\text{Aut}(g)_0^l)}) = \text{Ker}(\mathfrak{H}_{\text{ad}(g) \otimes_A k}); \]
see [Bourbaki 1989, Chapter I, Section 4.4, Proposition 6(b)]. As the Killing form $\mathfrak{H}_{\text{ad}(g) \otimes_A k}$ is perfect, we get $\text{Lie}(N_k) = 0$. Thus $N_k$ is the trivial subgroup of $\text{Aut}(g)_0^l$, and therefore the unipotent radical of $\text{Aut}(g)_0^l$ is trivial. Thus $\text{Aut}(g)_0^l$ is an affine, connected, smooth group over $\text{Spec} \ k$ whose unipotent radical is trivial. Therefore, $\text{Aut}(g)_0^l$ is a reductive group over $\text{Spec} \ k$; see [Borel 1991, Chapter IV, Section 11.21]. As $\text{Lie}(\text{Aut}(g)_0^l) = \text{ad}(g) \otimes_A k$ has trivial center, the group $\text{Aut}(g)_0^l$ is semisimple. Thus the smooth group scheme $\text{Aut}(g)_0^l$ of finite presentation over $Y$ has semisimple fibers. Therefore, $\text{Aut}(g)_0^l$ is a semisimple group scheme over $Y$; see [SGA 3, 1970, Exposé XVI, Theorem 5.2(ii)]. As $Z(\text{Aut}(g)_0^l)_k$ acts trivially on $\text{Lie}(\text{Aut}(g)_0^l) = \text{ad}(g) \otimes_A k$ and as $Z(\text{Aut}(g)_0^l)_k$ is a subgroup of $\text{Aut}(g)_k$, the group $Z(\text{Aut}(g)_0^l)_k$ is trivial. This implies that the finite, flat group scheme $Z(\text{Aut}(g)_0^l)$ is trivial and thus $\text{Aut}(g)_0^l$ is an adjoint group scheme.
The Lie subalgebras \( \text{Lie}(\text{Aut}(g)^0) \) and \( \text{ad}(g) \) of \( g_l \) are free \( A \)-submodules of the Lie subalgebra \( l \) of \( g_l \) formed by derivations of \( g \). As, for each point \( y \) of \( Y \), we have \( \text{Lie}(\text{Aut}(g)^0)_Y = \text{ad}(g) \otimes_A k = l \otimes_A k \), the Lie subalgebra \( l \) is locally generated by either \( \text{Lie}(\text{Aut}(g)^0) \) or \( \text{ad}(g) \). We easily get the identities \( \text{Lie}(\text{Aut}(g)^0) = \text{ad}(g) = l \).

The group scheme \( \text{Aut}(g) \) acts via inner conjugation on \( \text{Aut}(g)^0 \). As we have \( \text{Lie}(\text{Aut}(g)^0) = \text{ad}(g) \) and as \( \text{Aut}(g) \) is a closed subgroup scheme of \( \text{GL}_g \), the inner conjugation homomorphism \( \text{Aut}(g) \to \text{Aut}(\text{Aut}(g)^0) \) has trivial kernel. As \( \text{Aut}(\text{Aut}(g)^0) \) is a closed subgroup scheme of \( \text{Aut}(\text{Lie}(\text{Aut}(g)^0)) = \text{Aut}(\text{ad}(g)) \) (see Lemma 2.5.1), we can identify naturally \( \text{Aut}(\text{Aut}(g)^0) \) with a closed subgroup scheme of \( \text{Aut}(g) \). From the last two sentences, we get that \( \text{Aut}(\text{Aut}(g)^0) = \text{Aut}(g) \).

Thus both properties (i) and (ii) hold.

To check that property (iii) holds, let \( g = \text{Lie}(G) = \text{Lie}(H) \). It suffices to show that \( G \) and \( H \) are identified with \( \text{Aut}(g)^0 \). We will work only with \( G \). The adjoint representation \( G \to \text{GL}_g \) factors as composite closed embedding homomorphisms \( G \to \text{Aut}(g)^0 \to \text{Aut}(g) \to \text{GL}_g \) (see Lemma 2.5.1 and [SGA 3 1970, Exposé XXIV, Théorème 1.3]). We get a closed embedding homomorphism \( G \to \text{Aut}(g)^0 \) between adjoint group schemes that have the same Lie algebra \( g \) (see also property (i)). By reasons of dimensions, the geometric fibers of the closed embedding homomorphism \( G \to \text{Aut}(g)^0 \) are isomorphisms, and therefore \( G \to \text{Aut}(g)^0 \) is an isomorphism. Thus property (iii) holds as well. \( \square \)

The next proposition details the range of applicability of Theorem 1.2.

**Proposition 3.5.** (a) We recall that \( k \) is a field. Let \( H \) be a nontrivial semisimple group over \( \text{Spec} \ k \). Then the Killing form \( \mathfrak{K}_{\text{Lie}(H)} \) is perfect if and only if the following two conditions hold:

(i) either \( \text{char}(k) \) equals 0 or \( \text{char}(k) \) is an odd prime \( p \) and \( H^{\text{ad}} \) has no simple factor of isotypic \( A_{p-1}, B_{p-1}^+ \), \( C_{p-1} \), or \( D_{p+1} \) Dynkin type (here \( n \in \mathbb{N}^+ \));

(ii) if \( \text{char}(k) = 3 \) (resp. \( \text{char}(k) = 5 \)), then \( H^{\text{ad}} \) has no simple factor of isotypic \( E_6, E_7, E_8, F_4, G_2 \) (resp. isotypic \( E_7, E_8 \) Dynkin type.

(b) If \( \mathfrak{K}_{\text{Lie}(H)} \) is perfect, then the central isogenies \( H^{\text{sc}} \to H \to H^{\text{ad}} \) are étale; thus, by identifying tangent spaces at identity elements, \( \text{Lie}(H^{\text{sc}}) = \text{Lie}(H) = \text{Lie}(H^{\text{ad}}) \).

**Proof.** We can assume that \( k = \bar{k} \) and that \( \text{trdeg}(k) < \infty \). If \( \text{char}(k) = 0 \), then \( \text{Lie}(H) \) is a semisimple Lie algebra over \( k \) and therefore the proposition follows from [Humphreys 1972, Chapter II, Section 5.1, Theorem]. Thus we can assume \( \text{char}(k) \) is a prime \( p \in \mathbb{N}^+ \). If conditions (i) and (ii) hold, then \( p \) does not divide the order of the finite group scheme \( Z(H^{\text{sc}}) = \ker(H^{\text{sc}} \to H^{\text{ad}}) \) (see [Bourbaki 2002, Plates I to IX]) and therefore (a) implies (b).

Let \( W(k) \) be the ring of \( p \)-typical Witt vectors with coefficients in \( k \). Let \( H_{W(k)} \) be a semisimple group scheme over \( \text{Spec} W(k) \) that lifts \( H \); see [SGA 3 1970, Exposé...
XXIV, Proposition 1.21]. Then $\text{Lie}(H^\text{sc}_{W(k)}) = \text{Lie}(H^\text{ad}_{W(k)})$. This implies that:

(iii) the Killing form $\mathcal{K}_{\text{Lie}(H^\text{ad}_{W(k)})}$ is the composite of the natural $W(k)$-linear map $\text{Lie}(H^\text{ad}_{W(k)}) \times \text{Lie}(H^\text{ad}_{W(k)}) \to \text{Lie}(H^\text{ad}_{W(k)})$ with $\mathcal{K}_{\text{Lie}(H^\text{ad}_{W(k)})}$;

(iv) the Killing form $\mathcal{K}_{\text{Lie}(H^\text{sc}_{W(k)})}$ is the composite of the natural $W(k)$-linear map $\text{Lie}(H^\text{sc}_{W(k)}) \times \text{Lie}(H^\text{sc}_{W(k)}) \to \text{Lie}(H^\text{sc}_{W(k)})$ with $\mathcal{K}_{\text{Lie}(H^\text{sc}_{W(k)})}$.

We prove (a). We have $\text{Ker}(\text{Lie}(H) \to \text{Lie}(H^\text{ad})) \subseteq \text{Ker}(\mathcal{K}_{\text{Lie}(H)})$; see property (iii). If $\mathcal{K}_{\text{Lie}(H)}$ is perfect, then $\text{Ker}(\text{Lie}(H) \to \text{Lie}(H^\text{ad})) = 0$ and therefore $\text{Lie}(H) = \text{Lie}(H^\text{ad})$. Thus to prove (a) we can assume that $H = H^\text{ad}$ is adjoint. Even more, to prove (a) we can also assume that the adjoint group $H$ is simple; let $b$ be the Lie type of $H$. If $b$ is not of classical Lie type, then $\mathcal{K}_{\text{Lie}(H)}$ is perfect if and only if either $p > 5$ or $p = 5$ and $b \neq E_8$ (see [Humphreys 1995, Table, p. 49]). Thus to prove (a), we can assume that $b$ is a classical Lie type. We fix a morphism $\text{Spec } C \to \text{Spec } W(k)$.

Suppose that $b$ is either $A_n$ or $C_n$. By the standard trace form on $\text{Lie}(H^\text{sc}_{W(k)})$ (resp. $\text{Lie}(H^\text{sc}_{W(k)})$) we mean the trace form $\mathcal{T}$ (resp. $\mathcal{T}_{W(k)}$ or $\mathcal{T}_C$) associated to the faithful representation of $H^\text{sc}_{W(k)}$ (resp. $H^\text{sc}_{W(k)}$ or $H^\text{sc}_{C}$) of rank $n + 1$ if $b = A_n$ and of rank $2n$ if $b = C_n$. We have $\mathcal{K}_{\text{Lie}(H^\text{sc}_{W(k)})} = 2(n + 1)\mathcal{T}_C$; see [Helgason 1978, Chapter III, Section 8, (5) and (22)]. This identity implies that we also have $\mathcal{K}_{\text{Lie}(H^\text{sc}_{W(k)})} = 2(n + 1)\mathcal{T}_{W(k)}$ and thus $\mathcal{K}_{\text{Lie}(H^\text{sc}_{W(k)})} = 2(n + 1)\mathcal{T}$. If $p$ does not divide $2(n + 1)$, then $\text{Lie}(H^\text{sc}_{W(k)}) = \text{Lie}(H)$ and it is well known that $\mathcal{T}$ is perfect; thus $\mathcal{K}_{\text{Lie}(H^\text{sc}_{W(k)})} = \mathcal{K}_{\text{Lie}(H)} = 2(n + 1)\mathcal{T}$ is perfect. Suppose that $p$ divides $2(n + 1)$. This implies that $\mathcal{K}_{\text{Lie}(H^\text{sc}_{W(k)})}$ is the trivial bilinear form on $\text{Lie}(H^\text{sc}_{W(k)})$. From this and property (iv) we get that the restriction of $\mathcal{K}_{\text{Lie}(H)}$ to $\text{Im}(\text{Lie}(H^\text{sc}_{W(k)}) \to \text{Lie}(H))$ is trivial. As $\text{dim}_k(\text{Lie}(H)/\text{Im}(\text{Lie}(H^\text{sc}_{W(k)}) \to \text{Lie}(H))) = 1$ and $\text{dim}_k(\text{Lie}(H)) \geq 3$, we easily get that $\mathcal{K}_{\text{Lie}(H)}$ is degenerate.

Suppose that $b = B_n$ (resp. $b = D_n$ with $n \geq 4$). If $p > 2$, then we have $\text{Lie}(H^\text{sc}_{W(k)}) = \text{Lie}(H)$. Moreover, using [Helgason 1978, Chapter III, Section 8, (11) and (15)], as in the previous paragraph we argue that $\mathcal{K}_{\text{Lie}(H)}$ is perfect if $p$ does not divide $2(n - 1)$ (resp. if $p$ does not divide $2(n - 1)$) and is degenerate if $p$ divides $2n - 1$ (resp. if $p$ divides $2n - 1$).

We are left to show that $\mathcal{K}_{\text{Lie}(H)}$ is degenerate if $p = 2$ and $b = B_n$. The group $H$ is (isomorphic to) the SO-group of the quadratic form $x_0^2 + x_1x_{n+1} + \cdots + x_nx_{2n}$ on $W := k^{2n+1}$. Let $\{e_i, j | i, j \in \{0, 1, \ldots, n\}\}$ be the standard $k$-basis for $\mathfrak{gl}_W$. The direct sum $\mathfrak{n}_n := \bigoplus_{i=1}^{2n} k e_{0,i}$ is a nilpotent ideal of $\text{Lie}(H)$; see [Borel 1991, Chapter V, Section 23.6]. Thus $\mathfrak{n}_n \subseteq \text{Ker}(\mathcal{K}_{\text{Lie}(H)})$. by [Bourbaki 1989, Chapter I, Section 4.4, Proposition 6(b)] applied to the adjoint representation of $\text{Lie}(H)$. Therefore, $\mathcal{K}_{\text{Lie}(H)}$ is degenerate.
We conclude that $\mathcal{H}_{\text{Lie}(H)}$ is perfect if and only if both conditions (i) and (ii) hold. Therefore, (a) (and thus also (b)) holds.

Remark 3.6. Let $A$ and $g$ be as in the beginning of this section.

(a) Let $p \in \mathbb{N}^*$ be a prime. Suppose that $A$ is an algebraically closed field of characteristic $p$. Let $G$ be an adjoint group over $\text{Spec } A$ such that $g = \text{Lie}(G)$; see Theorem 1.2. We have $p \neq 2$; see Proposition 3.5. Let $G_\mathbb{Z}$ be the unique (up to isomorphism) split, adjoint group scheme over $\text{Spec } \mathbb{Z}$ such that $G$ is the pullback of $G_\mathbb{Z}$ to $\text{Spec } A$; see [SGA 3 \text{III} 1970, Exposé XXV, Corollary 1.3]. We have $g = \text{Lie}(G_\mathbb{Z}) \otimes_\mathbb{Z} A$, i.e., $g$ has a canonical model $\text{Lie}(G_\mathbb{Z})$ over $\mathbb{Z}$. For $p > 7$, this result was obtained in [Curtis 1957, Section 5, Theorem]. For $p > 3$, this result was obtained by Seligman [1967, Chapter II, Section 10], Mills [1957], Mills and Seligman [1957], and Block and Zassenhaus [1964]. For $p = 3$, this result was obtained in [Brown 1969, Theorem 4.1]. It seems to us that the fact that $p \neq 2$ (i.e., that all Killing forms of finite dimensional Lie algebras over fields of characteristic 2 are degenerate) is new.

(b) Let $B \twoheadrightarrow A$ be an epimorphism of commutative $\mathbb{Z}$-algebras whose kernel $j$ is a nilpotent ideal. Then $g$ has, up to isomorphisms, a unique lift to a Lie algebra over $B$ which as a $B$-module is projective and finitely generated. One can prove this statement using cohomological methods as in the proof of Theorem 3.3. The statement also follows from Theorem 1.2 and the fact that $\text{Aut}(g)^0$ has, up to isomorphisms, a unique lift to an adjoint group scheme over $\text{Spec } B$ (this can be easily checked at the level of torsors of adjoint group schemes; see [SGA 3 \text{III} 1970, Exposé XXIV, Corollaries 1.17 and 1.18]).

Corollary 3.7. Let $\text{Sc-perf}_Y$ be the category whose objects are simply connected semisimple group schemes over $Y$ with the property that their Lie algebra $\mathcal{O}_Y$-modules have perfect Killing forms and whose morphisms are isomorphisms of group schemes. Then the functor $\mathcal{L}^\text{sc}_Y : \text{Sc-perf}_Y \rightarrow \text{Lie-perf}_Y$, which associates to a morphism $f : G \to H$ of $\text{Sc-perf}_Y$ the morphism $df : \text{Lie}(G) \to \text{Lie}(H)$ of Lie-perf$_Y$ which is the differential of $f$, is an equivalence of categories.

Proof. The functor $\mathcal{L}^\text{sc}_Y$ is the composite of the canonical (‘division by the centers’) functor $\mathcal{L}_Y : \text{Sc-perf}_Y \rightarrow \text{Adj-perf}_Y$ with $\mathcal{L}_Y$; the functor $\mathcal{L}_Y$ makes sense (see Proposition 3.5(b)) and it is an equivalence of categories. Thus the corollary follows from Theorem 1.2.

Corollary 3.8. The category Lie-perf$_Y$ has a nonzero object if and only if $Y$ is a nonempty $\text{Spec } \mathbb{Z}[[\frac{1}{p}]]$-scheme.

Proof. The ‘if’ part is implied by the fact that an $\mathfrak{sl}_2$ Lie algebra $\mathcal{O}_Y$-module has perfect Killing form. The ‘only if’ part follows from the relation $p \neq 2$ of Remark 3.6(a).
4. Proof of Theorem 1.4

In this section we prove Theorem 1.4. See Sections 4.1 and 4.2 for the proofs of Theorem 1.4(a) and Theorem 1.4(b) (respectively). In Remark 4.3 we point out that the hypotheses of Theorem 1.4 are indeed needed in general. We will use the notation presented in Section 1.

4.1. Proof of Theorem 1.4(a). To prove Theorem 1.4 we can assume $Y$ is also integral. Let $K := K_Y$ be a field. If $H$ is a reductive group scheme over $Y$, then we have $\mathfrak{D}(H) = \mathfrak{D}(H_U)$ and thus the uniqueness parts of Theorem 1.4 follow from Proposition 2.2(b). Let $l$ be the Lie algebra $\mathcal{O}_Y$-module which extends $\text{Lie}(G_U)$.

We prove Theorem 1.4(a). Due to the uniqueness part, to prove Theorem 1.4(a) we can assume $Y = \text{Spec } A$ is also local and strictly henselian. Let $g := l(Y)$ be the Lie algebra over $A$ of global sections of $l$. As $U$ is connected, based on [SGA 3 III 1970, Exposé XXII, Proposition 2.8] we can speak about the split, adjoint group scheme $S$ over $Y$ of the same Lie type as all geometric fibers of $G_U$. Let $s := \text{Lie}(S)$. Let $\text{Aut}(S)$ be the group scheme over $Y$ of automorphisms of $S$. We have a short exact sequence $1 \to S \to \text{Aut}(S) \to C \to 1$, where $C$ is a finite, étale, constant group scheme over $Y$ (see [SGA 3 III 1970, Exposé XXIV, Theorem 1.3]). Let $\gamma \in H^1(U, \text{Aut}(S)_U)$ be the class that defines the form $G_U$ of $S_U$.

We recall that $\text{GL}_g$ and $\text{GL}_s$ are the reductive group schemes over $Y$ of linear automorphisms of $g$ and $s$ (respectively). The adjoint representations define closed embedding homomorphisms $j_U : G_U \hookrightarrow \text{GL}_g(U)$ and $i : S \hookrightarrow \text{GL}_s$ and, moreover, $i$ extends naturally to a closed embedding homomorphism $\text{Aut}(S) \hookrightarrow \text{GL}_s$; see Lemma 2.5.1. Let $\delta \in H^1(U, (\text{GL}_s)_U)$ be the image of $\gamma$ via the homomorphism $\text{Aut}(S)_U \hookrightarrow \text{GL}_s(U)$.

We recall that the quotient sheaf for the faithfully flat topology of $Y$ of the action of $S$ on $\text{GL}_s$ via right translations is representable by a $Y$-scheme $\text{GL}_s/S$ that is affine and that causes $\text{GL}_s$ to be a right torsor of $S$ over $\text{GL}_s/S$ (see [Colliot-Thélène and Sansuc 1979, Corollary 6.12]). Thus $\text{GL}_s/S$ is a smooth, affine $Y$-scheme. The finite, étale, constant group scheme $C$ acts naturally (from the right) on $\text{GL}_s/S$ and this action is free (see Lemma 2.5.1). From [SGA 3 I 1970, Exposé V, Theorem 4.1] we get that the quotient $Y$-scheme $(\text{GL}_s/S)/C$ is affine and that the quotient epimorphism $\text{GL}_s/S \twoheadrightarrow (\text{GL}_s/S)/C$ is a finite étale cover. Thus $(\text{GL}_s/S)/C$ is a smooth, affine scheme over $Y$ that represents the quotient sheaf for the faithfully flat topology of $Y$ of the action of $\text{Aut}(S)$ on $\text{GL}_s$ via right translations. From constructions we get that $\text{GL}_s$ is a right torsor of $\text{Aut}(S)$ over $\text{GL}_s/\text{Aut}(S) := (\text{GL}_s/S)/C$.

The twist of $i_U$ via the class $\gamma$ is $j_U$. This implies that the class $\delta$ defines the torsor that parametrizes isomorphisms between the pullbacks to $U$ of the vector group
We denote also by $\eta / H_{5115}$ which as an
which extends
Y \text{(b)} Suppose that $\eta$ is the restriction of a class in $H^0(Y, GL_{\eta, U} / Aut(S))_U$. But $H^0(U, GL_{\eta, U} / Aut(S))_U$ equals $H^0(Y, GL_{\eta} / Aut(S))$ (see Proposition 2.2(b)) and therefore $\gamma$ is the restriction of a class in $H^1(Y, Aut(S))$. As $Y$ is strictly henselian, each class in $H^1(Y, Aut(S))$ is trivial. Thus $\gamma$ is the trivial class. Therefore, the group schemes $G_U$ and $S_U$ are isomorphic. Thus $G_U$ extends to an adjoint group scheme $G$ over $Y$ isomorphic to $S$. \qed

\textbf{4.2. Proof of Theorem 1.4(b).} Let $\eta : \tilde{K} \to U$ be the geometric point of $U$ which is the composite of the natural morphisms $\text{Spec } \tilde{K} \to \text{Spec } K$ and $\text{Spec } K \to U$. We denote also by $\eta : \tilde{K} \to Y$ the resulting geometric point of $Y$. As $Y$ (resp. $U$) is normal and locally noetherian, from [SGA 3\text{è} 1970, Exposé X, Theorems 5.16 and 7.1] we get that there exists an antiequivalence of categories between the category of tori over $Y$ (resp. continuous $\pi_1(Y, \eta)$-representations) on free $\mathbb{Z}$-modules of finite rank. As the pair $(Y, Y \setminus U)$ is quasipure, we have a canonical identification $\pi_1(U, \eta) = \pi(Y, \eta)$. From the last two sentences we get that there exists a unique torus $H_{ab}$ over $Y$ which extends $H_{U, ab}$.

Let $H^{ad}$ be the adjoint group scheme over $Y$ that extends $H_{U, ad}$; see Theorem 1.4(a). Let $F \to H^{ad} \times_Y H_{ab}$ be the central isogeny over $Y$ that extends the central isogeny $H_K \to H_K^{ad} \times \text{Spec } K H_{ab}$; see Lemma 2.3.1(a). Both $F_U$ and $H_U$ are the normalization of $H_{ab}^{ad} \times_U H_{U, ad}$ in $H_K$; see Lemma 2.3.1(b). Thus $H_U = F_U$ extends uniquely to a reductive group scheme $H := F$ over $Y$ (see the first paragraph of Section 4.1 for the uniqueness part). \qed

\textbf{Remark 4.3.} (a) Let $Y_1 \to Y$ be a finite, nonétale morphism between normal, noetherian, integral Spec $\mathbb{Z}_{(2)}$-schemes such that there exists an open subscheme $U$ of $Y$ with the properties that: (i) $Y \setminus U$ has codimension in $Y$ at least 2, and (ii) $Y_1 \times_Y U \to U$ is a Galois cover of degree 2. Let $H_U$ be the rank-1 nonsplit torus over $U$ that splits over $Y_1 \times_Y U$. Then $H_U$ does not extend to a smooth, affine group scheme over $Y$. If, moreover, $Y = \text{Spec } A$ is an affine Spec $\mathbb{F}_2$-scheme, then we have $\text{Lie}(H_U)(U) = A$ and therefore $\text{Lie}(H_U)$ extends to a Lie algebra $\mathfrak{g}_Y$-module which as an $\mathfrak{g}_Y$-module is free. Thus the quasipure part of the hypotheses of Theorem 1.4(b) is needed in general.

(b) Suppose that $Y = \text{Spec } A$ is local, strictly henselian, regular, and of dimension $n \geq 3$. Let $K := K_Y$. Let $d \in \mathbb{N}^*$ be such that there exists an $A$-submodule $M$ of $K_d$ that contains $A^d$, is of finite type, is not free, and satisfies the identity $M = \bigcap_{V \in \mathfrak{g}(Y)} M \otimes_A V$ (inside $M \otimes_A K$). A typical example (communicated to us by Serre) is $d = n - 1$ and $M \xrightarrow{f} \text{Coker}(f)$, where the $A$-linear map $f : A \to A^n$ takes 1 to an $n$-tuple $(x_1, \ldots, x_n) \in A^n$ of regular parameters of $A$. 

\begin{thebibliography}{99}

\bibitem{A} Adrian Vasiu

\end{thebibliography}
Let $F$ be the coherent $\mathcal{O}_Y$-module defined by $M$. Let $U$ be an open subscheme of $Y$ such that $Y \setminus U$ has codimension in $Y$ at least 2 and the restriction $F_U$ of $F$ to $U$ is a locally free $\mathcal{O}_U$-module. Let $H_U$ be the reductive group scheme over $U$ of linear automorphisms of $F$. We recall the reason why the assumption that $H_U$ extends to a reductive group scheme $H$ over $Y$ leads to a contradiction. The group scheme $H$ is isomorphic to $\text{GL}_{d,A}$ (as $A$ is strictly henselian), and therefore there exists a free $A$-submodule $L$ of $K^d$ of rank $d$ such that $H = \text{GL}_L$. As $A$ is a unique factorization domain (being local and regular), it is easy to see that there exists an element $f \in K$ such that the identity $M \otimes_A V = fL \otimes_A V$ holds for each $V \in \mathfrak{D}(Y)$. This implies that $M = fL$. Thus $M$ is a free $A$-module. Contradiction.

As $H_U$ does not extend to a reductive group scheme over $Y$ and as the pair $(Y, Y \setminus U)$ is quasipure, from Section 4.2 we get that $H_U^{\text{ad}}$ also does not extend to an adjoint group scheme over $Y$. Thus the Lie part of the hypotheses of Theorem 1.4(a) is needed in general.

5. Extending homomorphisms via schematic closures

In this section we prove two results on extending homomorphisms of reductive group schemes via taking (normalizations of) schematic closures. Proposition 5.1 complements Theorem 1.4(b) and Proposition 2.5.2, and Proposition 5.2 refines [Vasiu 1999, Lemma 3.1.6].

**Proposition 5.1.** Let $Y$ be a normal, noetherian, integral scheme. Let $K := K_Y$. Let $U$ be an open subscheme of $Y$ such that the codimension of $Y \setminus U$ in $Y$ is at least 2. Let $H_U$ be a reductive group scheme over $U$ and let $G$ be a reductive group scheme over $Y$. We assume we have a finite homomorphism $\rho_U : H_U \to G_U$ whose generic fiber over $\text{Spec} K$ is a closed embedding. We assume that one of the following two properties holds:

(i) $H_U$ extends to a reductive group scheme $H$ over $Y$, or

(ii) $Y = \text{Spec} R$ is a local regular scheme of dimension 2 (thus $U$ is the complement in $Y$ of the closed point of $Y$).

Then the following three properties hold:

(a) There exists a unique reductive group scheme $H$ over $Y$ which extends $H_U$.

(b) The homomorphism $\rho_U$ extends uniquely to a finite homomorphism $\rho : H \to G$ between reductive group schemes over $Y$.

(c) If there exists a point of $Y \setminus U$ of characteristic 2, we assume that $H_K$ has no normal subgroup that is adjoint of isotypic $B_n$ Dynkin type for some $n \in \mathbb{N}^*$. Then $\rho : H \to G$ is a closed embedding.
Proof. If (i) holds, then the uniqueness of $H_U$ follows from Proposition 2.2(b). Thus to prove (a) we can assume that property (ii) holds. As (ii) holds, the pair $(Y, Y \setminus U)$ is quasipure (see Section 1) and the Lie algebra $\mathcal{O}_U$-module $\text{Lie}(H_U)$ extends to a Lie algebra $\mathcal{O}_Y$-module which is a free $\mathcal{O}_Y$-module (by Proposition 2.2(c) and the fact that $Y$ is local). Thus the hypotheses of Theorem 1.4(b) hold, and therefore from Theorem 1.4(b) we get that there exists a unique reductive group scheme $H$ over $Y$ that extends $H_U$. Thus (a) holds.

To prove (b) and (c) we can assume that $Y = \text{Spec } R$ is an affine scheme. We write $H = \text{Spec } R_H$ and $G = \text{Spec } R_G$. As $\mathfrak{d}(H) = \mathfrak{d}(H_U)$ and $\mathfrak{d}(G) = \mathfrak{d}(G_U)$, from Proposition 2.2(a) we get that $R_H$ and $R_G$ are the $R$-algebras of global functions of $H_U$ and $G_U$ (respectively). Let $R_G \to R_H$ be the $R$-homomorphism defined by $\rho_U$ and let $\rho : H \to G$ be the morphism of $Y$-schemes it defines. The morphism $\rho$ is a homomorphism, as it is so generically. To check that $\rho$ is finite, we can assume that $R$ is complete. Thus $R_H$ and $R_G$ are excellent rings; see [Matsumura 1980, Section 34]. Therefore, the normalization $H' = \text{Spec } R_{H'}$ of the schematic closure of $H_K$ in $G$ is a finite, normal $G$-scheme.

The identity components of the reduced geometric fibers of $\rho$ are trivial groups; see Proposition 2.5.2(a) or (b). Thus $\rho$ is a quasifinite morphism. From Zariski’s main theorem (see [Grothendieck 1966, Theorem 8.12.6]) we get that $H$ is an open subscheme of $H'$. But from Proposition 2.5.2(b) we get that the morphism $H \to H'$ satisfies the valuative criterion of properness with respect to discrete valuation rings which contain $R$. As each local ring of $H'$ is dominated by such a discrete valuation ring, we get that the morphism $H \to H'$ is surjective. Therefore, the open, surjective morphism $H \to H'$ is an isomorphism. Thus $\rho$ is finite, i.e., (b) holds.

We prove (c). The pullback of the homomorphism $\rho : H \to G$ via each dominant morphism $\text{Spec } V \to Y$, with $V$ a discrete valuation ring, is a closed embedding (see Proposition 2.5.2(c)). This implies that the fibers of $\rho$ are closed embeddings. Thus the homomorphism $\rho$ is a closed embedding; see Theorem 2.5. □

We have the following refinement of [Vasiu 1999, Lemma 3.1.6].

Proposition 5.2. Let $G$ be a reductive group scheme over a reduced, affine scheme $Y = \text{Spec } A$. Let $K$ be a localization of $A$. Let $s \in \mathbb{N}^*$. For $j \in\{1, \ldots , s\}$ let $G_{j,K}$ be a reductive, closed subgroup scheme of $G_K$. We assume that the group subschemes $G_{j,K}$ commute among themselves and that either

(i) the direct sum $\bigoplus_{j=1}^s \text{Lie}(G_{j,K})$ is a Lie subalgebra of $\text{Lie}(G_K)$, or

(ii) $s = 2$, $G_{1,K}$ is a torus, and $G_{2,K}$ is a semisimple group scheme.

Then the closed subgroup scheme $G_{0,K}$ of $G_K$ generated by the group subschemes $G_{j,K}$ exists and is reductive. Moreover:

(a) If condition (i) holds, then $\text{Lie}(G_{0,K}) = \bigoplus_{j=1}^s \text{Lie}(G_{j,K})$. 


(b) We assume that for each \( j \in \{1, \ldots, s\} \) the schematic closure \( G_j \) of \( G_{j,K} \) in \( G \) is a reductive group scheme over \( Y \). Then the schematic closure \( G_0 \) of \( G_{0,K} \) in \( G \) is a reductive, closed subgroup scheme of \( G \).

**Proof.** Let \( \Lambda \) be the category whose objects \( \text{Ob}(\Lambda) \) are finite subsets of \( K \) and whose morphisms are the inclusions of subsets. For \( \alpha \in \text{Ob}(\Lambda) \), let \( K_\alpha \) be the \( \mathbb{Z} \)-subalgebra of \( K \) generated by \( \alpha \) and let \( A_\alpha := A \cap K_\alpha \). We have \( K = \text{ind lim}_{\alpha \in \text{Ob}(\Lambda)} K_\alpha \) and \( A = \text{ind lim}_{\alpha \in \text{Ob}(\Lambda)} A_\alpha \). The reductive group schemes \( G_{j,K} \) are of finite presentation. Based on this and [Grothendieck 1966, Theorems 8.8.2 and 8.10.5], one gets that there exists a \( \beta \in \text{Ob}(\Lambda) \) such that each \( G_{j,K} \) is the pullback of a closed subgroup scheme \( G_{j,K_\beta} \) of \( G_{K_\beta} \). For \( \alpha \supseteq \beta \), the set \( C(\alpha) \) of points of \( \text{Spec} \, K_\alpha \) with the property that the fibers over them of all morphisms \( G_{j,K_\alpha} \to \text{Spec} \, K_\alpha \) are (geometrically) connected is a constructible set (see [Grothendieck 1966, Theorem 9.7.7]). We have \( \text{proj lim}_{\alpha \in \text{Ob}(\Lambda)} C(\alpha) = \text{Spec} \, K \). From this and [Grothendieck 1966, Theorem 8.5.2], we get that there exists a \( \beta_1 \in \text{Ob}(\Lambda) \) such that \( \beta_1 \supseteq \beta \) and \( C(\beta_1) = \text{Spec} \, K_{\beta_1} \). Thus, by replacing \( \beta \) with \( \beta_1 \), we can assume that the fibers of all morphisms \( G_{j,K_{\beta}} \to \text{Spec} \, K_{\beta} \) are connected. A similar argument shows that, by enlarging \( \beta \), we can assume that all morphisms \( G_{j,K_{\beta}} \to \text{Spec} \, K_{\beta} \) are smooth and that their fibers are reductive groups (the role of [Grothendieck 1966, Theorem 9.7.7] being replaced by [Grothendieck 1966, Corollary 9.9.5] applied to the \( \mathcal{O}_{G_j,K_\alpha} \)-module \( \text{Lie}(G_{j,K_\alpha}) \) and by [SGA 3, 2010, Exposé XIX, Corollary 2.6]). Thus each \( G_{j,K_\beta} \) is a reductive closed subgroup scheme of \( G_{K_\beta} \). The smooth group schemes \( G_{j,K_{\beta}} \) commute among themselves, as this is so after pullback through the dominant morphism \( \text{Spec} \, K \to \text{Spec} \, K_{\beta} \). By enlarging \( \beta \), we can also assume that either condition (i) or condition (ii) holds for the \( G_{j,K_{\beta}} \), and that \( K_{\beta} \) is a localization of \( A_{\beta} \). By replacing \( A \) with the local ring of \( \text{Spec} \, A_{\beta} \) dominated by \( A \), to prove the proposition we can assume that \( A \) is a localization of a reduced, finitely generated \( \mathbb{Z} \)-algebra.

Using induction on \( s \in \mathbb{N}^* \), it suffices to prove the proposition for \( s = 2 \). Moreover, we can assume that \( K = K_Y \). For the sake of flexibility, in what follows we will only assume that \( A \) is a reduced, noetherian \( \mathbb{Z} \)-algebra; thus \( K \) is a finite product of fields. As all the statements of the proposition are local for the étale topology of \( Y \), it suffices to prove the proposition under the extra assumption that \( G_1 \) and \( G_2 \) are split (see Proposition 2.3). Let \( C_K := G_{1,K} \cap G_{2,K} \) be a closed subgroup scheme of \( G_{j,K} \) that commutes with \( G_{j,K} \), \( j \in \{1, 2\} \). The Lie algebra \( \text{Lie}(C_K) \) is included in \( \text{Lie}(G_{1,K}) \cap \text{Lie}(G_{2,K}) \) and therefore it is trivial if condition (i) holds. Thus if condition (i) holds, then \( C_K \) is a finite, étale, closed subgroup scheme of \( Z(G_{j,K}) \). If condition (ii) holds, then \( C_K \) is a closed subgroup scheme of both \( G_{1,K} = Z(G_{1,K}) \) and \( Z(G_{2,K}) \) and thus (as \( K \) is a finite product of fields) it is a finite group scheme of multiplicative type.

Let \( C \) be the schematic closure of \( C_K \) in \( G \). Let \( T_j \) be a maximal torus of \( G_j \).
We have
\[ C_K \subseteq Z(G_{1,K}) \cap Z(G_{2,K}) \subseteq T_{1,K} \cap T_{2,K} \subseteq G_{1,K} \cap G_{2,K} = C_K \]
and thus \( C_K = T_{1,K} \cap T_{2,K} \). Let \( T_1 \times_Y T_2 \to G \) be the product homomorphism. The kernel \( \mathfrak{g} \) of this product homomorphism is a group scheme over \( Y \) of multiplicative type (see Lemma 2.3.2(a)) isomorphic to \( T_1 \cap T_2 \). But \( \mathfrak{g}_K \to C_K \) is a finite group scheme over Spec \( K \) and therefore \( \mathfrak{g} \) is a finite, flat group scheme over \( Y \) of multiplicative type (see Lemma 2.3.2(b)). Thus \( T_1 \cap T_2 \) is a finite, flat group scheme over \( Y \). From this, the identity \( C_K = (T_1 \cap T_2)_K \), and the definition of \( C \) we get that \( C = T_1 \cap T_2 \). We conclude that \( C \) is a finite, flat group scheme over \( Y \) of multiplicative type contained in the center of both \( G_1 \) and \( G_2 \). We embed \( C \) in \( G_1 \times Y G_2 \) via the natural embedding \( C \hookrightarrow G_1 \) and via the composite of the inverse isomorphism \( C \cong C \) with the natural embedding \( C \hookrightarrow G_2 \). Let \( G_{1,2} := (G_1 \times Y G_2)/C \) be a reductive group scheme over \( Y \). We have a natural product homomorphism \( q : G_{1,2} \to G \) whose pullback to Spec \( K \) can be identified with the closed embedding homomorphism \( G_{0,K} \to G_K \). Therefore, \( G_{0,K} \) is a reductive group scheme over Spec \( K \). Moreover, if condition (i) holds, then as \( C_K \) is étale we have natural identities
\[ \text{Lie}(G_{1,K}) \oplus \text{Lie}(G_{2,K}) = \text{Lie}(G_{1,2,K}) = \text{Lie}(G_{0,K}). \]
Thus (a) holds. If \( q \) is a closed embedding, then \( q \) induces an isomorphism \( G_{1,2} \cong G_0 \), and therefore \( G_0 \) is a reductive, closed subgroup scheme of \( G \). Thus to finish the proof of (b), we only have to show that the homomorphism \( q \) is a closed embedding.

To check that \( q \) is a closed embedding, it suffices to check that the fibers of \( q \) are closed embeddings (see Theorem 2.5). For this we can assume that \( A \) is a complete discrete valuation ring which has an algebraically closed residue field \( k \); this implies that \( G_0 \) is a flat, closed subgroup scheme of \( G \). Let \( n := \text{Lie}(\text{Ker}(q_k)) \). From Proposition 2.5.2(a) and Lemma 2.4 we get that either (iii) \( n = 0 \) or (iv) \( \text{char}(k) = 2 \) and there exists a normal subgroup \( F_k \) of \( G_{1,2,k} \) which is isomorphic to \( \text{SO}_{2n+1,k} \) for some \( n \in \mathbb{N}^* \) and for which we have \( \text{Lie}(F_k) \cap n \neq 0 \). We show that the assumption that condition (iv) holds leads to a contradiction. Let \( F \) be a normal, closed subgroup scheme of \( G_{1,2} \) that lifts \( F_k \) and that is isomorphic to \( \text{SO}_{2n+1,A} \) (see the last paragraph of the proof of Proposition 2.5.2(c)). Let \( j_0 \in \{1, 2\} \) be such that \( F \lhd G_{j_0} \lhd G_{1,2} \) (if condition (ii) holds, then \( j_0 = 2 \)). As \( G_{j_0} \) is a closed subgroup scheme of \( G \), we have \( \text{Lie}(G_{j_0,k}) \cap n = 0 \) and therefore also \( \text{Lie}(F_k) \cap n = 0 \). Contradiction. Thus condition (iv) does not hold, and therefore condition (iii) holds. Consequently, \( \text{Ker}(q_k) \) has a trivial Lie algebra, and so it is a finite, étale, normal subgroup of \( G_{1,2,k} \). Thus \( \text{Ker}(q_k) \) is a subgroup of \( Z(G_{1,2,k}) \) and therefore also of each maximal torus of \( G_{1,2,k} \). From this and Proposition 2.5.2(a) we get that \( \text{Ker}(q_k) \) is trivial. Therefore, \( q_k \) is a closed embedding; thus \( q \) is a closed embedding. \( \square \)
Acknowledgements

We would like to thank the University of Utah, University of Arizona, Binghamton University, and the Institut des Hautes Études Scientifiques for good working conditions. We thank J.-P. Serre and A. Langer for pointing out the reference [Colliot-Thélène and Sansuc 1979] and O. Gabber for an email that independently outlined parts of the same paper; these led to a shorter paper and to stronger forms of Theorems 1.2 and 1.4 and of Proposition 5.1, for those reductive group schemes whose adjoints have geometric fibers that have simple factors of exceptional Lie types. We would like to thank J.-L. Colliot-Thélène for some comments and J. E. Humphreys for pointing out the references quoted in Remark 3.6(a). We would like to thank O. Gabber also for many other valuable comments and suggestions; Section 3 is mainly based on these suggestions. This research was partially supported by NSF grant DMS #0900967.

References


114 Adrian Vasiu


Extension theorems for reductive group schemes


Communicated by Brian Conrad
Received 2015-01-06 Revised 2015-12-11 Accepted 2015-12-15

adrian@math.binghamton.edu *Department of Mathematical Sciences, Binghamton University, P.O. Box 6000, Binghamton, NY 13902-6000, United States*
Actions of some pointed Hopf algebras on path algebras of quivers

Ryan Kinser and Chelsea Walton

We classify Hopf actions of Taft algebras $T(n)$ on path algebras of quivers, in the setting where the quiver is loopless, finite, and Schurian. As a corollary, we see that every quiver admitting a faithful $\mathbb{Z}_n$-action (by directed graph automorphisms) also admits inner faithful actions of a Taft algebra. Several examples for actions of the Sweedler algebra $T(2)$ and for actions of $T(3)$ are presented in detail. We then extend the results on Taft algebra actions on path algebras to actions of the Frobenius–Lusztig kernel $u_q(sl_2)$, and to actions of the Drinfeld double of $T(n)$.

1. Introduction

Let $n$ be an integer $\geq 2$ and let $k$ be a field containing a primitive $n$-th root of unity $\zeta$. Both $k$ and $n$ will be fixed but arbitrary subject to this condition throughout the paper. Note that if char$(k) = p > 0$, this implies that $p$ and $n$ are coprime. All algebras in this work are associative $k$-algebras and let an unadorned $\otimes$ denote $\otimes_k$.

Generalizing the classical notion of a group acting on an algebra by automorphisms, one can consider actions of Hopf algebras (e.g., quantum groups). However, one obstacle is that the intricate structure of a Hopf algebra often prevents nontrivial actions on an algebra. When such actions exist, they can be difficult to construct and

MSC2010: primary 16T05; secondary 05C20, 16S99.
Keywords: Hopf action, module algebra, path algebra, Schurian quiver, Taft algebra.
are not generally well understood. This paper presents a case where a classification of these actions is achieved. Here, we consider actions of some finite dimensional, pointed Hopf algebras, namely actions of Taft algebras (Definition 2.1) as a start. Taft algebras can be thought of as Borel subalgebras of the Frobenius–Lusztig kernel $u_q(sl_2)$. The algebra being acted upon is the path algebra of a quiver, and actions are subject to Hypothesis 1.2. All necessary background, including definitions, is recalled in Section 2. In particular, we address the following question:

**Question 1.1.** When does the path algebra of a quiver admit a nontrivial action of a (finite dimensional, pointed) Hopf algebra? Specifically, of a Taft algebra?

Actions by Taft algebras are referred to as Taft actions for short. We give a complete answer to the question above for Taft actions, and extend Taft actions to actions of the quantum group $u_q(sl_2)$ and actions of the Drinfeld double of a Taft algebra, under the following conditions.

**Hypothesis 1.2.** Unless stated otherwise, we impose the assumptions below:

(a) The quiver $Q$ is finite, loopless, and Schurian.

(b) Hopf actions preserve the ascending filtration by path length of the path algebra $kQ$.

It is easy to see that $Q$ must at least admit a nontrivial action of the cyclic group $Z_n$ (namely, the group of grouplike elements of $T(n)$) to admit a nontrivial action of the $n$-th Taft algebra $T(n)$; see Example 3.13. Since Hopf algebras and quantum groups are generalizations of group algebras, we are interested in when a path algebra of a quiver that admits classical cyclic symmetry admits additional “quantum symmetry”, loosely speaking. Our strategy is to identify a class of quivers which is small enough so that we can explicitly describe all Taft actions on their path algebras, but large enough so that every quiver admitting a Taft action is a union of quivers in this class. We call these quivers the $Z_n$-minimal quivers (Definition 4.3). The reader may wish to look over Section 5 early on for a complete account of the case $n = 2$: actions of the Sweedler algebra $T(2)$ on $Z_2$-minimal quivers.

To begin, we first note that any action on a path algebra must restrict to an action on the subalgebra generated by the vertices, by Hypothesis 1.2(b). So we start by classifying Taft actions on products of fields in Proposition 3.5. Then, the form of actions on vertices places significant restrictions on actions on the arrows. The following theorem summarizes our results, with reference to more detailed statements in the body of the paper. Here, we let $g$ and $x$ be the standard generators of $T(n)$, where $g$ is grouplike, $x$ is $(1, g)$ skew-primitive, and $xg = \xi gx$ for $\xi$ some primitive $n$-th root of unity (see Section 2A). We identify the cyclic group generated by $g$ with $Z_n$. 
Theorem 1.3. Let $Q$ be a quiver, and suppose we have a Taft action on its path algebra $\mathbb{k}Q$.

(a) The Taft action determines an action of $\mathbb{Z}_n$ on $Q$ by quiver automorphisms (Lemma 3.2).

(b) Each $\mathbb{Z}_n$-orbit of vertices is stable under $T(n)$. If we let $\{e_1, \ldots, e_m\}$ be the collection of trivial paths corresponding to some orbit of vertices, numbered so that $g \cdot e_i = e_{i+1}$ with subscripts taken modulo $m$, then the action of $x$ on these is given by

$$x \cdot e_i = \gamma \xi^i e_i - \gamma \xi^{i+1} e_{i+1},$$

for any scalar $\gamma \in \mathbb{k}$ (Proposition 3.5).

(c) For each arrow $a$ of $Q$, the action of $x$ on $a$ is given by

$$x \cdot a = \alpha a + \beta (g \cdot a) + \lambda \sigma(a),$$

for some scalars $\alpha, \beta, \lambda$. Here, $\sigma(a)$ is an arrow or trivial path with the same source as $a$ and the same target as $g \cdot a$ (Notation 3.9, Proposition 3.10). Furthermore, when $Q$ is a $\mathbb{Z}_n$-minimal quiver, these scalars are determined explicitly by the formulae (6.2) and (6.4) (Theorems 6.1, 6.3).

With an explicit parametrization of Taft actions on path algebras of $\mathbb{Z}_n$-minimal quivers, it remains to show that this is sufficient to parametrize Taft actions on path algebras of quivers subject to Hypothesis 1.2. To do this, we introduce the notion of a $\mathbb{Z}_n$-component of a quiver with $\mathbb{Z}_n$-action (Definition 7.1). These are the smallest subquivers of $Q$ which have at least one arrow and are guaranteed to be stable under the action of $T(n)$ for any choice of parameters. Moreover, see Definition 7.4 for the notion of a compatible collection of Taft actions.

Theorem 1.4 (Lemmas 7.2, 7.3, Theorem 7.5, Corollary 7.6). Fix an action of $\mathbb{Z}_n$ on a quiver $Q$. Then, $Q$ decomposes uniquely into a union of its $\mathbb{Z}_n$ components, and any Taft action on $\mathbb{k}Q$ restricts to an action on each component. Moreover, this decomposition gives a bijection between Taft actions on $\mathbb{k}Q$ and compatible collections of Taft actions on the $\mathbb{Z}_n$-components of $Q$. In particular, any path algebra of a quiver with a faithful action of $\mathbb{Z}_n$ admits an inner faithful action of the $n$-th Taft algebra $T(n)$.

As mentioned above, we extend these results to get actions of other finite dimensional, pointed Hopf actions on path algebras of quivers.

Theorem 1.5 (Theorems 8.10, 8.21, Section 8C). Fix an action of $\mathbb{Z}_n$ on a quiver $Q$. Let $q \in \mathbb{k}$ be a $2n$-th root of unity. Additional restraints on parameters are determined so that the Taft actions on the path algebra of $Q$ produced in Theorems 1.3 and 1.4 extend to an action of the Frobenius–Lusztig kernel $u_q(\mathfrak{sl}_2)$ and to an action of the Drinfeld double of $T(n)$. 
As a consequence of the theorem above, we obtain that path algebras of quivers, that admit $\mathbb{Z}_n$-symmetry, are algebras in the category of Yetter–Drinfeld modules over $T(n)$ by Majid [1991]; see also [Radford 2012, Exercise 13.1.6]. Hence, motivated by the process of bosonization, or Radford’s biproduct construction to produce (potentially new) Hopf algebras (see [Majid 1994; Radford 2012, Theorems 11.6.7, 11.6.9]), we pose the following question.

**Question 1.6.** Let $Q$ be a quiver that admits $\mathbb{Z}_n$-symmetry. When does the path algebra $\mathbb{k} Q$ admit the structure of a Hopf algebra in the category of Yetter–Drinfeld modules over $T(n)$?

1A. **Comparisons to other work.** A path algebra $\mathbb{k} Q$ is naturally a coalgebra, where the comultiplication of a path is the sum of all splits of the path. There are previous studies on extending the coalgebra structure on $\mathbb{k} Q$ to a graded Hopf algebra, most notably Cibils and Rosso’s work [1997; 2002] on Hopf quivers. Here, when $\mathbb{k} Q$ admits the structure of a Hopf algebra, the group of grouplike elements of $\mathbb{k} Q$ consists of the vertex set $Q_0$ of $Q$. Moreover, any arrow $a \in Q_1$ is a skew-primitive element as $\Delta(a) = s(a) \otimes a + a \otimes t(a)$. One example of their theory is a construction of $T(n)$ from a Hopf quiver, and in this case it has the regular action on the path algebra of this quiver. Our study produces many more examples of Taft actions on path algebras, as our construction allows for nontrivial actions on path algebras of any quiver that admits $\mathbb{Z}_n$-symmetry.

Our work also has some intersection with [Gordienko 2015]. On the one hand, Gordienko works in the setting of Taft actions on arbitrary finite dimensional algebras, whereas path algebras of quivers are not always finite dimensional. For example, Gordienko’s Theorem 1 classifies Taft algebra actions on products of matrix algebras, while our Proposition 3.5 only classifies Taft algebra actions on products of fields (equivalently, path algebras of arrowless quivers). On the other hand, Gordienko’s classification [2015, Theorem 3] is restricted to actions giving $T(n)$-simple module-algebras, whereas we have classified all Taft actions on path algebras (subject to Hypothesis 1.2). With the exception of special parameter values, the path algebras in this work are not simple with respect to the Taft algebra action: one can easily see from our explicit formulas that the Jacobson radical (the ideal generated by the arrows of $Q$) is typically a nontrivial two-sided $T(n)$-invariant ideal.

There is an abundance of literature on both the study of quantum symmetry of graphs and group actions on directed graphs from the viewpoint of operator algebras, including [Banica 2005; Banica et al. 2007; Bates et al. 2012; Bichon 2003; Kumjian and Pask 1999]. Connections to our results merit further investigation.

Other works investigating relations between path algebras of quivers and Hopf algebras can be found in [Chen et al. 2004; Huang and Liu 2010; Huang et al. 2010; van Oystaeyen and Zhang 2004; Zhang 2006].
Montgomery and Schneider [2001] provide similar results for actions of Taft algebras, and extended actions of $u_q(sl_2)$ and of $D(T(n))$, on the commutative algebras: $\mathbb{k}(u)$, and $\mathbb{k}[u]/(u^n - \beta)$ with $\beta \in \mathbb{k}$.

2. Background

We begin by defining Taft algebras and Hopf algebra actions. We then discuss path algebras of quivers, which will be acted on by Taft algebras throughout this work.

2A. Taft algebras and Hopf algebra actions. Let $H$ be a Hopf algebra with coproduct $\Delta$, counit $\varepsilon$, and antipode $S$. A nonzero element $g \in H$ is grouplike if $\Delta(g) = g \otimes g$, and the set of grouplike elements of $H$ is denoted by $G(H)$. This forces $\varepsilon(g) = 1$ and $S(g) = g^{-1}$. An element $x \in H$ is $(g, g')$-skew-primitive, for grouplike elements $g, g'$ of $H$, when $\Delta(x) = g \otimes x + x \otimes g'$. In this case, $\varepsilon(x) = 0$ and $S(x) = -g^{-1}xg'^{-1}$. The following examples of Hopf algebras will be used throughout this work.

Definition 2.1 (Taft algebra $T(n)$, Sweedler algebra $T(2)$). The Taft algebra $T(n)$ is a $n^2$-dimensional Hopf algebra generated by a grouplike element $g$ and a $(1, g)$-skew-primitive element $x$, subject to relations:

$$g^n = 1, \quad x^n = 0, \quad xg = \zeta gx$$

for $\zeta$ a primitive $n$-th root of unity. The 4-dimensional Taft algebra $T(2)$ is known as the Sweedler algebra.

Note that $G(T(n))$ is isomorphic to the cyclic group $\mathbb{Z}_n$, generated by $g$.

We now recall basic facts about Hopf algebra actions; refer to [Montgomery 1993] for further details. A left $H$-module $M$ has left $H$-action structure map denoted by $\cdot : H \otimes M \to M$. We use Sweedler notation $\Delta(h) = \sum h_1 \otimes h_2$ for coproducts.

Definition 2.2 ($H$-action). Given a Hopf algebra $H$ and an algebra $A$, we say that $H$ acts on $A$ (from the left) if, for all $h \in H$ and $p, q \in A$,

(a) $A$ is a left $H$-module;

(b) $h \cdot (pq) = \sum (h_1 \cdot p)(h_2 \cdot q)$; and

(c) $h \cdot 1_A = \varepsilon(h)1_A$.

In this case, we say that $A$ is a left $H$-module algebra. Equivalently, the multiplication map $\mu_A : A \otimes A \to A$ and unit map $\eta_A : \mathbb{k} \to A$ are morphisms of $H$-modules, so $A$ is an algebra in the monoidal category of left $H$-modules.

For the Taft actions in this work, consider the following terminology.
Definition 2.3 (extending a $G$-action). Given an action of a group $G$ on an algebra $A$, we say that an action of a Hopf algebra $H$ on $A$ extends the $G$-action on $A$ if the restriction of the $H$-action to $G(H)$ agrees with the $G$-action via some isomorphism $G(H) \simeq G$.

In this paper, we are interested in the case where $G = \mathbb{Z}_n$ and $H = T(n)$ in the above definition. Moreover, it is useful to restrict to $H$-actions that do not factor through proper quotient Hopf algebras.

Definition 2.4 (inner faithful). A module $M$ over a Hopf algebra $H$ is inner faithful if the action of $H$ on $M$ does not factor through a quotient Hopf algebra of $H$; that is, $IM \neq 0$ for any nonzero Hopf ideal $I \subset H$. A Hopf action of $H$ on an algebra $A$ is inner faithful if $A$ is inner faithful as an $H$-module.

The following lemma is likely known to experts, but does not seem to be readily accessible in the literature, so we provide a proof.

Lemma 2.5. Every nonzero bi-ideal of $T(n)$ contains $x$. Therefore, a Taft action on an algebra $A$ is inner faithful if and only if $x \cdot A \neq 0$.

Proof. Writing $H := T(n)$, since $H \cong H^*$ as Hopf algebras it suffices to prove the dual statement. Namely, since $x$ generates the radical of $H$, the dual approach is to show that every proper sub-bialgebra of $H$ is contained in the coradical of $H$.

Suppose that $A \subseteq H$ is a nonzero Hopf sub-bialgebra of $H$ which is not contained in the coradical of $H$. We will show that $A = H$. Since the coradical $H_0$ of $H$ is the span of the grouplike elements $\{g^i | i = 0, \ldots, n - 1\}$, we have that $A$ contains a nonzero element $f = hx^j + \text{(terms of lower } x\text{-degree)}$, where $j \geq 1$ and $h \in H_0$. Say, $h = \sum_{d=0}^{j} v_d g^d$, for $v_d \in k$ with $v_1 \neq 0$. Since $A$ inherits the coproduct from $H$,

$$\Delta(v_i g^i x^j) = \sum_{\ell=0}^{j} \left[ \frac{j!}{\ell!} \xi \right] v_i g^i x^{j-\ell} \otimes v_i g^{j-\ell+i} x^{\ell},$$

which is in $A \otimes A$. Here, the equality above holds by [Radford 2012, Lemma 7.3.1]. By the maximality of $j$ and $i$ and by taking $\ell = 0$ above, we have that $f \in A$ implies $g^i x^j \in A$. Now applying $\Delta$ to $g^i x^j$ yields $g^i x^{j-\ell} \in A$ for $\ell = 0, \ldots, j$. So, we get $g^i x \in A$. Likewise, apply $\Delta$ to $g^i x$ to conclude that $g^i, g^{i+1} \in A$. Since $g$ has finite multiplicative order, $g^{-i} \in A$ as well. Thus, both $g$ and $x$ are in $A$, so $A = H$, as desired. \qed

So the extension of a faithful cyclic group ($\mathbb{Z}_n$) action on an algebra $A$ to a Taft algebra ($T(n)$) action is inner faithful if and only if $x \cdot A \neq 0$. To study module algebras of $T(n)$, the following standard fact will also be of use.
Lemma 2.6. For each \( T(n) \)-action on an algebra \( A \), there is a natural action of \( T(n) \) on the opposite algebra \( (A^{\text{op}}, \ast) \), say denoted by \( \diamond \), as follows:

\[
g \diamond p = g^{-1} \cdot p \quad \text{and} \quad x \diamond p = g^{-1}x \cdot p. \tag{2.7}
\]

This gives a bijection between \( T(n) \)-actions on \( A \) and on \( A^{\text{op}} \).

Proof. For any Hopf algebra \( H \) and algebra \( A \), we get that \( A \) is an \( H \)-module algebra if and only if \( (A^{\text{op}}, \ast) \) is an \( H^{\text{cop}} \)-module algebra. Here, \( H^{\text{cop}} \) is the co-opposite algebra of \( H \) and \( \Delta^{\text{cop}} = \tau \circ \Delta \) with \( \tau \left( \sum h_1 \otimes h_2 \right) = \sum h_2 \otimes h_1 \). Indeed, \( h \cdot (pq) = h \cdot (q \ast p) = \sum (h_2 \cdot q) \ast (h_1 \cdot p) = \sum (h_1 \cdot p)(h_2 \cdot q) \). The map sending \( g \) to \( g^{-1} \) and \( x \) to \( g^{-1}x \) gives an isomorphism \( T(n) \cong T(n)^{\text{cop}} \) as Hopf algebras, and the result follows. \( \square \)

2B. Path algebras of quivers. A quiver is another name for a directed graph, in the context where the directed graph is used to define an algebra. Here, we review the basic notions and establish notation. More detailed treatments can be found in the texts [Assem et al. 2006; Schiffler 2014]. Formally, a quiver \( Q = (Q_0, Q_1, s, t) \) consists of a set of vertices \( Q_0 \), a set of arrows \( Q_1 \), and two functions \( s, t : Q_1 \to Q_0 \) giving the source and target of each arrow, visualized as

\[
\begin{array}{ccc}
s(a) & \longrightarrow & t(a).
\end{array}
\]

A path \( p \) in \( Q \) is a sequence of arrows \( p = a_1a_2 \cdots a_\ell \) for which \( t(a_i) = s(a_{i+1}) \) for \( 1 \leq i \leq \ell - 1 \). (Note that we read paths in left-to-right order.) The length of \( p \) is the number of arrows \( \ell \). There is also a length 0 trivial path \( e_i \) at each vertex \( i \in Q_0 \), with \( s(e_i) = t(e_i) = i \).

A quiver \( Q \) has a path algebra \( \mathbb{k}Q \) whose basis consists of all paths in \( Q \), and multiplication of basis elements is given by composition of paths whenever it is defined, and 0 otherwise. More explicitly, we have the following definition.

Definition 2.8 (path algebra). The path algebra \( \mathbb{k}Q \) of a quiver \( Q \) is the \( \mathbb{k} \)-algebra presented by generators from the set \( Q_0 \cup Q_1 \) with the relations

\begin{itemize}
  \item[(a)] \( \sum_{i \in Q_0} e_i = 1 \);
  \item[(b)] \( e_i e_j = \delta_{ij} e_i \) for all \( e_i, e_j \in Q_0 \); and
  \item[(c)] \( a = e_{s(a)}a = ae_{t(a)} \) for all \( a \in Q_1 \).
\end{itemize}

Condition (a) is due to the assumption that \( |Q_0| < \infty \). Further, \( e_i \) is a primitive orthogonal idempotent in \( \mathbb{k}Q \) for all \( i \). So, \( \mathbb{k}Q \) is an associative algebra with unit, which is finite dimensional if and only if \( Q \) has no path of positive length which starts and ends at the same vertex. Notice that \( \mathbb{k}Q \) has a natural ascending filtration by path length. Namely, if we let \( F_i \) be the subspace of \( \mathbb{k}Q \) spanned by paths of length at most \( i \), then \( F_i \cdot F_j \subseteq F_{i+j} \).
Definition 2.9 (Schurian). A quiver \( Q \) is said to be \textit{Schurian} if, given any two vertices \( i \) and \( j \), there is at most one arrow which starts at \( i \) and ends at \( j \).

Note that the definition above does not exclude oriented 2-cycles.

Definition 2.10 (covering, gluing, \( \otimes \)). A quiver \( Q \) is \textit{covered} by a collection of subquivers \( Q^1, \ldots, Q^r \) if \( Q = \bigcup_i Q^i \). We say \( Q \) is obtained by \textit{gluing} the collection \( Q^1, \ldots, Q^r \) if the collection covers \( Q \), and in addition \( Q^i \cap Q^j \) consists entirely of vertices when \( i \neq j \); in this case, write

\[
Q = Q^1 \otimes \cdots \otimes Q^r.
\]

Remark 2.11. If a quiver \( Q \) is obtained by gluing subquivers \( Q^1, \ldots, Q^r \), then we get that \( kQ \) is the factor of the free product of path algebras \( kQ^1 \ast \cdots \ast kQ^r \) by the ideal generated by \( \{ e_{i,v} - e_{j,v} \} \), where for each pair \( (i, j) \), the index \( v \) varies over the vertices of \( Q^i \cap Q^j \). Here, \( e_{\ell,v} \) indicates the trivial path at \( v \), for \( v \in (Q^\ell)_0 \).

2C. \textbf{Group actions on path algebras of quivers.} Now we consider group actions on quivers and on their path algebras.

Definition 2.12 (quiver automorphism). Let \( Q, Q' \) be quivers, and consider two maps of sets \( f_0: Q_0 \to Q'_0 \) and \( f_1: Q_1 \to Q'_1 \).

(1) We say that the pair \( f = (f_0, f_1) \) is a \textit{quiver isomorphism} if each of these maps is bijective, and they form a commuting square with the source and target operations. That is, for all \( a \in Q_1 \) we have

\[
(f_0 \circ s)(a) = (s \circ f_1)(a) \quad \text{and} \quad (f_0 \circ t)(a) = (t \circ f_1)(a).
\]

(2) We say that the pair \( f = (f_0, f_1) \) is a \textit{quiver automorphism of \( Q \)} if \( f \) satisfies (1) and \( Q' = Q \).

(3) A group \( G \) is said to \textit{act on \( Q \)} or \( Q \) is \textit{\( G \)-stable}, if \( G \) acts on the sets \( Q_0 \) and \( Q_1 \) such that each element of \( G \) acts by a quiver automorphism.

Remark 2.13. A quiver automorphism induces an automorphism of its path algebra, but not every automorphism of a path algebra is of this form. See Lemma 3.2.

Remark 2.14. Given a quiver \( Q \), we can form the \textit{opposite quiver} \( Q^{\text{op}} \) by interchanging the source and target functions \( s \) and \( t \). It is clear from the definition of the path algebra that \( k(Q^{\text{op}}) \cong \langle kQ \rangle^{\text{op}} \). Hence, Lemma 2.6 implies there is a bijection between \( T(n) \)-actions on \( kQ \) and on \( kQ^{\text{op}} \) given by (2.7). See Remark 5.2 for an illustration.
In this section, we present preliminary results on Taft actions on path algebras of loopless, Schurian quivers. We begin by studying actions on vertices, first giving a simple lemma regarding group actions on \( kQ_0 \) (Lemma 3.1). Then, we extend this result to classifying Taft actions on \( kQ_0 \) (Proposition 3.5). Preliminary results on Taft actions on path algebras \( kQ \) are provided (Proposition 3.10), along with an example in the case where \( Q_0 \) is fixed by the group of grouplike elements of \( T(n) \) (Example 3.13).

The following lemma is elementary, but we provide the details in any case.

**Lemma 3.1 (G-action on \( kQ_0 \)).** Let \( G \) be a group, let \( Q_0 \) be a set of vertices, and let \( \{e_i\}_{i \in Q_0} \) be the corresponding primitive orthogonal idempotents in \( kQ_0 \). Then, any \( G \)-action on the set \( Q_0 \) induces an action on the ring \( kQ_0 \), given by \( g \cdot e_i = e_{g \cdot i} \) for each \( i \in Q_0 \) and \( g \in G \). Moreover, every \( G \)-action on \( kQ_0 \) arises in this way.

**Proof.** The first statement is clear, so suppose for the converse that we have a \( G \)-action on \( kQ_0 \). To act as a ring automorphism, each element of \( G \) must send a complete collection of primitive orthogonal idempotents to another such collection. But in this case, the set \( \{e_i\}_{i \in Q_0} \) is the unique such collection. So this set must be permuted by \( G \), defining an action of \( G \) on the set \( Q_0 \).

Now we turn our attention to group actions on arbitrary path algebras of quivers.

**Lemma 3.2 (G-action on \( kQ \)).** Let \( G \) be a group and \( Q \) a quiver which is loopless and Schurian, and suppose that \( G \) acts by automorphisms of \( kQ \), preserving the ascending filtration by path length. Then, the action of each \( g \in G \) on \( Q \) is given by

(i) a quiver automorphism \( \phi : Q \to Q \), along with

(ii) a collection of nonzero scalars \( \mu_a \in k^\times \), indexed by the arrows \( a \) of \( Q \),

such that \( g \cdot a = \mu_a \phi(a) \) for each \( a \in Q_1 \). (To be clear, both \( \phi \) and the collection \( \mu_a \) depend on \( g \).)

**Proof.** Since \( G \) preserves the path length filtration, it acts by permutations on the vertex set, by Lemma 3.1. Then for each \( g \in G \) and \( a \in Q_1 \), we have \( g \cdot a = g \cdot (e_{s(a)}a) = (g \cdot e_{s(a)})(g \cdot a) \), showing that \( g \cdot a \) lies in the span of arrows starting at \( g \cdot e_{s(a)} \). Similarly, we see that \( g \cdot a \) lies in the \( k \)-span of arrows with target \( g \cdot e_{t(a)} \). Since \( Q \) is Schurian, this determines a unique arrow \( \phi(a) \), with start \( s(g \cdot a) = g \cdot e_{s(a)} \) and target \( t(g \cdot a) = g \cdot e_{t(a)} \), such that \( g \cdot a \) is a scalar times \( \phi(a) \). It is immediate from the definition of \( \phi \) that \( \phi \) is a quiver automorphism.

**Convention 3.3.** We sometimes use \( g \cdot a \) to label an arrow in a diagram, or refer to \( g \cdot a \) as an arrow in exposition, in order to avoid introducing the extra notation \( \phi \). In these cases, it is understood that one must actually multiply by a scalar to get an arrow on the nose.
Next, we study the action of skew-primitive elements on arrowless quivers \( Q_0 \), which then leads to Taft actions on the semisimple algebra \( \mathbb{k} Q_0 \). Since the generator \( x \) of \( T(n) \) is \((1, g)\)-skew-primitive, the relation \( e_i^2 = e_i \) of \( \mathbb{k} Q_0 \) gives us that
\[
x \cdot e_i = x \cdot (e_i^2) = e_i(x \cdot e_i) + (x \cdot e_i)(g \cdot e_i) \in \text{span}_\mathbb{k}\{e_i, g \cdot e_i\}.
\] (3.4)
So, to study extensions of a \( G \)-action on \( \mathbb{k} Q_0 \) to a \( T(n) \)-action, we can restrict ourselves to a single \( G \)-orbit of vertices. From here on, we apply the above results to the case where \( G = G(T(n)) \) is the cyclic group generated by \( g \in T(n) \), which we identify with \( \mathbb{Z}_n \).

**Proposition 3.5** \((T(n)\)-actions on \( \mathbb{k} Q_0 \)). Let \( Q_0 = \{1, 2, \ldots, m\} \) be the vertex set of a quiver, where \( m \) divides \( n \), and \( \mathbb{Z}_n \) acts on \( \mathbb{k} Q_0 \) by \( g \cdot e_i = e_{i+1} \). Here, subscripts are always interpreted modulo \( m \).

(i) If \( m < n \) (so the \( \mathbb{Z}_n \)-action on \( Q_0 \) is not faithful), then \( x \) acts on \( \mathbb{k} Q_0 \) by 0.

(ii) If \( m = n \) (so \( \mathbb{Z}_n \) acts faithfully on \( Q_0 \)), then the action of \( x \) on \( \mathbb{k} Q_0 \) is exactly of the form
\[
x \cdot e_i = \gamma \xi^i (e_i - \zeta e_{i+1}) \quad \text{for all } i,
\] (3.6)
where \( \gamma \in \mathbb{k} \) can be any scalar.

In particular, we can extend the action of \( \mathbb{Z}_n \) on \( \mathbb{k} Q_0 \) to an inner faithful action of \( T(n) \) on \( \mathbb{k} Q_0 \) if and only if \( m = n \).

**Proof.** Assume that we have a \( T(n) \)-action on \( \mathbb{k} Q_0 \) extended from the \( \mathbb{Z}_n \)-action on \( Q_0 \) in Lemma 3.1. By (3.4), we know that \( x \cdot e_i = \alpha_i e_i + \beta_i e_{i+1} \) for some scalars \( \alpha_i, \beta_i \in \mathbb{k} \). Then, we have
\[
0 = x \cdot 1 = x \cdot \sum_{i=1}^m e_i = \sum_{i=1}^m \alpha_i e_i + \beta_i e_{i+1} = \sum_{i=1}^m (\alpha_i + \beta_{i-1}) e_i,
\] (3.7)
which gives \( \beta_{i-1} = -\alpha_i \). (Here, \( \sum_{i=1}^m \beta_i e_{i+1} = \sum_{i=1}^m \beta_{i-1} e_i \) by reindexing.) Now the relation \( xg = \xi gx \) applied to \( e_i \) gives
\[
\alpha_i e_{i+1} = \alpha_i e_{i+1} - \beta_i e_{i+2} = \zeta (\alpha_i e_{i+1} - \alpha_{i+2} e_{i+2}),
\] (3.8)
so that \( \alpha_i \) is \( \xi \alpha_i \) for all \( i \). Setting \( \gamma := \alpha_1 \xi^{-1} \) gives \( \alpha_i = \xi^i \gamma \), so that (3.6) holds whenever a \( T(n) \)-action exists. We have assumed that \( m \) divides \( n \), but on the other hand, \( x \cdot e_i = x \cdot e_{i+m} \) implies that \( \gamma \xi^i = \gamma \xi^{i+m} \). Thus, \( \gamma = \gamma \xi^m \). Hence, whenever \( m < n = \text{ord}(\xi) \), we have \( \gamma = 0 \), and \( x \) acts by 0. This establishes (i).

On the other hand, suppose that \( m = n \). We will show that Equation (3.6) defines a \( T(n) \)-action on \( \mathbb{k} Q_0 \) for any \( \gamma \in \mathbb{k} \). A simple substitution verifies that \( xg \cdot e_i = \xi gx \cdot e_i \). The fact that the \( x \)-action preserves the relations \( e_i e_j = \delta_{ij} e_i \) and \( \sum_{i=1}^n e_i = 1 \) is also easy to check by substitution. The only tedious part is to
show that $x^n$ acts on $kQ_0$ by 0, which is verified by Lemma 9.10 in the appendix of computations, using symmetric functions. Now, the $T(n)$-action on $kQ_0$ is inner faithful when $\gamma$ is nonzero, by Lemma 2.5. Therefore, (ii) holds.

Now to study Taft actions on path algebras of Schurian quivers in general, we set the following notation.

**Notation 3.9** ($\sigma(a)$). Suppose we have a quiver $Q$ and an action of $\mathbb{Z}_n \cong \langle g \rangle \subset T(n)$ on $kQ$. Given an arrow $a \in Q_1$, we know that there exists at most one path of length less than or equal to 1 from $s(a)$ to $t(g \cdot a)$ since $Q$ is Schurian and loopless. Denote this path by $\sigma(a)$ if it exists, and set $\sigma(a) = 0$ otherwise.

To be more explicit, consider the following case: when $g$ fixes neither $s(a)$ nor $t(a)$, and $g \cdot t(a) \neq s(a)$, then $\sigma(a)$ is either an arrow or 0, and can be visualized in the following diagram.

If $g \cdot t(a) = s(a)$, then $\sigma(a)$ is the trivial path at $s(a)$. Moreover, we have $\sigma(a) = g \cdot a$ whenever $g \cdot s(a) = s(a)$, and $\sigma(a) = a$ whenever $g \cdot t(a) = t(a)$.

We remind the reader of the standing assumptions made in Hypothesis 1.2. The following result determines the action of $x$ on any arrow of $Q$.

**Proposition 3.10.** Suppose we have an action of $T(n)$ on $kQ$, and let $a \in Q_1$ with $i_+ := s(a)$ and $j_- := t(a)$. Then, there exist scalars $\alpha, \beta, \lambda \in k$ such that

$$x \cdot a = \alpha a + \beta (g \cdot a) + \lambda \sigma(a).$$

(3.11)

Moreover, $\alpha, \beta, \lambda$ can be determined in special cases depending on the relative configuration of $a$ and $g \cdot a$, as described in Figures 1 and 2 below. Here, the dotted red arrows indicate the action of $g$ on $Q_0$.

**Proof.** Let $\gamma_+, \gamma_- \in k$ be the scalars from Proposition 3.5 such that

$$x \cdot e_{i_+} = (\gamma_+) \xi^i (e_{i_+} - \xi g \cdot e_{i_+}) \quad \text{and} \quad x \cdot e_{j_-} = (\gamma_-) \xi^j (e_{j_-} - \xi g \cdot e_{j_-}),$$

(3.12)

where $g \cdot e_{i_+} = e_{(\ell+1)_+}$ as usual.

From the relation $a = e_{i_+} a$, we can compute

$$x \cdot a = x \cdot (e_{i_+} a) = e_{i_+} (x \cdot a) + (\gamma_+) \xi^i (e_{i_+} - \xi g \cdot e_{i_+})(g \cdot a)$$

$$= \begin{cases} e_{i_+} (x \cdot a) + (\gamma_+) \xi^i (1 - \xi)(g \cdot a) & \text{if } e_{i_+} (g \cdot a) = g \cdot a, \\ e_{i_+} (x \cdot a) - (\gamma_+) \xi^{i+1} (g \cdot a) & \text{if } e_{i_+} (g \cdot a) = 0. \end{cases}$$
we have $\lambda \sigma(a)$ in (3.11) by absorbing a case. Similarly, if $\lambda \sigma(a)$ explicitly, but depend on the relative configuration of $x$ maps to $g\cdot j$. Suppose that $\gamma \zeta$ gives Equation (3.11). The coefficients $\alpha \beta$, $\gamma \zeta$ can be determined more explicitly, but depend on the relative configuration of $a$ and $g\cdot a$. We consider the various cases below.

Thus, $x\cdot a \in \text{span}_k\{\text{paths starting at } e_{i_+}, g\cdot a\}$. Similarly, the relation $a = \alpha e_{j_-}$ gives

\[
x \cdot a = x \cdot (\alpha e_{j_-}) = (\gamma_\alpha) \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+}) + (x \cdot a) (g \cdot e_{j_-})
\]

Thus, $x \cdot a \in \text{span}_k\{a, \text{paths ending at } g \cdot e_{j_-}\}$. Intersecting these two conditions on $x \cdot a$ gives Equation (3.11). The coefficients $\alpha$, $\beta$, $\gamma$ can be determined more explicitly, but depend on the relative configuration of $a$ and $g\cdot a$. We consider the various cases below.

Suppose that $e_{i_+} (g\cdot a) = g\cdot a$, then $g \cdot s(a) = s(a)$. So, as remarked in Notation 3.9, we have $\sigma(a) = g\cdot a$. We can omit this term in (3.11) by absorbing $\lambda$ into $\beta$ in this case. Similarly, if $a (g \cdot e_{j_-}) = a$, then $g \cdot t(a) = t(a)$ and we have $\sigma(a) = a$. We can omit this term in (3.11) by absorbing $\lambda$ into $\alpha$ in this case. On the other hand,

\[
x \cdot e_i = 0 \quad (\gamma_\alpha)(e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
x \cdot e_j = \alpha a = (\gamma_\alpha)(e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
x \cdot a = \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+}) + \lambda \sigma(a)
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]

\[
y \gamma \zeta^j (e_{i_+} - \zeta g \cdot e_{i_+})
\]
Thus in this case, \( \beta(\text{Example 3.13}) \) implies that

\[
\alpha a + \beta(g \cdot a) + \lambda \sigma(a) = e_{i+}[\alpha a + \beta(g \cdot a) + \lambda \sigma(a)] - (\gamma_+)\xi^{i+1}(g \cdot a) = \alpha a - (\gamma_+)\xi^{i+1}(g \cdot a) + \lambda \sigma(a).
\]

Thus in this case, \( \beta = -(\gamma_+)\xi^{i+1} \). Similarly, if \( a(g \cdot e_{j-}) = 0 \), then \( x \cdot a = x \cdot a e_{j-} \) implies that

\[
\alpha a + \beta(g \cdot a) + \lambda \sigma(a) = (\gamma_-)\xi^j a + [\alpha a + \beta(g \cdot a) + \lambda \sigma(a)](g \cdot e_{j-}) = (\gamma_-)\xi^j a + \beta(g \cdot a) + \lambda \sigma(a).
\]

So, \( \alpha = (\gamma_-)\xi^j \). These results are collected in Figure 1. In each case, the \( x \)-action on the vertices follow from Proposition 3.5. In Figure 2, the results are further specialized to the cases where \( g \cdot s(a) = t(a) \) or \( g \cdot t(a) = s(a) \). Since the \( \pm \) notation serves to distinguish the orbits of \( s(a) \) and \( t(a) \), the \( \pm \) notation is dropped in Figure 2.

As an illustration of the results above, we study Taft actions on a path algebra \( \mathbb{k} Q \), where \( \mathbb{Z}_n \) fixes \( Q_0 \).

**Example 3.13** \((T(n))-\)action on \( \mathbb{k} Q \), \( \mathbb{Z}_m \) fixes \( Q_0 \). If \( \mathbb{Z}_n \) fixes the vertices of a quiver \( Q \), then we claim that any extended action of \( T(n) \) on \( \mathbb{k} Q \) is not inner faithful. First, by Proposition 3.5(i) with \( m = 1 \), we get that \( x \cdot e_i = 0 \) for all \( i \in Q_0 \). For the arrows, we get that \( s(g \cdot a) = e_{s(a)} \) and \( t(g \cdot a) = e_{t(a)} \) by the assumption. So, \( \sigma(a) = g \cdot a \). Moreover, \( g \cdot a = \mu_a a \) by Lemma 3.2. Now, Proposition 3.10 implies that \( x \cdot a = \alpha a \) for some \( \alpha \in \mathbb{k} \). Finally, using the relation \( x^n = 0 \), we conclude that \( x \cdot a = 0 \), and the claim holds.

4. Minimal quivers

Given any Hopf algebra action on an algebra \( A \), it is natural to study the restricted Hopf action on a subalgebra of \( A \), if one exists. We introduce \( \mathbb{Z}_n \)-minimal quivers in this section, which will be the building blocks of Taft actions on path algebras of quivers in subsequent sections. The following definition serves to fix notation for specific quivers that will be used throughout the rest of the paper.

**Definition 4.1** \((K_m, K_{m,m'}, a^i_j, b^i_j)\). Let \( m \) and \( m' \) be positive integers.

1. The **complete quiver** \( K_m \) (or **complete digraph**) has vertex set \{1, 2, \ldots, m\}, with an arrow \( a^i_j \) from \( i \) to \( j \) for every ordered pair of distinct vertices \((i, j)\). For uniformity in the formulas, we also take the symbol \( a^i_i \) to mean the trivial path \( e_i \) at vertex \( i \) (rather than a loop, which we have excluded).

2. The **complete bipartite quiver** \( K_{m,m'} \) has a top row of \( m \) vertices and a bottom row of \( m' \) vertices, labeled by \{1+, \ldots, m+\} and \{1-, \ldots, m'\}, respectively.
There is an arrow $b_{ij}^j$ from each vertex $i_+$ in the top row to each vertex $j_-$ in the bottom row; that is, $s(b_{ij}^j) = i_+$ and $t(b_{ij}^j) = j_-$. An example of each type is given below, without vertex or arrow labels.

$$K_4 = \quad K_{2,3} =$$

To illustrate the arrow labels, the first diagram below shows some arrows in $K_4$ and the second diagram shows all arrows starting at $1^+$ in $K_{2,3}$.

See Figure 5 for further illustrations.

Suppose we have an action of $\mathbb{Z}_n$ on the path algebra of a subquiver of $K_m$ or $K_{m,m'}$. Let $g$ be a generator of $\mathbb{Z}_n$. After possibly relabeling, we can assume $g$ acts on the trivial paths by $g \cdot e_i = e_{i+1}$ (for $K_m$) and $g \cdot e_{i_+} = e_{(i+1)_+}$, $g \cdot e_{i_-} = (i+1)_-$ (for $K_{m,m'}$), where the indices are taken modulo $m$ or $m'$ as appropriate. By Lemma 3.2, there is a collection of nonzero scalars $\mu_{i,j}$ such that $g \cdot a_{ij}^i = \mu_{i,j} a_{i+1,j+1}$ (for $K_m$), or $g \cdot b_{ij}^i = \mu_{i,j} b_{i+1,j+1}$ (for $K_{m,m'}$). Again, subscripts and superscripts are interpreted modulo $m$ or $m'$ as appropriate. Since $g^n$ is the identity and $g \cdot e_i = e_{i+1}$ (for $K_m$), these scalars $\mu_{i,j}$ satisfy

$$\mu_{i,i} = 1 \quad \text{in the } K_m \text{ case},$$

$$\prod_{\ell=0}^{n-1} \mu_{i+i_+\ell,j+i_+\ell} = 1 \quad \text{in both cases.} \tag{4.2}$$

Since an arrow (or, more specifically, the source and target of an arrow) of a quiver $Q$ can only be part of one or two $\mathbb{Z}_n$-orbits of $Q_0$, we make the following definition.

**Definition 4.3 ($\mathbb{Z}_n$-minimal, Type A, Type B).** Let $\mathbb{Z}_n$ act on a quiver $Q$. Say that $Q$ is $\mathbb{Z}_n$-minimal of

- **Type A** if $Q$ is a $\mathbb{Z}_n$-stable subquiver of $K_m$, where $m > 1$ is a positive integer dividing $n$; or
Pointed Hopf actions on path algebras of quivers

• Type B if \( Q \) is a \( \mathbb{Z}_n \)-stable subquiver of \( K_{m,m'} \), where \( m, m' \geq 1 \) and \( m, m' \) divide \( n \).

Example 4.4. The following quiver \( Q \) admits both a \( \mathbb{Z}_4 \)-action and a \( \mathbb{Z}_2 \)-action illustrated by the dotted red arrows below. Observe that \( Q \) is \( \mathbb{Z}_4 \)-minimal of Type B. Further, we see it is \( \mathbb{Z}_{4d} \)-minimal of Type B for any integer \( d \geq 1 \). However, it is not \( \mathbb{Z}_2 \)-minimal of Type B.

Now we list the \( \mathbb{Z}_n \)-minimal quivers for small \( n \). For Type A, note that there is only one \( \mathbb{Z}_2 \)-minimal quiver that admits a transitive action of \( \mathbb{Z}_2 \) on vertices; see (I) of Figure 5. See Figure 3 for the three \( \mathbb{Z}_3 \)-minimal quivers of Type A. The dotted, red arrow indicates the action of \( \mathbb{Z}_3 \) on \( Q_0 \) (clockwise rotation in each case). For Type B, there are five Type B \( \mathbb{Z}_2 \)-minimal quivers; see Figure 5 (II)–(VI) for an illustration. Moreover, see Figure 4 for the six \( \mathbb{Z}_3 \)-minimal quivers of Type B.

5. Sweedler algebra actions on path algebras of minimal quivers

In this section, we study the action of the Sweedler algebra \( T(2) \) on path algebras of quivers. This is achieved by first computing the action of the Sweedler algebra on \( \mathbb{Z}_2 \)-minimal quivers (Theorem 5.1). Later, in Section 7, we present results on gluing such actions to yield Sweedler actions on more general quivers.

Recall from Definition 2.1 that the Sweedler algebra is the 4-dimensional Taft algebra \( T(2) \) generated by a grouplike element \( g \) and a \((1, g)\)-skew-primitive element \( x \), subject to relations:

\[
g^2 = 1, \quad x^2 = 0, \quad xg + gx = 0.
\]

Note that \( G(T(2)) \simeq \mathbb{Z}_2 \).
Theorem 5.1. Recall Hypothesis 1.2. For each quiver $Q$ in Figure 5 (where the dotted red arrow illustrates the $g$-action on vertices), the $\mathbb{Z}_2$-action on $Q$ extends to an action of the Sweedler algebra $T(2)$ on $\mathbb{k}Q$ precisely as follows. We take $\gamma$
(resp. \(\gamma_+\) and \(\gamma_-\)) to be the scalar from the \(x\)-action on \(e_i\) (resp. on \(e_{i_+}\) and on \(e_{j_-}\)) in (3.6) (resp. in (3.12)).

<table>
<thead>
<tr>
<th>(Q)</th>
<th>(x)-action on (Q_0)</th>
<th>(Q)</th>
<th>(x)-action on (Q_0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I)</td>
<td>(x \cdot e_1 = -\gamma (e_1 + e_2)) (x \cdot e_2 = \gamma (e_1 + e_2))</td>
<td>(IV)</td>
<td>(x \cdot e_{1_+} = -(\gamma_+)(e_{1_+} + e_{2_+})) (x \cdot e_{2_+} = (\gamma_+)(e_{1_+} + e_{2_+})) (x \cdot e_{1_-} = 0)</td>
</tr>
<tr>
<td>(II)</td>
<td>(x \cdot e_{1_+} = x \cdot e_{1_-} = 0)</td>
<td>(V)</td>
<td>(x \cdot e_{1_+} = -(\gamma_+)(e_{1_+} + e_{2_+})) (x \cdot e_{2_+} = (\gamma_+)(e_{1_+} + e_{2_+})) (x \cdot e_{1_-} = -(\gamma_-)(e_{1_-} + e_{2_-})) (x \cdot e_{2_-} = (\gamma_-)(e_{1_-} + e_{2_-}))</td>
</tr>
<tr>
<td>(III)</td>
<td>(x \cdot e_{1_+} = 0) (x \cdot e_{1_-} = -(\gamma_-)(e_{1_-} + e_{2_-})) (x \cdot e_{2_-} = (\gamma_-)(e_{1_-} + e_{2_-}))</td>
<td>(VI)</td>
<td>(x \cdot e_{1_+} = -(\gamma_+)(e_{1_+} + e_{2_+})) (x \cdot e_{2_+} = (\gamma_+)(e_{1_+} + e_{2_+})) (x \cdot e_{1_-} = -(\gamma_-)(e_{1_-} + e_{2_-})) (x \cdot e_{2_-} = (\gamma_-)(e_{1_-} + e_{2_-}))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Q)</th>
<th>(x)-action on (Q_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I)</td>
<td>(x \cdot a_1^1 = \gamma a_1^1 - \gamma \mu a_1^2 + \lambda e_1) (x \cdot a_1^2 = \gamma \mu^{-1} a_1^1 - \gamma a_1^2 - \lambda \mu^{-1} e_2) (\text{for } \lambda \in \mathbb{k}, \mu \in \mathbb{k}^\times)</td>
</tr>
<tr>
<td>(II)</td>
<td>(x \cdot b_1^1 = 0)</td>
</tr>
<tr>
<td>(III)</td>
<td>(x \cdot b_1^1 = -(\gamma_-)b_1^1 + \beta \mu b_2^2) (x \cdot b_2^1 = -\beta \mu^{-1} b_1^1 + (\gamma_-)b_2^2) (\text{for } \beta^2 = (\gamma_-)^2, \mu \in \mathbb{k}^\times)</td>
</tr>
<tr>
<td>(IV)</td>
<td>(x \cdot b_1^1 = \alpha b_1^1 - (\gamma_+)\mu b_2^2) (x \cdot b_2^1 = (\gamma_+)\mu^{-1} b_1^1 - \alpha b_2^2) (\text{for } \alpha^2 = (\gamma_+)^2, \mu \in \mathbb{k}^\times)</td>
</tr>
<tr>
<td>(V)</td>
<td>(x \cdot b_1^1 = -(\gamma_-)b_1^1 - (\gamma_-)\mu b_2^2) (x \cdot b_2^1 = (\gamma_-)\mu^{-1} b_1^1 + (\gamma_-)b_2^2) (\text{for } (\gamma_-)^2 = (\gamma_-)^2, \mu \in \mathbb{k}^\times)</td>
</tr>
<tr>
<td>(VI)</td>
<td>(x \cdot b_1^1 = -(\gamma_-)b_1^1 - (\gamma_-)\mu b_2^2 + \lambda b_1^1) (x \cdot b_2^1 = (\gamma_-)\mu^{-1} b_1^1 + (\gamma_-)b_2^2 - \lambda \mu^{-1} \mu' b_1^2) (\text{for } (\gamma_-)^2 = (\gamma_-)^2 + \lambda \lambda') (\text{with } \lambda, \lambda' \in \mathbb{k}) (\text{and } \mu, \mu' \in \mathbb{k}^\times)</td>
</tr>
</tbody>
</table>

**Proof.** The \(x\)-action on vertices follows from (3.6) and (3.12).

(I) We are in the situation illustrated in the second column of Figure 2, where additionally \(g \cdot i = j\). So, we have

\[
x \cdot a_2^1 = \gamma a_2^1 - \gamma \mu_1 a_1^2 + \lambda e_1
\]
from Proposition 3.10 and Lemma 3.2. Then using the relation \( xg = -gx \) in \( T(2) \), we find that
\[
\mu_{1,2}x \cdot a_1^2 = x \cdot (g \cdot a_1^1) = -g \cdot (x \cdot a_1^1) = -\gamma \mu_{1,2}a_1^2 + \gamma \mu_{1,2}a_1^1 + \lambda \mu_{1,1}e_2.
\]
Using that \( \mu_{1,2}^{-1} = \mu_{2,1} \) and \( \mu_{1,1} = 1 \) from Equation (4.2), we obtain
\[
x \cdot a_1^2 = \gamma \mu_{2,1}a_1^1 - \gamma a_1^2 - \lambda \mu_{2,1}e_2.
\]

The \( x \)-action on all other relations in \( \mathbb{k}Q \) yields tautologies, and the relations \( x^2 = xg + gx = 0 \) are satisfied when applied to the arrows \( a_1^1 \) and \( a_1^2 \). Taking \( \mu = \mu_{1,2} \) and \( \mu^{-1} = \mu_{2,1} \), we see that the action is of the claimed form.

(II) We apply Example 3.13.

(III) Examining Figure 1, we see that \( x \cdot b_1^1 = -(\gamma_-)b_1^1 + \beta \mu_{1,1}b_1^2 \) for some \( \beta \in \mathbb{k} \). The relation \( xg = -gx \) in \( T(2) \) then gives
\[
\mu_{1,1}x \cdot b_1^2 = x \cdot (g \cdot b_1^1) = -g \cdot (x \cdot b_1^1) = -\beta \mu_{1,1} \mu_{2,2}^1 b_1^1 + (\gamma_-) \mu_{1,1} b_1^2.
\]
Now applying the relation \( x^2 = 0 \) of \( T(2) \) to the arrow \( b_1^1 \) implies that \( (\gamma_-)^2 = \beta^2 \), and we take \( \mu = \mu_{1,1} \).

(IV) We get from Figure 1 that \( x \cdot b_1^1 = \alpha b_1^1 - (\gamma_+) \mu_{1,1} b_1^2 \) for some \( \alpha \in \mathbb{k} \). The relation \( xg = -gx \) in \( T(2) \) then gives
\[
\mu_{1,1}x \cdot b_1^2 = x \cdot (g \cdot b_1^1) = -g \cdot (x \cdot b_1^1) = -\alpha \mu_{1,1} b_1^2 + (\gamma_+) \mu_{1,1} \mu_{2,2} b_1^1.
\]
Now applying the relation \( x^2 = 0 \) to the arrow \( b_1^1 \) implies that \( (\gamma_+)^2 = \alpha^2 \), and again we take \( \mu = \mu_{1,1} \).

(V) Examining Figure 1 we find that \( x \cdot b_1^1 = -(\gamma_-)b_1^1 - (\gamma_+) \mu_{1,1} b_1^2 \). Then the relation \( xg = -gx \) in \( T(2) \) gives that \( \mu_{1,1}x \cdot b_1^2 = (\gamma_+) \mu_{1,1} \mu_{2,2} b_1^1 + (\gamma_-) \mu_{1,1} b_1^2 \). The relation \( x^2 = 0 \) applied to the arrows \( b_1^1 \) and \( b_1^2 \) implies that \( (\gamma_+)^2 = (\gamma_-)^2 \), and again we take \( \mu = \mu_{1,1} \).

(VI) We can see from Figure 1 that \( x \cdot b_1^1 = -(\gamma_-)b_1^1 - (\gamma_+) \mu_{1,1} b_1^2 + \lambda b_1^2 \) for some \( \lambda \in \mathbb{k} \). Similarly, we can see from this figure that \( x \cdot b_2^1 = (\gamma_-)b_2^1 - (\gamma_+) \mu_{1,2} b_2^1 + \lambda' b_1^1 \) for some \( \lambda' \in \mathbb{k} \). The relation \( xg = -gx \) gives the formulas for \( x \cdot b_2^1 \) and \( x \cdot b_1^2 \) as claimed. Finally, the relation \( x^2 = 0 \) implies that \( (\gamma_+)^2 = (\gamma_-)^2 + \lambda \lambda' \). We have taken \( \mu = \mu_{1,1} \) and \( \mu' = \mu_{1,2} \).

**Remark 5.2.** Since \( Q(III)^{op} = Q(IV) \), we can get the Sweedler action on the path algebra of \( Q(IV) \) from its action on the path algebra of \( Q(III) \) using (2.7); see Remark 2.14. For instance, take \( b_1^1 \in Q(IV) \):
\[
x \cdot b_1^1 = g^{-1} x \cdot b_1^1 = g^{-1} \cdot (-(\gamma_-)b_1^1 + \beta \mu b_2^1) = -(\gamma_-) \mu b_2^1 + \beta b_1^1.
\]
Now by identifying \( \gamma_- \) with \( \gamma_+ \), and \( \beta \) with \( \alpha \), we get the desired action.
6. Taft algebra actions on path algebras of minimal quivers

In this section, we give a complete description of $T(n)$-actions on path algebras of $\mathbb{Z}_n$-minimal quivers. Since the action of the grouplike elements of $T(n)$ has already been determined in Lemma 3.2, and the $T(n)$-action on the vertices has already been determined in Proposition 3.5, all that remains is to describe possible actions of $x$ on the arrows.

6A. Taft actions on Type A $\mathbb{Z}_n$-minimal quivers. Let $Q$ be a $\mathbb{Z}_n$-minimal quiver of Type A, so that $Q$ is a subquiver of $K_m$ for some $m | n$ by Definition 4.3. Recall that the path algebra $\mathbb{k} K_m$ has basis $a_i^j$ for $1 \leq i, j \leq m$, with $a_i^i = e_i$ being the trivial path at vertex $i$.

**Theorem 6.1.** Recall Hypothesis 1.2 and retain the notation above. Any $T(n)$-action on the path algebra of Type A $\mathbb{Z}_n$-minimal quiver $Q$ is given by

$$x \cdot a_j^i = \gamma(\xi^j a_j^i - \xi^{i+1} \mu_{i,j} a_{j+1}^{i+1}) + \lambda_{i,j} a_{j+1}^i,$$

where $\gamma \in \mathbb{k}$ from (3.6), $\mu_{i,j} \in \mathbb{k}^\times$ from (4.2), and $\lambda_{i,j} \in \mathbb{k}$ are all scalars satisfying

- $\mu_{i,i} = 1$ for all $i$, and $\prod_{\ell=0}^{n-1} \mu_{i+i, j+i+\ell} = 1$;
- $\lambda_{i,j} = 0$ if either $i = j$ or the arrow $a_j^i$ does not exist in $Q$; and
- $\lambda_{i+1,j+1} \mu_{i,j} = \xi \lambda_{i,j} \mu_{i,j+1}$.

The superindices and subindices of arrows and scalars are taken modulo $m$.

**Proof.** We have already used the relations of the path algebra to find that

$$x \cdot a_j^i = \gamma(\xi^j a_j^i - \xi^{i+1} \mu_{i,j} a_{j+1}^{i+1}) + \lambda_{i,j} a_{j+1}^i,$$

for some scalars $\mu_{i,j} \in \mathbb{k}^\times$ and $\lambda_{i,j} \in \mathbb{k}$; see Lemma 3.2 and Propositions 3.5 and 3.10. Namely, the case of $i = j$ is Proposition 3.5, which gives $\lambda_{i,i} = 0$ and $\mu_{i,j} = 1$ for all $i$. Moreover, the case of $i \neq j$ is Proposition 3.10, which gives that $\lambda_{i,j} = 0$ if the arrow $a_j^i$ does not exist in $Q$. The condition on the $\{\mu_{i,j}\}$ is from (4.2).

The relation $xg = \xi g x$ of $T(n)$ applied to $a_j^i$ gives on the one hand that

$$xg \cdot a_j^i = x \cdot \mu_{i,j} a_{j+1}^{i+1},$$

while on the other hand that

$$\xi g x \cdot a_j^i = \xi \gamma(\xi^j \mu_{i,j} a_{j+1}^{i+1} - \xi^{i+1} \mu_{i,j} \mu_{i+1,j+1} a_{j+2}^{i+2}) + \zeta_{i,j} \mu_{i,j+1} a_{j+2}^{i+1}.$$

Thus we see that $\lambda_{i+1,j+1} \mu_{i,j} = \xi \lambda_{i,j} \mu_{i,j+1}$, which is the third condition on the scalars. We also obtain that the relation $x^n = 0$ applied to $a_j^i$ imposes no further restrictions on the $x$-action on $a_j^i$; this is verified in Lemma 9.11 of the appendix. □
6B. Taft actions on Type B $\mathbb{Z}_n$-minimal quivers. In this subsection, let $Q$ be a $\mathbb{Z}_n$-minimal quiver of Type B. By Definition 4.3, we know that $Q$ is a subquiver of $K_{m,m'}$ for some positive integers $m, m'$ both dividing $n$. Recall that the path algebra $K_{m,m'}$ has basis $e_i + e_j - b_{ij}$ where $1 \leq i \leq m$ and $1 \leq j \leq m'$.

Theorem 6.3. Recall Hypothesis 1.2 and retain the notation above. Any $T(n)$-action on the path algebra of Type B $\mathbb{Z}_n$-minimal quiver $Q$ is given by

$$x \cdot b_{ij} = (\gamma_+ - \gamma_-)\zeta_j b_{ij} - (\gamma_+ + \gamma_-)\mu_{i,j}b_{ij+1} + \lambda_{i,j}b_{i+1,j+1},$$

(6.4)

where $\gamma_+, \gamma_- \in \mathbb{k}$ from (3.12), $\mu_{i,j} \in \mathbb{k}^\times$ from (4.2), and $\lambda_{i,j} \in \mathbb{k}$ are all scalars satisfying

- $\prod_{\ell=0}^{n-1} \mu_{i+\ell,j+\ell} = 1$;
- $\lambda_{i,j} = 0$ if the arrow $b_{ij+1}$ does not exist in $Q$;
- $\lambda_{i+1,j+1}\mu_{i,j} = \xi \lambda_{i,j}\mu_{i,j+1}$;
- $(\gamma_+)^n = (\gamma_-)^n + \prod_{\ell=0}^{n-1} \lambda_{i,j+\ell}$.

The superindices of arrows are taken modulo $m$ and the subindices of arrows are taken modulo $m'$.

Proof. The formula (6.4) and first three conditions on the scalars are derived exactly as in the proof of Theorem 6.1, replacing $a^* \star \star$ with $b^* \star \star$ and $\gamma$ with $\gamma_\pm$, appropriately. It just remains to check that $x^n$ acts by 0; this is equivalent to the last condition on the scalars, as shown in Lemma 9.12. $\square$

7. Gluing Taft algebra actions on minimal quivers

7A. Gluing actions from components. In this section, we provide a recipe for gluing actions of Taft algebras on minimal quivers. We also show that, given any quiver with $\mathbb{Z}_n$-symmetry, one can construct an inner faithful extended action of $T(n)$ on the path algebra of this quiver.

Definition 7.1 ($\mathbb{Z}_n$-component). Let $Q$ be a quiver with an action of $\mathbb{Z}_n$, and consider the set of $\mathbb{Z}_n$-minimal subquivers of $Q$, partially ordered by inclusion. We say that a $\mathbb{Z}_n$-minimal subquiver of $Q$ is a $\mathbb{Z}_n$-component of $Q$ if it is maximal in the given ordering.

Lemma 7.2. Fix an action of $\mathbb{Z}_n$ on a quiver $Q$. Then, there exists a collection of $\mathbb{Z}_n$-components of $Q$, unique up to relabeling, such that $Q$ is obtained by gluing this collection.

Proof. The $\mathbb{Z}_n$-components of $Q$ exist and are uniquely determined by the definition because they are the maximal elements of a finite poset. Each arrow of $Q$ lies in some $\mathbb{Z}_n$-minimal subquiver, so the $\mathbb{Z}_n$-components cover $Q$. So, it suffices to show
that the intersection of two distinct $\mathbb{Z}_n$-components consists entirely of vertices. If $Q^1$ and $Q^2$ are two $\mathbb{Z}_n$-minimal subquivers such that $Q^1 \cap Q^2$ contains an arrow, then we can see from the definition of minimality that $Q^1$ and $Q^2$ have the same set of vertices. Thus, $Q^1 \cup Q^2$ is a $\mathbb{Z}_n$-minimal subquiver of $Q$. Repeat this process to conclude that, by maximality, any two distinct $\mathbb{Z}_n$-components can only have vertices in their intersection. □

A visualization of the result above can be found in Step 1 of Examples 7.8 and 7.9 below.

**Lemma 7.3.** Any $T(n)$-action on a path algebra $\mathbb{k}Q$ restricts to an action on the path algebra of each $\mathbb{Z}_n$-component of $Q$.

**Proof.** Let $Q^i$ be a $\mathbb{Z}_n$-component of $Q$. Since $Q^i$ is $\mathbb{Z}_n$-minimal, by definition $\mathbb{k}Q^i$ is stable under the action of $g$, so it suffices to show that $\mathbb{k}Q^i$ is stable under the action of $x$. From Proposition 3.10, it is enough to see that $\sigma(a) \in Q^i$ when $a \in Q^i$. Suppose not, that is, $\sigma(a) \in Q^j$ for some $j \neq i$. Then, $Q^i \cup Q^j$ is a $\mathbb{Z}_n$-minimal quiver, which contradicts the maximality of the $\mathbb{Z}_n$-component $Q^i$. □

**Definition 7.4** (compatibility). Let $Q$ be a quiver and $Q^1, \ldots, Q^r \subseteq Q$ a collection of subquivers of $Q$. Suppose that we have a $T(n)$-action on each path algebra $\mathbb{k}Q^i$. We say that this collection of $T(n)$-actions is compatible if, for each pair $(i, j)$, the restriction of the actions on $\mathbb{k}Q^i$ and on $\mathbb{k}Q^j$ to $\mathbb{k}[Q^i \cap Q^j]$ are the same.

Now we have our main result of the section.

**Theorem 7.5.** Let $Q$ be a quiver with $\mathbb{Z}_n$-action. The $T(n)$-actions on the path algebra of $Q$ extending the given $\mathbb{Z}_n$-action are in bijection with compatible collections of $T(n)$-actions on path algebras of the $\mathbb{Z}_n$-components of $Q$.

**Proof.** Given a $T(n)$-action on $\mathbb{k}Q$, it restricts to a $T(n)$-action on each path algebra of a $\mathbb{Z}_n$-component of $Q$ by Lemma 7.3. On the other hand, suppose we have a collection of compatible $T(n)$-actions on the path algebras of the $\mathbb{Z}_n$-components of $Q$. This uniquely determines an action of $T(n)$ on $\mathbb{k}Q$ as follows:

- The action on each arrow of $\mathbb{k}Q$ is uniquely determined since each arrow lies in a unique $\mathbb{Z}_n$-component of $Q$.
- The action on any vertex is determined because every vertex lies in at least one $\mathbb{Z}_n$-component.
- Suppose that a vertex lies in multiple $\mathbb{Z}_n$-components. Then, the action on this vertex is uniquely determined because the actions on different $\mathbb{Z}_n$-components restrict to the same action on vertices in their intersection, due to compatibility. □

**Corollary 7.6.** If a quiver $Q$ admits a faithful $\mathbb{Z}_n$-action, then $\mathbb{k}Q$ admits an inner faithful action of the Taft algebra $T(n)$, extending the given $\mathbb{Z}_n$-action on $Q$. 
Proof. This follows immediately from Theorem 7.5 and Proposition 3.5 since \( Q \) admits an orbit of vertices of size \( n \) in this case. \( \square \)

Example 7.7. This example shows that a faithful \( \mathbb{Z}_n \)-action on \( Q \) is not necessary for \( \mathbb{k}Q \) to admit an inner faithful action of \( T(n) \). Fix \( \zeta \), a primitive fourth root of unity. Consider the action of \( T(4) \) on \( \mathbb{k}K_2 \) (see (1) of Figure 5) given by

\[
\begin{align*}
g \cdot e_i &= e_{i+1}, \quad g \cdot a_{ij}^i = \zeta a_{j+1,i}^{i+1}, \quad x \cdot e_i = 0, \quad x \cdot a_{ij}^i = \lambda e_i,
\end{align*}
\]

for \( i \neq j \), where subscripts and superscripts are taken modulo 2, and \( \lambda \in \mathbb{k}^* \) is an arbitrary nonzero scalar. By Theorem 6.1, this defines an action of \( T(4) \) on the path algebra \( \mathbb{k}K_2 \), which is inner faithful even though the induced action of \( \mathbb{Z}_4 \) on the quiver \( K_2 \) is not faithful.

7B. Algorithm to explicitly parametrize \( T(n) \)-actions on \( \mathbb{k}Q \). Let \( Q \) be a quiver that admits an action of \( \mathbb{Z}_n \). We construct all actions of \( T(n) \) on \( \mathbb{k}Q \) which extend the given \( \mathbb{Z}_n \)-action via the following steps. Those for which \( x \) does not act by 0 are inner faithful by Lemma 2.5. The reader may wish to refer to one of the examples in the next subsection for an illustration of the algorithm below.

First, we know that \( Q \) decomposes uniquely into the union of certain \( \mathbb{Z}_n \)-minimal quivers \( \{Q^\ell\}_{\ell=1}^r \) (namely, \( \mathbb{Z}_n \)-components) so that \( Q = Q^1 \oplus \cdots \oplus Q^r \), due to Lemma 7.2.

\begin{center}
\textbf{Step 1}
\end{center}

Decompose \( Q \) into this unique union of \( \mathbb{Z}_n \)-components \( \{Q^\ell\} \).

Next, we define the extended action of \( T(n) \) on the path algebra of each component \( Q^\ell \). Let \( m, m' \) be positive divisors of \( n \). Recall that Type A and Type B \( \mathbb{Z}_n \)-minimal quivers are subquivers of \( K_m \) with \( m > 1 \) and of \( K_m, m' \), respectively.

\begin{center}
\textbf{Step 2}
\end{center}

For each \( \mathbb{Z}_n \)-component \( Q^\ell \) of Type A, label its vertices by \( \{1^{(\ell)}, \ldots, m_i^{(\ell)}\} \). For each \( \mathbb{Z}_n \)-component \( Q^\ell \) of Type B, label its sources and sinks by \( \{1_+^{(\ell)}, \ldots, m_+^{(\ell)}\} \) and \( \{1_-^{(\ell)}, \ldots, (m')_-^{(\ell)}\} \), respectively.

Recall that \( \zeta \) is the primitive \( n \)-th root of unity from the definition of \( T(n) \). Now invoke Proposition 3.5 in the following step.

\begin{center}
\textbf{Step 3}
\end{center}

Take scalars \( \gamma^{(\ell)}, \gamma_+^{(\ell)}, \gamma_-^{(\ell)} \in \mathbb{k} \) and define

\[
\begin{align*}
x \cdot e_{i^{(\ell)}} &= \gamma^{(\ell)} \zeta^i (e_{i^{(\ell)}} - \zeta e_{(i+1)^{(\ell)}}) \quad \text{for Type A}, \\
x \cdot e_{i_+^{(\ell)}} &= \gamma_+^{(\ell)} \zeta^i (e_{i_+^{(\ell)}} - \zeta e_{(i+1)_+^{(\ell)}}) \quad \text{for Type B}, \\
x \cdot e_{i_-^{(\ell)}} &= \gamma_-^{(\ell)} \zeta^i (e_{i_-^{(\ell)}} - \zeta e_{(i+1)_-^{(\ell)}}) \quad \text{for Type B},
\end{align*}
\]
where the indices are taken modulo $m$ for Type A, and are taken modulo $m$ or $m'$ for Type B. Here, $\gamma_{+}^{(\ell)} = 0$ if $m < n$ for Type A and $\gamma_{+}^{(\ell)} = 0$ or $\gamma_{-}^{(\ell)} = 0$ if $m < n$ or $m' < n$ respectively, for Type B. To obtain compatibility, impose relations amongst the scalars by identifying vertices.

Finally, we invoke Theorems 6.1 and 6.3 to get all actions of $T(n)$ on the arrows of $Q$.

**STEP 4**

Label the arrows of each Type A and Type B component of $Q$ by arrows $(a_{ij}^{(\ell)})$ and $(b_{ij}^{(\ell)})$ respectively.

**STEP 5**

Given the scalars $\gamma^{(\ell)}, \gamma_{+}^{(\ell)}, \gamma_{-}^{(\ell)} \in \mathbb{k}$ of Step 3, we have, for all arrows $a_{ij}^{(\ell)}$ of a component of Type A and $b_{ij}^{(\ell)}$ of a component of Type B, that

$$x \cdot (a_{ij}^{(\ell)}) = \gamma^{(\ell)}(\xi^{i}(a_{ij}^{(\ell)}) - \xi^{i+1}\mu_{i,j}^{(\ell)}(a_{i,j+1}^{(\ell)}) + \lambda_{i,j}^{(\ell)}(a_{i,j+1}^{(\ell)}),$$

$$x \cdot (b_{ij}^{(\ell)}) = \gamma_{-}^{(\ell)}\xi^{i+1}(b_{ij}^{(\ell)}) - \gamma_{+}^{(\ell)}\xi^{i+1}\mu_{i,j}^{(\ell)}(b_{i,j+1}^{(\ell)}) + \lambda_{i,j}^{(\ell)}(b_{i,j+1}^{(\ell)}),$$

where $\mu_{i,j}^{(\ell)}, \lambda_{i,j}^{(\ell)} \in \mathbb{k}$ are scalars satisfying the conditions of Theorems 6.1 and 6.3.

Thus, the desired extended action of the $n$-th Taft algebra on $\mathbb{k}Q$ is complete by the five steps above.

**7C. Examples.** In this section, we illustrate the algorithm of the previous section to get actions of $T(n)$ on $\mathbb{k}Q$ for various quivers $Q$ having $\mathbb{Z}_n$ symmetry.

For ease of exposition, we take all parameters $\mu_{i,j}$ equal to 1 in this section.

**Example 7.8.** Consider the following quiver $Q$:

![Diagram](image)

which has $\mathbb{Z}_2$-symmetry by reflection over the central vertical axis and exchanging arrows $f_1$ and $f_2$. So, we decompose $Q$ into the four $\mathbb{Z}_2$-components as follows.
Here, the dotted red arrows indicate the action of $\mathbb{Z}_2$ on $Q_0$.

Now choose scalars $\gamma^{(1)}, \gamma^{(2)}_+, \gamma^{(2)}_-, \gamma^{(3)}_+, \gamma^{(3)}_-, \gamma^{(4)}_+, \gamma^{(4)}_- \in k$ to execute Steps 2 and 3.

We have

\[
x \cdot e_1^{(1)} = -\gamma^{(1)}(e_1^{(1)} + e_2^{(1)}), \quad x \cdot e_2^{(1)} = \gamma^{(1)}(e_2^{(1)} + e_1^{(1)}),
\]

while, for $\ell = 2, 3, 4$,

\[
x \cdot e_1^{(\ell)} = -\gamma^{(\ell)}_+(e_1^{(\ell)} + e_2^{(\ell)}), \quad x \cdot e_2^{(\ell)} = \gamma^{(\ell)}_+(e_2^{(\ell)} + e_1^{(\ell)}),
\]

\[
x \cdot e_1^{(\ell)} = -\gamma^{(\ell)}_-(e_1^{(\ell)} + e_2^{(\ell)}), \quad x \cdot e_2^{(\ell)} = \gamma^{(\ell)}_-(e_2^{(\ell)} + e_1^{(\ell)}).
\]

By using the identification of vertices

\[
v_1 := 1^{(1)} = 1^{(2)}_+ = 1^{(4)}_+, \quad v_2 := 2^{(1)} = 2^{(2)}_+ = 2^{(4)}_+,
\]

\[
v_3 := 1^{(2)}_- = 1^{(3)}_-, \quad v_4 := 2^{(2)}_- = 2^{(3)}_-,
\]

\[
v_5 := 1^{(3)}_+ = 1^{(4)}_-, \quad v_6 := 2^{(3)}_+ = 2^{(4)}_-,
\]

we have

\[
\gamma := \gamma^{(1)} = \gamma^{(2)}_+ = \gamma^{(4)}_+, \quad \gamma' := \gamma^{(2)}_- = \gamma^{(3)}_-, \quad \gamma'' := \gamma^{(3)}_+ = \gamma^{(4)}_-.
\]
Pointed Hopf actions on path algebras of quivers 141

Now, we need to label the arrows of $Q$ appropriately, and invoke Theorems 6.1 and 6.3 to get the desired action of $T(2)$ on $kQ_1$.

**Steps 4 and 5**

![Diagram showing paths and actions](attachment:diagram.png)

$x \cdot (a_j^i)^{(1)} = \gamma^{(1)}(\xi^j(a_j^i)^{(1)}) - \xi^{i+1}(a_{j+1}^{i+1})^{(1)} + \lambda_{i,j}^{(1)}(a_{j+1}^i)^{(1)},$

$x \cdot (b_j^i)^{(\ell)} = \gamma^{(\ell)}(\xi^j(b_j^i)^{(\ell)}) - \gamma^{(\ell)}(\xi^{i+1}(b_{j+1}^{i+1})^{(\ell)}) + \lambda_{i,j}^{(\ell)}(b_{j+1}^i)^{(\ell)}, \quad \ell = 2,3,4.$

Using the conditions on the scalars from Theorems 6.1 and 6.3, we find that the $x$-action on the arrows is controlled by additional parameters $\lambda := \lambda_{1,2}^{(1)}$ as $\sigma(a)$ exists for component (1), and $\lambda' := \lambda_{1,2}^{(4)}$, $\lambda'' := \lambda_{1,2}^{(4)}$ as $\sigma(a)$ exists for component (4). Putting this all together, the action of $\mathbb{Z}_2$ on $Q$ (where by abuse of notation, $v_i$ denotes the trivial path at $v_i$) given by

$g \cdot v_1 = v_2, \quad g \cdot v_2 = v_1, \quad g \cdot f_1 = f_2, \quad g \cdot f_2 = f_1,$

$g \cdot v_3 = v_4, \quad g \cdot v_4 = v_3, \quad g \cdot f_3 = f_4, \quad g \cdot f_4 = f_3,$

$g \cdot v_5 = v_6, \quad g \cdot v_6 = v_5, \quad g \cdot f_5 = f_6, \quad g \cdot f_6 = f_5,$

$g \cdot f_7 = f_8, \quad g \cdot f_8 = f_7, \quad g \cdot f_9 = f_{10}, \quad g \cdot f_{10} = f_9,$

extends to an action of the Sweedler algebra on $kQ$, as follows:

$x \cdot v_1 = -\gamma(v_1 + v_2), \quad x \cdot v_2 = \gamma(v_2 + v_1),$

$x \cdot v_3 = -\gamma'(v_3 + v_4), \quad x \cdot v_4 = \gamma'(v_4 + v_3),$

$x \cdot v_5 = -\gamma''(v_5 + v_6), \quad x \cdot v_6 = \gamma''(v_6 + v_5),$

$x \cdot f_1 = \gamma f_1 - \gamma f_2 + \lambda e_1, \quad x \cdot f_2 = \gamma f_1 - \gamma f_2 - \lambda e_2,$

$x \cdot f_3 = -\gamma' f_3 - \gamma f_4, \quad x \cdot f_4 = \gamma' f_4 + \gamma f_3,$

$x \cdot f_5 = -\gamma' f_5 - \gamma'' f_6, \quad x \cdot f_6 = \gamma' f_6 + \gamma'' f_5,$

$x \cdot f_7 = -\gamma'' f_7 - \gamma f_8 + \lambda' f_9, \quad x \cdot f_8 = \gamma'' f_8 + \gamma f_7 - \lambda' f_{10},$

$x \cdot f_9 = \gamma'' f_9 - \gamma f_{10} + \lambda'' f_7, \quad x \cdot f_{10} = -\gamma'' f_{10} + \gamma f_9 - \lambda'' f_8,$

for any scalars $\gamma$, $\gamma'$, $\gamma''$, $\lambda$, $\lambda'$, $\lambda'' \in k$ which satisfy $(\gamma)^2 = (\gamma')^2 + \lambda' \lambda''$.

Note that the Sweedler algebra action restricted to the path algebras of respective components (1), (2), (3), (4) is the same as the Sweedler algebra action on the path.
algebra of respective quivers (I), (V), (V), (VI) from Theorem 5.1, as expected; namely, $Q = Q(I) \otimes Q(V) \otimes Q(V) \otimes Q(VI)$.

**Example 7.9.** Consider the following quiver $Q$:

![Quiver Diagram]

which has $\mathbb{Z}_3$-symmetry. We decompose $Q$ into three $\mathbb{Z}_3$-components as follows. Here, the dotted red arrows indicate the action of $\mathbb{Z}_3$ on $Q_0$.

**Step 1**

Now choose scalars $\gamma_+^{(1)}, \gamma_-^{(1)}, \gamma_+^{(2)}, \gamma_-^{(2)}, \gamma^{(3)} \in \mathbb{k}$ to execute Steps 2 and 3. Further, let $\omega$ be a primitive third root of unity.

**Steps 2 and 3**

$1_-, 1_+, 1^{(3)}$

$3_-, 3_+, 3^{(3)}$

(the middle vertex is labeled by $2_-, 2_+, 2^{(3)}$)
We have $x \cdot e_{1+}^{(1)} = x \cdot e_{1-}^{(2)} = 0$ since these vertices are fixed by $g$, and furthermore

$$
\begin{align*}
    x \cdot e_{1-}^{(1)} &= \gamma_{-}^{(1)} \omega (e_{1-}^{(1)} - \omega e_{2-}^{(1)}), &
    x \cdot e_{1+}^{(2)} &= \gamma_{+}^{(2)} \omega (e_{1+}^{(2)} - \omega e_{2+}^{(2)}), \\
    x \cdot e_{2-}^{(1)} &= \gamma_{-}^{(1)} \omega^2 (e_{2-}^{(1)} - \omega e_{3-}^{(1)}), &
    x \cdot e_{2+}^{(2)} &= \gamma_{+}^{(2)} \omega^2 (e_{2+}^{(2)} - \omega e_{3+}^{(2)}), \\
    x \cdot e_{3-}^{(1)} &= \gamma_{-}^{(1)} (e_{3-}^{(1)} - \omega e_{1-}^{(1)}), &
    x \cdot e_{3+}^{(2)} &= \gamma_{+}^{(2)} (e_{3+}^{(2)} - \omega e_{1+}^{(2)}), \\
    x \cdot e_{1}^{(3)} &= \gamma^{(3)} \omega (e_{1}^{(3)} - \omega e_{2}^{(3)}), \\
    x \cdot e_{2}^{(3)} &= \gamma^{(3)} \omega^2 (e_{2}^{(3)} - \omega e_{3}^{(3)}), \\
    x \cdot e_{3}^{(3)} &= \gamma^{(3)} (e_{3}^{(3)} - \omega e_{1}^{(3)}).
\end{align*}
$$

By using the identification of vertices,

$$
v_1 = 1_{-}^{(1)} = 1_{+}^{(2)} = 1^{(3)}, \quad v_2 = 2_{-}^{(1)} = 2_{+}^{(2)} = 2^{(3)}, \quad v_3 = 3_{-}^{(1)} = 3_{+}^{(2)} = 3^{(3)},
$$

we have

$$
\gamma := \gamma_{-}^{(1)} = \gamma_{+}^{(2)} = \gamma^{(3)}, \quad \gamma_{+}^{(1)} = 0, \quad \gamma_{-}^{(2)} = 0.
$$

Now, we need to label the arrows of $Q$ appropriately, and invoke Theorems 6.1 and 6.3 to get the desired action of $T(3)$ on $\mathbb{k}Q_1$.

**Steps 4 and 5**

$$
\begin{align*}
    x \cdot (b_{j}^{1})^{(1)} &= \gamma_0 \omega^0 (b_{j}^{1})^{(1)} + \lambda_0 \omega^{-1} (b_{j+1}^{1})^{(1)}, \\
    x \cdot (b_{j}^{1})^{(2)} &= -\gamma_0 \omega^{j+1} (b_{j+1}^{1})^{(2)} + \lambda_0 \omega^{-1} (b_{j}^{1})^{(2)}, \\
    x \cdot (a_{j}^{i})^{(3)} &= \gamma_i \omega^j (a_{j}^{i})^{(3)} - \omega^{i+1} (a_{i+1}^{j+1})^{(3)} + \lambda_i \omega (a_{i+1}^{j})^{(3)},
\end{align*}
$$

where $\lambda := \lambda_{1,1}^{(1)}$ and $\lambda' := \lambda_{1,1}^{(2)}$.

Moreover, let $\lambda'' := \lambda_{1,2}^{(3)}$ and $\lambda''' := \lambda_{1,3}^{(3)}$ for component (3). Putting this all together, the action of $\mathbb{Z}_3$ on $Q$ given by Step 1 extends to an action of $T(3)$ on $\mathbb{k}Q$, as
follows:
\[ x \cdot v_1 = \gamma \omega (v_1 - \omega v_2), \quad x \cdot v_2 = \gamma \omega^2 (v_2 - \omega v_3), \quad x \cdot v_3 = \gamma (v_3 - \omega v_1), \]
\[ x \cdot v_0 = 0, \quad x \cdot v_4 = 0, \]
\[ x \cdot f_1 = \gamma \omega f_1 + \lambda f_2, \quad x \cdot f_4 = \lambda' f_4 - \gamma \omega^2 f_5, \]
\[ x \cdot f_2 = \gamma \omega^2 f_2 + \lambda \omega f_3, \quad x \cdot f_5 = \lambda' \omega f_5 - \gamma f_6, \]
\[ x \cdot f_3 = \gamma f_3 + \lambda \omega^2 f_1, \quad x \cdot f_6 = \lambda' \omega^2 f_6 - \gamma \omega f_4, \]
\[ x \cdot f_7 = \gamma (\omega^2 f_7 - \omega^2 f_8) + \lambda'' f_{10}, \quad x \cdot f_{10} = \gamma (f_{10} - \omega^2 f_{12}) + \lambda''' f_7, \]
\[ x \cdot f_8 = \gamma (f_8 - f_9) + \lambda''\omega f_{12}, \quad x \cdot f_{11} = \gamma (\omega^2 f_{11} - \omega f_{10}) + \lambda'''\omega f_9, \]
\[ x \cdot f_9 = \gamma (\omega f_9 - \omega f_7) + \lambda''\omega^2 f_{11}, \quad x \cdot f_{12} = \gamma (\omega f_{12} - f_{11}) + \lambda'''\omega^2 f_8, \]
for any scalars \(\gamma, \lambda, \lambda', \lambda'', \lambda''' \in \mathbb{k}\).

8. Extended actions of other pointed Hopf algebras on path algebras

In this section, we extend the results in the previous sections to actions of the Frobenius–Lusztig kernel (Section 8A) and to actions of the Drinfeld double of a Taft algebra (Section 8B). We remind the reader of the standing assumptions made in Hypothesis 1.2.

**Notation 8.1** \((q, T(n, \xi))\). Let \(q \in \mathbb{k}\) be a primitive \(2n\)-th root of unity, and let \(\xi \in \mathbb{k}\) be a primitive \(n\)-th root of unity. Moreover, let \(T(n, \xi)\) be the Taft algebra generated by a grouplike element \(g\) and a \((1, g)\)-skew primitive element \(x\), subject to the relations \(g^n = 1, x^n = 0, xg = \xi gx\). Note that \(T(n, \xi) = T(n)\) as in Definition 2.1, for \(\xi\) the fixed primitive \(n\)-th root of unity of this work.

8A. Actions of \(u_q(\mathfrak{sl}_2)\)

**Definition 8.2** \((u_q(\mathfrak{sl}_2), Borel subalgebras \(u_q^{\geq 0}(\mathfrak{sl}_2)\) and \(u_q^{\leq 0}(\mathfrak{sl}_2))\). The Frobenius–Lusztig kernel \(u_q(\mathfrak{sl}_2)\) is the Hopf algebra generated by a grouplike element \(K\), a \((1, K)\)-skew-primitive element \(E\), and a \((K^{-1}, 1)\)-skew-primitive element \(F\), subject to the relations
\[ KE = q^2 KE, \quad KF = q^{-2} FK, \quad K^n = 1, \quad E^n = F^n = 0, \]
\[ EF - FE = \frac{K - K^{-1}}{q - q^{-1}}. \quad (8.3) \]

Note that \(u_q(\mathfrak{sl}_2)\) is pointed, of dimension \(n^3\). Let \(u_q^{\geq 0}(\mathfrak{sl}_2)\) be the Hopf subalgebra of \(u_q(\mathfrak{sl}_2)\) generated by \(K, E,\) and let \(u_q^{\leq 0}(\mathfrak{sl}_2)\) be the Hopf subalgebra of \(u_q(\mathfrak{sl}_2)\) generated by \(K, F\); we refer to these as Borel subalgebras.

**Lemma 8.4.** There are isomorphisms of Hopf algebras
\[ u_q^{\geq 0}(\mathfrak{sl}_2) \simeq T(n, q^{-2}) \quad and \quad u_q^{\leq 0}(\mathfrak{sl}_2) \simeq T(n, q^{2}). \quad (8.5) \]
Proof. This is easy to check: identify $K$ with $g$ for both $u_q^{\geq 0}(sl_2)$ and $u_q^{\leq 0}(sl_2)$, and $E$ with $x$ for $u_q^{\geq 0}(sl_2)$, and $F$ with $g^{-1}x$ for $u_q^{\leq 0}(sl_2)$. \hfill $\square$

Since $u_q(sl_2)$ is generated by the Hopf subalgebras $u_q^{\geq 0}(sl_2)$ and $u_q^{\leq 0}(sl_2)$, an action of $u_q(sl_2)$ on a path algebra $\mathbb{k}Q$ is determined by its restriction to these Hopf subalgebras. Since these actions are given by the results on Taft actions on $\mathbb{k}Q$ in the previous sections, we can classify actions of $u_q(sl_2)$ by determining some additional compatibility conditions on the scalar parameters from these Taft actions. We begin with the result below.

**Proposition 8.6.** Let $u_q(sl_2)$ act on the path algebra of a set of vertices $\{1, \ldots, m\}$, labeled so that $K \cdot e_i = e_{i+1}$ with subscripts taken modulo $m$.

(i) If $m < n$, then $m = 1$ or $m = 2$ and the action factors through the quotient $u_q(sl_2)/\langle E, F \rangle$.

(ii) If $m = n$, then
\[
E \cdot e_i = \gamma^E q^{-2i} (e_i - q^{-2} e_{i+1}), \quad F \cdot e_i = \gamma^F q^{2i} (e_{i-1} - q^2 e_i),
\]
for some $\gamma^E, \gamma^F \in \mathbb{k}$. If furthermore $n \geq 3$, then the action is subject to the restriction
\[
-\gamma^E \gamma^F q^{-1} (q^2 - 1)^2 = 1. \tag{8.8}
\]

Proof. Part (i) is a result of Proposition 3.5, translated through the isomorphisms of Lemma 8.4. The condition $m = 1$ or $m = 2$ comes from the relation (8.3), which implies that $K - K^{-1}$ acts by zero in this case. For (ii), Proposition 3.5 also gives the formulas (8.7), and the first row of relations in the definition of $u_q(sl_2)$ then hold. On the other hand, substituting (8.7) into the relation (8.3) of $u_q(sl_2)$ yields
\[
(EF - FE) \cdot e_i = \gamma^E \gamma^F (q^2 - 1)(e_i - e_{i+1}) \quad \text{while} \quad (K - K^{-1}) \cdot e_i = e_{i+1} - e_{i-1}.
\]
If $n = m = 2$, then $e_{i-1} = e_{i+1}$ and the relation (8.3) imposes no restriction on the action (8.7). If $n = m \geq 3$, then $e_{i+1} \neq e_{i-1}$ and the relation (8.3) imposes the restriction (8.8) on the action (8.7). \hfill $\square$

We get the following immediate consequence.

**Corollary 8.9.** If $n \geq 3$, then either one of $\gamma^E$ or $\gamma^F$ determines the other. So, an action of $u_q(sl_2)$ on a $\mathbb{Z}_n$-orbit of vertices is completely determined by the action of a Borel subalgebra, $u_q^{\geq 0}(sl_2)$ or $u_q^{\leq 0}(sl_2)$. \hfill $\square$

We now consider $u_q(sl_2)$-actions on Type A and Type B $\mathbb{Z}_n$-minimal quivers. Note that $u_q(sl_2)$ can only act on subquivers of $K_m$ and $K_{m,m'}$ for $m, m' \in \{1, 2, n\}$, by Proposition 8.6. We only give theorems describing the case $m = m' = n \geq 3$ here, but the other cases can be worked out similarly.
Theorem 8.10. Retain the notation of Section 4 with \( K \) in place of \( g \), so the action of \( K \) is fixed. Namely, \( K \cdot a^i_j = \mu_{i,j} a^i_{j+1} \) for Type A and \( K \cdot b^i_j = \mu_{i,j} b^i_{j+1} \) for Type B. Assume \( n \geq 3 \). Then any \( u_q(sl_2) \)-action on the path algebra of a \( \mathbb{Z}_n \)-minimal subquiver \( Q \) of \( K \) is given by

\[
E \cdot a^i_j = \gamma^E (q^{-2j} a^i_j - q^{-2i-2} a^i_{j+1}) + \lambda^E_{i,j} a^i_{j+1},
\]

\[
F \cdot a^i_j = \gamma^F (q^{2j} a^i_{j-1} - q^{2i+2} a^i_j) + \lambda^F_{i,j} a^i_{j-1},
\]

subject to the restriction (8.8) along with

- \( \lambda^E_{i,j} = 0 \), if \( i = j \) or the arrow \( a^i_{j+1} \) does not exist in \( Q \);
- \( \lambda^F_{i,j} = 0 \), if \( i = j \) or the arrow \( a^i_{j-1} \) does not exist in \( Q \);
- \( \lambda^E_{i+1,j+1} = q^{-2} \lambda^E_{i,j} \), \( \lambda^F_{i+1,j+1} = q^2 \lambda^F_{i,j} \), and \( \lambda^F_{i,j} \lambda^E_{i-1,j} = \lambda^E_{i,j} \lambda^F_{i-1,j+1} \).

Any \( u_q(sl_2) \)-action on the path algebra of a \( \mathbb{Z}_n \)-minimal subquiver \( Q \) of \( K_{n,n} \) is given by

\[
E \cdot b^i_j = (\gamma^E)^{-1} b^i_j - (\gamma^E) a^i_{j+1} + \lambda^E_{i,j} b^i_{j+1},
\]

\[
F \cdot a^i_j = (\gamma^F)^{-1} b^i_{j-1} - (\gamma^F) a^i_{j-1} + \lambda^F_{i,j} b^i_{j-1},
\]

subject to restrictions of the form (8.8) on \( \gamma^E \) and \( \gamma^F \), along with

- \( \lambda^E_{i,j} = 0 \), if \( i = j \) or the arrow \( b^i_{j+1} \) does not exist in \( Q \);
- \( \lambda^F_{i,j} = 0 \), if \( i = j \) or the arrow \( b^i_{j-1} \) does not exist in \( Q \);
- \( \lambda^E_{i+1,j+1} = q^{-2} \lambda^E_{i,j} \), \( \lambda^F_{i+1,j+1} = q^2 \lambda^F_{i,j} \), and \( \lambda^F_{i,j} \lambda^E_{i-1,j} = \lambda^E_{i,j} \lambda^F_{i-1,j+1} \);
- \( (\gamma^E)^n = (\gamma^E)^n = \prod_{\ell=0}^{n-1} \lambda^E_{i,j+\ell} \) and \( (\gamma^F)^n = (\gamma^F)^n = \prod_{\ell=0}^{n-1} \lambda^F_{i,j+\ell} \).

Proof. First consider the Type A case. Proposition 8.6 gives the \( u_q(sl_2) \)-action on the vertices, imposing the restriction (8.8). Theorem 6.1, translated through the isomorphisms of Lemma 8.4, initially gives expressions for \( E \cdot a^i_j \) and \( F \cdot a^i_j \) which involve parameters \( \mu_{i,j} \). However, we examine the coefficients of various arrows in the relation (8.3) applied to \( a^i_j \), and find that the coefficient of \( a^i_{j-1} \) gives

\[
\gamma^E \gamma^F (q^2 - 1) + \mu_{i,j} (q - q^{-1})^{-1} = 0.
\]

By applying (8.8), this simplifies to \( (\mu_{i,j} - 1)(q - q^{-1})^{-1} = 0 \). So, \( \mu_{i,j} = 1 \) for all \( i, j \), which gives our formulas (8.11). Similarly, the coefficient of \( a^i_{j+1} \) in the relation (8.3) applied to \( a^i_j \) implies that

\[
\lambda^F_{i,j} \lambda^E_{i-1,j} = \lambda^E_{i,j} \lambda^F_{i,j+1} = 0.
\]

Theorem 6.1 gives the remaining restrictions on the scalars, and the coefficients of other arrows yield restrictions that are already implied by those above.

For Type B, the action is derived exactly as in the Type A case, by replacing \( a^i_j \) with \( b^i_j \), and \( \gamma \) with \( \gamma^\pm \), appropriately. \( \square \)
8B. Actions of the double $D(T(n))$. We take the presentation of the Drinfeld double of the $n$-th Taft algebra in [Chen 1999, Definition-Theorem 3.1]. Namely, by [Chen 1999, Theorem 3.3], the double is isomorphic to the Hopf algebra $H_n(p, q)$ generated by $a, b, c, d$. We take $p = 1$, $q = \zeta^{-1}$, $a = x$, $b = g$, $c = G$, and $d = X$ to get that:

**Definition 8.15 ($D(T(n))$).** The Drinfeld double $D(T(n))$ of the $n$-th Taft algebra is generated by $g, x, G, X$, subject to relations

$$
\begin{align*}
    xg &= \zeta gx, \quad gX = \zeta Xg, \quad g^n = G^n = 1, \\
xG &= \zeta XG, \quad xG = \zeta Gx, \quad x^n = X^n = 0, \quad gG = Gg,
\end{align*}
$$

and

$$
    xX - \zeta XX = \zeta(gG - 1). \tag{8.16}
$$

Here, $g$ and $G$ are grouplike elements, $x$ is $(1, g)$-a skew primitive element, and $X$ is a $(1, G)$-skew primitive element. Here, $g, x$ generate the Hopf subalgebra $T(n)$, and $G = g^*, X = x^*$ generate $(T(n)^{op})^* = (T(n)^* \otimes p)$. The following lemma is easy to check.

**Lemma 8.17.** The Hopf subalgebra of $D(T(n))$ generated by $g, x$ is isomorphic to $T(n) = T(n, \zeta)$, as Hopf algebras. Moreover, the Hopf subalgebra of $D(T(n))$ generated by $G, X$ is isomorphic to $T(n, \zeta^{-1})$, as Hopf algebras. \hfill \Box

To give the complete classification of $D(T(n))$-actions on path algebras of quivers, one would need to define $\mathbb{Z}_n \times \mathbb{Z}_n$-minimal quivers and consider the $D(T(n))$-action on these, since $\mathbb{Z}_n \times \mathbb{Z}_n$ is isomorphic to the group of grouplike elements of $D(T(n))$. This is the subject of future work. Our aim for now is to exhibit an action of $D(T(n))$ on $\mathbb{K}Q$, where $Q$ admits an action of $\mathbb{Z}_n$. We begin by providing an action of $D(T(n))$ on a $\mathbb{Z}_n$-orbit of vertices.

**Proposition 8.18.** An action of $D(T(n))$ on $\mathbb{K}Q_0$ where $Q_0 = \{1, \ldots, n\}$ is given by

$$
\begin{align*}
g \cdot e_i &= e_{i+1}, \quad x \cdot e_i &= \gamma^x \zeta^i (e_i - \zeta e_{i+1}), \\
g \cdot e_i &= e_{i+1}, \quad X \cdot e_i &= \gamma^X \zeta^{-i} (e_i - \zeta^{-1} e_{i+1}) \tag{8.19}
\end{align*}
$$

for some $\gamma^x, \gamma^X \in \mathbb{K}$. If $n \geq 3$, then these scalars are subject to the restriction

$$
    \gamma^x \gamma^X (1 - \zeta^{-1}) = 1. \tag{8.20}
$$

**Proof.** The formulas (8.19) satisfy the first two rows of relations of $D(T(n))$. On the other hand, the restriction (8.20) comes from substituting (8.19) into the relation (8.16) of $D(T(n))$ when $n \geq 3$. If $n = 2$, the restriction from (8.16) is vacuous. \hfill \Box
Theorem 6.1 gives the expressions in (8.22), along with some restrictions on $n ≥ \mu_i$ implying, respectively, that $D$ satisfy the relations of the theorem. Now we need to check that the expressions, this imposes the restriction (8.20) when $n ≥ 3$, then we also impose the condition $\gamma \gamma^X (1 - \zeta^{-1}) = 1$.

Proof. First, Proposition 8.18 gives an action of $D(T(n))$ on Type A and Type B $\mathbb{Z}_n$-minimal quivers. As in the $u_q(\mathfrak{sl}_2)$ case, we only give theorems describing the cases when each orbit of vertices has $n$ elements here.

Theorem 8.21. Retain the notation of Section 4. An action of $D(T(n))$ on the path algebra of a Type A $\mathbb{Z}_n$-minimal subquiver $Q$ of $K_n$ is given by

$$g \cdot a^i_j = \mu^g_{i,j} a^{i+1}_{j+1}, \quad x \cdot a^i_j = \gamma^x (\zeta^i a^i_j - \zeta^{-i+1} \mu^g_{i,j} a^{i+1}_{j+1}) + \lambda^x_{i,j} a^i_j,$$
$$G \cdot a^i_j = \mu^G_{i,j} a^{i+1}_{j+1}, \quad X \cdot a^i_j = \gamma^X (\zeta^{-j} a^{-j}_i - \zeta^{-i-1} \mu^G_{i,j} a^{i+1}_{j+1}) + \lambda^X_{i,j} a^i_j,$$

subject to the restrictions

$$\mu^g_{i,i} = \mu^G_{i,i} = 1 \text{ for all } i,$$
$$\prod_{\ell=0}^{n-1} \mu^g_{i+\ell,j+\ell} = \prod_{\ell=0}^{n-1} \mu^G_{i+\ell,j+\ell} = 1, \quad \mu^G_{i,j} \mu^g_{i+1,j+1} = \mu^G_{i,j} \mu^g_{i+1,j+1},$$
$$\zeta \mu^g_{i,j+1} \lambda^X_{i,j} = \mu^g_{i,j+1} \lambda^X_{i,j+1}, \quad \zeta \mu^G_{i,j+1} \lambda^X_{i,j} = \mu^G_{i,j+1} \lambda^X_{i,j+1},$$
$$\zeta \mu^g_{i,j} \lambda^X_{i+1,j+1} = \mu^g_{i,j} \lambda^X_{i+1,j+1}, \quad \zeta \mu^G_{i,j} \lambda^X_{i+1,j+1} = \mu^G_{i,j} \lambda^X_{i+1,j+1},$$
$$\lambda^X_{i,j} \lambda^X_{i,j+1} - \zeta \lambda^X_{i,j+1} \lambda^X_{i,j} = 0,$$
$$\lambda^X_{i,j} = \lambda^X_{i,j} = 0 \text{ if either } i = j, \text{ or the arrow } a^i_{j+1} \text{ does not exist in } Q.$$

If $n ≥ 3$, then we also impose the condition $\gamma \gamma^X (1 - \zeta^{-1}) = 1$.

An action of $D(T(n))$ on the path algebra of a Type B minimal subquiver $Q$ of $K_{n,n}$ is given by

$$g \cdot b^i_j = \mu^g_{i,j} b^{i+1}_{j+1}, \quad G \cdot b^i_j = \mu^G_{i,j} b^{i+1}_{j+1},$$
$$x \cdot b^i_j = (\gamma^x) \zeta^i b^i_j - (\gamma^x) \zeta^{-i+1} \mu^g_{i,j} b^{i+1}_{j+1} + \lambda^x_{i,j} b^i_j,$$
$$X \cdot b^i_j = (\gamma^X) \zeta^{-j} b^{-j}_i - (\gamma^X) \zeta^{-i-1} \mu^G_{i,j} b^{i+1}_{j+1} + \lambda^X_{i,j} b^i_j,$$

subject to the restrictions (8.23). If $n ≥ 3$, then we also impose the condition $\gamma^X (1 - \zeta^{-1}) = 1$.

Proof. First, Proposition 8.18 gives an $D(T(n))$-action on a $\mathbb{Z}_n$-orbit of vertices; this imposes the restriction (8.20) when $n ≥ 3$. Now, consider the Type A case. Theorem 6.1 gives the expressions in (8.22), along with some restrictions on parameters, which are listed in the first two rows of restrictions in the statement of the theorem. Now we need to check that the expressions, $g \cdot a^i_j, G \cdot a^i_j, x \cdot a^i_j, X \cdot a^i_j,$ satisfy the relations of $D(T(n))$. The relations $gG = Gg, gX = Xg, xG = \zeta Gx$, imply, respectively, that

$$\mu^G_{i,j} \mu^g_{i+1,j+1} = \mu^g_{i,j} \mu^G_{i+1,j+1}, \quad \mu^g_{i,j} \lambda^X_{i+1,j+1} = \mu^g_{i,j} \lambda^X_{i,j+1}, \quad \mu^G_{i,j} \lambda^X_{i+1,j+1} = \mu^G_{i,j} \lambda^X_{i,j+1},$$
With these restrictions, it is easy to check that the relation \( xX - \zeta XX - \zeta (gG - 1) = 0 \) yields
\[
\lambda_{i,j}^X \lambda_{i,j+1}^X - \zeta \lambda_{i,j}^X \lambda_{i,j+1}^X = 0,
\]
and we are done with the Type A case. For Type B, the action is derived exactly as in the Type A case, by replacing \( a_i^* \) with \( b_i^* \), and \( \gamma \) with \( \gamma \pm \), appropriately. \(\square\)

8C. Gluing. Gluing is achieved in the following manner. Let \( Q \) be a quiver that admits an action of \( \mathbb{Z}_n \). Since the actions of \( u_q(\mathfrak{sl}_2) \) and \( D(T(n)) \) on \( kQ \) in Theorems 8.10 and 8.21 are determined by Taft actions, Theorem 7.5 and the algorithm in Section 7B applies. In other words, we can glue the actions of \( u_q(\mathfrak{sl}_2) \) or of \( D(T(n)) \) on \( kQ \) and \( \mathbb{Z}_n \)-minimal quivers. Hence, we obtain an action of \( u_q(\mathfrak{sl}_2) \) (resp. \( D(T(n)) \)) extending a Taft action on any path algebra \( kQ \), where each path algebra of a \( \mathbb{Z}_n \)-minimal component of \( Q \) admits an action of \( u_q(\mathfrak{sl}_2) \) (resp. \( D(T(n)) \)).

9. Appendix

In this appendix, we show that the relation \( x^n = 0 \) imposes no restrictions on the \( x \)-action on the vertex \( e_i \) and on the arrows \( a_i^j, b_i^j \) of \( kQ \), which is used in the proofs of Proposition 3.5 and Theorems 6.1 and 6.3. Recall that \( \zeta \) is a primitive \( n \)-th root of unity, with \( n \geq 2 \). An easy computation shows that
\[
\prod_{\ell=1}^{n} \zeta^\ell = \zeta^{\frac{n(n+1)}{2}} = \begin{cases} 1, & n \text{ odd}, \\ -1, & n \text{ even}, \end{cases} \tag{9.1}
\]
a fact we will use several times here.

This appendix uses some basic properties of symmetric polynomials and \( q \)-binomial coefficients, which we recall here. For integers \( a, b \), the \textit{complete symmetric polynomial of degree} \( a \), denoted \( h_a(x_0, \ldots, x_b) \), is defined as the sum of all monomials of total degree \( a \) in the variables \( x_0, \ldots, x_b \). We interpret the variable set as being empty when \( b < 0 \). When \( a = 0 \), we have \( h_0(x_0, \ldots, x_b) = 1 \) for any integer \( b \). When \( a \neq 0 \), we have \( h_a(x_0, \ldots, x_b) = 0 \) if \( a < 0 \) or \( b < 0 \). Also, a product \( \prod_{i=a}^b \) is taken to be 1 whenever \( b < a \). These polynomials are homogeneous, meaning that \( h_a(cx_0, \ldots, cx_b) = c^a h_a(x_0, \ldots, x_b) \) for any scalar \( c \). We have immediately from the definition that
\[
h_a(x_0, \ldots, x_{b-1}) + x_b h_{a-1}(x_0, \ldots, x_b) = h_a(x_0, \ldots, x_b) \tag{9.2}
\]
for any integers \( a, b \).

**Lemma 9.3.** For nonnegative integers \( a, b \), we have \( h_a(1, \zeta, \zeta^2, \ldots, \zeta^b) = 0 \) when \( n \) divides \( a + b \) and \( a, b \neq 0 \).
Proof. Recall that for a nonnegative integer \(k\), the corresponding \(q\)-integer is defined as \([k]_q = 1 + q + q^2 + \cdots + q^{k-1}\), and that its evaluation \([k]_{q=\zeta}\) is 0 if and only if \(n\) divides \(k\). The \(q\)-binomial coefficient \(\binom{i}{j}_q\) is defined by a factorial formula analogous to that of binomial coefficients, so the evaluation \(\binom{i}{j}_{q=\zeta}\) vanishes when \(n\) divides \(i\) and \(j \neq 0\), \(i\).

The principal specialization from [Stanley 1999, Proposition 7.8.3], with \(q = \zeta\), gives us that

\[
h_a(1, \zeta, \zeta^2, \ldots, \zeta^b) = \left[\frac{a+b}{a}\right]_{q=\zeta},
\]

so by the paragraph above this quantity vanishes when \(n\) divides \(a + b\) and \(a, b \neq 0\).

We prove a general lemma about a linear operator \(X\) acting as the element \(x\) does.

**Lemma 9.4.** Let \(V\) be a vector space, and consider a collection of vectors \(\{v_j^i\}\) in \(V\), where \(1 \leq i \leq m\) and \(1 \leq j \leq m'\). Interpret \(i\) and \(j\) modulo \(m\) and \(m'\) respectively, so that they lie inside their ranges. Let \(X\) be a linear operator acting on \(V\) such that

\[
Xv_j^i = \eta_j v_j^i + \theta_{i,j} v_{j+1}^{i+1} + \tau_{i,j} v_j^{i+1}
\]

for some scalars \(\eta_j, \theta_{i,j}, \tau_{i,j} \in \mathbb{k}\). Furthermore, assume that the scalars satisfy the relation

\[
\tau_{i+1,j+1} = \zeta \theta_{i+1,j+1} \tau_{i,j}
\]

(9.5)

for all pairs \((i, j)\). Then, for all positive integers \(k\), we have

\[
X^k(v_j^i) = \sum_{0 \leq s \leq t \leq k} \psi_{k,s,t} v_{j+t}^{i+s},
\]

where

\[
\psi_{k,s,t} = \left(\prod_{\ell=0}^{s-1} \theta_{i+\ell,j+\ell+t-s}\right)\left(\prod_{\ell=0}^{t-s-1} \tau_{i,j+\ell}\right)
\]

\[
\times h_{k-t}(\eta_j, \eta_{j+1}, \ldots, \eta_{j+t}) h_s(1, \zeta, \ldots, \zeta^{t-s}).
\]

(9.6)

Proof. We proceed by induction on \(k\). The base case of \(k = 1\) follows from simple substitution. So assume that the statement is true when \(k\) is replaced by \(k - 1\); that is, we have a formula for \(X^{k-1}(v_j^i)\) for all pairs \((i, j)\). Now, when applying \(X\) to \(X^{k-1}(v_j^i)\), the coefficient \(\psi_{k,s,t}\) is the sum of contributions from three terms:

- \(X(\psi_{k-1,s,t} v_{j+t}^{i+s})\) contributes \(\eta_{j+t} \psi_{k-1,s,t}\);
- \(X(\psi_{k-1,s-1,t-1} v_{j+t}^{i+s-1})\) contributes \(\theta_{i+s-1,j+t-1} \psi_{k-1,s-1,t-1}\);
- \(X(\psi_{k-1,s,t-1} v_{j+t}^{i+s})\) contributes \(\tau_{i+s,j+t-1} \psi_{k-1,s,t-1}\).

So we obtain
\[ \psi_{k,s,t} = \left( \prod_{\ell=0}^{s-1} \theta_{i+\ell,j+\ell+t-s} \right) \left( \prod_{\ell=0}^{t-s-1} \tau_{i,j+\ell} \right) \times \left[ \eta_{j+t} h_{k-1-t}(\eta_j, \ldots, \eta_{j+t}) h_s(1, \zeta, \ldots, \zeta^{t-s}) \\
+ h_{k-t}(\eta_j, \ldots, \eta_{j+t-1}) h_s(1, \zeta, \ldots, \zeta^{t-s}) \right] \\
+ \tau_{i+s,j+t-1} \left( \prod_{\ell=0}^{s-1} \theta_{i+\ell,j+\ell+t-s-1} \right) \left( \prod_{\ell=0}^{t-s-2} \tau_{i,j+\ell} \right) \times h_{k-t}(\eta_j, \ldots, \eta_{j+t-1}) h_s(1, \zeta, \ldots, \zeta^{t-s-1}). \tag{9.7} \]

To simplify the last summand, we use (9.5) to compute

\[ \tau_{i+s,j+t-1} \prod_{\ell=0}^{s-1} \theta_{i+\ell,j+\ell+t-1-s} = (\tau_{i+s,j+t-1}) \theta_{i+s-1,j+t-2} \theta_{i+s-2,j+t-3} \cdots \theta_{i,j+t-1-s} = \xi^{s} (\tau_{i,j+t-s-1}) \prod_{\ell=0}^{s-1} \theta_{i+\ell,j+\ell+t-s}. \tag{9.8} \]

By (9.7) and (9.8), we now have

\[ \psi_{k,s,t} = \left( \prod_{\ell=0}^{s-1} \theta_{i+\ell,j+\ell+t-s} \right) \left( \prod_{\ell=0}^{t-s-1} \tau_{i,j+\ell} \right) \times \left[ \eta_{j+t} h_{k-1-t}(\eta_j, \ldots, \eta_{j+t}) h_s(1, \zeta, \ldots, \zeta^{t-s}) \\
+ h_{k-t}(\eta_j, \ldots, \eta_{j+t-1}) h_s(1, \zeta, \ldots, \zeta^{t-s}) \\
+ h_{k-t}(\eta_j, \ldots, \eta_{j+t-1}) \xi^{s} h_s(1, \zeta, \ldots, \zeta^{t-s-1}) \right]. \tag{9.9} \]

The factor of \( \xi^{s} \) in the last term can be absorbed into the (homogeneous) polynomial \( h_s \) to make the sum of the last two terms equal to

\[ h_{k-t}(\eta_j, \ldots, \eta_{j+t-1})(h_{s-1}(1, \zeta, \ldots, \zeta^{t-s}) + h_s(\xi, \xi^2, \ldots, \xi^{t-s})). \]

Now, taking \( x_{t-s-i} = \xi^i \), (9.2) implies that the two terms in the parenthesis simplify to \( h_s(1, \zeta, \ldots, \zeta^{t-s}) \). Thus, (9.9) yields

\[ \psi_{k,s,t} = \left( \prod_{\ell=0}^{s-1} \theta_{i+\ell,j+\ell+t-s} \right) \left( \prod_{\ell=0}^{t-s-1} \tau_{i,j+\ell} \right) h_s(1, \zeta, \ldots, \zeta^{t-s}) \times \left[ \eta_{j+t} h_{k-1-t}(\eta_j, \ldots, \eta_{j+t}) + h_{k-t}(\eta_j, \ldots, \eta_{j+t-1}) \right]. \]
Again applying (9.2), we find that (9.6) holds, and the induction is complete. □

We can apply this result to show that \( x^n \) acts by 0 in the following cases.

**Lemma 9.10.** In the notation of Proposition 3.5, we have \( x^n \cdot e_i = 0 \).

**Proof.** Fix \( i \). We will apply Lemma 9.4 in the case \( k = n, m = m' \), with \( X = x \) and \( v^j_i = e_i \) for \( 1 \leq i \leq m \). The \( v^j_i \) do not play a role in the proof for \( i \neq j \), so we set \( \theta_{i,j} = \tau_{i,j} = 0 \) when \( i \neq j \) and assume \( i = j \) for the remainder of the proof. From the proof of Proposition 3.5, we have \( \eta_i = \gamma \xi^i, \theta_{i,i} = -\gamma \xi^{i+1}, \) and \( \tau_{i,i} = 0 \) for all \( i \). So, (9.5) holds and we can apply Lemma 9.4. Making these substitutions into (9.6), we immediately see that \( \psi_{n,s,t} = 0 \) unless \( s = t \), by considering the \( \tau \) factor. In case \( s = t \), the \( \tau \) factor and the last factor are equal to 1. This gives

\[
\psi_{n,s,s} = \left( \prod_{\ell=0}^{s-1} \theta_{i+\ell,i+\ell} \right) h_{n-s}(\eta_i, \eta_{i+1}, \ldots, \eta_{i+s}).
\]

Since \( h_{n-s}(\eta_i, \eta_{i+1}, \ldots, \eta_{i+s}) \) is a scalar multiple of \( h_{n-s}(1, \xi, \ldots, \xi^s) \), this vanishes by Lemma 9.3 unless \( s = 0 \) or \( s = n \).

So we have \( x^n \cdot e_i = (\psi_{n,0,0} + \psi_{n,n,n}) e_i \). These are easily computed by substitution to be

\[
\psi_{n,0,0} = h_n(\eta_i) = (\eta_i)^n = \gamma^n
\]

\[
\psi_{n,n,n} = \prod_{\ell=0}^{n-1} \theta_{i+\ell,i+\ell} = (-1)^n \gamma^n \prod_{\ell=0}^{n-1} \xi^{i+\ell+1} = (-1)^n \gamma^n \prod_{\ell=0}^{n-1} \xi^{\ell} = -\gamma^n,
\]

where the last equality invokes Equation (9.1). Thus we have shown that \( x^n \cdot e_i = 0 \). □

**Lemma 9.11.** In the notation of Theorem 6.1, we have \( x^n \cdot a^j_i = 0 \) for all \( (i, j) \).

**Proof.** We will apply Lemma 9.4 in the case \( k = n \) with \( X = x \) and \( v^j_i = a^j_i \). From the proof of Theorem 6.1, we have \( \eta_j = \gamma \xi^j, \theta_{i,j} = -\gamma \mu_{i,j} \xi^{i+1}, \) and \( \tau_{i,j} = \lambda_{i,j} \), and that these satisfy (9.5). Now making these substitutions into (9.6), we see that \( \psi_{n,s,i} \) is a scalar times

\[
h_s(1, \xi, \ldots, \xi^{t-s}) h_{n-t}(1, \xi, \ldots, \xi^t).
\]

By Lemma 9.3, the second factor vanishes except when \( t = 0 \) or \( t = n \). Then, again by Lemma 9.3, the first factor vanishes except when \( s = 0 \) or \( s = t = n \). Therefore, we need only consider the cases when \( (s, t) \) equals \( (0, 0) \), \( (0, n) \), and \( (n, n) \), in which these factors both equal 1. Notice that we have in all of these cases, \( a^j_{i+s} = a^j_i \), because the indices for arrows are taken modulo \( n \). So it suffices to show
that $\psi_{n,0,0} + \psi_{n,0,n} + \psi_{n,n,n} = 0$. Substitute and compute, omitting factors equal to 1. We get

\[
\psi_{n,0,0} = h_n(\eta_j) = (\eta_j)^n = \gamma^n \quad \text{and} \quad \psi_{n,0,n} = \prod_{\ell=0}^{n-1} \tau_{i,j+\ell} = 0.
\]

The latter follows by the second condition of Theorem 6.1, because there will be a factor with $i = j + \ell$ modulo $n$. Further,

\[
\psi_{n,n,n} = \prod_{\ell=0}^{n-1} \theta_{i+\ell,j+\ell} = (-1)^n \gamma^n \prod_{\ell=0}^{n-1} \mu_{i+\ell,j+\ell} \xi^{\ell+\ell+1} = (-1)^n \gamma^n \prod_{\ell=1}^{n} \xi^\ell = -\gamma^n,
\]

where the penultimate equality uses (4.2) and the last equality invokes Equation (9.1). Thus we have shown that $x^n \cdot a^i_j = 0$. □

**Lemma 9.12.** In the notation of Theorem 6.3, we have $x^n \cdot b^i_j = 0$ for all $(i, j)$.

**Proof.** Again, we apply Lemma 9.4 with $k = n$, $X = x$, and $v^i_j = b^i_j$. The proof is the same as for Lemma 9.11, except in this case we find that

\[
\psi_{n,0,0} = \gamma^n_-, \quad \psi_{n,n,n} = -\gamma^n_+, \quad \text{and} \quad \psi_{n,0,n} = \prod_{\ell=0}^{n-1} \tau_{i,j+\ell} = \prod_{\ell=0}^{n-1} \lambda_{i,j+\ell}.
\]

Thus, we see that $x^n$ acts by 0 if and only if $(\gamma^+_n)^n = (\gamma^-_n)^n + \prod_{\ell=0}^{n-1} \lambda_{i,j+\ell}$, which is the condition assumed on the scalars in Theorem 6.3. □

**Acknowledgments**

We thank Steven Sam for directing us to the reference for principal specializations of symmetric polynomials. We also thank Mio Iovanov for helping us with the proof of Lemma 2.5, and Susan Montgomery for suggesting the extensions of our results in Section 8. We are grateful to the referee for careful comments and suggestions. Walton was supported by the National Science Foundation: NSF-grant DMS-1550306.

**References**


154 Ryan Kinser and Chelsea Walton


Communicated by Susan Montgomery

Received 2015-01-06 Revised 2015-07-30 Accepted 2015-09-19

ryan-kinser@uiowa.edu Department of Mathematics, University of Iowa, Iowa City, IA 52242, United States

notlaw@temple.edu Department of Mathematics, Temple University, Philadelphia, PA 19122, United States

mathematical sciences publishers
On the image of the Galois representation associated to a non-CM Hida family

Jaclyn Lang

Fix a prime $p > 2$. Let $\rho : \text{Gal}(\overline{Q}/Q) \to \text{GL}_2(\mathcal{O})$ be the Galois representation coming from a non-CM irreducible component $\mathcal{O}$ of Hida’s $p$-ordinary Hecke algebra. Assume the residual representation $\bar{\rho}$ is absolutely irreducible. Under a minor technical condition we identify a subring $\mathcal{O}_0$ of $\mathcal{O}$ containing $\mathbb{Z}_p[[T]]$ such that the image of $\rho$ is large with respect to $\mathcal{O}_0$. That is, $\text{Im} \rho$ contains $\ker(\text{SL}_2(\mathcal{O}_0) \to \text{SL}_2(\mathcal{O}_0/\mathfrak{a}))$ for some nonzero $\mathcal{O}_0$-ideal $\mathfrak{a}$. This paper builds on recent work of Hida who showed that the image of such a Galois representation is large with respect to $\mathbb{Z}_p[[T]]$. Our result is an $\mathcal{O}$-adic analogue of the description of the image of the Galois representation attached to a non-CM classical modular form obtained by Ribet and Momose in the 1980s.

1. Introduction

A Hida family $F$ that is an eigenform and has coefficients in a domain $\mathcal{O}$ has an associated Galois representation $\rho_F : \text{Gal}(\overline{Q}/Q) \to \text{GL}_2(Q(\mathcal{O}))$, where $Q(\mathcal{O})$ is the field of fractions of $\mathcal{O}$. A fundamental problem is to understand the image of such a representation. One expects the image to be “large” in an appropriate sense, so long as $F$ does not have any extra symmetries; that is, as long as $F$ does not have CM. (In the CM case there is a nontrivial character $\eta$ such that $\rho_F \cong \rho_F \otimes \eta$. This forces the image of $\rho_F$ to be “small”.) This notion of “largeness” can be defined relative to any subring $\mathcal{O}_0$ of $\mathcal{O}$, and one can then ask whether $\text{Im} \rho_F$ is large with respect to $\mathcal{O}_0$. Even when $F$ does not have CM it might happen that there is an automorphism $\sigma$ of $\mathcal{O}$ and a nontrivial character $\eta$ such that $\rho_F^\sigma \cong \rho_F \otimes \eta$. Such automorphisms, called conjugate self-twists of $F$, can be thought of as weak symmetries of $F$. In this paper we explain how conjugate self-twists constrict the image of $\rho_F$. In particular, let $\mathcal{O}_0$ be the subring of $\mathcal{O}$ fixed by all conjugate self-twists of $F$. Our main result is that $\text{Im} \rho_F$ is “large” with respect to $\mathcal{O}_0$.

The study of the image of the Galois representation attached to a modular form,
and showing that it is large in the absence of CM, was first carried out by Serre [1973] and Swinnerton-Dyer [1973]. They studied the Galois representation attached to a modular form of level 1 with integral coefficients. Ribet [1980; 1985] and Momose [1981] generalized the work of Serre and Swinnerton-Dyer to cover all Galois representations coming from classical modular forms. Ribet’s work dealt with the weight two case, and Momose proved the general case. The main theorem in this paper is an analogue of their results in the $\ell$-adic setting. In fact, their work is a key input for our proof.

Shortly after Hida constructed the representations $\rho_F$, Mazur and Wiles [1986] showed that if $I = \mathbb{Z}_p[[T]]$ and the image of the residual representation $\bar{\rho}_F$ contains $\text{SL}_2(\mathbb{F}_p)$ then $\text{Im} \rho_F$ contains $\text{SL}_2(\mathbb{Z}_p[[T]])$. Under the assumptions that $I$ is a power series ring in one variable and the image of the residual representation $\bar{\rho}_F$ contains $\text{SL}_2(\mathbb{F}_p)$, our main result was proved by Fischman [2002]. Fischman’s work is the only previous work that considers the effect of conjugate self-twists on $\text{Im} \rho_F$. Hida [2015] has shown under some technical hypotheses that if $F$ does not have CM then $\text{Im} \rho_F$ is large with respect to the ring $\mathbb{Z}_p[[T]]$, even when $I \supseteq \mathbb{Z}_p[[T]]$. The methods he developed play an important role in this paper. The local behavior of $\rho_F$ at $p$ was studied by Ghate and Vatsal [2004] and later by Hida [2013]. They showed, under some assumptions later removed by Zhao [2014], that $\rho_F|_{D_p}$ is indecomposable, where $D_p$ denotes the decomposition group at $p$ in $G_{\mathbb{Q}}$. We will make use of this result later. Finally, Hida and Tilouine [2015] showed that certain $\text{GSp}_4$-representations associated to Siegel modular forms have large image.

Our result is the first to describe the effect of conjugate self-twists on the image of $\rho_F$ without any assumptions on $I$ and without assuming that the image of $\bar{\rho}_F$ contains $\text{SL}_2(\mathbb{F}_p)$. We do need an assumption on $\bar{\rho}_F$, namely that $\bar{\rho}_F$ is absolutely irreducible and another small technical condition, but this is much weaker than assuming $\text{Im} \bar{\rho}_F \supseteq \text{SL}_2(\mathbb{F}_p)$.

2. Main theorems and structure of paper

We begin by fixing notation that will be in place throughout the paper. Let $p > 2$ be prime. Fix algebraic closures $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ and $\overline{\mathbb{Q}}_p$ of $\mathbb{Q}_p$ as well as an embedding $i_p : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$. Let $G_{\mathbb{Q}} = \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ be the absolute Galois group of $\mathbb{Q}$. Let $\mathbb{Z}^+$ denote the set of positive integers. Fix $N_0 \in \mathbb{Z}^+$ prime to $p$; it will serve as our tame level. Let $N = N_0p^r$ for some fixed $r \in \mathbb{Z}^+$. Fix a Dirichlet character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \overline{\mathbb{Q}}^\times$ which will serve as our nebentypus. Let $\chi_1$ be the product of $\chi|_{(\mathbb{Z}/N_0\mathbb{Z})^\times}$ with the tame $p$-part of $\chi$, and write $c(\chi)$ for the conductor of $\chi$. During the proof of the main theorem we will assume that the order of $\chi$ is a power of 2 and that $2c(\chi)|N$. The fact that we can assume these restrictions on $\chi$ for the purpose of demonstrating $I_0$-fullness is shown in Proposition 3.9.
For a valuation ring $W$ over $\mathbb{Z}_p$, let $\Lambda_W = W[[T]]$. Let $\mathbb{Z}_p[\chi]$ be the extension of $\mathbb{Z}_p$ generated by the values of $\chi$. When $W = \mathbb{Z}_p[\chi]$ we write $\Lambda_\chi$ for $\Lambda_W$. When $W = \mathbb{Z}_p$ then we let $\Lambda = \Lambda_{\mathbb{Z}_p}$. For any valuation ring $W$ over $\mathbb{Z}_p$, an arithmetic prime of $\Lambda_W$ is a prime ideal of the form

$$P_{k,\varepsilon} := (1 + T - \varepsilon(1+p)(1+p)^k)$$

for an integer $k \geq 2$ and character $\varepsilon : 1 + p\mathbb{Z}_p \to W^\times$ of $p$-power order. We shall write $r(\varepsilon)$ for the nonnegative integer such that $p^{r(\varepsilon)}$ is the order of $\varepsilon$. If $R$ is a finite extension of $\Lambda_W$, then we say a prime of $R$ is arithmetic if it lies over an arithmetic prime of $\Lambda_W$.

For a Dirichlet character $\psi : (\mathbb{Z}/M\mathbb{Z})^\times \to \overline{\mathbb{Q}}^\times$, let $S_k(\Gamma_0(M), \psi)$ be the space of classical cusp forms of weight $k$, level $\Gamma_0(M)$, and nebentypus $\psi$. Let $h_k(\Gamma_0(M), \psi)$ be the Hecke algebra of $S_k(\Gamma_0(M), \psi)$, and let $h_k^{ord}(\Gamma_0(M), \psi)$ denote the $p$-ordinary Hecke algebra. Let $\omega$ be the $p$-adic Teichmüller character. We can describe Hida’s big $p$-ordinary Hecke algebra $h^{ord}(N, \chi; \Lambda_\chi)$ as follows [Hida 2015]. It is the unique $\Lambda_\chi$-algebra that is

1. free of finite rank over $\Lambda_\chi$,
2. equipped with Hecke operators $T(n)$ for all $n \in \mathbb{Z}^+$, and
3. satisfies the following specialization property: for every arithmetic prime $P_{k,\varepsilon}$ of $\Lambda_\chi$ there is an isomorphism

$$h^{ord}(N, \chi; \Lambda_\chi)/P_{k,\varepsilon} h^{ord}(N, \chi; \Lambda_\chi) \cong h_k^{ord}(\Gamma_0(Np^{r(\varepsilon)}), \chi_1 \varepsilon \omega^{-k})$$

that sends $T(n)$ to $T(n)$ for all $n \in \mathbb{Z}^+$.

For a commutative ring $R$, we use $Q(R)$ to denote the total ring of fractions of $R$. Hida [1986a] has shown that there is a Galois representation

$$\rho_{N_0,\chi} : G_\mathbb{Q} \to \text{GL}_2(Q(h^{ord}(N_0, \chi; \Lambda_\chi)))$$

that is unramified outside $N$ and satisfies $\text{tr} \rho_{N_0,\chi}(\text{Frob}_\ell) = T(\ell)$ for all primes $\ell$ not dividing $N$. Let $\text{Spec} \ p$ be an irreducible component of $\text{Spec} h^{ord}(N_0, \chi; \Lambda_\chi)$. Assume further that $\mathfrak{p}$ is primitive in the sense of [Hida 1986b, Section 3]. Let $\lambda_F : h^{ord}(N_0, \chi; \Lambda_\chi) \to \mathfrak{p}$ be the natural $\Lambda_\chi$-algebra homomorphism coming from the inclusion of spectra. By viewing

$$Q(h^{ord}(N_0, \chi; \Lambda_\chi)) = h^{ord}(N_0, \chi; \Lambda_\chi) \otimes_{\Lambda_\chi} Q(\Lambda_\chi)$$

and composing $\rho_{N_0,\chi}$ with $\lambda_F \otimes 1$ we obtain a Galois representation

$$\rho_F : G_\mathbb{Q} \to \text{GL}_2(Q(\mathfrak{p}))$$
that is unramified outside \( N \) and satisfies
\[
\text{tr} \, \rho_F(\text{Frob}_\ell) = \lambda_F(T(\ell))
\]
for all primes \( \ell \) not dividing \( N \).

Henceforth for any \( n \in \mathbb{Z}^+ \) we shall let \( a(n, F) \) denote \( \lambda_F(T(n)) \). Let \( F \) be the formal power series in \( q \) given by
\[
F = \sum_{n=1}^{\infty} a(n, F)q^n.
\]

Let \( \mathfrak{l}' = \Lambda_X\{a(\ell, F) : \ell \nmid N\} \) which is an order in \( Q(\mathfrak{l}) \) since \( F \) is primitive. We shall consider the Hida family \( F \) and the associated ring \( \mathfrak{l}' \) to be fixed throughout the paper. For a local ring \( R \) we will use \( \mathfrak{m}_R \) to denote the unique maximal ideal of \( R \). Let \( \mathbb{F} := \mathfrak{l}'/\mathfrak{m}_R \), the residue field of \( \mathfrak{l}' \). We exclusively use the letter \( \mathfrak{P} \) to denote a prime of \( \mathfrak{l} \), and \( \mathfrak{P}' \) shall always denote \( \mathfrak{P} \cap \mathfrak{l}' \). Conversely, we exclusively use \( \mathfrak{P}' \) to denote a prime of \( \mathfrak{l}' \) in which case we are implicitly fixing a prime \( \mathfrak{P} \) of \( \mathfrak{l} \) lying over \( \mathfrak{P}' \).

If \( \mathfrak{P} \) is a height one prime of \( \mathfrak{l} \) then we write \( f_{\mathfrak{P}} \) for the \( p \)-adic modular form obtained by reducing the coefficients of \( F \) modulo \( \mathfrak{P} \). In particular, if \( \mathfrak{P} \) is an arithmetic prime lying over \( P_{k, e} \) then \( f_{\mathfrak{P}} \in S_k(\Gamma_0(Np^r(\ell)), \varepsilon \chi_1 \omega^{-k}) \).

Recall that Hida [1986a] has shown that there is a well defined residual representation \( \tilde{\rho}_F : G_\mathbb{Q} \to \text{GL}_2(\mathfrak{l}/\mathfrak{m}_\mathfrak{l}) \) of \( \rho_F \). Throughout this paper we impose the following assumption.

Assume that \( \tilde{\rho}_F \) is absolutely irreducible. \( \text{(abs)} \)

By the Chebotarev density theorem, we see that \( \text{tr} \, \tilde{\rho}_F \) is valued in \( \mathbb{F} \). Under \( \text{(abs)} \) we may use pseudorepresentations to find a \( \text{GL}_2(\mathfrak{l}') \)-valued representation that is isomorphic to \( \rho_F \) over \( Q(\mathfrak{l}) \). Thus we may (and do) assume that \( \rho_F \) takes values in \( \text{GL}_2(\mathfrak{l}') \).

**Definition 2.1.** Let \( g = \sum_{n=1}^{\infty} a(n, g)q^n \) be either a classical Hecke eigenform or a Hida family of such forms. Let \( K \) be the field generated by \( \{a(n, g) : n \in \mathbb{Z}^+\} \) over either \( \mathbb{Q} \) in the classical case or \( Q(\Lambda_X) \) in the \( \Lambda_X \)-adic case. We say a pair \( (\sigma, \eta_\sigma) \) is a **conjugate self-twist** of \( g \) if \( \eta_\sigma \) is a Dirichlet character, \( \sigma \) is an automorphism of \( K \), and
\[
\sigma(a(\ell, g)) = \eta_\sigma(\ell)a(\ell, g)
\]
for all but finitely many primes \( \ell \). If there is a nontrivial character \( \eta \) such that \((1, \eta)\) is a conjugate self-twist of \( g \), then we say that \( g \) has **complex multiplication** or CM. Otherwise, \( g \) does not have CM.
If a modular form does not have CM then a conjugate self-twist is uniquely determined by the automorphism.

We shall always assume that our fixed Hida family $F$ does not have CM. Let

$$\Gamma = \{ \sigma \in \text{Aut}(\mathbb{Q}(\mathfrak{f})) : \sigma \text{ is a conjugate self-twist of } F \}.$$ 

Under the assumption (abs) it follows from a lemma of Carayol and Serre [Hida 2000b, Proposition 2.13] that if $\sigma \in \Gamma$ then $\rho_F^\sigma \cong \rho_F \otimes \eta_\sigma$ over $\mathfrak{f}$. As $\rho_F$ is unramified outside $N$ we see that in fact $\sigma(a(\ell, F)) = \eta_\sigma(\ell)a(\ell, F)$ for all primes $\ell$ not dividing $N$. Therefore $\sigma$ restricts to an automorphism of $I'$. Let $I_0 = (\mathfrak{f}^\prime)^\Gamma$.

Define

$$H_0 := \bigcap_{\sigma \in \Gamma} \ker \eta_\sigma$$

and

$$H := H_0 \cap \ker(\det(\bar{\rho}_F)).$$

These open normal subgroups of $G_\mathbb{Q}$ play an important role in our proof.

For a commutative ring $B$ and ideal $b$ of $B$, write

$$\Gamma_B(b) := \ker(\text{SL}_2(B) \to \text{SL}_2(B/b)).$$

We call $\Gamma_B(b)$ a congruence subgroup of $\text{GL}_2(B)$ if $b \neq 0$. We can now define what we mean when we say a representation is “large” with respect to a ring.

**Definition 2.2.** Let $G$ be a group, $A$ a commutative ring, and $r : G \to \text{GL}_2(A)$ a representation. For a subring $B$ of $A$, we say that $r$ is $B$-full if there is some $\gamma \in \text{GL}_2(A)$ such that $\gamma(\text{Im } r)\gamma^{-1}$ contains a congruence subgroup of $\text{GL}_2(B)$.

Let $D_p$ be the decomposition group at $p$ in $G_\mathbb{Q}$. That is, $D_p$ is the image of $G_{\mathbb{Q}_p} := \text{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ under the embedding $G_{\mathbb{Q}_p} \hookrightarrow G_\mathbb{Q}$ induced by $\iota_p$. Recall that over $\mathbb{Q}(\mathfrak{f})$ the local representation $\rho_F|_{D_p}$ is isomorphic to $\left( \begin{smallmatrix} \epsilon & \delta \\ 0 & \eta \end{smallmatrix} \right)$ [Hida 2000a, Theorem 4.3.2]. Let $\bar{\epsilon}$ and $\bar{\delta}$ denote the residual characters of $\epsilon$ and $\delta$, respectively.

**Definition 2.3.** For any open subgroup $G_0 \leq G_\mathbb{Q}$ we say that $\rho_F$ is $G_0$-regular if $\bar{\epsilon}|_{D_p \cap G_0} \neq \bar{\delta}|_{D_p \cap G_0}$.

The main result of this paper is the following.

**Theorem 2.4.** Assume $p > 2$ and let $F$ be a primitive non-CM $p$-adic Hida family. Assume $|F| \neq 3$ and that the residual representation $\bar{\rho}_F$ is absolutely irreducible and $H_0$-regular. Then $\rho_F$ is $I_0$-full.

The strategy of the proof is to exploit the results of Ribet [1980; 1985] and Momose [1981]. Since an arithmetic specialization of a non-CM Hida family cannot be CM, their work implies that if $\mathfrak{P}$ is an arithmetic prime of $\mathfrak{f}$ then there is a certain subring $\mathcal{O} \subseteq \mathfrak{f}/\mathfrak{P}$ for which $\rho_F$ mod $\mathfrak{P}$ is $\mathcal{O}$-full. To connect
their ring $\mathcal{O}$ with $\mathfrak{l}_0$. In Section 6 we show that $Q(\mathcal{O}) = Q(\mathfrak{l}_0 / \mathfrak{Q})$, where $\mathfrak{Q} = \mathfrak{l}_0 \cap \mathfrak{P}'$. The proof that $Q(\mathcal{O}) = Q(\mathfrak{l}_0 / \mathfrak{Q})$ relies on establishing a relationship between conjugate self-twists of $F$ and conjugate self-twists of the arithmetic specializations of $F$. As this may be of independent interest we state the result here.

**Theorem 2.5.** Let $\mathfrak{P}$ be an arithmetic prime of $\mathfrak{l}$ and $\sigma$ be a conjugate self-twist of $f_{\mathfrak{P}}$ that is also an automorphism of the local field $\mathbb{Q}_p((a(n, f_{\mathfrak{P}}) : n \in \mathbb{Z}^+))$. Then $\sigma$ can be lifted to $\tilde{\sigma} \in \Gamma$ such that $\tilde{\sigma}(\mathfrak{P}') = \mathfrak{P}'$, where $\mathfrak{P}' = \mathfrak{P} \cap \mathfrak{l}'$.

The proof, in Section 3, uses a combination of abstract deformation theory and automorphic techniques. Deformation theory is used to lift $\sigma$ to an automorphism of the universal deformation ring of $\bar{\rho}_F$. Then we use automorphic methods to show that this lift preserves the irreducible component $\text{Spec} \mathfrak{l}$.

The key technical input is that $h_{\text{ord}}(N, \chi; \Lambda, \chi)$ is étale over arithmetic points of $\Lambda$.

The remainder of the paper consists of a series of reduction steps that allow us to deduce our theorem from the aforementioned results of Ribet and Momose. Our methods make it convenient to modify $\rho_F$ to a related representation $\rho : H \to \text{SL}_2(\mathfrak{l}_0)$ and show that $\rho$ is $\mathfrak{l}_0$-full. We axiomatize the properties of $\rho$ at the beginning of Section 4 and use $\rho$ in the next three sections to prove Theorem 2.4. Then in Section 7 we explain how to show the existence of $\rho$ with the desired properties.

The task of showing that $\rho$ is $\mathfrak{l}_0$-full is done in three steps. In Section 4 we consider the projection of $\text{Im} \rho$ to $\prod_{Q|P} \text{SL}_2(\mathfrak{l}_0 / \mathfrak{Q})$, where $P$ is an arithmetic prime of $\Lambda$ and $Q$ runs over all primes of $\mathfrak{l}_0$ lying over $P$. We show that if the image of $\text{Im} \rho$ in $\prod_{Q|P} \text{SL}_2(\mathfrak{l}_0 / \mathfrak{Q})$ is open then $\rho$ is $\mathfrak{l}_0$-full. This uses Pink’s theory of Lie algebras for $p$-profinite subgroups of $\text{SL}_2$ over $p$-profinite semilocal rings [Pink 1993] and the related techniques developed by Hida [2015].

In Section 5 we show that if the image of $\text{Im} \rho$ in $\text{SL}_2(\mathfrak{l}_0 / \mathfrak{Q})$ is $\mathfrak{l}_0 / \mathfrak{Q}$-full for all primes $Q$ of $\mathfrak{l}_0$ lying over $P$, then the image of $\text{Im} \rho$ is indeed open in $\prod_{Q|P} \text{SL}_2(\mathfrak{l}_0 / \mathfrak{Q})$. The argument is by contradiction and uses Goursat’s lemma. It was inspired by an argument of Ribet [1975]. This is the only section where we make use of the assumption that $|F| \neq 3$.

The final step showing that the image of $\text{Im} \rho$ in $\text{SL}_2(\mathfrak{l}_0 / \mathfrak{Q})$ is $\mathfrak{l}_0 / \mathfrak{Q}$-full for every $Q$ lying over $P$ is done in Section 6. The key input is Theorem 2.5 from Section 3 together with the work of Ribet and Momose on the image of the Galois representation associated to a non-CM classical modular form. We give a brief exposition of their work and a precise statement of their result at the beginning of Section 6. We reiterate the structure of the proof of Theorem 2.4 at the end of Section 6.
3. Lifting twists

Let \( \mathfrak{P}_1 \) and \( \mathfrak{P}_2 \) be (not necessarily distinct) arithmetic primes of \( \mathfrak{l} \), and let \( \mathfrak{P}_i' = \mathfrak{P}_i \cap \mathfrak{l}' \). We shall often view \( \mathfrak{P}_i \) as a geometric point in \( \text{Spec}(\mathfrak{l})(\overline{\mathbb{Q}}_p) \). Suppose there is an isomorphism \( \sigma : \mathfrak{l}/\mathfrak{P}_1 \cong \mathfrak{l}/\mathfrak{P}_2 \) and a Dirichlet character \( \eta : G_{\mathbb{Q}} \to \mathbb{Q}(\mathfrak{l}/\mathfrak{P}_2)^\times \) such that

\[
\sigma(a(\ell, f_{\mathfrak{P}_1})) = \eta(\ell) a(\ell, f_{\mathfrak{P}_2})
\]

for all primes \( \ell \) not dividing \( N \). We may (and do) assume without loss of generality that \( \eta \) is primitive since the above relation holds even when \( \eta \) is replaced by its primitive character. In this section we show that \( \sigma \) can be lifted to a conjugate self-twist of \( \mathfrak{F} \).

**Theorem 3.1.** Assume that \( \eta \) takes values in \( \mathbb{Z}_p[\chi] \) and that the order of \( \chi \) is a power of 2. If \( \eta \) is ramified at 2, assume further that \( 2c(\chi)|N \). Then there is an automorphism \( \tilde{\sigma} : \mathfrak{l}' \to \mathfrak{l}' \) such that

\[
\tilde{\sigma}(a(\ell, F)) = \eta(\ell) a(\ell, F)
\]

for all but finitely many primes \( \ell \) and \( \sigma \circ \mathfrak{P}_1' = \mathfrak{P}_2' \circ \tilde{\sigma} \). In particular, \( \mathfrak{P}_1' \) and \( \mathfrak{P}_2' \) necessarily lie over the same prime of \( \mathfrak{l}_0 \).

**Remark.** The condition that the order of \( \chi \) be a power of 2 looks restrictive. However in Proposition 3.9 we show that for the purpose of proving \( \mathfrak{l}_0 \)-fullness we may replace \( F \) with a family whose nebentypus has order a power of 2. The same proposition shows that the condition that \( 2c(\chi)|N \) is not restrictive when proving \( \mathfrak{l}_0 \)-fullness.

There are two steps in the proof of Theorem 3.1. First we use abstract deformation theory to construct a lift \( \Sigma \) of \( \sigma \) to the universal deformation ring of \( \bar{\rho}_F \) (or some base change of that ring). This allows us to show that \( \eta \) is necessarily quadratic. Then we show that the induced map on spectra \( \Sigma^* \) sends the irreducible component \( \text{Spec} \mathfrak{l}' \) to another modular component of the universal deformation ring. Since \( \sigma \) is an isomorphism between \( \mathfrak{l}/\mathfrak{P}_1 \) and \( \mathfrak{l}/\mathfrak{P}_2 \), it follows that the arithmetic point \( \mathfrak{P}_1' \) lies on both \( \text{Spec} \mathfrak{l}' \) and \( \Sigma^*(\text{Spec} \mathfrak{l}') \). Since the Hecke algebra is étale over arithmetic points of \( \Lambda \), it follows that \( \Sigma^*(\text{Spec} \mathfrak{l}') = \text{Spec} \mathfrak{l}' \) and hence \( \Sigma \) descends to the desired automorphism of \( \mathfrak{l}' \).

**Lifting \( \sigma \) to the universal deformation ring.** Let \( W \) be the ring of Witt vectors of \( \mathbb{F} \). Let \( \mathbb{Q}^N \) be the maximal subfield of \( \overline{\mathbb{Q}} \) unramified outside \( N \) and infinity, and let \( G_{\mathbb{Q}}^N := \text{Gal}(\mathbb{Q}^N/\mathbb{Q}) \). Note that \( \rho_F \) factors through \( G_{\mathbb{Q}}^N \). For the remainder of this section we shall consider \( G_{\mathbb{Q}}^N \) to be the domain of \( \rho_F \) and \( \bar{\rho}_F \).

We set up the notation for deformation theory. For our purposes universal deformation rings of pseudorepresentations are sufficient. However, since we are
assuming that \( \tilde{\rho}_F \) is absolutely irreducible, we use universal deformation rings of representations to avoid introducing extra notation for pseudorepresentations.

Let \( C \) denote the category of complete local \( p \)-profinite \( W \)-algebras with residue field \( F \). Let \( \tilde{\pi} : G^N_Q \to GL_n(\bar{F}) \) be an absolutely irreducible representation. We say an object \( R_{\tilde{\pi}} \in C \) and representation \( \tilde{\pi}^{\text{univ}} : G^N_Q \to GL_n(R_{\tilde{\pi}}) \) is a universal couple for \( \tilde{\pi} \) if: \( \tilde{\pi}^{\text{univ}} \mod m_{R_{\tilde{\pi}}} \cong \tilde{\pi} \) and for every \( A \in C \) and representation \( r : G^N_Q \to GL_n(A) \) such that \( r \mod m_A \cong \tilde{\pi} \), there exists a unique \( W \)-algebra homomorphism \( \alpha(r) : R_{\tilde{\pi}} \to A \) such that \( r \cong \alpha(r) \circ \tilde{\pi}^{\text{univ}} \). Mazur [1989] proved that a universal couple always exists (and is unique) when \( \tilde{\pi} \) is absolutely irreducible. Since \( \eta \) takes values in \( \mathbb{Z}_p[\chi] \) which may not be contained in \( W \), we need to extend scalars. Let \( \mathcal{O} = W[\eta] \). We recommend the reader assume \( \mathcal{O} = W \) on the first read. In fact, in Proposition 3.4 we will use deformation theory to conclude that \( \eta \) is quadratic, but we cannot assume that from the start. For a commutative \( W \)-algebra \( A \), let \( \mathcal{O}A := \mathcal{O} \otimes_W A \). It will be important that we are tensoring on the left by \( \mathcal{O} \) as we will sometimes want to view \( \mathcal{O}A \) as a right \( W \)-algebra.

Let \( \tilde{\sigma} \) denote the automorphism of \( F \) induced by \( \sigma \) and \( \tilde{\eta} \) the projection of \( \eta \) to \( F \). The automorphism \( \tilde{\sigma} \) of \( F \) induces an automorphism \( W(\tilde{\sigma}) \) on \( W \). For any \( W \)-algebra \( A \), let \( A^{\tilde{\sigma}} := A \otimes_W W(\tilde{\sigma}) \), where \( W \) is considered as a \( W \)-algebra via \( W(\tilde{\sigma}) \). Note that \( A^{\tilde{\sigma}} \) is a \( W \)-bimodule with different left and right actions. Namely there is the left action given by \( w(a \otimes w') = aw \otimes w' \), which may be different from the right action given by \( (a \otimes w')w = a \otimes ww' \). In particular, \( \mathcal{O}A^{\tilde{\sigma}} = \mathcal{O} \otimes_W W A \otimes_W W(\tilde{\sigma}) \). Let \( \iota(\tilde{\sigma}, A) : A \to A^{\tilde{\sigma}} \) be the usual map given by \( \iota(\tilde{\sigma}, A)(a) = a \otimes 1 \). It is an isomorphism of rings with inverse given by \( \iota(\tilde{\sigma}^{-1}, A) \). Furthermore, \( \iota(\tilde{\sigma}, A) \) is a left \( W \)-algebra homomorphism.

The next lemma describes the relationship between the deformation rings arising from the universal couples \((R_{\tilde{\rho}_F}, \tilde{\rho}^{\text{univ}}_F), (\tilde{\rho}^{\text{univ}}_F, (\tilde{\rho}^{\text{univ}}_F \otimes \tilde{\rho}^{\text{univ}}_F)\), and \((R_{\tilde{\eta} \otimes \tilde{\rho}_F}, (\tilde{\eta} \otimes \tilde{\rho}_F)^{\text{univ}})\).

**Lemma 3.2.**

1. If \( \tilde{\rho}^{\text{univ}}_F \cong \tilde{\eta} \otimes \tilde{\rho}_F \), then the universal couples \((R_{\tilde{\rho}^{\text{univ}}_F}, (\tilde{\rho}^{\text{univ}}_F)^{\text{univ}})\) and \((R_{\tilde{\eta} \otimes \tilde{\rho}^{\text{univ}}_F}, (\tilde{\eta} \otimes \tilde{\rho}^{\text{univ}}_F)^{\text{univ}})\) are canonically isomorphic.

2. There is a canonical isomorphism \( \varphi : R_{\tilde{\rho}^{\text{univ}}_F} \to R_{\tilde{\rho}^{\text{univ}}_F} \) of right \( W \)-algebras such that \( (\tilde{\rho}^{\text{univ}}_F)^{\text{univ}} \cong \varphi \circ \iota(\tilde{\sigma}, R_{\tilde{\rho}^{\text{univ}}_F}) \circ \tilde{\rho}_F^{\text{univ}} \).

3. Viewing \((\tilde{\eta} \otimes \tilde{\rho}_F)^{\text{univ}}\) as a representation valued in \( GL_2(\mathcal{O}R_{\tilde{\eta} \otimes \tilde{\rho}^{\text{univ}}_F}) \) via the natural map \( R_{\tilde{\eta} \otimes \tilde{\rho}^{\text{univ}}_F} \to \mathcal{O}R_{\tilde{\eta} \otimes \tilde{\rho}^{\text{univ}}_F} \), there is a natural \( W \)-algebra homomorphism \( \psi : R_{\tilde{\eta} \otimes \tilde{\rho}^{\text{univ}}_F} \to \mathcal{O}R_{\tilde{\rho}_F} \) such that \( \eta \otimes \tilde{\rho}_F^{\text{univ}} \cong (1 \otimes \psi) \circ (\tilde{\eta} \otimes \tilde{\rho}_F)^{\text{univ}} \).

**Proof.** The first statement follows directly from the definition of universal couples.

For (2), we show that the right \( W \)-algebra \( R_{\tilde{\rho}^{\text{univ}}_F} \) satisfies the universal property for \( \tilde{\rho}^{\text{univ}}_F \). Let \( A \in C \) and \( r : G^N_Q \to GL_2(A) \) be a deformation of \( \tilde{\rho}^{\text{univ}}_F \). Then \( \iota(\tilde{\sigma}^{-1}, A) \circ r \)
is a deformation of $\tilde{\rho}_F$, viewing $A^{\tilde{\sigma}}$ as a right $W$-algebra. By universality there is a unique right $W$-algebra homomorphism $\alpha((\tilde{\sigma}^{-1}, A) \circ r): R_{\tilde{\rho}_F} \to A^{\tilde{\sigma}^{-1}}$ such that $t((\tilde{\sigma}^{-1}, A) \circ r) \cong \alpha((\tilde{\sigma}^{-1}, A) \circ r) \circ \rho_F^{univ}$. Tensoring $\alpha((\tilde{\sigma}^{-1}, A) \circ r)$ with $W$ over $W(\tilde{\sigma})$ gives a homomorphism of right $W$-algebras $\alpha((\tilde{\sigma}^{-1}, A) \circ r) \otimes W(\tilde{\sigma}) 1: R_{\tilde{\rho}_F} \to A$ such that $r \cong (\alpha((\tilde{\sigma}^{-1}, A) \circ r) \otimes W(\tilde{\sigma}) 1) \circ t((\tilde{\sigma}, \tilde{\rho}_F) \circ \rho_F^{univ})$. This shows that the right $W$-algebra $R_{\tilde{\rho}_F}$ satisfies the universal property for $\tilde{\rho}_F$. With notation as above, when $r = (\tilde{\rho}_F)^{univ}$ we set $\varphi = \alpha((\tilde{\sigma}^{-1}, R_{\tilde{\rho}_F}) \circ (\tilde{\rho}_F)^{univ}) \otimes W(\tilde{\sigma}) 1$, so

$$
(\tilde{\rho}_F)^{univ} \cong \varphi \circ t((\tilde{\sigma}, \tilde{\rho}_F) \circ \rho_F^{univ}).
$$

(1)

In particular, $\varphi$ is a right $W$-algebra homomorphism.

Finally, let $i : R_{\tilde{\eta} \otimes \tilde{\rho}_F} \to \mathcal{O} R_{\tilde{\eta} \otimes \tilde{\rho}_F}$ be the map given by $x \mapsto 1 \otimes x$. If $A$ is a $W$-algebra and $r : G^N_{\tilde{\sigma}} \to \mathbb{GL}_2(\mathcal{O})$ is a deformation of $\tilde{\rho}_F$, then $\tilde{\eta} \otimes r : G^N_{\tilde{\sigma}} \to \mathbb{GL}_2(\mathcal{O})$ is a deformation of $\tilde{\eta} \otimes \tilde{\rho}_F$. Hence there is a unique $W$-algebra homomorphism $\alpha(\eta \otimes r) : R_{\tilde{\eta} \otimes \tilde{\rho}_F} \to \mathcal{O} A$ such that $\eta \otimes r \cong \alpha(\eta \otimes r) \circ (\tilde{\eta} \otimes \tilde{\rho}_F)^{univ}$. We can extend $\alpha(\eta \otimes r)$ to an $\mathcal{O}$-algebra homomorphism $1 \otimes \alpha(\eta \otimes r) : \mathcal{O} R_{\tilde{\eta} \otimes \tilde{\rho}_F} \to \mathcal{O} A$ by sending $x \otimes y$ to $(x \otimes 1) \alpha(\eta \otimes r)(y)$. In particular, $\eta \otimes r \cong (1 \otimes \alpha(\eta \otimes r)) \circ i \circ (\tilde{\eta} \otimes \tilde{\rho}_F)^{univ}$. When $r = \tilde{\rho}_F^{univ}$, let $\psi$ denote $\alpha(\eta \otimes \tilde{\rho}_F^{univ})$, so

$$
\eta \otimes \tilde{\rho}_F^{univ} \cong (1 \otimes \psi) \circ i \circ (\tilde{\eta} \otimes \tilde{\rho}_F)^{univ}.
$$

Let $A$ be a $W$-algebra. We would like to define a ring homomorphism $m(\tilde{\sigma}, A) : A^{\tilde{\sigma}} \to A$ such that $m(\tilde{\sigma}, A) \circ t(\tilde{\sigma}, A)$ is a lift of $\tilde{\sigma}$. When $A = \mathbb{F}$ we can do this by defining $m(\tilde{\sigma}, \mathbb{F})(x \otimes y) = t(\tilde{\sigma})(x) y$. Similarly, when $A = W$ we can define $m(\tilde{\sigma}, W)(x \otimes y) = W(\tilde{\sigma})(x) y$. If $A = W[T]$ or $W[[T]]$ then $A^{\tilde{\sigma}} = W^{\tilde{\sigma}}[T]$ or $W^{\tilde{\sigma}}[[T]]$, and we can define $m(\tilde{\sigma}, A)$ by simply applying $m(\tilde{\sigma}, W)$ to the coefficients of the polynomials or power series. However, for a general $W$-algebra $A$ it is not necessarily possible to define $m(\tilde{\sigma}, A)$ or to lift $\tilde{\sigma}$. (If $A$ happens to be smooth over $W$ then it is always possible to lift $\tilde{\sigma}$ to $A$.) Note that by Nakayama’s lemma, if $m(\tilde{\sigma}, A)$ exists then $m(\tilde{\sigma}, A) \circ t(\tilde{\sigma}, A)$ is a ring automorphism of $A$.

Fortunately, we do not need $m(\tilde{\sigma}, A)$ to exist for all $W$-algebras; just for $\psi$. Our strategy is to prove that if $\tilde{\rho}_F^{\tilde{\sigma}} \cong \tilde{\eta} \otimes \tilde{\rho}_F$, then the ring homomorphism $m(\tilde{\sigma}, \mathcal{O} R_{\tilde{\rho}_F})$ exists.

**Lemma 3.3.** If $\tilde{\rho}_F$ is absolutely irreducible and $\tilde{\rho}_F^{\tilde{\sigma}} \cong \tilde{\eta} \otimes \tilde{\rho}_F$, then there is a ring homomorphism $m(\tilde{\sigma}, \mathcal{O} R_{\tilde{\rho}_F}) : \mathcal{O} R_{\tilde{\rho}_F} \to \mathcal{O} \tilde{\rho}_F$ that is a lift of $m(\tilde{\sigma}, \mathbb{F})$. In particular, $m(\tilde{\sigma}, \mathcal{O} R_{\tilde{\rho}_F}) \circ t(\tilde{\sigma}, \mathcal{O} R_{\tilde{\rho}_F})$ is a lift of $\tilde{\sigma}$.

**Proof.** With notation as in Lemma 3.2 define $m(\tilde{\sigma}, \mathcal{O} R_{\tilde{\rho}_F}) = (1 \otimes \psi) \circ (1 \otimes \varphi)$. We will show that $1 \otimes \varphi$ induces $m(\tilde{\sigma}, \mathbb{F})$ and $1 \otimes \psi$ induces the identity on $\mathbb{F}$. Note that $\mathbb{F}$ is the residue field of $\mathcal{O}$ since $\tilde{\chi}$, and hence $\tilde{\eta}$, takes values in $\mathbb{F}$. Therefore all of the tensor products with $\mathcal{O}$ residually disappear. Hence it suffices to show that $\varphi$ induces $m(\tilde{\sigma}, \mathbb{F})$ and $\psi$ acts trivially on $\mathbb{F}$. 

By definition $\mathbb{F}$ is generated by $\{a(\ell, F) : \ell \nmid N\}$. Therefore it suffices to check that $\psi$ acts trivially on $a(\ell, F)$ for any prime $\ell$ not dividing $N$. But $\psi \circ (\tilde{\eta} \otimes \tilde{\rho}_F)^{\text{univ}} \cong \eta \otimes \tilde{\rho}_F^{\text{univ}}$. Evaluating at Frobl, taking traces, and reducing to the residue field shows that $\psi$ induces the identity on $\mathbb{F}$.

Let $\tilde{\varphi} : \mathbb{F} \otimes_{\mathbb{Q}} \mathbb{F} \to \mathbb{F}$ be the residual map induced by $\varphi$. By reducing (1) to the residue field we find that $\sigma \circ \tilde{\rho}_F \cong \tilde{\varphi} \circ \iota(\tilde{\sigma}, \mathbb{F}) \circ \tilde{\rho}_F$. By universality we conclude that $\sigma = \tilde{\varphi} \circ \iota(\tilde{\sigma}, \mathbb{F})$. But $\sigma = m(\tilde{\sigma}, \mathbb{F}) \circ \iota(\tilde{\sigma}, \mathbb{F})$ and hence $\tilde{\varphi} = m(\tilde{\sigma}, \mathbb{F})$, as desired. □

Define $\Sigma = (1 \otimes \psi) \circ (1 \otimes \varphi) \circ (1 \otimes \iota(\tilde{\sigma}, R_{\tilde{\rho}_F}))$. By the proof of Lemma 3.3 we see that $\Sigma$ is a lift of $\sigma$ to $\mathcal{O}R_{\tilde{\rho}_F}$. In the next subsection we use automorphic techniques to descend $\Sigma$ to $\mathbb{F}$. In order to do so we need the following properties of $\Sigma$.

**Proposition 3.4.**

(1) For all $w \in W$ we have $\Sigma(1 \otimes w) = 1 \otimes W(\tilde{\sigma})(w)$.

(2) For all $x \in \mathcal{O}$ we have $\Sigma(x \otimes 1) = x \otimes 1$.

(3) The automorphism $\tilde{\sigma}$ of $\mathbb{F}$ is necessarily trivial and hence, under the assumption that the order of $\chi$ is a power of 2 and $p \neq 2$, it follows that $\eta$ is a quadratic character.

(4) The automorphism $\Sigma$ of $R_{\tilde{\rho}_F}$ is a lift of $\sigma$.

**Proof.** The first point is the most subtle. The key point is that $\varphi$ is a right $W$-algebra homomorphism. Let $w \in W$. Then

$$(1 \otimes \iota(\tilde{\sigma}, \mathcal{O}R_{\tilde{\rho}_F}))(1 \otimes w) = 1 \otimes w \otimes 1 = 1 \otimes 1 \otimes W(\tilde{\sigma})(w).$$

Since $\varphi$ is a right $W$-algebra homomorphism and $\psi$ is a $W$-algebra homomorphism we see that $\Sigma(1 \otimes w) = 1 \otimes W(\tilde{\sigma})(w)$, as claimed.

The fact that $\Sigma(x \otimes 1) = x \otimes 1$ for all $x \in \mathcal{O}$ follows directly from the definition of $\Sigma$.

The first two facts imply that $W(\tilde{\sigma})$ is trivial. Indeed, for any $w \in W$ we have $w \otimes 1 = 1 \otimes w \in \mathcal{O}R_{\tilde{\rho}_F}$. Therefore by the first two facts, in $\mathcal{O}R_{\tilde{\rho}_F}$ we have

$$w \otimes 1 = \Sigma(w \otimes 1) = \Sigma(1 \otimes w) = 1 \otimes W(\tilde{\sigma})(w) = W(\tilde{\sigma})(w) \otimes 1.$$ 

The ring homomorphism $\mathcal{O} \to \mathcal{O}R_{\tilde{\rho}_F}$ is injective since $R_{\tilde{\rho}_F}$ covers $\mathbb{I}'$ and $\mathbb{I}' \supset \mathcal{O}$. Therefore $W(\tilde{\sigma})$ and hence $\tilde{\sigma}$ must be trivial.

Therefore $\tilde{\rho}_F \cong \tilde{\eta} \otimes \tilde{\rho}_F$. Taking determinants we find that $\det \tilde{\rho}_F = \tilde{\eta}^2 \det \tilde{\rho}_F$ and hence $\tilde{\eta}$ is quadratic. Therefore the values of $\eta$ are of the form $\pm \zeta$, where $\zeta$ is a $p$-power root of unity. But by assumption $\eta$ takes values in $\mathbb{Z}_p[\chi]$ and $\chi$ has 2-power order. Since $p \neq 2$ it follows that $\eta$ must be quadratic.

In view of the previous parts of the current proposition we see that $\mathcal{O} = W$ and hence $\mathcal{O}R_{\tilde{\rho}_F} = R_{\tilde{\rho}_F}$. Furthermore, the first two maps in the definition of $\Sigma$ become trivial and hence $\Sigma = \psi$. By definition of $\psi$ we have $\psi \circ \tilde{\rho}_F^{\text{univ}} \cong \eta \otimes \tilde{\rho}_F^{\text{univ}}$. 

Let $\alpha = \alpha(\rho_F) : R_{\rho_F} \to \mathcal{L}$ and regard $\mathcal{P}_i : \mathcal{L} \to \mathcal{O}_p$ as an algebra homomorphism. Since $\rho_0^\sigma \cong \eta \otimes \rho_0$ it follows from the definitions of all maps involved that

$$\sigma \circ \mathcal{P}_1 \circ \alpha \circ \bar{\rho}_F^{\text{univ}} \cong \mathcal{P}_2 \circ \alpha \circ \Sigma \circ \bar{\rho}_F^{\text{univ}}.$$ 

By universality $\sigma \circ \mathcal{P}_1 \circ \alpha = \mathcal{P}_2 \circ \alpha \circ \Sigma$ and thus $\Sigma$ is a lift of $\sigma$. \qed

**Descending $\Sigma$ to $\mathcal{L}$ via automorphic methods.** To prove Theorem 3.1 it now remains to show that $\Sigma$ descends to an automorphism of $\mathcal{L}$. Let us describe the strategy of proof before proceeding. We begin by showing that the character $\eta$ is unramified at $p$. Once we know this, it is fairly straightforward to check that the irreducible component $\Sigma^*(\text{Spec } \mathcal{L})$ is modular in the sense that it is an irreducible component of an ordinary Hecke algebra of some tame level and nebentypus. We then verify that the tame level and nebentypus of $\Sigma^*(\text{Spec } \mathcal{L})$ match those for $\text{Spec } \mathcal{L}'$, so we have two irreducible components of the same Hecke algebra. Finally, $\mathcal{P}_1'$ is an arithmetic point on both $\text{Spec } \mathcal{L}'$ and $\Sigma^*(\text{Spec } \mathcal{L})$. As the ordinary Hecke algebra is étale over $\Lambda$ at arithmetic points [Hida 2006, Proposition 3.78], the two irreducible components $\Sigma^*(\text{Spec } \mathcal{L})$ and $\text{Spec } \mathcal{L}'$ must coincide. In other words, $\Sigma$ descends to $\mathcal{L}'$ as desired. There is a technical point that $\text{Spec } \mathcal{L}$ and $\Sigma^*(\text{Spec } \mathcal{L})$ are only irreducible components of the algebra generated by Hecke operators away from $N$, so in order to use étaleness we must associate to $\Sigma^*(\text{Spec } \mathcal{L})$ a primitive irreducible component $\text{Spec } \mathcal{J}$ of the full Hecke algebra. See the discussion after Corollary 3.7.

**Lemma 3.5.** Let $\rho_1$, $\rho_2 : G_{\mathbb{Q}_p} \to \text{GL}_2(\mathcal{O}_p)$ be ordinary representations such that the inertia group acts by an infinite order character on the kernel of the unique $p$-unramified quotient of each $\rho_i$. Assume there is an automorphism $\sigma \in G_{\mathbb{Q}_p}$ and a finite order character $\eta$ such that $\rho_1^\sigma \cong \eta \otimes \rho_2$. Then $\eta$ is unramified at $p$.

**Proof.** Since $\rho_i$ is $p$-ordinary, by choosing bases appropriately we may assume $\rho_i = (\begin{smallmatrix} \epsilon_i & \ast \\ 0 & \delta_i \end{smallmatrix})$ with $\delta_i$ unramified. By assumption $\epsilon_i |_{I_p}$ has infinite order. As $\rho_1^\sigma \cong \eta \otimes \rho_2$, it follows that for some $x \in \text{GL}_2(\mathcal{O}_p)$ we have $\rho_1^\sigma = x(\eta \otimes \rho_2)x^{-1}$. Write $x = (\begin{smallmatrix} a & b \\ c & d \end{smallmatrix})$ and $\eta \otimes \rho_2 = (\begin{smallmatrix} \eta \epsilon_2 & \ast \\ 0 & \eta \delta_2 \end{smallmatrix})$. A straightforward matrix computation shows that on $I_p$ we have

$$\begin{pmatrix} \epsilon_1^\sigma & \ast \\ 0 & 1 \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} (ad \epsilon_2 - bc) \eta - acu & \ast \\ c(d \eta (\epsilon_2 - 1) - cu) & (ad - bc \epsilon_2) \eta + acu \end{pmatrix}.$$ 

Hence either $c = 0$ or $cu = d \eta (\epsilon_2 - 1)$.

If $c = 0$ then on $I_p$ we have

$$\begin{pmatrix} \epsilon_1^\sigma & \ast \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \eta \epsilon_2 & \ast \\ 0 & \eta \delta_2 \end{pmatrix},$$

and so $\eta |_{I_p} = 1$, as desired. If $cu = d \eta (\epsilon_2 - 1)$ then on $I_p$ we have

$$\begin{pmatrix} \epsilon_1^\sigma & \ast \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \eta & \ast \\ 0 & \eta \epsilon_2 \end{pmatrix}.$$
Therefore we have $\varepsilon_1|_{I_p} = \eta|_{I_p} = \varepsilon_2^{-1}|_{I_p}$. But this is impossible since $\varepsilon_i|_{I_p}$ has infinite order by assumption while $\eta$ has finite order. Therefore $\eta$ must be unramified.

In what follows we use Wiles’s interpretation of Hida families [Wiles 1988]. Namely for a finite extension $\mathcal{J}$ of $\Lambda$, a formal power series $G = \sum_{n=1}^{\infty} a(n, G)q^n$ is a $\mathcal{J}$-adic cusp form of level $\Gamma_0(N)$ and character $\chi$ if for almost all arithmetic primes $\mathfrak{P}$ of $\mathcal{J}$, the specialization of $G$ at $\mathfrak{P}$ gives the $q$-expansion of an element $g_{\mathfrak{P}}$ of $S_k(\Gamma_0(Np^{r(\varepsilon)}), \varepsilon \chi \omega^{-k})$, where $\mathfrak{P}$ lies over $P_{k, \varepsilon}$. One defines the Hecke operators by the usual formulae on coefficients of $q$-expansions. We say $G$ is ordinary if it is an eigenform for the Hecke operators whose eigenvalue under $U(p)$ is in $\mathcal{J}^\times$. Let $\mathfrak{S}(N, \chi; \mathcal{J})$ be the $\mathcal{J}$-submodule of $\mathfrak{B}[q]$ spanned by all $\mathcal{J}$-adic cusp forms of level $\Gamma_0(N)$ and character $\chi$ that are also Hecke eigenforms. Let $\mathfrak{S}_{\text{ord}}(N, \chi; \mathcal{J})$ denote the $\mathcal{J}$-subspace of $\mathfrak{S}(N, \chi; \mathcal{J})$ spanned by all ordinary $\mathcal{J}$-adic cusp forms.

For each Dirichlet character $\psi$, we shall write $c(\psi) \in \mathbb{Z}^+$ for the conductor of $\psi$. Let $\psi : (\mathbb{Z}/L\mathbb{Z})^\times \rightarrow \mathbb{Q}^\times$ be a Dirichlet character. Let $\eta$ be a primitive Dirichlet character with values in $\mathbb{Z}[\chi]$. (Every twist character of $F$ has this property by Lemma 3.11.) Denote by $M(\psi, \eta)$ the least common multiple of $L, c(\eta)^2$, and $c(\psi)c(\eta)$. By [Shimura 1971, Proposition 3.64], there is a linear map

$$R_{\psi, \eta} : S_k(\Gamma_0(M(\psi, \eta)), \psi) \rightarrow S_k(\Gamma_0(M(\psi, \eta)), \eta^2 \psi)$$

$$f = \sum_{n=1}^{\infty} a(n, f)q^n \mapsto \eta f = \sum_{n=1}^{\infty} \eta(n)a(n, f)q^n.$$

We now show that there is an analogous map in the $\mathcal{J}$-adic setting.

**Lemma 3.6.** There is a well defined $\mathcal{J}$-linear map

$$\mathbb{R}_{\chi, \eta} : \mathfrak{S}(M(\chi, \eta), \chi; \mathcal{J}) \rightarrow \mathfrak{S}(M(\chi, \eta), \eta^2 \chi; \mathcal{J})$$

$$G = \sum_{n=1}^{\infty} a(n, G)q^n \mapsto \eta G = \sum_{n=1}^{\infty} \eta(n)a(n, G)q^n.$$

If $p \nmid c(\eta)$ then $\mathbb{R}_{\chi, \eta}$ sends $\mathfrak{S}_{\text{ord}}(M(\chi, \eta), \chi; \mathcal{J})$ to $\mathfrak{S}_{\text{ord}}(M(\chi, \eta), \eta^2 \chi; \mathcal{J})$.

**Proof.** Let $\mathfrak{P}$ be an arithmetic prime of $\mathcal{J}$, and let $P_{k, \varepsilon}$ be the arithmetic prime of $\Lambda$ lying under $\mathfrak{P}$. If $G \in \mathfrak{S}_{\text{ord}}(M(\chi, \eta), \chi; \mathcal{J})$ then

$$g_{\mathfrak{P}} \in S_k(\Gamma_0(M(\chi, \eta)p^{r(\varepsilon)}), \varepsilon \chi \omega^{-k}).$$

Let $\psi = \varepsilon \chi \omega^{-k}$. It follows easily from the definitions that $M(\psi, \eta) = M(\chi, \eta)p^{r(\varepsilon)}$. Therefore

$$\eta g_{\mathfrak{P}} = R_{\psi, \eta}(g_{\mathfrak{P}}) \in S_k(\Gamma_0(M(\psi, \eta)), \eta^2 \psi) = S_k(\Gamma_0(M(\chi, \eta)p^{r(\varepsilon)}), \eta^2 \varepsilon \chi \omega^{-k}),$$

and hence $\mathbb{R}_{\chi, \eta}$ is well defined.
so \( \eta G \in \mathfrak{S}(M(\chi, \eta), \eta^2 \chi; J) \).

For the statement about ordinarity, we may assume \( G \) is a normalized eigenform, so \( a(p, G) \) is the eigenvalue of \( G \) under the \( U(p) \) operator. If \( G \) is ordinary then \( a(p, G) \in J^\times \). Hence \( \eta(p)a(p, G) = a(p, \eta G) \in J^\times \) if and only if \( \eta(p) \neq 0 \).

**Corollary 3.7.** The representation associated to \( \Sigma^*(\text{Spec} \mathfrak{l}) \) is modular of level \( M(\chi, \eta) \) and nebentypus \( \chi \).

**Proof.** The representation associated to \( \Sigma^*(\text{Spec} \mathfrak{l}) \) is isomorphic to \( \eta \otimes \rho_F \). Consider the formal \( q \)-expansion \( \eta F := \sum_{n=1}^{\infty} \eta(n)a(n, F)q^n \in \mathbb{Z}[[q]] \). By Lemma 3.6 and Lemma 3.5 we see that \( \eta F \) is a Hida family of level \( \Gamma_0(M(\chi, \eta)) \) and nebentypus \( \eta^2 \chi \). Clearly the Galois representation of \( \eta F \) is isomorphic to \( \eta \otimes \rho_F \) since their traces on Frobenius elements agree on all but finitely many primes. Since \( \eta \otimes \rho_F \cong \alpha \ast \Sigma \ast \tilde{\rho}_F^{\text{univ}} \), it follows that \( \Sigma^*(\text{Spec} \mathfrak{l}) \) is modular of level \( M(\chi, \eta) \) and nebentypus \( \eta^2 \chi \). By Proposition 3.4 we know that \( \eta \) is quadratic and hence \( \eta^2 \chi = \chi \).

For any integer multiple \( M \) of \( N \), let \( h^\text{ord}(M, \chi; \Lambda \chi) \) be the \( \Lambda \chi \)-subalgebra of \( h^\text{ord}(M, \chi; \Lambda \chi) \) generated by \( \{ T(n) : (n, N) = 1 \} \). Corollary 3.7 shows that \( \Sigma^*(\text{Spec} \mathfrak{l}) \) is an irreducible component of \( \text{Spec} h^\text{ord}(M(\chi, \eta), \chi; \Lambda \chi)' \). There is a natural map \( \beta : \text{Spec} h^\text{ord}(M, \chi; \Lambda \chi) \rightarrow \text{Spec} h^\text{ord}(M, \chi; \Lambda \chi)' \) coming from the natural inclusion of algebras. An irreducible component \( \text{Spec} \mathfrak{J}' \) of \( h^\text{ord}(M, \chi; \Lambda \chi)' \) essentially corresponds to the data of the Fourier coefficients away from \( N \). The preimage \( \beta^{-1}(\text{Spec} \mathfrak{J}') \) is a union of irreducible components whose Fourier coefficients agree with those of \( \mathfrak{J}' \) away from \( N \). By the theory of newforms we know that there is a unique irreducible component \( \text{Spec} \mathfrak{J} \) of \( h^\text{ord}(M, \chi; \Lambda \chi) \) that projects to \( \text{Spec} \mathfrak{J}' \) under \( \beta \). Let \( \text{Spec} \mathfrak{J} \) be the primitive component of \( h^\text{ord}(M(\chi, \eta), \chi; \Lambda \chi) \) that projects to \( \Sigma^*(\text{Spec} \mathfrak{l}) \) under \( \beta \). By the proof of Corollary 3.7, \( \mathfrak{J} \) is the primitive form associated to \( \eta F \) and so \( \rho_{\mathfrak{J}} \cong \eta \otimes \rho_F \).

Since \( N | M(\chi, \eta) \) there is a natural inclusion

\[
\text{Spec} h^\text{ord}(N, \chi; \Lambda \chi) \hookrightarrow \text{Spec} h^\text{ord}(M(\chi, \eta), \chi; \Lambda \chi).
\]

We wish to show that \( \text{Spec} \mathfrak{J} \) is an irreducible component of \( \text{Spec} h^\text{ord}(N, \chi; \Lambda \chi) \). We do this locally by computing the level of \( \text{Spec} \mathfrak{J} \) at each prime \( \ell \). Let \( v_\ell \) denote the usual \( \ell \)-adic valuation on the integers, normalized such that \( v_\ell(\ell) = 1 \).

**Proposition 3.8.** The primitive component \( \text{Spec} \mathfrak{J} \) is an irreducible component of \( \text{Spec} h^\text{ord}(N, \chi; \Lambda \chi) \).

**Proof.** First note that if \( \ell \nmid c(\eta) \) then \( v_\ell(M(\chi, \eta)) = v_\ell(N) \) since \( c(\chi)|N \). In particular, by Lemma 3.5 we have \( v_p(M(\chi, \eta)) = v_p(N) \).

Fix a prime \( \ell \neq p \) at which \( \eta \) is ramified. For a pro-\( p \) ring \( A \) and representation \( \pi : G_{\mathbb{Q}_\ell} \rightarrow \text{GL}_2(A) \), let \( C_\ell(\pi) \) denote the \( \ell \)-conductor of \( \pi \). See [Hida 2015, p. 659]
for the precise definition. When π is the representation associated to a classical form f, the ℓ-conductor of π is related to the level of f by the proof of the local Langlands conjecture for GL2. Indeed, when f is a classical newform of level N we have \( C_\ell(f) = \ell^{v_\ell(N)}. \) If \( f \) is new away from \( p \) and \( \ell \neq p \) then we still have \( C_\ell(f) = \ell^{v_\ell(N)}. \)

First suppose that \( \rho_F|_{I_1} \) is not reducible indecomposable. Then \( (\eta \otimes \rho_F)|_{I_1} \) is not reducible indecomposable either. Therefore \( C_\ell(\rho_F) = C_\ell(\rho_{f_{\mathfrak{P}_1}}) \) and \( C_\ell(\eta \otimes \rho_F) = C_\ell(\eta \otimes \rho_{f_{\mathfrak{P}_2}}) \) [Hida 2015, Lemma 10.2(2)]. Since Galois action does not change conductors we have

\[
C_\ell(\rho_F) = C_\ell(\rho_{f_{\mathfrak{P}_1}}) = C_\ell(\rho_{f_{\mathfrak{P}_2}}) = C_\ell(\eta \otimes \rho_{f_{\mathfrak{P}_2}}) = C_\ell(\eta \otimes \rho_F).
\]

Since \( F \) is a primitive form we have that \( f_{\mathfrak{P}_1} \) is new away from \( p \) and hence \( C_\ell(\rho_{f_{\mathfrak{P}_1}}) = \ell^{v_\ell(N)}. \) On the other hand since \( \mathfrak{J} \) is primitive we have \( C_\ell(\rho_{f_{\mathfrak{P}_1}}) = C_\ell(\eta \otimes \rho_F) \) is equal to the \( \ell \)-part of the level of \( \mathfrak{J} \), which gives the desired result at \( \ell \).

Now assume that \( \rho_F|_{I_1} \) is reducible indecomposable. By Lemma 10.1(4) of [Hida 2015] we have a character \( \psi : G_{\mathfrak{O}_\ell} \to \mathbb{I}^\times \) such that \( \rho_F|_{G_{\mathfrak{O}_\ell}} \cong \left( \begin{array}{cc} \mathcal{N} & \psi \\ 0 & 1 \end{array} \right) \), where \( \mathcal{N} \) is the unramified cyclotomic character acting on \( p \)-power roots of unity and \( \psi|_{I_1} \) has finite order. Note that since \( \eta \) is a quadratic character, \( c(\eta) \) is squarefree away from 2. Similarly, since \( \chi \) has 2-power order it follows that \( c(\chi) \) is a power of 2 times a product of distinct odd primes. Therefore, for odd primes \( \ell \) it is enough to show that \( c_\ell(\eta)^2|N. \)

We use the description of the conductor of a locally reducible indecomposable representation given on page 660 of [Hida 2015]. Let \( \psi_1 = \psi \mod \mathfrak{P}_1. \) Then \( \rho_{f_{\mathfrak{P}_2}}|_{I_1} \cong \left( \begin{array}{cc} \eta^{-1} \psi_1^\sigma & * \\ 0 & \eta^{-1} \psi_1 \end{array} \right). \) If \( \psi_1 \) is unramified then \( \eta^{-1} \psi_1^\sigma \) is ramified and hence

\[
c_\ell(\eta)^2 = c_\ell(\eta^{-1})^2 = c_\ell(\eta^{-1} \psi_1^\sigma)^2 = c_\ell(\rho_{f_{\mathfrak{P}_2}}).
\]

Since \( \rho_{f_{\mathfrak{P}_2}} \) is a specialization of \( \rho_F \) we have \( C_\ell(\rho_{f_{\mathfrak{P}_2}})|_N \) giving the desired result.

Now suppose that \( \psi_1 \) is ramified. Then \( c_\ell(\eta) = \ell|c_\ell(\psi_1) \) and \( C_\ell(\rho_{f_{\mathfrak{P}_2}}) = c_\ell(\psi_1)^2. \) Again, since \( \rho_{f_{\mathfrak{P}_1}} \) is a specialization of \( \rho_F \) we see that \( c_\ell(\eta)^2|N. \)

Finally the case \( \ell = 2 \) follows from the assumption that \( 2c(\chi)|N. \) We are able to make this hypothesis by Proposition 3.9.

Therefore \( \text{Spec} \mathfrak{J} \) is an irreducible component of \( h^\text{ord}(N, \chi; \Lambda_\chi) \), as desired. \( \square \)

**Proof of Theorem 3.1.** We first lift \( \sigma \) to an automorphism \( \Sigma \) of \( {}^O R_{\bar{\rho}_F} \) by Lemma 3.3. We are able to use the definition of \( \Sigma \) to show that \( {}^O R_{\bar{\rho}_F} = R_{\bar{\rho}_F} \) and that \( \eta \) is quadratic in Proposition 3.4. By Proposition 3.8 we see that \( \Sigma^*(\text{Spec } \mathfrak{P}) \) is a component of \( \text{Spec } h^\text{ord}(N, \chi; \Lambda_\chi)' \). Since \( \rho_{\mathfrak{P}_1}^\sigma \cong \eta \otimes \rho_{f_{\mathfrak{P}_2}} \) it follows that the arithmetic point \( \mathfrak{P}_1' \) is a point on both \( \text{Spec } \mathfrak{P} \) and \( \Sigma^* (\text{Spec } \mathfrak{P}) \). We claim that in fact \( \mathfrak{P}_1' \in \text{Spec } \mathfrak{P} \cap \text{Spec } \mathfrak{J} \).

Note that \( \mathfrak{J} \) is the primitive family passing through \( f_{\mathfrak{P}_1}^\sigma. \) (We know \( f_{\mathfrak{P}_1}^\sigma \) is primitive since \( f_{\mathfrak{P}_1} \) is an arithmetic specialization of the primitive family \( F \), and
Galois conjugation does not change the level. Indeed, \( \mathcal{J} \) is the primitive form of \( \eta F \). Let \( \mathfrak{P} \in \text{Spec } \mathcal{J} \) such that \( \mathfrak{J} \mod \mathfrak{P} = f^\sigma_{\mathfrak{P}_1} \). On the other hand the kernel of the specialization map giving rise to \( f^\sigma_{\mathfrak{P}_1} \) is \( \mathfrak{P}_1 \), since \( f^\sigma_{\mathfrak{P}_1} = \sigma (F \mod \mathfrak{P}_1) \). Therefore \( \mathfrak{P} = \mathfrak{P}_1 \in \text{Spec } \mathcal{J} \cap \text{Spec } \mathcal{J} \).

Since \( h^{\text{ord}}(N, \chi; \Lambda \chi) \) is étale over arithmetic points of \( \Lambda \) by [Hida 2006, Proposition 3.78] it follows that the irreducible components \( \text{Spec } \mathcal{J} \) and \( \text{Spec } \mathcal{J} \) must coincide and hence \( \Sigma^*(\text{Spec } \mathcal{J}) = \text{Spec } \mathcal{J} \). Therefore \( \Sigma \) descends to the desired automorphism \( \tilde{\sigma} \) of \( \mathcal{J} \). The fact that \( \tilde{\sigma} (a(\ell, F)) = \eta(\ell) a(\ell, F) \) for almost all primes \( \ell \) follows from specializing \( \Sigma \circ \tilde{\rho}_F^{\text{univ}} \cong \eta \otimes \tilde{\rho}_F^{\text{univ}} \) to \( \mathcal{J} \) and taking traces. Finally, \( \sigma \circ \mathfrak{P}'_1 = \mathfrak{P}'_2 \circ \tilde{\sigma} \) since \( \Sigma \) is a lift of \( \sigma \) by Proposition 3.4.

\section*{Nebentypus and twist characters}

We end this section with some information about twist characters. In particular Proposition 3.9 shows that we may assume from the beginning that \( \chi \) has 2-power order with \( 2c(\chi)|N \).

Note that the ring \( \mathfrak{l}_0 \) depends on \( F \). However, if \( \psi \) is a character then \( \psi F \) has the same group of conjugate self-twists as that of \( F \), and thus the same fixed ring \( \mathfrak{l}_0 \). Indeed, if \( \sigma \) is a conjugate self-twist of \( F \) with character \( \eta \), then a straightforward calculation shows that \( \psi^\sigma \eta \psi^{-1} \) is the twist character of \( \sigma \) on \( \psi F \).

\begin{proposition}
There is a Dirichlet character \( \psi \) such that the nebentypus \( \psi^2 \chi \) of \( \psi F \) has order a power of 2 and \( 2c(\psi^2 \chi)|M(\chi, \psi) \). Furthermore, \( \rho_F \) is \( \mathfrak{l}_0 \)-full if and only if \( \rho_{\psi F} \) is \( \mathfrak{l}_0 \)-full.
\end{proposition}

\begin{proof}
It is well known that the nebentypus of \( \psi F \) is \( \psi^2 \chi \) [Shimura 1971, Proposition 3.64]. Write \( \chi = \chi_2 \xi \), where \( \chi_2 \) is a character whose order is a power of 2 and \( \xi \) is an odd order character. Let \( 2n - 1 \) denote the order of \( \xi \). Then \( \xi^{2n} = \xi \), so taking \( \psi_{\text{odd}} = \xi^{-n} \) we see that \( \psi_{\text{odd}}^2 \chi = \chi_2 \xi^{2n} = \chi_2 \) is a character whose order is a power of 2.

Let \( 2^t - 1 \) be the order of \( \psi_{\text{odd}}^2 \chi \), and let \( \psi_2 : (\mathbb{Z}/2^t \mathbb{Z})^\times \rightarrow \mathbb{Q}^\times \) be the associated primitive character. Let \( \psi = \psi_2 \psi_{\text{odd}} \). Then \( 2^{2t}|M(\chi, \psi) \) whereas \( c_2(\psi^2 \chi)|2^{t-1} \). Since \( t \geq 1 \) we see that \( 2c_2(\psi^2 \chi)|2^t|M(\chi, \psi) \), as desired.

Suppose that \( \rho_{\psi F} \) is \( \mathfrak{l}_0 \)-full. Since \( \psi \) is a finite order character, ker \( \psi \) is an open subgroup of \( G_{\mathfrak{Q}} \). Thus \( \rho_{\psi F}|_{\ker \psi} \) is also \( \mathfrak{l}_0 \)-full. Note that \( \rho_{\psi F}|_{\ker \psi} = \rho_F|_{\ker \psi} \). Thus \( \rho_F \) is \( \mathfrak{l}_0 \)-full.
\end{proof}

We finish this section by recalling a lemma of Momose that shows that twist characters are valued in \( \mathbb{Z}_p[\chi] \). Thus Theorem 3.1 says that whenever a conjugate self-twist of a classical specialization \( f_\mathfrak{P} \) of \( F \) induces an automorphism of \( \mathbb{Q}_p(f_\mathfrak{P}) \), that conjugate self-twist can be lifted to a conjugate self-twist of the whole family \( F \).
Lemma 3.10 [Momose 1981, Lemma 1.5]. If \( \sigma \) is a conjugate self-twist of \( f \in \mathcal{S}_1(\Gamma_0(N), \chi) \), then \( \eta_\sigma \) is the product of a quadratic character with some power of \( \chi \). In particular, \( \eta_\sigma \) takes values in \( \mathbb{Z}[\chi] \).

The proof of Lemma 3.10 is not difficult and goes through without change in the \( \ell \)-adic setting. For completeness, we give the proof in that setting.

Lemma 3.11. If \( \sigma \) is a conjugate self-twist of \( F \) then \( \eta_\sigma \) is the product of a quadratic character with some power of \( \chi \). In particular, \( \eta_\sigma \) has values in \( \mathbb{Z}[\chi] \).

Proof. As \( \tilde{\rho}_F \) is absolutely irreducible, \( \rho_\sigma^F \cong \eta_\sigma \otimes \rho_F \). Thus \( \sigma(\det \rho_F) = \eta_\sigma^2 \det \rho_F \).

Define \( \kappa : 1 + p\mathbb{Z}_p \rightarrow \Lambda^\times \) by \( \kappa((1 + p)^s) = (1 + T)^s \) for \( s \in \mathbb{Z}_p \). Recall that for all primes \( \ell \) not dividing \( N \) we have

\[
\det \rho_F(\text{Frob}_\ell) = \chi(\ell)\kappa((\ell))\ell^{-1}.
\]

Substituting this expression for \( \det \rho_F \) into \( \sigma(\det \rho_F) = \eta_\sigma^2 \det \rho_F \) yields \( \eta_\sigma^2 = \chi^\sigma \chi^{-1} \).

Recall that \( \chi^\sigma = \chi^\alpha \) for some integer \( \alpha > 0 \). To prove the result it suffices to show that there is some \( i \in \mathbb{Z} \) such that \( \eta_\sigma^2 = \chi^{2i} \). If \( \chi \) has odd order then there is a positive integer \( j \) for which \( \chi = \chi^{2j} \). Thus \( \eta_\sigma^2 = \chi^{\sigma^{-1}} = \chi^{2j(\alpha^{-1})} \). If \( \chi \) has even order then \( \chi^\sigma \) also has even order since \( \sigma \) is an automorphism. Thus \( \alpha \) must be odd. Then \( \alpha - 1 \) is even and \( \eta_\sigma^2 = \chi^\sigma \chi^{-1} = \chi^{\alpha^{-1}} \), as desired. \( \square \)

4. Sufficiency of open image in product

Recall that \( H_0 = \bigcap_{\sigma \in \Gamma} \ker(\eta_\sigma) \) and \( H = H_0 \cap \ker(\det \tilde{\rho}_F) \). For a variety of reasons, our methods work best for representations valued in \( \text{SL}_2(\mathfrak{l}_0) \) rather than \( \text{GL}_2(\mathfrak{g}') \). Therefore, for the next three sections we assume the following theorem, the proof of which is given in Section 7.

Theorem 4.1. Assume that \( \tilde{\rho}_F \) is absolutely irreducible and \( H_0 \)-regular. If \( V = \mathfrak{g}'^2 \) is the module on which \( G_\mathfrak{Q} \) acts via \( \rho_F \), then there is a basis for \( V \) such that all of the following happen simultaneously:

1. \( \rho_F \) is valued in \( \text{GL}_2(\mathfrak{g}') \).
2. \( \rho_F |_{D_p} \) is upper triangular.
3. \( \rho_F |_{H_0} \) is valued in \( \text{GL}_2(\mathfrak{l}_0) \).
4. There is a matrix \( j = (\begin{smallmatrix} \xi & 0 \\ 0 & \xi' \end{smallmatrix}) \), where \( \xi \) and \( \xi' \) are roots of unity, such that \( j \) normalizes the image of \( \rho_F \) and \( \xi \not\equiv \xi' \mod p \).

Let \( H' = \ker(\det \tilde{\rho}_F) \). For any \( h \in H' \) we have \( \det \rho_F(h) \in 1 + \mathfrak{m}_V \). Since \( p \neq 2 \) and \( \mathfrak{g}' \) is \( p \)-adically complete, we have

\[
\sqrt{\det \rho_F(h)} = \sum_{n=0}^{\infty} \left( \frac{1}{2} \right)^n (\det \rho_F(h) - 1)^n \in \mathfrak{g}'^\times.
\]
Since $\rho_F$ is a 2-dimensional representation $\rho_F|_{H'} \otimes \left(\sqrt{\det \rho_F|_{H'}}\right)^{-1}$ takes values in $\text{SL}_2(I')$. Restricting further it follows from Theorem 4.1 that

$$\rho := \rho_F|_{H} \otimes \left(\sqrt{\det \rho_F|_{H}}\right)^{-1}$$

takes values in $\text{SL}_2(I_0)$. Note that the image of $\rho$ is still normalized by the matrix $j$ of Theorem 4.1 since we only modified $\rho_F$ by scalars, which commute with $j$. In Proposition 4.10 we show that $\rho_F$ is $I_0$-full if and only if $\rho$ is $I_0$-full. The proof of Proposition 4.10 is postponed until the end of the current section since it uses the theory of Pink–Lie algebras developed below. In the next three sections we prove that $\rho$ is $I_0$-full.

The purpose of the current section is to make the following reduction step in the proof of Theorem 2.4.

**Proposition 4.2.** Assume there is an arithmetic prime $P$ of $\Lambda$ such that the image of $\text{Im} \rho$ in $\prod_{Q|P} \text{SL}_2(I_0/Q)$ is open in the product topology. Then $\rho$ (and hence $\rho_F$) is $I_0$-full.

In the proof we use a result of Pink [1993] that classifies $p$-profinite subgroups of $\text{SL}_2(A)$ for a complete semilocal $p$-profinite ring $A$. (Our assumption that $p > 2$ is necessary for Pink’s theory.) We give a brief exposition of the relevant parts of his work for the sake of establishing notation. Define

$$\Theta : \text{SL}_2(A) \to \mathfrak{sl}_2(A), \quad x \mapsto x - \frac{1}{2} \text{tr}(x),$$

where we consider $\frac{1}{2} \text{tr}(x)$ as a scalar matrix. Let $G$ be a $p$-profinite subgroup of $\text{SL}_2(A)$. Define $L_1(G)$ to be the closed subgroup of $\mathfrak{sl}_2(A)$ that is topologically generated by $\Theta(G)$. Let $L_1 \cdot L_1$ be the closed (additive) subgroup of $M_2(A)$ topologically generated by $\{xy : x, y \in G\}$. Let $C$ denote $\text{tr}(L_1 \cdot L_1)$. Sometimes we will view $C \subset M_2(A)$ as a set of scalar matrices. For $n \geq 2$ define $L_n(G)$ to be the closed (additive) subgroup of $\mathfrak{sl}_2(A)$ generated by

$$[L_1(G), L_{n-1}(G)] := \{xy - yx : x \in L_1(G), y \in L_{n-1}(G)\}.$$

**Definition 4.3.** The Pink–Lie algebra of a $p$-profinite group $G$ is $L_2(G)$. Whenever we write $L(G)$ without a subscript we shall always mean $L_2(G)$.

As an example one can compute that for an ideal $a$ of $A$, the $p$-profinite subgroup $G = \Gamma_A(a)$ has Pink–Lie algebra $L_2(G) = a^2\mathfrak{sl}_2(A)$. This example plays an important role in what follows.

As for $n \geq 1$, define

$$\mathcal{M}_n(G) = C \oplus L_n(G) \subset M_2(A)$$

$$\mathcal{H}_n(G) = \{x \in \text{SL}_2(A) : \Theta(x) \in L_n(G) \text{ and } \text{tr}(x) - 2 \in C\}.$$
Pink proves that $\mathcal{M}_n(\mathcal{G})$ is a closed $\mathbb{Z}_p$-Lie algebra of $M_2(A)$ and that $\mathcal{H}_n = \text{SL}_2(A) \cap (1 + \mathcal{M}_n)$ for all $n \geq 1$. Furthermore, write

$$\mathcal{G}_1 = \mathcal{G}, \quad \mathcal{G}_{n+1} = (\mathcal{G}, \mathcal{G}_n),$$

where $(\mathcal{G}, \mathcal{G}_n)$ is the closed subgroup of $\mathcal{G}$ topologically generated by the commutators $\{gg_n g^{-1} g_n^{-1} : g \in \mathcal{G}, g_n \in \mathcal{G}_n\}$.

**Theorem 4.4** [Pink 1993]. With notation as above, $\mathcal{G}$ is a closed normal subgroup of $\mathcal{H}_1(\mathcal{G})$. Furthermore, $\mathcal{H}_n(\mathcal{G}) = (\mathcal{G}, \mathcal{G}_n)$ for $n \geq 2$.

There are two important functoriality properties of the correspondence $\mathcal{G} \mapsto L(\mathcal{G})$ that we will use. First, since $\Theta$ is constant on conjugacy classes of $\mathcal{G}$ it follows that $L_n(\mathcal{G})$ is stable under the adjoint action of the normalizer $N_{\text{SL}_2(A)}(\mathcal{G})$ of $\mathcal{G}$ in $\text{SL}_2(A)$. That is, for $g \in N_{\text{SL}_2(A)}(\mathcal{G})$, $x \in L_n(\mathcal{G})$ we have $gxg^{-1} \in L_n(\mathcal{G})$. If $\mathfrak{a}$ is an ideal of $A$ such that $A/\mathfrak{a}$ is $p$-profinite, then we write $\mathcal{G}_\mathfrak{a}$ for the $p$-profinite group $\mathcal{G} \cdot \Gamma_A(a)/\Gamma_A(\mathfrak{a}) \subseteq \text{SL}_2(A/a)$. The second functoriality property is that the canonical linear map $L(\mathcal{G}) \to L(\mathcal{G}_\mathfrak{a})$ induced by $x \mapsto x \mod \mathfrak{a}$ is surjective.

Let $\mathfrak{m}_0$ be the maximal ideal of $\mathfrak{l}_0$, and let $\mathfrak{g}$ denote the $p$-profinite group $\text{Im} \rho \cap \Gamma_{\mathfrak{l}_0}(\mathfrak{m}_0)$. The proof of Proposition 4.2 consists of showing that if $\overline{\mathcal{G}}_{\mathfrak{P}\mathfrak{l}_0}$ is open in $\prod_{\mathfrak{Q}|\mathfrak{P}} \text{SL}_2(\mathfrak{l}_0/\mathfrak{Q})$ then $\mathfrak{g}$ contains $\Gamma_{\mathfrak{l}_0}(\mathfrak{a}_0)$ for some nonzero $\mathfrak{l}_0$-ideal $\mathfrak{a}_0$. Let $L = L(\mathcal{G})$ be the Pink–Lie algebra of $\mathcal{G}$. Since $\overline{\mathcal{G}}_{\mathfrak{P}\mathfrak{l}_0}$ is open, for every prime $\mathfrak{Q}$ of $\mathfrak{l}_0$ lying over $\mathfrak{P}$ there is a nonzero $\mathfrak{l}_0/\mathfrak{Q}$-ideal $\mathfrak{a}_\mathfrak{Q}$ such that

$$\overline{\mathcal{G}}_{\mathfrak{P}\mathfrak{l}_0} \supseteq \prod_{\mathfrak{Q}|\mathfrak{P}} \Gamma_{\mathfrak{l}_0/\mathfrak{Q}}(\mathfrak{a}_\mathfrak{Q}).$$

Thus $L(\overline{\mathcal{G}}_{\mathfrak{P}\mathfrak{l}_0}) \supseteq \bigoplus_{\mathfrak{Q}|\mathfrak{P}} \mathfrak{a}_\mathfrak{Q}^2 \mathfrak{s}\mathfrak{l}_2(\mathfrak{l}_0/\mathfrak{Q})$.

Recall from Theorem 4.1 that we have roots of unity $\zeta$ and $\zeta'$ such that $\zeta \neq \zeta'$ mod $p$ and the matrix $j := \begin{pmatrix} \zeta & 0 \\ 0 & \zeta' \end{pmatrix}$ normalizes $\mathfrak{g}$. Let $\alpha = \zeta \zeta'^{-1}$. A straightforward calculation shows that the eigenvalues of $\text{Ad}(j)$ acting on $\mathfrak{s}\mathfrak{l}_2(\mathfrak{l}_0)$ are $\alpha$, $1$, $\alpha^{-1}$. Note that since $\zeta \neq \zeta'$ either all of $\alpha$, $1$, $\alpha^{-1}$ are distinct or else $\alpha = -1$. For $\lambda \in \{\alpha, 1, \alpha^{-1}\}$ let $L[\lambda]$ be the $\lambda$-eigenspace of $\text{Ad}(j)$ acting on $L$. One computes that $L[1]$ is the set of diagonal matrices in $L$. If $\alpha = -1$ then $L[-1]$ is the set of antidiagonal matrices in $L$. If $\alpha \neq -1$ then $L[\alpha]$ is the set of upper nilpotent matrices in $L$, and $L[\alpha^{-1}]$ is the set of lower nilpotent matrices in $L$. Regardless of the value of $\alpha$, let $u$ denote the set of upper nilpotent matrices in $L$ and $u'$ denote the set of lower nilpotent matrices in $L$. Let $\mathcal{L}$ be the $\mathbb{Z}_p$-Lie algebra generated by $u$ and $u'$ in $\mathfrak{s}\mathfrak{l}_2(\mathfrak{l}_0)$.

**Lemma 4.5.** The matrix $J := \begin{pmatrix} 1+T & 0 \\ 0 & 1 \end{pmatrix}$ normalizes $\text{Im} \rho$, and $\mathcal{L}$ is a $\Lambda$-submodule of $\mathfrak{s}\mathfrak{l}_2(\mathfrak{l}_0)$.

**Proof.** First we show that $\mathcal{L}$ is a $\Lambda$-module assuming that $J$ normalizes $\text{Im} \rho$. Since $\mathcal{L}$ is a $\mathbb{Z}_p$-Lie algebra and $\Lambda = \mathbb{Z}_p[\!\![T]\!\!]$, it suffices to show that $x \in \mathcal{L}$ implies $Tx \in \mathcal{L}$. Therefore
Theorem 4.1 normalizes $\text{Im } \rho$. The proof is nearly identical to the proof of [Hida 2015, Lemma 1.4] except we do not require $\xi, \zeta' \in \mathbb{Z}_p$. As in the proof of Proposition 4.10, we know there is an element $\tau = \begin{pmatrix} 1 + T & u \\ 0 & 1 \end{pmatrix} \in \text{Im } \rho_F$. A straightforward matrix calculation shows that $\tau \in \text{Im } \rho_F|_H$. Writing $t = (1 + T)^{1/2}$ and $u' = t^{-1}u$ we see that $\tau' = \begin{pmatrix} t & u' \\ 0 & t^{-1} \end{pmatrix} \in \text{Im } \rho$. Since $\rho_F|_H$ and $\rho$ differ only by a character, their images have the same normalizer. In particular, the matrix $j$ from Theorem 4.1 normalizes $\text{Im } \rho$. Hence the commutator $(\tau', j) \in \text{Im } \rho$ and we can compute

$$(\tau', j) = \begin{pmatrix} 1 & u't(1 - \alpha) \\ 0 & 1 \end{pmatrix}.$$ 

Let $v = \{ x \in \mathbb{L}_0 : \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \in u \}$. Then $v$ is a $\mathbb{Z}_p[\alpha]$-module. Indeed, it is a $\mathbb{Z}_p$-module since we can raise unipotent matrices to $\mathbb{Z}_p$-powers, so it suffices to show that $v$ is closed under multiplication by $\alpha$. This follows by conjugating unipotent elements by $j$. Since $\alpha \not \equiv 1 \mod p$ we have that $1 - \alpha$ is a unit in $\mathbb{Z}_p[\alpha]$. Therefore $u't \in v$. Let $\beta = \tau'^{-1}(1 \ u') \tau' \in \text{Im } \rho$. Then $t^{-1}J = \tau'\beta^{-1}$ (and hence $J$) normalizes $\text{Im } \rho$. 

The proof of Proposition 4.2 is easier when $\alpha \neq -1$, so we start with that case.

**Proof of Proposition 4.2 when $\alpha \neq -1$.** We will show that the finitely generated $\Lambda$-module

$$X := \mathfrak{sl}_2(\mathbb{L}_0)/\mathcal{L}$$

is a torsion $\Lambda$-module. From this it follows that there is a nonzero $\Lambda$-ideal $a$ such that $a\mathfrak{sl}_2(\mathbb{L}_0) \subseteq \mathcal{L}$. Thus

$$(a\mathbb{L}_0)^2\mathfrak{sl}_2(\mathbb{L}_0) \subseteq \mathcal{L} \subseteq L$$

since $\mathbb{L}_0\mathfrak{sl}_2(\mathbb{L}_0) = \mathfrak{sl}_2(\mathbb{L}_0)$. But $(a\mathbb{L}_0)^2\mathfrak{sl}_2(\mathbb{L}_0)$ is the Pink–Lie algebra of $\Gamma_{\mathbb{L}_0}(a\mathbb{L}_0)$ and so $\Gamma_{\mathbb{L}_0}(a\mathbb{L}_0) \subseteq G_2 \subseteq G$, as desired.

To show that $X$ is a finitely generated $\Lambda$-module, recall that the arithmetic prime $P$ in the statement of Proposition 4.2 is a height one prime of $\Lambda$. By Nakayama’s lemma it suffices to show that $X/PX$ is $\Lambda/P$-torsion. The natural epimorphism $\mathfrak{sl}_2(\mathbb{L}_0)/P\mathfrak{sl}_2(\mathbb{L}_0) \twoheadrightarrow X/PX$ has kernel $\mathcal{L} \cdot P\mathfrak{sl}_2(\mathbb{L}_0)/P\mathfrak{sl}_2(\mathbb{L}_0)$, so

$$X/PX \cong (\mathfrak{sl}_2(\mathbb{L}_0)/P\mathbb{L}_0)/(\mathcal{L} \cdot P\mathfrak{sl}_2(\mathbb{L}_0)/P\mathfrak{sl}_2(\mathbb{L}_0)).$$

We use the following notation:

$$\mathcal{L} = L(\mathbb{G}_P \mathbb{L}_0) = \text{the Pink–Lie algebra of } \mathbb{G}_P \mathbb{L}_0.$$
By the functoriality of Pink’s construction, the canonical surjection \( \mathbb{I}_0 \to \mathbb{I}_0/P\mathbb{I}_0 \) induces surjections

\[
L[\lambda] \to \overline{L}[\lambda]
\]

for all \( \lambda \in \{\alpha, 1, \alpha^{-1}\} \). Therefore the canonical linear map \( L \to \overline{L} \) is also a surjection. That is, \( L \cdot P\mathfrak{sl}_2(\mathbb{I}_0)/P\mathfrak{sl}_2(\mathbb{I}_0) = \overline{L} \) and so \( X/PX \cong \mathfrak{sl}_2(\mathbb{I}_0/P\mathbb{I}_0)/\overline{L} \).

Since each \( \mathfrak{a}_Q^2 \) is a nonzero \( \mathbb{I}_0/Q \)-ideal, it follows that \( \mathfrak{a}_Q^2 \mathfrak{sl}_2(\mathbb{I}_0/Q) \subseteq \mathfrak{sl}_2(\mathbb{I}_0/P\mathbb{I}_0) \subseteq \mathfrak{a}_Q^2 \mathfrak{sl}_2(\mathbb{I}_0/Q) \). Finally, the inclusions

\[
\mathfrak{a}_Q^2 \mathfrak{sl}_2(\mathbb{I}_0/Q) \subseteq \overline{L} \subseteq \mathfrak{sl}_2(\mathbb{I}_0/P\mathbb{I}_0) \subseteq \mathfrak{a}_Q^2 \mathfrak{sl}_2(\mathbb{I}_0/Q)
\]

show that \( \mathfrak{sl}_2(\mathbb{I}_0/P\mathbb{I}_0)/\overline{L} \cong X/PX \) is \( \Lambda/P \)-torsion.

Let

\[
v = \left\{ v \in \mathbb{I}_0 : \begin{pmatrix} 0 & v \\ 0 & 0 \end{pmatrix} \in u \right\} \quad \text{and} \quad v' = \left\{ v \in \mathbb{I}_0 : \begin{pmatrix} 0 & 0 \\ v & 0 \end{pmatrix} \in u' \right\}.
\]

**Definition 4.6.** A \( \Lambda \)-lattice in \( Q(\mathbb{I}_0) \) is a finitely generated \( \Lambda \)-submodule \( M \) of \( Q(\mathbb{I}_0) \) such that the \( Q(\Lambda) \)-span of \( M \) is equal to \( Q(\mathbb{I}_0) \). If in addition \( M \) is a subring of \( \mathbb{I}_0 \) then we say \( M \) is a \( \Lambda \)-order.

**Proof of Proposition 4.2 when \( \alpha = -1 \).** We show in Lemmas 4.7 and 4.8 that \( v \) and \( v' \) are \( \Lambda \)-lattices in \( Q(\mathbb{I}_0) \). To do this we use the fact that the local Galois representation \( \rho_F \mid_{D_{\mathbb{I}_0}} \) is indecomposable [Ghate and Vatsal 2004; Zhao 2014].

We then show in Proposition 4.9 that any \( \Lambda \)-lattice in \( Q(\mathbb{I}_0) \) contains a nonzero \( \mathbb{I}_0 \)-ideal. Let \( b \) and \( b' \) be nonzero \( \mathbb{I}_0 \)-ideals such that \( b \subseteq v \) and \( b' \subseteq v' \). Let \( a_0 = bb' \).

Then from the definitions of \( v \), \( v' \), and \( L \), we find that

\[
L \supseteq \mathfrak{a}_Q^2 \mathfrak{sl}_2(\mathbb{I}_0).
\]

By Pink’s theory it follows that \( \mathfrak{G} \supseteq \Gamma_{\mathbb{I}_0}(a_0) \).
Finally, we prove the three key facts used in the proof of Proposition 4.2 when \( \alpha = -1 \).

**Lemma 4.7.** With notation as above, \( \nu \) is a \( \Lambda \)-lattice in \( Q(\llbracket 0 \rrbracket) \).

**Proof.** Let \( \bar{L} = L(\bar{\mathbb{G}} \rho |_{\bar{0}}) \). Recall that \( L[1] \) surjects onto \( \bar{L}[1] \). Now \( \bar{L}[1] \) contains

\[
\left\{ \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix} : a \in \oplus_{Q/P} \bar{a}^2_Q \right\},
\]

and \( \oplus_{Q/P} \bar{a}^2_Q \) is a \( \Lambda / P \)-lattice in \( Q(\llbracket 0 \rrbracket/P(\llbracket 0 \rrbracket)) \). It follows from Nakayama’s lemma that the set of entries in the matrices of \( L[1] \) contains a \( \Lambda \)-lattice \( \alpha \) for \( Q(\llbracket 0 \rrbracket) \).

By a theorem of Ghate and Vatsal [2004] (later generalized by Hida [2013] and Zhao [2014]) we know that \( \rho_F |_{D_p} \) is indecomposable. Hence there is a matrix in the image of \( \rho \) whose upper right entry is nonzero. This produces a nonzero nilpotent matrix in \( L_1 \). Taking the Lie bracket of this matrix with a nonzero element of \( L[1] \) produces a nonzero nilpotent matrix in \( L \) which we will call \( \begin{pmatrix} 0 & v \\ 0 & 0 \end{pmatrix} \). Note that for any \( a \in \alpha \) we have

\[
\begin{pmatrix} 0 & 2av \\ 0 & 0 \end{pmatrix} = \left[ \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}, \begin{pmatrix} 0 & v \\ 0 & 0 \end{pmatrix} \right] \in L.
\]

Thus the lattice \( \alpha \nu \) is contained in \( \nu \), so \( Q(\Lambda) \nu = Q(\llbracket 0 \rrbracket) \). The fact that \( \nu \) is finitely generated follows from the fact that \( \Lambda \) is noetherian and \( \nu \) is contained in the finitely generated \( \Lambda \)-module \( \llbracket 0 \rrbracket \). \( \square \)

**Lemma 4.8.** With notation as above, \( \nu' \) is a \( \Lambda \)-lattice in \( Q(\llbracket 0 \rrbracket) \).

**Proof.** Let \( \bar{c} \in \oplus_{Q/P} \bar{a}^2_Q \). Since \( L[-1] \) surjects to \( \bar{L}[-1] \) there is some \( \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \) \( \in L \) such that \( b \in P(\llbracket 0 \rrbracket) \) and \( c \mod P(\llbracket 0 \rrbracket) \equiv \bar{c} \). Since \( \nu \) is a \( \Lambda \)-lattice in \( Q(\llbracket 0 \rrbracket) \) by Lemma 4.7, it follows that there is some nonzero \( \alpha \in \Lambda \) such that \( \alpha \beta \in \nu \).

We claim that there is some nonzero \( \beta \in \Lambda \) for which \( \begin{pmatrix} 0 & \alpha b \\ \beta c & 0 \end{pmatrix} \) \( \in L \). Assuming the existence of \( \beta \), since \( \alpha \beta \in \nu \) it follows that \( \beta \nu \in \nu' \). That is, \( \nu \in Q(\Lambda) \nu' \). Since \( \bar{c} \) runs over \( \oplus_{Q/P} \bar{a}^2_Q \), it follows from Nakayama’s lemma that \( \nu' \) is a \( \Lambda \)-lattice in \( Q(\llbracket 0 \rrbracket) \).

To see that \( \beta \) exists, recall that \( L \) is normalized by the matrix \( J = \begin{pmatrix} 1 + T & 0 \\ 0 & 1 \end{pmatrix} \) by Lemma 4.5. Thus

\[
\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} + \begin{pmatrix} 0 & Tc \\ ((1 + T)^{-1} - 1)c & 0 \end{pmatrix} = \begin{pmatrix} 1 + T & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \begin{pmatrix} (1 + T)^{-1} & 0 \\ 0 & 1 \end{pmatrix} \in L.
\]

Write \( \alpha = f(T) \) as a power series in \( T \). Since \( (1 + T)^{-1} - 1 \) is divisible by \( T \), we can evaluate \( f \) at \( (1 + T)^{-1} - 1 \) to get another element of \( \mathbb{Z}_p \llbracket T \rrbracket \). Taking \( \beta = f((1 + T)^{-1} - 1) \), the calculation above shows the desired inclusion:

\[
\begin{pmatrix} 0 & \alpha b \\ \beta c & 0 \end{pmatrix} \in L. \]
Proposition 4.9. Every $\Lambda$-lattice in $Q(\mathfrak{l}_0)$ contains a nonzero $\mathfrak{l}_0$-ideal.

Proof. Let $M$ be a $\Lambda$-lattice in $Q(\mathfrak{l}_0)$. Define

$$R = \{ x \in \mathfrak{l}_0 : xM \subseteq M \}.$$ 

Then $R$ is a subring of $\mathfrak{l}_0$ that is also a $\Lambda$-lattice for $Q(\mathfrak{l}_0)$. Thus $R$ is a $\Lambda$-order in $\mathfrak{l}_0$, and $M$ is an $R$-module. Therefore

$$c := \{ x \in \mathfrak{l}_0 : x\mathfrak{l}_0 \subseteq R \}$$

is a nonzero $\mathfrak{l}_0$-ideal. Note that $Q(R) = Q(\mathfrak{l}_0) = Q(\Lambda)M$. Since $M$ is a finitely generated $\Lambda$-module there is some nonzero $r \in \mathfrak{l}_0$ such that $rM \subseteq R$. As $rM$ is still a $\Lambda$-lattice for $Q(\mathfrak{l}_0)$, by replacing $M$ with $rM$ we may assume that $M$ is an $R$-ideal.

Now consider $a = c \cdot (M \mathfrak{l}_0)$, where $M \mathfrak{l}_0$ is the ideal generated by $M$ in $\mathfrak{l}_0$. Note that $a$ is a nonzero $\mathfrak{l}_0$-ideal since both $c$ and $M \mathfrak{l}_0$ are nonzero $\mathfrak{l}_0$-ideals. To see that $a \subseteq M$, let $x \in \mathfrak{l}_0$ and $c \in c$. Then $xc \in R$ by definition of $c$. If $a \in M$ then $xca \in M$ since $M$ is an $R$-ideal. Thus $xca \in M$, so $a \subseteq M$. □

Remark. Note that the only property of $\mathfrak{l}_0$ that is used in the proof of Proposition 4.9 is that $\mathfrak{l}_0$ is a $\Lambda$-order in $Q(\mathfrak{l}_0)$. Thus, once we have shown that $\rho$ (or $\rho_F$) is $\mathfrak{l}_0$-full, it follows that the representation is $R$-full for any $\Lambda$-order $R$ in $Q(\mathfrak{l}_0)$. In particular, if $\mathfrak{l}_0$ is the maximal $\Lambda$-order in $Q(\mathfrak{l}_0)$ then $\rho_F$ is $\mathfrak{l}_0$-full.

Finally, we show that for the purposes of proving $\mathfrak{l}_0$-fullness it suffices to work with $\rho$ instead of $\rho_F$.

Proposition 4.10. The representation $\rho_F$ is $\mathfrak{l}_0$-full if and only if $\rho$ is $\mathfrak{l}_0$-full.

Proof. Note that $\text{Im} \rho_F |_{H_0} \cap \text{SL}_2(\mathfrak{l}_0) \subseteq \text{Im} \rho$ by definition. Thus if $\rho_F$ is $\mathfrak{l}_0$-full then so is $\rho$.

Now assume that $\rho$ is $\mathfrak{l}_0$-full. As in the proof of [Hida 2015, Theorem 8.2], let $\Gamma = \{(1 + T)^s : s \in \mathbb{Z}_p\}$ and

$$\mathbb{K} = \{ x \in \rho_F(H_0) : \det x \in \Gamma \}.$$ 

Note that $\mathbb{K}$ is a finite index subgroup of $\text{Im} \rho_F$. Since $F$ is ordinary and non-CM we can find an element of the form $\tau = \begin{pmatrix} 1 + T & a \\ 0 & 1 \end{pmatrix} \in \text{Im} \rho_F$ [Hida 2000a, Theorem 4.3.2]. Let $n = [G_Q : H_0]$. By replacing $\Gamma$ with $\{(1 + T)^{ns} : s \in \mathbb{Z}_p\}$ and $\tau$ with $\tau^n$, we may assume that $\tau \in \mathbb{K}$.

Let $\mathbb{S} = \mathbb{K} \cap \text{SL}_2(\mathfrak{l}_0)$ and $\mathcal{T} = \{ \tau^s : s \in \mathbb{Z}_p\}$. We can write $\mathbb{K}$ as a semidirect product

$$\mathbb{K} = \mathcal{T} \ltimes \mathbb{S}.$$ 

Indeed, given $x \in \mathbb{K}$ there is a unique $s \in \mathbb{Z}_p$ such that $\det x = (1 + T)^s$. Thus we identify $x$ with $(\tau^s, \tau^{-s}x) \in \mathcal{T} \ltimes \mathbb{S}$. 

\[176\] Jaclyn Lang
Let \( K' \) be the image of \( K \) under the natural map

\[ \Phi : K \rightarrow \text{Im} \rho, \quad x \mapsto x (\det x)^{-1/2}. \]

Then \( K' \) is a finite index subgroup of \( \text{Im} \rho \) and therefore contains \( \Gamma_{l_0}(a) \) for some nonzero \( l_0 \)-ideal \( a \) since \( \rho \) is \( l_0 \)-full. Note that \( \ker \Phi \) is precisely the set of scalar matrices in \( K \). Therefore, for some \( 0 \leq r \leq \infty \),

\[ \ker \Phi \cong \{(1 + T)^s : s \in p^r \mathbb{Z}_p\}, \]

where \( r = \infty \) means \( \ker \Phi = \{1\} \). If \( r \neq \infty \) then by passing to finite index subgroups of \( K, K' \), and \( \Gamma \) we may assume that \( \ker \Phi = \Gamma \). Thus, given any \( y \in \Gamma_{l_0}(a) \) we can find \( x \in K \) such that \( \Phi(x) = y \). Let \( s \in \mathbb{Z}_p \) such that \( \det x = (1 + T)^{s/2} \). Then the scalar matrix \( (1 + T)^{-s/2} \) is in \( \Gamma \) hence in \( K \). Hence \( x (1 + T)^{-s/2} \in \mathbb{S} \) and \( \Phi(x (1 + T)^{-s/2}) = y \). But \( \Phi \) is the identity on \( \mathbb{S} \), so \( y = x (1 + T)^{-s/2} \in \mathbb{S} \). Therefore \( \Gamma_{l_0}(a) \subseteq \mathbb{S} \) and \( \rho_F \) is \( l_0 \)-full.

It remains to deal with the case when \( \ker \Phi = \{1\} \). In this case \( \Phi \) is an isomorphism onto \( K' \) and we can use \( \Phi^{-1} \) to get a continuous group homomorphism from \( K' \) onto \( \mathbb{Z}_p \):

\[ s : K' \cong K \cong \mathbb{T} \times \mathbb{S} \twoheadrightarrow \mathbb{T} \cong \mathbb{Z}_p. \]

Note that \( \ker s = \mathbb{S} \), so we want to show that \( \ker s \) is \( l_0 \)-full. By assumption there is a nonzero \( l_0 \)-ideal \( a_0 \) such that \( \Gamma_{l_0}(a_0) \subseteq K' \). Let \( v = \{ b \in a_0 : \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in \ker s \} \) and \( v' = \{ c \in a_0 : \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \in \ker s \} \). Both \( v \) and \( v' \) are \( \Lambda \)-lattices in \( Q(l_0) \). We shall prove this for \( v \); the proof for \( v' \) is similar. Note that \( v \) is a \( \mathbb{Z}_p \)-module: if \( \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in \mathbb{S} \) then \( \begin{pmatrix} 1 & sb \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}^s \in \mathbb{S} \) since \( \mathbb{S} \) is closed (as it is the determinant 1 image of a Galois representation). To see that \( v \) is a \( \Lambda \)-module, recall that \( \mathbb{S} \) is normalized by \( J = \begin{pmatrix} 1 + T & 0 \\ 0 & 1 \end{pmatrix} \) by the proof of Lemma 4.5. Therefore conjugation by \( J \) gives an action of \( T \) on \( v \) as in the proof of Lemma 4.5. Now we consider the \( \Lambda \)-module \( a_0/v \) which, as a group, is isomorphic to a closed subgroup of \( \mathbb{Z}_p \). Therefore \( a_0/v \) is a torsion \( \Lambda \)-module. Since \( a_0 \) is a \( \Lambda \)-lattice in \( Q(l_0) \) it follows that \( v \) must also be a \( \Lambda \)-lattice in \( Q(l_0) \), as claimed.

We have shown that there are nonzero \( l_0 \)-ideals \( b \subseteq v \) and \( b' \subseteq v' \) such that the Pink–Lie algebra \( L(\mathbb{S}) \) contains

\[ \left\{ \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} : b \in b \right\} \cup \left\{ \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} : c \in b' \right\}. \]

By letting \( c = bb' \) and taking Lie brackets of the upper and lower nilpotent matrices above we find that \( L(\mathbb{S}) \cong c^2 \mathfrak{sl}_2(l_0) \). Therefore \( \mathbb{S} \) is \( l_0 \)-full, as desired. \( \Box \)
5. Open image in product

The purpose of this section is to prove the following reduction step in the proof of Theorem 2.4.

**Proposition 5.1.** Assume that $|\mathbb{F}| \neq 3$. Fix an arithmetic prime $P$ of $\Lambda$. Assume that for every prime $Q$ of $\mathbb{Q}$ lying over $P$, the image of $\text{Im} \rho$ in $\text{SL}_2(\mathbb{Q}_0/Q)$ is open. Then the image of $\text{Im} \rho$ in $\prod_{Q|P} \text{SL}_2(\mathbb{Q}_0/Q)$ is open in the product topology.

Thus if we can show that there is some arithmetic prime $P$ of $\Lambda$ satisfying the hypothesis of Proposition 5.1, then combining the above result with Proposition 4.2 yields Theorem 2.4.

Fix an arithmetic prime $P$ of $\Lambda$ satisfying the hypothesis of Proposition 5.1. Note that $\mathbb{Z}_p$ does not contain any $p$-power roots of unity since $p > 2$. Therefore $P = P_{k,1}$ for some $k \geq 2$. Recall that $G = \text{Im} \rho \cap \Gamma(m_0)$, and write $\mathcal{G}$ for the image of $G$ in $\prod_{Q|P} \text{SL}_2(\mathbb{Q}_0/Q)$. We begin our proof of Proposition 5.1 with the following lemma of Ribet which allows us to reduce to considering products of only two copies of $\text{SL}_2$.

**Lemma 5.2 [Ribet 1975, Lemma 3.4].** Let $S_1, \ldots, S_t$ be profinite groups. Assume for each $i$ that the following condition is satisfied: for each open subgroup $U$ of $S_i$, the closure of the commutator subgroup of $U$ is open in $S_i$. Let $\mathcal{G}$ be a closed subgroup of $S = S_1 \times \cdots \times S_t$ that maps to an open subgroup of each group $S_i \times S_j (i \neq j)$. Then $\mathcal{G}$ is open in $S$.

Apply this lemma to our situation with $\{S_1, \ldots, S_t\} = \{\text{SL}_2(\mathbb{Q}_0/Q) : Q|P\}$ and $G = \mathcal{G}$. The lemma implies that it is enough to prove that for all primes $Q_1 \neq Q_2$ of $\mathbb{Q}_0$ lying over $P$, the image $G$ of $\mathcal{G}$ under the projection to $\text{SL}_2(\mathbb{Q}_0/Q_1) \times \text{SL}_2(\mathbb{Q}_0/Q_2)$ is open. We shall now consider what happens when this is not the case. Indeed, the reader should be warned that the rest of this section is a proof by contradiction.

**Proposition 5.3.** Let $P$ be an arithmetic prime of $\Lambda$ satisfying the hypotheses of Proposition 5.1, and assume $|\mathbb{F}| \neq 3$. Let $Q_1$ and $Q_2$ be distinct primes of $\mathbb{Q}_0$ lying over $P$. Let $\mathfrak{p}_1$ be a prime of $\mathbb{Q}$ lying over $Q_1$. If $G$ is not open in $\text{SL}_2(\mathbb{Q}_0/Q_1) \times \text{SL}_2(\mathbb{Q}_0/Q_2)$ then there is an isomorphism $\sigma : \mathbb{Q}_0/Q_1 \cong \mathbb{Q}_0/Q_2$ and a character $\varphi : H \to \mathbb{Q}(\mathbb{Q}_0/Q_2)^\times$ such that

$$\sigma(a(\ell, f_{\mathfrak{p}_1})) = \varphi(\ell)a(\ell, f_{\mathfrak{p}_2})$$

for all primes $\ell$ for which $\text{Frob}_\ell \in H$.

**Proof.** Our strategy is to mimic the proof of [Ribet 1975, Theorem 3.5]. Let $G_i$ be the projection of $G$ to $\text{SL}_2(\mathbb{Q}_0/Q_i)$, so $G \subseteq G_1 \times G_2$. By hypothesis $G_i$ is open in $\text{SL}_2(\mathbb{Q}_0/Q_i)$. Let $\pi_i : G \to G_i$ be the projection maps and set $N_1 = \ker\pi_2$ and $N_2 = \ker\pi_1$. Though a slight abuse of notation, we view $N_i$ as a subset of $G_i$. Therefore
Goursat’s lemma implies that the image of $G$ in $G_1/N_1 \times G_2/N_2$ is the graph of an isomorphism

$$\alpha : G_1/N_1 \cong G_2/N_2.$$ 

Since $G$ is not open in $G_1 \times G_2$ by hypothesis, either $N_1$ is not open in $G_1$ or $N_2$ is not open in $G_2$. (Otherwise $N_1 \times N_2$ is open and hence $G$ is open.) Without loss of generality we may assume that

$$\eta \in \text{Hom}(G_1, Q(I_0/Q_2)^\times), \ y \in \text{GL}_2(Q(I_0/Q_2))\text{ and } \sigma : I_0/Q_1 \cong I_0/Q_2\text{ is a ring isomorphism. By } \sigma(x)\text{ we mean that we apply } \sigma\text{ to each entry of the matrix } x.$$ 

For any $g \in G$ we can write $g = (x, y)$ with $x \in G_1, y \in G_2$. Since $G$ is the graph of $\alpha$ we have $\alpha(x) = y$. By definition of $G$ there is some $h \in H$ such that $x = \mathcal{P}_1(\rho(h))$ and $y = \mathcal{P}_2(\rho(h))$. Recall that for almost all primes $\ell$ for which $\text{Frob}_\ell \in H$ we have $\text{tr}(\rho(\text{Frob}_\ell)) = (\sqrt{\det \rho_F(\text{Frob}_\ell)})^{-1} a(\ell, F)$. Furthermore $\det \rho_F(\text{Frob}_\ell) \mod P = \chi(\ell)\ell^{k-1}$ since $P = P_{k,1}$. Using these facts together with Equation (2) we see that for almost any $\text{Frob}_\ell \in H$ we have

$$\sigma(a(\ell, f_{\mathcal{P}_1})) = \varphi(\ell)a(\ell, f_{\mathcal{P}_2}),$$ 

where

$$\varphi(\ell) := \eta^{-1}(\mathcal{P}_1(\rho(\text{Frob}_\ell))) \frac{\sigma(\sqrt{\chi(\ell)\ell^{k-1}})}{\sqrt{\chi(\ell)\ell^{k-1}}},$$ 

as claimed. □

To finish the proof of Proposition 5.1 we need to remove the condition that $\text{Frob}_\ell \in H$ from the conclusion of Proposition 5.3. That is, we would like to
show that there is an isomorphism $\tilde{\sigma} : \mathcal{P} / \mathcal{P}_1' \cong \mathcal{P} / \mathcal{P}_2' \times \mathcal{P}_2''$ extending $\sigma$ and a character $\tilde{\varphi} : G_{\mathcal{Q}} \to Q(\mathcal{P} / \mathcal{P}_2') \times$ extending $\varphi$ such that

$$\tilde{\sigma}(a(\ell, f_{\mathcal{P}_1})) = \tilde{\varphi}(\ell)a(\ell, f_{\mathcal{P}_2})$$

for almost all primes $\ell$. If we can do this, then applying Theorem 3.1 allows us to lift $\tilde{\sigma}$ to an element of $\Gamma$ that sends $\mathcal{P}_1'$ to $\mathcal{P}_2'$. (We also need to verify that $\tilde{\varphi}$ takes values in $\mathbb{Z}_p[\chi]$ in order to apply Theorem 3.1.) But this is a contradiction since $\mathcal{P}_1'$ and $\mathcal{P}_2'$ lie over different primes of $\mathfrak{p}_0$. Hence it follows from Proposition 5.3 that $G$ must be open in $\text{SL}_2(\mathfrak{p}_0 / \mathcal{Q}_1) \times \text{SL}_2(\mathfrak{p}_0 / \mathcal{Q}_2)$ and Lemma 5.2 implies Proposition 5.1.

We show the existence of $\tilde{\sigma}$ and $\tilde{\varphi}$ using obstruction theory as developed in [Hida 2000b, §4.3.5]. For the sake of notation, we briefly recall the theory here. For the proofs we refer the reader to [Hida 2000b]. Let $K$ be a finite extension of $\mathbb{Q}_p$, $n \in \mathbb{Z}^+$, and $r : H \to \text{GL}_n(K)$ be an absolutely irreducible representation. For all $g \in G_{\mathcal{Q}}$ define a twisted representation on $H$ by $r^g(h) := r(ghg^{-1})$. Assume the following condition:

$$r \cong r^g \text{ over } K \text{ for all } g \in G_{\mathcal{Q}}. \hspace{1cm} (3)$$

Under Hypothesis (3) it can be shown that there is a function $c : G_{\mathcal{Q}} \to \text{GL}_n(K)$ with the following properties:

1. $r = c(g)^{-1}r^g c(g)$ for all $g \in G_{\mathcal{Q}}$;
2. $c(hg) = r(h)c(g)$ for all $h \in H, g \in G_{\mathcal{Q}}$;
3. $c(1) = 1$.

As $r$ is absolutely irreducible, it follows that $b(g, g') := c(g)c(g')c(gg')^{-1}$ is a 2-cocycle with values in $K^\times$. In fact $b$ factors through $\Delta := G_{\mathcal{Q}} / H$ and hence represents a class in $H^2(\Delta, K^\times)$. We call this class $\text{Ob}(r)$. It is independent of the function $c$ satisfying the above three properties. The class $\text{Ob}(r)$ measures the obstruction to lifting $r$ to a representation of $G_{\mathcal{Q}}$. We say a continuous representation $\tilde{r} : G_{\mathcal{Q}} \to \text{GL}_n(K)$ is an extension of $r$ to $G_{\mathcal{Q}}$ if $\tilde{r}|_H = r$.

**Proposition 5.4.** (1) There is an extension $\tilde{r}$ of $r$ to $G_{\mathcal{Q}}$ if and only if $\text{Ob}(r) = 0 \in H^2(\Delta, K^\times)$.

(2) If $\text{Ob}(r) = 0$ and $\tilde{r}$ is an extension of $r$ to $G_{\mathcal{Q}}$, then all other extensions of $r$ to $G_{\mathcal{Q}}$ are of the form $\tilde{r} \otimes \psi$ for some character $\psi : \Delta \to K^\times$.

For ease of notation we shall write $K_i = Q(\mathcal{P}_i)$ and $E_i = Q(\mathfrak{p}_0 / \mathcal{Q}_i)$. Write $\rho_i : G_{\mathcal{Q}} \to \text{GL}_2(K_i)$ for $\rho_{f_{\mathcal{P}_i}}$. By Theorem 4.1 we see that $\rho_i|_H$ takes values in $\text{GL}_2(E_i)$. Proposition 5.3 gives an isomorphism $\sigma : E_1 \cong E_2$ and a character $\varphi : H \to E_2^\times$ such that

$$\text{tr}(\rho_1|_H) = \text{tr}(\rho_2|_H \otimes \varphi).$$
In order to use obstruction theory to show the existence of \( \tilde{\sigma} \) and \( \tilde{\varphi} \) we must show that all of the representations in question satisfy Hypothesis (3).

**Lemma 5.5.** Let \( L_i \) be a finite extension of \( K_i \). View \( \rho_1 \) as a representation over \( L_1 \) and \( \rho_2|_H, \rho_1|_H, \rho_2|_H \otimes \varphi \), and \( \varphi \) as representations over \( L_2 \). Then \( \rho_1|_H, \rho_1|_H, \rho_2|_H \otimes \varphi \), and \( \varphi \) all satisfy Hypothesis (3). Furthermore we have \( \text{Ob}(\rho_1|_H) = 0 \), \( \text{Ob}(\rho_1|_H^m) = \text{Ob}(\rho_2|_H \otimes \varphi) \), and

\[
\text{Ob}(\rho_2|_H \otimes \varphi) = \text{Ob}(\rho_2|_H) + \text{Ob}(\varphi) \in H^2(\Delta, (L_2)^\times).
\]

**Proof.** Recall that a continuous representation of a compact group over a field of characteristic 0 is determined up to isomorphism by its trace. Therefore to verify (3) it suffices to show that if \( r \) is any of the representations listed in the statement of the lemma, then

\[
\text{tr} r = \text{tr} r^g
\]

for all \( g \in G_Q \). This is obvious when \( r = \rho_1|_H \) or \( \rho_2|_H \) since both extend to representations of \( G_Q \) and hence

\[
\text{tr} \rho_i^g(h) = \text{tr} \rho_i(g) \rho_i(h) \rho_i(g)^{-1} = \text{tr} \rho_i(h).
\]

Since \( \rho_i \) is an extension of \( \rho_1|_H \) and \( L_i \supseteq K_i \) we have \( \text{Ob}(\rho_1|_H) = 0 \).

When \( r = \rho_1|_H^m \), let \( \tau : K_1 \hookrightarrow \overline{Q}_p \) be an extension of \( \sigma \). Then \( \rho_1^\tau \) is an extension of \( \rho_1|_H^m \) and hence we can use the same argument as above to conclude that \( \text{tr} \rho_1|_H^m = \text{tr} (\rho_1|_H^m)^g \). (Note that for this particular purpose, we do not care about the field in which \( \tau \) takes values.)

When \( r = \rho_2|_H \otimes \varphi \), recall that \( \text{tr} \rho_1|_H^m = \varphi \text{ tr} \rho_2|_H \). Since both \( \rho_1|_H^m \) and \( \rho_2|_H \) satisfy Hypothesis (3) so does \( \rho_2|_H \otimes \varphi \). Furthermore, \( \text{tr} \rho_1|_H^m = \text{tr}(\rho_2|_H \otimes \varphi) \) implies that \( \rho_1|_H^m \cong \rho_2|_H \otimes \varphi \) and hence \( \text{Ob}(\rho_1|_H^m) = \text{Ob}(\rho_2|_H \otimes \varphi) \).

Since \( (\rho_1|_H^m)^g \cong \rho_2|_H^m \otimes \varphi^g \) for any \( g \in G_Q \) and since both \( \rho_i|_H \) satisfy (3) we see that

\[
\varphi^g \text{ tr} \rho_2|_H = \varphi \text{ tr} \rho_2|_H. \tag{4}
\]

Thus if we know \( \text{tr} \rho_2|_H \) is nonzero sufficiently often then we can deduce that \( \varphi \) satisfies (3). More precisely, let \( m \in \mathbb{Z}^+ \) be the conductor for \( \varphi \), so \( \varphi : (\mathbb{Z}/m\mathbb{Z})^\times \to \overline{Q}^\times \). Then we have a surjection

\[
H \to \text{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q}) \cong (\mathbb{Z}/m\mathbb{Z})^\times \text{ with kernel } \kappa.
\]

Choose a set \( S \) of coset representatives of \( \kappa \) in \( H \), so \( H = \bigsqcup_{s \in S} s \kappa \). If we can show that \( \text{tr} \rho_2(s \kappa) \neq \{0\} \) for all \( s \in S \), then it follows from Equation (4) that \( \varphi^g \varphi = \varphi \) for all \( g \in G_Q \). Recall that \( \rho_2 \) is a Galois representation attached to a classical modular form, and so by Ribet [1980; 1985] and Momose’s [1981] result we know that its image is open. (See Theorem 6.1 for a precise statement of their result.) Then the restriction of \( \rho_2 \) to any open subset of \( G_Q \) also has open image and hence \( \text{tr} \rho_2 \) is not identically zero. Each \( s \kappa \) is open in \( G_Q \), so \( \varphi^g \varphi = \varphi \).
Finally, note that if \( c : G_\mathbb{Q} \to \text{GL}_2(L_2) \) is a function satisfying conditions 1–3 above for \( r = \rho_2|_H \) and \( \eta : G_\mathbb{Q} \to L_2^\times \) is a function satisfying conditions 1–3 above for \( \varphi \), then \( \eta \circ c \) is a function satisfying conditions 1–3 for \( \rho_2|_H \otimes \varphi \). From this it follows that \( \text{Ob}(\rho_2|_H \otimes \varphi) = \text{Ob}(\rho_2|_H) + \text{Ob}(\varphi) \). \( \square \)

With \( L_i \) as in the previous lemma, suppose there is an extension \( \tilde{\sigma} : L_1 \cong L_2 \) of \( \sigma \) and an extension \( \tilde{\psi} : G_\mathbb{Q} \to L_2^\times \) of \( \varphi \). We now show that this gives us the desired relation among traces.

**Lemma 5.6.** If there exists extensions \( \tilde{\sigma} \) of \( \sigma \) and \( \tilde{\psi} \) of \( \varphi \), then there exists a character \( \eta : G_\mathbb{Q} \to L_2^\times \) that is also a lift of \( \varphi \) such that \( \rho_1^\tilde{\sigma} \cong \rho_2 \otimes \eta \).

**Proof.** Note that since \( F \) does not have CM, \( \rho_1|_H \) and \( \rho_2|_H \) are absolutely irreducible by results of Ribet [1977]. For any absolutely irreducible representation \( \pi : G_\mathbb{Q} \to \text{GL}_2(L_2) \) Frobenius reciprocity gives

\[
\langle \pi, \text{Ind}_{G_H}^G(\rho_1|_H^\sigma) \rangle_{G_\mathbb{Q}} = \langle \pi|_H, \rho_1^\sigma \rangle_H = \langle \pi|_H, \rho_2|_H \otimes \varphi \rangle_H. \tag{5}
\]

Thus if \( \pi \) is a 2-dimensional irreducible constituent of \( \text{Ind}(\rho_1|_H^\varphi) \) then \( \rho_1^\sigma \) is a constituent of \( \pi|_H \). As both are 2-dimensional, it follows that \( \rho_1^\sigma \cong \pi|_H \) and thus \( \pi \) is an extension of \( \rho_1^\sigma \). Since \( \tilde{\sigma} \) exists by hypothesis, we know that \( \rho_1^\tilde{\sigma} \) is also an extension of \( \rho_1^\sigma \).

Since \( \tilde{\psi} \) exists by hypothesis, we can take \( \pi = \rho_2 \otimes \tilde{\psi} \). Then (5) implies that \( \pi \) is an irreducible constituent of \( \text{Ind}_{G_H}^G(\rho_1|_H^\tilde{\sigma}) \). By Proposition 5.4 there is a character \( \psi : \Delta \to L_2^\times \) such that \( \rho_2 \otimes \tilde{\psi} \cong \rho_1^\tilde{\sigma} \otimes \psi \). That is,

\[
\rho_1^\tilde{\sigma} \cong \rho_2 \otimes (\tilde{\psi} \psi^{-1}).
\]

Setting \( \eta = \tilde{\psi} \psi^{-1} \) gives the desired conclusion. \( \square \)

Finally, we turn to showing the existence of \( \tilde{\sigma} \) and \( \tilde{\psi} \). With notation as in Lemma 5.5, suppose there exists \( \tilde{\sigma}^{-1} : L_2 \cong L_1 \) that lifts \( \sigma^{-1} \). Then \( \tilde{\sigma}^{-1} \) induces an isomorphism \( H^2(\Delta, L_2^\times) \cong H^2(\Lambda, L_1^\times) \) that sends \( \text{Ob}(\rho_1|_H^\sigma) \) to \( \text{Ob}(\rho_1|_H) \). It follows from Lemma 5.5 that \( \text{Ob}(\rho_1|_H^\sigma) = 0 \) and hence \( \text{Ob}(\rho_2|_H \otimes \varphi) = 0 \). But \( 0 = \text{Ob}(\rho_2|_H \otimes \varphi) = \text{Ob}(\rho_2|_H) + \text{Ob}(\varphi) = \text{Ob}(\varphi) \), and thus we can extend \( \varphi \) to \( \tilde{\psi} : G_\mathbb{Q} \to L_2^\times \).

The above argument requires that we find \( L_i \supseteq K_i \) such that \( L_1 \) is isomorphic to \( L_2 \) via a lift of \( \sigma \). We can achieve this as follows. Let \( \tau : K_1 \hookrightarrow \overline{\mathbb{Q}}_p \) be an extension of \( \sigma \). Let \( \tau^\tau_2 = K_2 \tau(K_1) \). Let \( \tilde{\sigma}^{-1} : L_2 \hookrightarrow \overline{\mathbb{Q}}_p \) be an extension of \( \tau^{-1} \) and set \( L_1 = \tilde{\sigma}^{-1}(L_2) \). This construction satisfies the desired properties. Applying Lemma 5.6 we see that there is a character \( \eta : G_\mathbb{Q} \to L_2^\times \) such that

\[
\text{tr} \rho_1^\tilde{\sigma} = \text{tr} \rho_2 \otimes \eta. \tag{6}
\]

This is almost what we want. Note that by (6) it follows that \( \tilde{\sigma} \) restricts to an isomorphism from \( (\mathcal{L}/\mathcal{O})[\eta] \) to \( (\mathcal{L}/\mathcal{O})[\eta] \). The only problem is that \( \tilde{\sigma} \) may not
send $\mathcal{V}'/\mathfrak{P}'_1$ to $\mathcal{V}'/\mathfrak{P}'_2$ and $\eta$ may have values in $L_2$ that are not in $(\mathcal{V}'/\mathfrak{P}'_2)^\times$. We shall show that this cannot be the case.

Recall that $\chi$ is the nebentypus of $F$, and $\mathfrak{P}_1$ and $\mathfrak{P}_2$ lie over the arithmetic prime $P_{k,1}$ of $\Lambda$. Thus for almost all primes $\ell$ we have $\det \rho_\ell(\text{Frob}_\ell) = \chi(\ell)\ell^{k-1}$. Applying this to Equation (6) we find that

$$\chi^\sigma(\ell)\ell^{k-1} = \eta^2(\ell)\chi(\ell)\ell^{k-1}.$$  

Recall that $\chi(\ell)$ is a root of unity and hence $\chi^\sigma(\ell)$ is just a power of $\chi(\ell)$. Thus $\eta^2(\ell) \in \mathbb{Z}_p[\chi] \subseteq \mathcal{V}'/\mathfrak{P}'_1$ and hence $[(\mathcal{V}'/\mathfrak{P}'_1)[\eta]:\mathcal{V}'/\mathfrak{P}'_1] \leq 2$. Thus we may assume that $L_2 = \mathbb{K}_2[\eta]$, which is at most a quadratic extension of $\mathbb{K}_2$.

Note that since $\eta^2$ takes values in $\mathcal{V}'/\mathfrak{P}'_1$ we can obtain $(\mathcal{V}'/\mathfrak{P}'_1)[\eta]$ from $\mathcal{V}'/\mathfrak{P}'_2$ by adjoining a 2-power root of unity. (Write $\eta$ as the product of a 2-power order character and an odd order character and note that any odd order root of unity is automatically a square in any ring in which it is an element.)

**Lemma 5.7.** We have $(\mathcal{V}'/\mathfrak{P}'_1)[\eta] = \mathcal{V}'/\mathfrak{P}'_2$ for $i = 1, 2$. Therefore $\tilde{\sigma} : \mathcal{V}'/\mathfrak{P}'_1 \cong \mathcal{V}'/\mathfrak{P}'_2$ and $\eta$ takes values in $\mathbb{Z}_p[\chi]$.

**Proof.** Suppose first that $\mathcal{V}'/\mathfrak{P}'_2 = (\mathcal{V}'/\mathfrak{P}'_2)[\eta]$ but $[(\mathcal{V}'/\mathfrak{P}'_1)[\eta]:\mathcal{V}'/\mathfrak{P}'_1] = 2$. Then we have that $\tilde{\sigma} : (\mathcal{V}'/\mathfrak{P}'_1)[\eta] \cong \mathcal{V}'/\mathfrak{P}'_2$. Note that $(\mathcal{V}'/\mathfrak{P}'_1)[\eta]$ is unramified over $\mathcal{V}'/\mathfrak{P}'_1$ since it is obtained by adjoining a prime-to-$p$ root of unity (namely a 2-power root of unity). Thus the residue field of $(\mathcal{V}'/\mathfrak{P}'_1)[\eta]$ must be a quadratic extension of the residue field $\mathbb{F}$ of $\mathcal{V}'/\mathfrak{P}'_1$. But $\mathbb{F}$ is also the residue field of $\mathcal{V}'/\mathfrak{P}'_2$ and since $(\mathcal{V}'/\mathfrak{P}'_1)[\eta] \cong \mathcal{V}'/\mathfrak{P}'_2$ they must have the same residue field, a contradiction. Therefore we must have $(\mathcal{V}'/\mathfrak{P}'_1)[\eta] = \mathcal{V}'/\mathfrak{P}'_1$.

It remains to deal with the case when $[(\mathcal{V}'/\mathfrak{P}'_1)[\eta]:\mathcal{V}'/\mathfrak{P}'_1] = [(\mathcal{V}'/\mathfrak{P}'_2)[\eta]:\mathcal{V}'/\mathfrak{P}'_2] = 2$. As noted above, these extensions must be unramified and hence the residue field of $(\mathcal{V}'/\mathfrak{P}'_1)[\eta]$ must be the unique quadratic extension $\mathbb{E} = \mathbb{F}[\tilde{\eta}]$ of $\mathbb{F}$. Note that $\tilde{\sigma}$ induces an automorphism $\hat{\sigma}$ of $\mathbb{E}$ that necessarily restricts to an automorphism of $\mathbb{F}$. From $\chi^{\hat{\sigma}} = \eta^2\chi$ we find that

$$\tilde{\chi}^{\hat{\sigma}} = \tilde{\eta}^2\tilde{\chi}.$$  

On the other hand $\hat{\sigma}$ is an automorphism of $\mathbb{F}$ and hence is equal to some power of Frobenius. So we see that for some $s \in \mathbb{Z}$ we have $\tilde{\eta}^2 = \tilde{\chi}^{p^s-1}$. Since $p$ is odd, $p^s - 1$ is even and hence $\tilde{\eta}^2$ takes values in $\mathbb{F}_p[\tilde{\chi}^2]$. Thus $\tilde{\eta}$ takes values in $\mathbb{F}_p[\tilde{\chi}] \subseteq \mathbb{F}$, a contradiction to the assumption that $[\mathbb{F}[\tilde{\eta}]:\mathbb{F}] = 2$.

Since $\eta^2$ takes values in $\mathbb{Z}_p[\chi]$ and $\mathbb{F}_p[\tilde{\eta}] \subseteq \mathbb{F}_p[\tilde{\chi}]$, it follows that in fact $\eta$ must take values in $\mathbb{Z}_p[\chi]$. Hence we may take $L_i = \mathbb{K}_i$ and $\tilde{\sigma} : \mathcal{V}'/\mathfrak{P}'_1 \cong \mathcal{V}'/\mathfrak{P}'_2$. \hfill $\square$

**Proof of Proposition 5.1.** By Lemma 5.2 it suffices to show that, for any two primes $\mathcal{Q}_1 \neq \mathcal{Q}_2$ of $\mathfrak{I}_0$ lying over $P_{k,1}$, the image of $\text{Im} \rho$ in $\text{SL}_2(\mathfrak{I}_0/\mathcal{Q}_1) \times \text{SL}_2(\mathfrak{I}_0/\mathcal{Q}_2)$ is open. Proposition 5.3 says that if that is not the case, then there is an isomorphism
σ : \langle 0 \rangle / Q_1 \cong \langle 0 \rangle / Q_2 \text{ and a character } \varphi : H \to Q(\langle 0 \rangle / Q_2)^\times \text{ such that } \operatorname{tr} \rho_{f,\varphi_1}|_H = \operatorname{tr} \rho_{f,\varphi_2}|_H \otimes \varphi. \text{ The obstruction theory arguments allow us to lift } \sigma \text{ and } \varphi \text{ to } \tilde{\sigma} : \mathcal{P}_1' \cong \mathcal{P}_2' \text{ and } \tilde{\varphi} : G_Q \to Q(\langle 0 \rangle / Q_2)^\times \text{ such that } \operatorname{tr} \rho_{f,\varphi_1} = \operatorname{tr} \rho_{f,\varphi_2} \otimes \tilde{\varphi}. \text{ Theorem 3.1 allows us to lift } \tilde{\sigma} \text{ to an element of } \Gamma \text{ that sends } \mathcal{P}_1' \text{ to } \mathcal{P}_2'. \text{ But } \mathcal{P}_1' \text{ and } \mathcal{P}_2' \text{ lie over different primes of } \langle 0 \rangle \text{ and } \Gamma \text{ fixes } \langle 0 \rangle, \text{ so we reach a contradiction. Therefore the image of } \operatorname{Im} \rho \text{ in the product } \text{SL}_2(\langle 0 \rangle / Q_1) \times \text{SL}_2(\langle 0 \rangle / Q_2) \text{ is open.} \square

6. Proof of main theorem

In this section we use the compatibility between the conjugate self-twists of } F \text{ and those of its classical specializations established in Section 3 to relate } \langle 0 \rangle / Q \text{ to the ring appearing in the work of Ribet [1980; 1985] and Momose [1981]. This allows us to use their results to finish the proof of Theorem 2.4.}

We begin by working the Ribet and Momose. We follow [Ribet 1985] closely. Let } f = \sum_{n=1}^{\infty} a(n, f)q^n \text{ be a classical eigenform of weight } k. \text{ Let } K = \mathbb{Q}((a(n, f) : n \in \mathbb{Z}^+)) \text{ with ring of integers } \mathcal{O}. \text{ Denote by } \Gamma_f \text{ the group of conjugate self-twists of } f. \text{ Let } E = K^{\Gamma_f} \text{ and } H_f = \bigcap_{\sigma \in \Gamma_f} \ker \eta_\sigma. \text{ For any character } \psi, \text{ let } G(\psi) \text{ denote the Gauss sum of the primitive character of } \psi. \text{ For } \sigma, \tau \in \Gamma_f \text{ Ribet defined}

c(\sigma, \tau) := \frac{G(\eta^{-1}_\sigma)G(\eta^{-1}_\tau)}{G(\eta^{-1}_{\sigma\tau})}.

One shows that } c \text{ is a 2-cocycle on } \Gamma_f \text{ with values in } K^\times.

Let } \mathfrak{X} \text{ be the central simple } E \text{-algebra associated to } c. \text{ Then } K \text{ is the maximal commutative semisimple subalgebra of } \mathfrak{X}. \text{ It can be shown that } \mathfrak{X} \text{ has order two in the Brauer group of } E, \text{ and hence there is a 4-dimensional } \text{E-algebra } D \text{ that represents the same element as } \mathfrak{X} \text{ in the Brauer group of } E. \text{ Namely, if } \mathfrak{X} \text{ has order one then } D = M_2(E) \text{ and otherwise } D \text{ is a quaternion division algebra over } E.

For a prime } p, \text{ recall that we have a Galois representation}

\rho_{f,p} : G_Q \to \text{GL}_2(\mathcal{O}_K \otimes \mathbb{Z} / p\mathbb{Z})

associated to } f. \text{ The following theorem is due to Ribet [1980] in the case when } f \text{ has weight 2.}

Theorem 6.1 [Momose 1981]. \text{ We may view } \rho_{f,p}|_{H_f} \text{ as a representation valued in } (D \otimes Q_p)^\times. \text{ Furthermore, letting } n \text{ denote the reduced norm map on } D, \text{ the image of } \rho_{f,p}|_{H_f} \text{ is open in}

\{ x \in (D \otimes Q_p)^\times : nx \in Q_p^\times \}.

In particular, when } D \otimes Q_p \text{ is a matrix algebra, the above theorem tell us that}

\text{Im} \rho_{f,p}|_{H_f} \text{ is open in}

\{ x \in \text{GL}_2(\mathcal{O}_E \otimes \mathbb{Z} / p\mathbb{Z}) : \det x \in (\mathbb{Z} / p\mathbb{Z})^{k-1} \}.\]
Let $p$ be a prime of $\mathcal{O}_E$ lying over $p$, and let $\rho_{f,p}$ be the representation obtained by projecting $\rho_{f,p}|_{H_f}$ to the $\mathcal{O}_{E_p}$-component. Under the assumption that $D \otimes \mathcal{O}_p$ is a matrix algebra Theorem 6.1 implies that $\rho_{f,p}$ is $\mathcal{O}_{E_p}$-full. Finally, Brown and Ghate [2003, Theorem 3.3.1] proved that if $f$ is ordinary at $p$, then $D \otimes \mathcal{O}_p$ is a matrix algebra.

Thus, the Galois representation associated to each classical specialization of our $\mathbb{F}$-adic form $F$ is $\mathcal{O}_{E_p}$-full with respect to the appropriate ring $\mathcal{O}_{E_p}$. We must show that $E_p$ is equal to $Q(\mathbb{I}/\mathbb{Q})$, where $\mathbb{Q}$ corresponds to $p$ in a way we will make precise below.

Recall that we have a fixed embedding $\iota_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Let $\mathfrak{P} \in \text{Spec}(\mathbb{I})(\overline{\mathbb{Q}}_p)$ be an arithmetic prime of $\mathbb{I}$, and let $\mathfrak{Q}$ be the prime of $\mathbb{I}_0$ lying under $\mathfrak{P}$. As usual, let $\mathfrak{P}' = \mathfrak{P} \cap \mathfrak{I}'$. Let $D(\mathfrak{P}'|\mathbb{Q}) \subseteq \Gamma'$ be the decomposition group of $\mathfrak{P}'$ over $\mathbb{Q}$. Let

$$K_{\mathfrak{P}} = \mathbb{Q}(\{\iota_p^{-1}(a(n, f_{\mathfrak{P}})) : n \in \mathbb{Z}^+\}) \subset \overline{\mathbb{Q}},$$

and let $\Gamma_{\mathfrak{P}}$ be the group of all conjugate self-twists of the classical modular form $f_{\mathfrak{P}}$.

Set $E_{\mathfrak{P}} = K_{\mathfrak{P}}^{\Gamma_{\mathfrak{P}}}$. Let $q_{\mathfrak{P}}$ be the prime of $K_{\mathfrak{P}}$ corresponding to the embedding $\iota_p|_{K_{\mathfrak{P}}}$, and set $p_{\mathfrak{P}} = q_{\mathfrak{P}} \cap E_{\mathfrak{P}}$. Let $D(q_{\mathfrak{P}}|q_{\mathfrak{P}}) \subseteq \Gamma_{\mathfrak{P}}$ be the decomposition group of $q_{\mathfrak{P}}$ over $p_{\mathfrak{P}}$. Thus we have that the completion $K_{\mathfrak{P},q_{\mathfrak{P}}}$ of $K_{\mathfrak{P}}$ at $q_{\mathfrak{P}}$ is equal to $Q(\mathfrak{I}/\mathfrak{P})$ and $\text{Gal}(K_{\mathfrak{P},q_{\mathfrak{P}}} / E_{\mathfrak{P},q_{\mathfrak{P}}}) = D(q_{\mathfrak{P}}|p_{\mathfrak{P}})$. Thus we may view $D(q_{\mathfrak{P}}|p_{\mathfrak{P}})$ as the set of all automorphisms of $K_{\mathfrak{P},q_{\mathfrak{P}}}$ that are conjugate self-twists of $f_{\mathfrak{P}}$.

With this in mind, we see that there is a natural group homomorphism

$$\Phi : D(\mathfrak{P}'|\mathbb{Q}) \rightarrow D(q_{\mathfrak{P}}|p_{\mathfrak{P}})$$

since any element of $D(\mathfrak{P}'|\mathbb{Q})$ stabilizes $\mathfrak{P}'$ and hence induces an automorphism of $Q(\mathfrak{I}/\mathfrak{P}') = Q(\mathfrak{I}/\mathfrak{P}) = K_{\mathfrak{P},q_{\mathfrak{P}}}$. The induced automorphism will necessarily be a conjugate self-twist of $f_{\mathfrak{P}}$ since we started with a conjugate self-twist of $F$. Thus we get an element of $D(q_{\mathfrak{P}}|p_{\mathfrak{P}})$. The main compatibility result is that $\Phi$ is an isomorphism.

**Proposition 6.2.** The natural group homomorphism $\Phi$ is an isomorphism. Hence $Q(\mathbb{I}_0/\mathbb{Q}) = E_{\mathfrak{P},q_{\mathfrak{P}}}$. 

**Proof.** The fact that $\Phi$ is injective is easy. Namely, if $\sigma \in D(\mathfrak{P}'|\mathbb{Q})$ acts trivially on $K_{\mathfrak{P},q_{\mathfrak{P}}}$ then for almost all $\ell$ we have

$$a(\ell, f_{\mathfrak{P}}) = a(\ell, f_{\mathfrak{P}})^{\sigma} = \eta_{\sigma}(\ell)a(\ell, f_{\mathfrak{P}}).$$

Since $F$ (and hence its arithmetic specialization $f_{\mathfrak{P}}$) does not have CM it follows that $\eta_{\sigma} = 1$. Hence $\sigma = 1$ and $\Phi$ is injective.

To see that $\Phi$ is surjective, let $\sigma \in D(q_{\mathfrak{P}}|p_{\mathfrak{P}})$. By Theorem 3.1 we see that there is $\tilde{\sigma} \in \text{Aut} \mathfrak{I}'$ that is a conjugate self-twist of $F$ and $\sigma \circ \mathfrak{P} = \mathfrak{P} \circ \tilde{\sigma}$. That is,
\( \tilde{\sigma} \in D(\mathfrak{P}'|Q) \) and \( \Phi(\tilde{\sigma}) = \sigma \). We have

\[
E_{\mathfrak{P}, \mathfrak{P}'}^0 = K^{D(\mathfrak{P}'|\mathfrak{P}')}_{\mathfrak{P}, \mathfrak{P}'} = Q(\mathfrak{l}'/\mathfrak{P}')^{D(\mathfrak{P}'|Q)}.
\]

A general fact from commutative algebra [Bourbaki 1972, Theorem V.2.2.2] tells us that \( Q(\mathfrak{l}'/\mathfrak{P}')^{D(\mathfrak{P}'|Q)} = Q(\mathfrak{l}_0/Q) \), as desired. \( \square \)

**Corollary 6.3.** Let \( Q \) be a prime of \( \mathfrak{l}_0 \) lying over an arithmetic prime of \( \Lambda \). There is a nonzero \( \mathfrak{l}_0/Q \)-ideal \( \tilde{\mathfrak{a}}_Q \) such that

\[
\Gamma_{l_0/Q}(\tilde{\mathfrak{a}}_Q) \subseteq \text{Im}(\rho_F \mod \mathfrak{l}') \subseteq \prod_{\mathfrak{P}'|Q} \text{GL}_2(\mathfrak{l}'/\mathfrak{P}'),
\]

where the inclusion of \( \Gamma_{l_0/Q}(\tilde{\mathfrak{a}}_Q) \) in the product is via the diagonal embedding \( \text{GL}_2(\mathfrak{l}_0/Q) \hookrightarrow \prod_{\mathfrak{P}'|Q} \text{GL}_2(\mathfrak{l}'/\mathfrak{P}') \). Hence the image of \( \text{Im} \rho \) in \( \text{SL}_2(\mathfrak{l}_0/Q) \) is open.

**Proof.** For a prime \( \mathfrak{P} \) of \( \mathfrak{l} \), write \( \mathcal{O}_{\mathfrak{P}} \) for the ring of integers of \( E_{\mathfrak{P}, \mathfrak{P}''} \). By Theorem 6.1 and the remarks following it, for each prime \( \mathfrak{P} \) of \( \mathfrak{l} \) lying over \( Q \) we have that \( \text{Im} \rho_{\mathfrak{P}_Q} \) contains \( \Gamma_{\mathfrak{P}_Q}(\tilde{\mathfrak{a}}_Q) \) for some nonzero \( \mathcal{O}_{\mathfrak{P}} \)-ideal \( \tilde{\mathfrak{a}}_Q \). While \( \mathfrak{l}_0/Q \) need not be integrally closed, by Proposition 6.2 we see that \( \tilde{\mathfrak{a}}_Q \cap (\mathfrak{l}_0/Q) \) is a nonzero \( \mathfrak{l}_0/Q \)-ideal.

Thus we have

\[
\Gamma_{l_0/Q}(\tilde{\mathfrak{a}}_Q \cap \mathfrak{l}_0/Q) \subseteq \Gamma_{\mathfrak{P}_Q}(\tilde{\mathfrak{a}}_Q) \subseteq \text{Im} \rho_{\mathfrak{P}_Q} = \text{Im} \rho_F \mod \mathfrak{P}' \subseteq \text{GL}_2(\mathfrak{l}'/\mathfrak{P}').
\]

Let \( \tilde{\mathfrak{a}}_Q = \bigcap_{\mathfrak{P}_Q|Q} \tilde{\mathfrak{a}}_Q \cap \mathfrak{l}_0/Q \). This is a finite intersection of nonzero \( \mathfrak{l}_0/Q \)-ideals and hence is nonzero. The first statement follows from the above inclusions.

For the statement about \( \rho \), recall that \( \rho_F|_{H_0} \) is valued in \( \text{GL}_2(\mathfrak{l}_0) \) and consequently \( \text{Im} \rho_F|_{H_0} \mod \mathfrak{P}' \) lies in the diagonally embedded copy of \( \text{GL}_2(\mathfrak{l}_0/Q) \) in \( \prod_{\mathfrak{P}'|Q} \text{GL}_2(\mathfrak{l}'/\mathfrak{P}') \). Since \( H \) is open in \( G_Q \), by replacing \( \tilde{\mathfrak{a}}_Q \) with a smaller \( \mathfrak{l}_0/Q \)-ideal if necessary, we may assume that \( \Gamma_{l_0/Q}(\tilde{\mathfrak{a}}_Q) \) is contained in the image of \( \rho_F|_H \) in \( \text{GL}_2(\mathfrak{l}_0/Q) \). Since \( \rho \) and \( \rho_F \) are equal on elements of determinant 1 and \( \Gamma_{l_0/Q}(\tilde{\mathfrak{a}}_Q) \subseteq \text{SL}_2(\mathfrak{l}_0/Q) \), it follows that \( \Gamma_{l_0/Q}(\tilde{\mathfrak{a}}_Q) \) is contained in the image of \( \text{Im} \rho \) in \( \text{SL}_2(\mathfrak{l}_0/Q) \). That is, the image of \( \text{Im} \rho \) in \( \text{SL}_2(\mathfrak{l}_0/Q) \) is open. \( \square \)

**Summary of Proof of Theorem 2.4.** Theorem 4.1, which will be proved in the next section, allows us to create a representation \( \rho : H \to \text{SL}_2(\mathfrak{l}_0) \) with the property that if \( \rho \) is \( \mathfrak{l}_0 \)-full then so is \( \rho_F \). This is important for the use of Pink’s theory in Section 4 as well as for the techniques of Section 5. Proposition 4.2 shows that it is sufficient to prove that the image of \( \text{Im} \rho \) in \( \prod_{\mathfrak{P}_Q|Q} \text{SL}_2(\mathfrak{l}_0/Q) \) is open for some arithmetic prime \( P \) of \( \Lambda \). Proposition 5.1 further reduces the problem to showing that the image of \( \rho \mod \mathfrak{P} \) is open in \( \text{SL}_2(\mathfrak{l}_0/Q) \) for all primes \( \mathfrak{P} \) of \( \mathfrak{l} \) lying over a fixed arithmetic prime \( P \) of \( \Lambda \).

This reduces the problem to studying the image of a Galois representation attached to one of the classical specializations of \( F \) (twisted by the inverse square
The Galois representation associated to a non-CM Hida family

[187]

root of the determinant). Hence we can apply the work of Ribet and Momose, but only after we show that $Q(I_0/Q)$ is the same field that occurs in their work. This is done in Proposition 6.2, though the main input is Theorem 3.1.

□

7. Obtaining an $SL_2(I_0)$-valued representation

In this section we prove:

**Theorem 4.1.** Assume that $\tilde{\rho}_F$ is absolutely irreducible and $H_0$-regular. If $V = \mathfrak{l}^2$ is the module on which $G_\mathbb{Q}$ acts via $\rho_F$, then there is a basis for $V$ such that all of the following happen simultaneously:

1. $\rho_F$ is valued in $GL_2(I')$;
2. $\rho_F|_{D_p}$ is upper triangular;
3. $\rho_F|_{H_0}$ is valued in $GL_2(I_0)$;
4. There is a matrix $j = \begin{pmatrix} \zeta & 0 \\ 0 & \zeta' \end{pmatrix}$, where $\zeta$ and $\zeta'$ are roots of unity, such that $j$ normalizes the image of $\rho_F$ and $\zeta \not\equiv \zeta' \mod p$.

It is well known that so long as $\tilde{\rho}_F$ is absolutely irreducible we may assume that $\rho_F$ has values in $GL_2(I')$ and the local representation $\rho_F|_{D_p}$ is upper triangular [Hida 2000a, Theorem 4.3.2]. To show that $\rho_F|_{H_0}$ has values in $GL_2(I_0)$ we begin by investigating the structure of $\Gamma$.

**Proposition 7.1.** The group $\Gamma$ is a finite abelian 2-group.

**Proof.** Let $S$ be the set of primes $\ell$ for which $a(\ell, F)^{\sigma} = \eta_{\sigma}(\ell) a(\ell, F)$ for all $\sigma \in \Gamma$, so $S$ excludes only finitely many primes. For $\ell \in S$, let

$$b_\ell := \frac{a(\ell, F)^2}{\det \rho_F(\text{Frob}_\ell)}.$$ 

It turns out that $b_\ell \in I_0$. To see this, note that since $\tilde{\rho}_F$ is absolutely irreducible, for any $\sigma \in \Gamma$ we have $\rho_F^{\sigma} \cong \eta_{\sigma} \otimes \rho_F$ over $\mathfrak{l}$. Taking determinants we find that $\det \rho_F^{\sigma-1} = \eta_{\sigma}^2$. Thus we have

$$(a(\ell, F)^{\sigma})^2 = \eta_{\sigma}(\ell)^2 a(\ell, F)^2 = \det \rho_F(\text{Frob}_\ell) \sigma^{-1} a(\ell, F)^2,$$ 

from which it follows that $b_\ell^2 = b_\ell$. Solving for $a(\ell, F)$ in the definition of $b_\ell$ we find that

$$Q(\mathfrak{l}) = Q(I_0)\left[\sqrt{b_\ell} \det \rho_F(\text{Frob}_\ell) : \ell \in S\right].$$

Recall that for $\ell \in S$ we have $\det \rho_F(\text{Frob}_\ell) = \chi(\ell) \kappa((\ell)) \ell^{-1}$, where $\kappa((\ell)) \in 1 + m_\Lambda$. (Currently all that matters is that $\kappa$ is valued in $1 + m_\Lambda$. For a precise definition of $\kappa$, see the proof of Lemma 3.11.) In particular, $\sqrt{\kappa((\ell))} \in \Lambda$. Similarly, we can write $\ell = (\ell) \omega(\ell)$ with $(\ell) \in 1 + p\mathbb{Z}_p$ and $\omega(\ell) \in \mu_{p-1}$. So $\sqrt{\ell} \in \Lambda$ as well.
Let
\[
K = Q(\ell_0)\left[\sqrt{b_\ell}, \sqrt{\det \rho_F(\text{Frob}_\ell)} : \ell \in S\right],
\]
which is an abelian extension of \(Q(\ell_0)\) since it is obtained by adjoining square roots. The above argument shows that in fact \(K\) is obtained from \(Q(\ell_0)[\sqrt{b_\ell} : \ell \in S]\) by adjoining finitely many roots of unity, namely the square roots of the values of \(\chi\) and the square roots of \(\mu_{p-1}\). As odd order roots of unity are automatically squares, we can write \(K = Q(\ell_0)[\sqrt{b_\ell} : \ell \in S][\mu_{2s}]\) for some \(s \in \mathbb{Z}^+\). Thus we have
\[
\text{Gal}(K/Q(\ell_0)) \cong \text{Gal}(Q(\ell_0)[\sqrt{b_\ell} : \ell \in S]/Q(\ell_0)) \times \text{Gal}(Q(\ell_0)[\mu_{2s}]/Q(\ell_0)).
\]
By Kummer theory the first group is an elementary abelian 2-group. The second group is isomorphic to \((\mathbb{Z}/2^s\mathbb{Z})^\times\) and hence is a 2-group. As \(\Gamma\) is a quotient of \(\text{Gal}(K/Q(\ell_0))\) it follows that \(\Gamma\) is a finite abelian 2-group, as claimed.

For ease of notation let \(\pi = \tilde{\rho}_F|_{H_0} : H_0 \to \text{GL}_2(\mathbb{F})\). Let \(D\) be a nonsquare in \(\mathbb{F}\), and let \(\mathbb{E} = \mathbb{F}[\sqrt{D}]\) be the unique quadratic extension of \(\mathbb{F}\).

**Lemma 7.2.** Let \(K\) be a field and \(S \subset \text{GL}_n(K)\) a set of nonconstant semisimple operators that can be simultaneously diagonalized over \(\bar{K}\). If \(y \in \text{GL}_n(\bar{K})\) such that \(yS y^{-1} \subset \text{GL}_n(K)\), then there is a matrix \(z \in \text{GL}_n(K)\) such that \(z S z^{-1} = y S y^{-1}\). In particular, if \(\pi\) is irreducible over \(\mathbb{F}\) but not absolutely irreducible, then \(\mathbb{E}\) is the splitting field for \(\pi\).

**Proof.** Let \(\sigma \in G_K := \text{Gal}(\bar{K}/K)\). Then for any \(x \in S\) we have \(y^\sigma x y^{-\sigma} = (y x y^{-1})^\sigma = y x y^{-1}\), so \(y^{-1} y^\sigma\) centralizes \(x\). As elements in \(S\) are simultaneously diagonalizable, they have the same centralizer in \(\text{GL}_n(\bar{K})\). Since elements of \(S\) are semisimple, their centralizer is a torus and hence isomorphic to \((\bar{K}^\times)^{\otimes n}\). It’s not hard to show that \(a : G_K \to (\bar{K}^\times)^{\otimes n}\) given by \(\sigma \mapsto y^{-1} y^\sigma\) is a 1-cocycle. (Here we view \((\bar{K}^\times)^{\otimes n}\) as a \(G_K\)-module by letting elements of \(G_K\) act componentwise.) By Hilbert’s theorem 90 we have \(H^1(G_K, (\bar{K}^\times)^{\otimes n}) = H^1(G_K, \bar{K}^\times)^{\otimes n} = 0\). Hence \(a\) is a coboundary. That is, there is some \(\alpha \in (\bar{K}^\times)^{\otimes n}\) such that
\[
a_\sigma = y^{-1} y^\sigma = \alpha^{-1} \alpha^\sigma
\]
for all \(\sigma \in G_K\). Thus \((y \alpha^{-1})^\sigma = y \alpha^{-1}\) for all \(\sigma \in G_K\), so \(z := y \alpha^{-1} \in \text{GL}_n(K)\). But \(\alpha\) commutes with \(S\) and so \(z S z^{-1} = y S y^{-1}\), as claimed.

To deduce the claim about \(\pi\), let \(S = \text{Im} \pi\). The fact that \(S\) is semisimple follows from Clifford’s theorem since \(\tilde{\rho}_F\) is absolutely irreducible [Isaacs 1976, Theorem 6.5, Corollary 6.6]. If \(\pi\) is not absolutely irreducible then there is a matrix \(y \in \text{GL}_2(\mathbb{F})\) that simultaneously diagonalizes \(S\). Note that every matrix in \(\text{Im} \pi\) has eigenvalues in \(\mathbb{E}\). Indeed every matrix has a quadratic characteristic polynomial and \(\mathbb{E}\) is the unique quadratic extension of \(\mathbb{F}\). Thus, taking \(K = \mathbb{E}\) we see that \(y S y^{-1} \subset \text{GL}_2(K)\). The first statement of the lemma tells us that \(\text{Im} \pi\) is
diagonalizable over $\mathbb{E}$. Since $\pi$ is irreducible over $\mathbb{F}$ and $[\mathbb{E} : \mathbb{F}] = 2$, it follows that $\mathbb{E}$ is the smallest extension of $\mathbb{F}$ over which $\text{Im} \pi$ is diagonalizable. □

Let $Z$ be the centralizer of $\text{Im} \pi$ in $M_2(\mathbb{F})$. Since $\tilde{\rho}_F$ is $H_0$-regular, exactly one of the three cases must occur:

1. The representation $\pi$ is absolutely irreducible. In this case $Z$ consists of scalar matrices over $\mathbb{F}$.
2. The representation $\pi$ is not absolutely irreducible, but $\pi$ is irreducible over $\mathbb{F}$. In this case we may assume $Z = \left\{ \begin{pmatrix} \alpha & \beta D \\ \beta & \alpha \end{pmatrix} : \alpha, \beta \in \mathbb{F} \right\} \cong \mathbb{E}$.
3. The representation $\pi$ is reducible over $\mathbb{F}$. In this case we may assume that $Z$ consists of diagonal matrices over $\mathbb{F}$.

Recall that since $\tilde{\rho}_F$ is absolutely irreducible, for any $\sigma \in \Gamma$ we have $\rho_F^\sigma \cong \eta_\sigma \otimes \rho_F$. That is, there is some $t_\sigma \in \text{GL}_2(\mathbb{I}')$ such that $\rho_F(g)^\sigma = \eta_\sigma(g)t_\sigma \rho_F(g)t_\sigma^{-1}$ for all $g \in G_\mathbb{Q}$. Then for all $\sigma, \tau \in \Gamma$, $g \in G_\mathbb{Q}$ we have

$$\eta_{\sigma \tau}(g)t_{\sigma \tau} \rho_F(g)t_{\sigma \tau}^{-1} = \rho(g)^{\sigma \tau} = \eta_\sigma(g)\eta_\tau(g)t_\sigma^T t_\tau \rho_F(g)t_\tau^{-1} t_\sigma^{-1} t_\tau^{-1}.$$

Using the fact that $\eta_{\sigma \tau} = \eta_\sigma^T \eta_\tau$ we see that $c(\sigma, \tau) := t_{\sigma \tau}^{-1} t_\sigma^T t_\tau$ commutes with the image of $\rho_F$. As $\rho_F$ is absolutely irreducible, $c(\sigma, \tau)$ must be a scalar. Hence $c$ represents a 2-cocycle of $\Gamma$ with values in $\mathbb{I}'^\times$.

We will need to treat case 2 ($\pi$ is irreducible over $\mathbb{F}$ but not absolutely irreducible) a bit differently, so we establish notation that will unify the proofs that follow. For a finite extension $M$ of $\mathbb{Q}_p$, let $\mathcal{O}_M$ denote the ring of integers of $M$. Let $K$ be the largest finite extension of $\mathbb{Q}_p$ for which $\mathcal{O}_K[[T]]$ is contained in $\mathbb{I}'$. So $K$ has residue field $\mathbb{F}$. Let $L$ be the unique unramified quadratic extension of $K$. Write $\mathbb{J} = \Lambda_{\mathcal{O}_L}[[a(\ell, F) : \ell \nmid N]]$. Note that the residue field of $\mathbb{J}$ is the unique quadratic extension of $\mathbb{F}$. Let

$$A = \begin{cases} \mathbb{J} & \text{in case 2,} \\ \mathbb{I}' & \text{else.} \end{cases}$$

Let $\kappa$ be the residue field of $A$, so $\kappa = \mathbb{E}$ in case 2 and $\kappa = \mathbb{F}$ otherwise.

Since $L$ is obtained from $K$ by adjoining some prime-to-$p$ root of unity, in case 2 it follows that $Q(A)$ is Galois over $Q(\mathbb{I}_0)$ with Galois group isomorphic to $\Gamma \times \mathbb{Z}/2\mathbb{Z}$. In particular, we have an action of $\Gamma$ on $A$ in all cases. Let $B = A^\Gamma$. In case 2, $A$ is a quadratic extension of $B$ and $B \cap \mathbb{I}' = \mathbb{I}_0$. Otherwise $B = \mathbb{I}_0$. We may consider the 2-cocycle $c$ in $H^2(\Gamma, A^\times)$. 
Lemma 7.3. With notation as above, \([c] = 0 \in H^2(\Gamma, A^\times)\). Thus there is a function \(\zeta : \Gamma \to A^\times\) such that \(c(\sigma, \tau) = \zeta((\sigma \tau)^{-1}(\sigma)^\tau \zeta(\tau))\) for all \(\sigma, \tau \in \Gamma\).

Proof. Consider the exact sequence \(1 \to 1 + m_A \to A^\times \to \kappa^\times \to 1\). Note that for \(j > 0\) we have \(H^j(\Gamma, 1 + m_A) = 0\) since \(1 + m_A\) is a \(p\)-profinite group for \(p > 2\) and \(\Gamma\) is a 2-group by Proposition 7.1. Thus the long exact sequence in cohomology gives isomorphisms

\[
H^j(\Gamma, A^\times) \cong H^j(\Gamma, \kappa^\times)
\]

for all \(j > 0\). Hence it suffices to prove that \([\tilde{c}] = 0 \in H^2(\Gamma, \kappa^\times)\).

Let \(\sigma \in \Gamma\) and \(h \in H_0\). Recall that \(\Gamma\) acts trivially on \(\mathbb{F}\) by Proposition 3.4. Since \(\rho^\sigma_F(h) = \eta_\sigma(h)t_\sigma \rho_F(h)t_\sigma^{-1}\) and \(\eta_\sigma(h) = 1\) it follows that \(\tilde{t}_\sigma \in Z\).

We now split into the three cases depending on the irreducibility of \(\pi\). Suppose we are in case 1, so \(\pi\) is absolutely irreducible and \(\kappa = \mathbb{F}\). Then \(\tilde{t}_\sigma\) must be a scalar in \(\mathbb{F}^\times\). Call it \(\tilde{\zeta}(\sigma)\). Then \(\tilde{\zeta}(\sigma, \tau) = \tilde{\zeta}(\sigma \tau)^{-1} \tilde{\zeta}(\sigma)^\tau \tilde{\zeta}(\tau)\), and so \([\tilde{c}] = 0 \in H^2(\Gamma, \mathbb{F}^\times)\).

In case 2, using the description of \(Z\) above we see that \(\tilde{t}_\sigma = (\alpha_\sigma, \beta_\sigma, D_\sigma)\) for some \(\alpha_\sigma, \beta_\sigma \in \mathbb{F}\). This becomes a scalar, say \(\tilde{\zeta}(\sigma) = \alpha_\sigma + \beta_\sigma \sqrt{D}\), over \(\tilde{E} = \kappa\). Thus \(\tilde{t}_\sigma = \tilde{\zeta}(\sigma)\). As above \(\tilde{\zeta}(\sigma, \tau) = \tilde{\zeta}(\sigma \tau)^{-1} \tilde{\zeta}(\sigma)^\tau \tilde{\zeta}(\tau)\), and thus \([\tilde{c}] = 0 \in H^2(\Gamma, \kappa^\times)\).

Finally, in case 3 we have that \(\tilde{t}_\sigma\) is a diagonal matrix. The diagonal map \(\mathbb{F} \leftarrow \mathbb{F} \oplus \mathbb{F}\) induces an injection \(H^2(\Gamma, \mathbb{F}^\times) \hookrightarrow H^2(\Gamma, \mathbb{F}^\times \oplus \mathbb{F}^\times)\). The fact that \(\tilde{t}_\sigma\) is a diagonal matrix allows us to calculate that the image of \([\tilde{c}]\) in \(H^2(\Gamma, \mathbb{F}^\times \oplus \mathbb{F}^\times)\) is 0. Since the map is an injection, it follows that \([\tilde{c}] = 0 \in H^2(\Gamma, \mathbb{F}^\times)\), as desired. \(\square\)

Replace \(t_\sigma \in \text{GL}_2(\mathbb{F})\) by \(t_\sigma \zeta(\sigma)^{-1} \in \text{GL}_2(A)\). Then we still have \(\rho^\sigma_F = \eta_\sigma(t_\sigma \rho_F t_\sigma^{-1})\), and now \(t_\sigma \tau = t_\sigma \tau\). That is, \(\sigma \mapsto t_\sigma\) is a nonabelian 1-cocycle with values in \(\text{GL}_2(A)\). Since \(F\) is primitive we have \(Q(\mathbb{F}) = Q(\mathbb{F})\). Thus by [Hida 2000a, Theorem 4.3.2] we see that \(\rho^\sigma_F|_{D_p}\) is isomorphic to an upper triangular representation over \(Q(\mathbb{F})\).

Under the assumptions that \(\rho^\sigma_F\) is absolutely reducible and \(H_0\)-regular, the proof of [Hida 2000a, Theorem 4.3.2] goes through with \(\mathbb{F}\) in place of \(\mathbb{F}\). That is, \(\rho^\sigma_F|_{D_p}\) is isomorphic to an upper triangular representation over \(\mathbb{F}\). Let \(V = \mathbb{F}^2\) be the representation space for \(\rho^\sigma_F\) with basis chosen such that

\[
\rho^\sigma_F|_{D_p} = \begin{pmatrix} \varepsilon & u \\ 0 & \delta \end{pmatrix},
\]

and assume \(\bar{\varepsilon} \neq \bar{\delta}\). Let \(V[\varepsilon] \subset V\) be the free direct summand of \(V\) on which \(D_p\) acts by \(\varepsilon\) and \(V[\delta]\) be the quotient of \(V\) on which \(D_p\) acts by \(\delta\). Let \(V_A = V \otimes_{\mathbb{F}} A\). Similarly for \(\lambda \in \{\varepsilon, \delta\}\) let \(V_A[\lambda] := V[\lambda] \otimes_{\mathbb{F}} A\). For \(v \in V_A\), define

\[
v^{[\sigma]} := t_\sigma^{-1} v^\sigma,
\]

where \(\sigma\) acts on \(v\) componentwise. Note that in case 2 we are using the action of \(\Gamma\) on \(A\) described prior to Lemma 7.3.
Lemma 7.4. For all $\sigma, \tau \in \Gamma$ we have $(v^{[\sigma\tau]})^{[\tau]} = v^{[\sigma\tau]}$, so this defines an action of $\Gamma$ on $V_A$. Furthermore, this action stabilizes $V_A[\varepsilon]$ and $V_A[\delta]$.

Proof. The formula (7) defines an action since $\sigma \mapsto t_\sigma$ is a nonabelian 1-cocycle. Let $\lambda$ be either $\varepsilon$ or $\delta$. Let $v \in V_A[\lambda]$ and $\sigma \in \Gamma$. We must show that $v^{[\sigma]} \in V_A[\lambda]$. Let $d \in D_p$. Using the fact that $v \in V_A[\lambda]$ and $\rho^\sigma_F = \eta^\sigma(t_\sigma \rho_F t^{-1}_\sigma)$ we find that $\rho^\sigma_F(d)v^{[\sigma]} = \eta^{-1}^\sigma(d)\lambda^\sigma(d)v^{[\sigma]}$.

Note that for all $d \in D_p$

$$
\begin{pmatrix}
\varepsilon^\sigma(d) & u^\sigma(d) \\
0 & \delta^\sigma(d)
\end{pmatrix}
= \rho^\sigma_F(d) = \eta^\sigma(d)t_\sigma \rho_F(d)t^{-1}_\sigma = \eta^\sigma(d)t_\sigma \begin{pmatrix}
\varepsilon(d) & u(d) \\
0 & \delta(d)
\end{pmatrix}t^{-1}_\sigma. \tag{8}
$$

Using the fact that $\varepsilon \neq \delta$ and that $\rho^\sigma_F|_{D_p}$ is indecomposable [Ghate and Vatsal 2004; Zhao 2014] we see that $u/(\varepsilon - \delta)$ cannot be a constant. (If $u/(\varepsilon - \delta) = \alpha$ is a constant, then conjugating by $\begin{pmatrix}1 & a \\ 0 & 1\end{pmatrix}$ makes $\rho^\sigma_F|_{D_p}$ diagonal.) Hence $t_\sigma$ must be upper triangular. Therefore (8) implies that $\lambda^\sigma(d) = \eta^\sigma(d)\lambda(d)$, and thus

$$
\rho^\sigma_F(d)v^{[\sigma]} = \eta^{-1}^\sigma(d)\lambda^\sigma(d)v^{[\sigma]} = \lambda(d)v^{[\sigma]}.
$$

We are now ready to show that $\rho^\sigma_F|_{H_0}$ takes values in $GL_2(\mathbb{F}_0)$.

Theorem 7.5. Let $\bar{\rho}_F : G_{\mathbb{Q}} \to GL_2(\mathbb{F})$ such that $\rho^\sigma_F|_{D_p}$ is upper triangular. Assume that $\bar{\rho}_F$ is absolutely irreducible and $H_0$-regular. Then $\rho^\sigma_F|_{H_0}$ takes values in $GL_2(\mathbb{F}_0)$.

Proof. We have an exact sequence of $A[D_p]$-modules

$$
0 \to V_A[\varepsilon] \to V_A \to V_A[\delta] \to 0 \tag{9}
$$

that is stable under the new action of $\Gamma$ defined in Lemma 7.4. Tensoring with $\kappa$ over $A$ we get an exact sequence of $\kappa$-vector spaces

$$
V_\kappa[\tilde{\varepsilon}] \to V_\kappa \to V_\kappa[\tilde{\delta}] \to 0. \tag{10}
$$

Since $V_A[\varepsilon]$ is a direct summand of $V_A$, the first arrow is injective. Since $V_A[\varepsilon]$ and $V_A$ are free $A$-modules, it follows that $\dim_\kappa V_\kappa[\tilde{\varepsilon}] = 1$ and $\dim_\kappa V_\kappa = 2$. Counting dimensions in (10) now tells us that $\dim_\kappa V_\kappa[\tilde{\delta}] = 1$.

Going back to the exact sequence (9) we can take $\Gamma$-invariants since all of the modules are stable under the new action of $\Gamma$. This gives an exact sequence of $B[D_p \cap H_0]$-modules

$$
0 \to V_A[\varepsilon]^\Gamma \to V_A^\Gamma \to V_A[\delta]^\Gamma \to H^1(\Gamma, V_A[\varepsilon]).
$$

Since $\Gamma$ is a 2-group by Proposition 7.1 and $V_A[\varepsilon] \cong A$ is $p$-profinite, we find that $H^1(\Gamma, V_A[\varepsilon]) = 0$. Tensoring with $\kappa^\Gamma$ over $B$ we get an exact sequence

$$
V_A[\varepsilon]^\Gamma \otimes_B \kappa^\Gamma \to V_A^\Gamma \otimes_B \kappa^\Gamma \to V_A[\delta]^\Gamma \otimes_B \kappa^\Gamma \to 0.
$$
If \( \dim_{\kappa^r} V_A[\lambda] \Gamma \otimes_B \kappa^r \) = 1 for \( \lambda \in \{\varepsilon, \delta\} \), then it follows from Nakayama’s lemma that \( V_A[\lambda] \Gamma \) is a free \( B \)-module of rank 1. Hence \( V_A^\Gamma \) is a free \( B \)-module of rank 2. In all the cases except case 2, this completes the proof. In case 2 the above argument tells us that if we view \( \rho_F \) as a \( \GL_2(A) \)-valued representation, then \( \rho_F|_{H_0} \) takes values in \( \GL_2(B) \). We know that \( \rho_F \) actually has values in \( \GL_2(\ell') \) and hence \( \rho_F|_{H_0} \) has values in \( \GL_2(B \cap \ell') = \GL_2(\ell_0) \).

Thus we must show that for \( \lambda \in \{\varepsilon, \delta\} \) we have \( \dim_{\kappa^r} V_A[\lambda] \Gamma \otimes_B \kappa^r = 1 \). Note that \( V_A[\lambda] \Gamma \otimes_B \kappa^r = V_\kappa[\lambda] \Gamma \). When we are not in case 2, \( \Gamma \) acts trivially on \( \kappa \) and hence

\[
\dim_{\ell} V_{\ell} \bar{\lambda} \Gamma = \dim_{\ell} V_{\ell} \bar{\lambda} = 1.
\]

Now assume we are in case 2, so \( \kappa = \ell \). Write \( \bar{\Gamma} \) for the quotient of \( \Gamma \) that acts on \( \ell \). That is, \( \bar{\Gamma} = \mathrm{Gal}(\ell/\bar{\ell}) \). Let \( \sigma \in \bar{\Gamma} \) be a generator. Since \( \dim_{\ell} V_{\ell}[\bar{\lambda}] = 1 \) we can choose some nonzero \( v \in V_{\ell}[\bar{\lambda}] \). We would like to show that

\[
v + v^{[\sigma]} \neq 0
\]

since the right hand side is \( \bar{\Gamma} \)-invariant.

Since \( V_{\ell}[\bar{\lambda}] \) is 1-dimensional, there is some \( \alpha \in \ell^\times \) such that \( v^{[\sigma]} = \alpha v \). Thus \( v + v^{[\sigma]} = (1 + \alpha) v \). If \( \alpha \neq -1 \) then we are done. Otherwise we can change \( v \) to \( a v \) for any \( a \in \ell^\times \). It is easy to see that \( (a v)^{[\sigma]} = a^\sigma a a^{-1} (a v) \) and thus changing \( v \) to \( a v \) changes \( \alpha \) to \( a^\sigma a^{-1} \). So we need to show that there is some \( a \in \ell^\times \) such that \( a^\sigma a^{-1} \neq 1 \). But clearly this holds for any \( a \in \ell \setminus \ell \). Therefore \( \dim_{\ell} V_{\ell}[\bar{\lambda}] \Gamma \geq 1 \).

To get equality, let \( 0 \neq w \in V_{\ell}[\bar{\lambda}] \Gamma \). Since \( V_{\ell}[\bar{\lambda}] \Gamma \subseteq V_{\ell}[\bar{\lambda}] \) and \( \dim_{\ell} V_{\ell}[\bar{\lambda}] = 1 \), any element of \( V_{\ell}[\bar{\lambda}] \Gamma \) is an \( \ell \)-multiple of \( w \). If \( \beta \in \ell \setminus \ell \) then \( \sigma \) does not fix \( \beta \). Thus

\[
(\beta w)^{[\sigma]} = \beta^\sigma w^{[\sigma]} = \beta^\sigma w \neq \beta w.
\]

Hence \( V_{\ell}[\bar{\lambda}] \Gamma = \ell w \) and \( \dim_{\ell} V_{\ell}[\bar{\lambda}] \Gamma = 1 \), as desired. \( \square \)

Finally, we modify \( \rho_F \) to obtain the normalizing matrix \( j \) in the last part of Theorem 4.1.

**Lemma 7.6.** Suppose \( \rho_F : G_Q \to \GL_2(\ell') \) such that \( \rho_F|_{D_0} \) is upper triangular and \( \rho_F|_{H_0} \) is valued in \( \GL_2(\ell_0) \). Assume \( \tilde{\rho}_F \) is absolutely irreducible and \( H_0 \)-regular. Then there is an upper triangular matrix \( x \in \GL_2(\ell_0) \) and roots of unity \( \zeta \) and \( \zeta' \) such that \( j := \begin{pmatrix} \zeta & 0 \\ 0 & \zeta' \end{pmatrix} \) normalizes the image of \( x \rho_F x^{-1} \) and \( \zeta \equiv \varepsilon(h) \mod \ell_0 \) and \( \zeta' \equiv \delta(h) \mod \ell_0 \). By our choice of \( h \) we have \( \zeta \neq \zeta' \mod p \).

**Proof.** This argument is due to Hida [2000a, Lemma 4.3.20]. As \( \tilde{\rho}_F \) is \( H_0 \)-regular there is an \( h \in H_0 \) such that \( \varepsilon(h) \neq \delta(h) \). Let \( \xi \) and \( \xi' \) be the roots of unity in \( \ell_0 \) satisfying \( \xi \equiv \varepsilon(h) \mod \ell_0 \) and \( \xi' \equiv \delta(h) \mod \ell_0 \). By our choice of \( h \) we have \( \xi \neq \xi' \mod p \).
Let \( q = |F| \). Then for some \( u \in I_0 \)

\[
\lim_{n \to \infty} \rho_F(h)^{q^n} = \begin{pmatrix} \xi & u \\ 0 & \xi' \end{pmatrix}.
\]

Conjugating \( \rho_F \) by \( \begin{pmatrix} 1 & u/\xi - \xi' \\ 0 & 1 \end{pmatrix} \) preserves all three of the desired properties, and the image of the resulting representation is normalized by \( j = \begin{pmatrix} \xi & 0 \\ 0 & \xi' \end{pmatrix} \).

\[\square\]

Acknowledgements

I am grateful to my advisor, Haruzo Hida, for suggesting this problem to me and for his endless patience and insights as I worked on it. I would like to thank Ashay Burungale for many helpful and encouraging conversations about this project. Richard Pink and Jacques Tilouine provided me with insights into the larger context of this problem. The referee provided many useful comments that prompted me to make major corrections to earlier versions of this paper. Tomomi Ozawa pointed out many typos. Finally, I am grateful for the financial support provided by UCLA and National Science Foundation that allowed me to complete this work.

References


Communicated by Barry Mazur
Received 2015-01-07 Revised 2015-10-06 Accepted 2015-11-27

jaclynlang@math.ucla.edu
UCLA Mathematics Department,
University of California, Los Angeles, Box 951555,
Los Angeles, CA 90095-1555, United States

mathematical sciences publishers
Linear relations in families of powers of elliptic curves

Fabrizio Barroero and Laura Capuano

Motivated by recent work of Masser and Zannier on simultaneous torsion on the Legendre elliptic curve $E_\lambda$ of equation $Y^2 = X(X-1)(X-\lambda)$, we prove that, given $n$ linearly independent points $P_1(\lambda), \ldots, P_n(\lambda)$ on $E_\lambda$ with coordinates in $\mathbb{Q}(\lambda)$, there are at most finitely many complex numbers $\lambda_0$ such that the points $P_1(\lambda_0), \ldots, P_n(\lambda_0)$ satisfy two independent relations on $E_{\lambda_0}$. This is a special case of conjectures about unlikely intersections on families of abelian varieties.

1. Introduction

Let $n \geq 2$ be an integer and let $E_\lambda$ denote the elliptic curve in the Legendre form defined by

$$Y^2 = X(X-1)(X-\lambda). \quad (1-1)$$

Masser and Zannier [2010; see also 2008] showed that there are at most finitely many complex numbers $\lambda_0 \neq 0, 1$ such that the two points

$$(2, \sqrt{2(2-\lambda_0)}), \quad (3, \sqrt{6(3-\lambda_0)})$$

both have finite order on the elliptic curve $E_{\lambda_0}$. Stoll [2014] recently noted that there is actually no such $\lambda_0$. Later, Masser and Zannier [2012] proved that one can replace 2 and 3 with any two distinct complex numbers ($\neq 0, 1$) or even choose distinct $X$-coordinates ($\neq \lambda$) defined over an algebraic closure of $\mathbb{C}(\lambda)$.

In his book, Zannier [2012] asks if there are finitely many $\lambda_0 \in \mathbb{C}$ such that two independent relations between the points $(2, \sqrt{2(2-\lambda_0)}), (3, \sqrt{6(3-\lambda_0)})$ and $(5, \sqrt{20(5-\lambda_0)})$ hold on $E_{\lambda_0}$.

In this article we prove that this question has a positive answer, as Zannier expected in view of very general conjectures. We actually prove a more general

The authors are supported by the European Research Council, grant number 267273.

MSC2010: primary 11G05; secondary 11G50, 11U09, 14K05.

Keywords: linear relations, elliptic curves, unlikely intersections.
result, analogous to the one in [Masser and Zannier 2012] but, at the moment, we are only able to replace 2, 3 and 5 with any three pairwise distinct algebraic numbers, or choose \(X\)-coordinates defined over an algebraic closure of \(\mathbb{Q}(\lambda)\), with the obvious exceptions 0, 1 and \(\lambda\) since the corresponding points are identically 2-torsion. Moreover, our method allows us to deal with arbitrarily many points since we consider a curve \(C \subseteq \mathbb{A}^{2n+1}\) with coordinate functions \((x_1, y_1, \ldots, x_n, y_n, \lambda)\), where \(\lambda\) is nonconstant, such that, for every \(j = 1, \ldots, n\), the points \(P_j = (x_j, y_j)\) lie on the elliptic curve \(E_\lambda\). As the point \(c\) varies on the curve \(C\), the specialized points \(P_j(c) = (x_j(c), y_j(c))\) will lie on the specialized elliptic curve \(E_\lambda(c)\). We implicitly exclude the finitely many \(c\) with \(\lambda(c) = 0\) or 1, since in that case \(E_\lambda(c)\) is not an elliptic curve.

**Theorem 1.1.** Let \(C \subseteq \mathbb{A}^{2n+1}\) be an irreducible curve defined over \(\overline{\mathbb{Q}}\) with coordinate functions \((x_1, y_1, \ldots, x_n, y_n, \lambda)\), where \(\lambda\) is nonconstant. Suppose that, for every \(j = 1, \ldots, n\), the points \(P_j = (x_j, y_j)\) lie on \(E_\lambda\) and there are no integers \(a_1, \ldots, a_n \in \mathbb{Z}\), not all zero, such that

\[
a_1 P_1 + \cdots + a_n P_n = O, \tag{1-2}
\]

identically on \(C\). Then there are at most finitely many \(c \in C\) such that the points \(P_1(c), \ldots, P_n(c)\) satisfy two independent relations on \(E_\lambda(c)\).

Note that the case \(n = 2\) is covered by the main proposition of [Masser and Zannier 2012] in the more general setting of a curve defined over \(\overline{\mathbb{C}}\).

Moreover, Rémond and Viada [2003] proved an analogue of Theorem 1.1 for a power of a constant elliptic curve with complex multiplication, where one must allow the coefficients \(a_1, \ldots, a_n\) in (1-2) to lie in the larger endomorphism ring. For the general case of powers of a constant elliptic curve, the result follows from works of Viada [2008] and Galateau [2010]. If \(n = 2\) this is nothing but Raynaud’s theorem [1983], also known as the Manin–Mumford conjecture.

We already mentioned the example of the three points with fixed abscissas 2, 3 and 5. It is easy to see that this will follow from Theorem 1.1 once we show that there is no identical relation between the three points on the generic curve \(E_\lambda\). Indeed, the minimal fields of definition of these three points are disjoint quadratic extensions of \(\mathbb{Q}(\lambda)\), and by conjugating one can see that the points would be identically torsion on \(E_\lambda\). This is not possible, as it can be seen in different ways (see [Zannier 2012, p. 68]). For instance, applying the Lutz–Nagell theorem [Silverman 2009, Corollary 7.2], one can show that the point of abscissa 2 is not torsion on \(E_6\).

One may ask if finiteness holds if we impose only one relation. This is not the case. Indeed, there are infinitely many \(\lambda_0\) such that a point with fixed algebraic abscissa is torsion (see [Zannier 2012, Notes to Chapter 3]). On the other hand,
the values of $\lambda$ such that at least one relation holds are “sparse”, as follows from [Masser 1989b]. Actually, a well-known theorem of Silverman [1983] implies that the absolute Weil height of such values is bounded. A direct effective proof of this can be found in Masser’s Appendix C of [Zannier 2012]. In particular, there are at most finitely many $\lambda_0$ yielding one relation in a given number field or of bounded degree over $\mathbb{Q}$.

Our proof follows the general strategy introduced in [Pila and Zannier 2008] and used in [Masser and Zannier 2008; 2010; 2012]. In particular, we consider the elliptic logarithms $z_1, \ldots, z_n$ of $P_1, \ldots, P_n$ and the equations

$$z_j = u_j f + v_j g,$$

for $j = 1, \ldots, n$, where $f$ and $g$ are suitably chosen basis elements of the period lattice of $E_\lambda$. If we consider the coefficients $u_j$, $v_j$ as functions of $\lambda$ and restrict them to a compact set, we obtain a subanalytic surface $S$ in $\mathbb{R}^{2n}$. The points of $C$ that yield two independent relations on the elliptic curve will correspond to points of $S$ lying on linear varieties defined by equations of some special form and with integer coefficients. In the case $n = 2$, one faces the simpler problem of counting rational points with bounded denominator in $S$. For this, a previous result of Pila [2004] suffices together with the fact that the surface is “sufficiently” transcendental. In the general case we adopt ideas of Pila (see [Capuano et al. 2016, Appendix]) to obtain an upper bound of order $T^\epsilon$ for the number of points of $S$ lying on subspaces of the special form mentioned above and integral coefficients of absolute value at most $T$, provided $S$ does not contain a semialgebraic curve segment. Under the hypothesis that no identical relation holds on $C$, using a result of Bertrand [1990], we are able to show that there are no such semialgebraic curve segments.

Now, we use [Masser 1988; 1989a; David 1997] and exploit the boundedness of the height to show that the number of points of $S$ considered above is of order at least $T^\delta$ for some $\delta > 0$. Comparing the two estimates leads to an upper bound for $T$ and thus for the coefficients of the two relations, concluding the proof.

With similar methods, a toric analogue of Theorem 1.1 was proved in [Capuano 2014] and [Capuano et al. 2016], giving an alternative proof of a result appearing in [Bombieri et al. 1999] and generalized in [Maurin 2008] (see also [Bombieri et al. 2008]).

We will use $\gamma_1, \gamma_2, \ldots$ to denote positive constants. The indices are reset at the end of each section.

### 2. The Zilber–Pink conjectures

In this section we see how our theorem relates to the so-called Zilber–Pink conjectures on unlikely intersections.
First, let us examine the objects we are investigating from the point of view of dimensions. We consider our elliptic curve $E_λ$ as an elliptic scheme over $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. Our ambient space is then the fiber power of $n$ copies of this elliptic scheme and has dimension $n + 1$. Now, for any choice of linearly independent vectors $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbb{Z}^n$, imposing the two corresponding conditions yields an $(n - 1)$-fold. Therefore, the intersection of a curve and an $(n - 1)$-fold in a space of dimension $n + 1$ is indeed unlikely to be nonempty and one expects finiteness for varying integer vectors.

Our result fits in the framework of very general conjectures formulated by Zilber [2002] and Bombieri, Masser and Zannier [Bombieri et al. 2007] in the toric case and by Pink [2005] in a more general setting, also known as the Zilber–Pink conjectures.

In a series of papers by Masser and Zannier [2010; 2012; 2014; 2015], the authors proved a variant of Pink’s conjecture in the case of a curve in an abelian surface scheme over $\overline{\mathbb{Q}}$, and over $\mathbb{C}$ in the nonsimple case. On the other hand, Pink’s conjecture concerns families of semiabelian varieties. However, Bertrand [2011] found a counterexample to this, for a suitable nonsplit extension of a CM elliptic constant family $E_0 \times B$ (over a curve $B$) by $\mathbb{G}_m$. This situation is rather “special”; in fact, as it is shown in [Bertrand et al. 2016], the possible presence of the so-called “Ribet sections” is the only obstruction to the validity of the conjecture in the case of semiabelian surface schemes.

Now, let us see how our Theorem 1.1 implies a statement in the spirit of the conjectures mentioned above. In particular, we translate our result in the language of schemes, borrowing some terminology and results from a work of Habegger [2013].

Let $S$ be an irreducible and nonsingular quasiprojective curve defined over $\overline{\mathbb{Q}}$ and let $\mathcal{E} \to S$ be an elliptic scheme over $S$, i.e., a group scheme whose fibers are elliptic curves. Let $n \geq 2$. We define $\mathcal{A}$ to be the $n$-fold fibered power $\mathcal{E} \times_S \cdots \times_S \mathcal{E}$ with the structural morphism $\pi : \mathcal{A} \to S$. We suppose that $\mathcal{E}$ is not isotrivial. In other words, $\mathcal{E} \to S$ cannot become a constant family after a finite étale base change.

A subgroup scheme $G$ of $\mathcal{A}$ is a closed subvariety, possibly reducible, which contains the image of $G \times_S G$ under the addition morphism and the image of the zero section $S \to \mathcal{A}$, and is mapped to itself by the inversion morphism. A subgroup scheme $G$ is called flat if $\pi_{\text{lg}} : G \to S$ is flat, i.e., all irreducible components of $G$ dominate the base curve $S$ (see [Hartshorne 1977, Chapter III, Proposition 9.7]).

**Theorem 2.1.** Let $\mathcal{A}$ be as above and let $\mathcal{A}^{[2]}$ be the union of the flat subgroup schemes of $\mathcal{A}$ with codimension at least 2. Let $\mathcal{C}$ be a curve in $\mathcal{A}$ defined over $\overline{\mathbb{Q}}$ and suppose $\pi(\mathcal{C})$ dominates $S$. Then $\mathcal{C} \cap \mathcal{A}^{[2]}$ is contained in a finite union of flat subgroup schemes of positive codimension.
In order to prove that this theorem is a consequence of Theorem 1.1, we need some notation and facts from [Habegger 2013].

For every \( a = (a_1, \ldots, a_n) \in \mathbb{Z}^n \) we have a morphism \( a : A \to E \) defined by
\[
a(P_1, \ldots, P_n) = a_1 P_1 + \cdots + a_n P_n.
\]

We identify the elements of \( \mathbb{Z}^n \) with the morphisms they define. The fibered product
\[
\alpha = a_1 \times_S \cdots \times_S a_r, \text{ for } a_1, \ldots, a_r \in \mathbb{Z}^n,
\]
defines a morphism \( A \to B \) over \( S \) where \( B \) is the \( r \)-fold fibered power of \( E \). The kernel of \( \alpha \), denoted by \( \ker \alpha \), indicates the fibered product of \( \alpha : A \to B \) with the zero section \( S \to B \). We consider it as a closed subscheme of \( A \).

**Lemma 2.2.** Let \( G \) be a codimension-\( r \) flat subgroup scheme of \( A \) with \( 1 \leq r \leq n \). Then there exist independent \( a_1, \ldots, a_r \in \mathbb{Z}^n \) such that \( G \subseteq \ker(a_1 \times_S \cdots \times_S a_r) \). Moreover, \( \ker(a_1 \times_S \cdots \times_S a_r) \) is a flat subgroup scheme of \( A \) of codimension \( r \).

**Proof.** This follows from Lemma 2.5 of [Habegger 2013] and its proof. \( \square \)

Consider the Legendre family defined by (1-1). This gives an example of an elliptic scheme, which we call \( E_L \), over the modular curve \( Y(2) = \mathbb{P}^1 \setminus \{0, 1, \infty\} \). We write \( A_L \) for the \( n \)-fold fibered power of \( E_L \).

**Lemma 2.3** [Habegger 2013, Lemma 5.4]. Let \( A \) be as above. After possibly replacing \( S \) by a Zariski open, nonempty subset, there exists an irreducible, nonsingular, quasiprojective curve \( S' \) defined over \( \overline{\mathbb{Q}} \) such that we have a commutative diagram
\[
\begin{array}{ccc}
A & \xleftarrow{f} & A' & \xrightarrow{e} & A_L \\
\pi \downarrow & & \downarrow & & \downarrow \pi_L \\
S & \xleftarrow{l} & S' & \xrightarrow{\lambda} & Y(2)
\end{array}
\]
where \( l \) is finite, \( \lambda \) is quasifinite, \( A' \) is the abelian scheme \( A \times_S S' \), \( f \) is finite and flat and \( e \) is quasifinite and flat. Moreover, the restriction of \( f \) and \( e \) to any fiber of \( A' \to S' \) is an isomorphism of abelian varieties.

**Lemma 2.4.** If \( G \) is a flat subgroup scheme of \( A \), then \( e(f^{-1}(G)) \) is a flat subgroup scheme of \( A_L \) of the same dimension. Moreover, let \( X \) be a subvariety of \( A \) dominating \( S \) and not contained in a proper flat subgroup scheme of \( A \), let \( X'' \) be an irreducible component of \( f^{-1}(X) \) and let \( X' \) be the Zariski closure of \( e(X'') \) in \( A_L \). Then \( X' \) has the same dimension as \( X \), dominates \( Y(2) \) and is not contained in a proper flat subgroup scheme of \( A_L \).

**Proof.** This follows from the proof of Lemma 5.5 of [Habegger 2013]. \( \square \)
Proof of Theorem 2.1. First, we can assume that \( C \) is not contained in a flat subgroup scheme of \( A \) of positive codimension. Therefore, it is enough to prove that \( C \cap \bigcup G \) is finite where the union is taken over all flat subgroup schemes of \( A \) of codimension at least 2.

Consider the Zariski closure \( C' \) of \( e(C'') \) for a component \( C'' \) of \( f^{-1}(C) \). By Lemma 2.4, \( C' \) is a curve in \( A_L \) dominating \( Y(2) \) and not contained in a proper flat subgroup scheme.

Now, since \( e \) is quasifinite, if \( e(f^{-1}(C \cap A[2])) \) is finite then \( C \cap A[2] \) is finite and by Lemma 2.4 we have

\[
e(f^{-1}(C \cap A[2])) \subseteq e(f^{-1}(C)) \cap A[2].
\]

Therefore, we can reduce to proving our claim for the Legendre family and for \( C' \).

By Lemma 2.2, each flat subgroup scheme of codimension at least 2 of \( A_L \) is contained in \( \ker(a_1 \times Y(2) a_2) \) for some independent \( a_1, a_2 \in \mathbb{Z}^n \). Therefore, it is enough to show that \( C' \cap \bigcup \ker(a_1 \times Y(2) a_2) \) is finite, where the union is taken over all pairs of independent \( a_1, a_2 \in \mathbb{Z}^n \). The claim follows by applying Theorem 1.1 since \( C' \) is not contained in a proper flat subgroup scheme. \( \square \)

3. O-minimal structures and a result of Pila

In this section we introduce the notion of an o-minimal structure, recall some definitions and properties we will need later and state a result from [Pila 2011]. For the basic properties of o-minimal structures we refer to [van den Dries 1998] and [van den Dries and Miller 1996].

**Definition 3.1.** A *structure* is a sequence \( S = (S_N) \), \( N \geq 1 \), where each \( S_N \) is a collection of subsets of \( \mathbb{R}^N \) such that, for each \( N, M \geq 1 \):

1. \( S_N \) is a boolean algebra (under the usual set-theoretic operations);
2. \( S_N \) contains every semialgebraic subset of \( \mathbb{R}^N \);
3. if \( A \in S_N \) and \( B \in S_M \) then \( A \times B \in S_{N+M} \);
4. if \( A \in S_{N+M} \) then \( \pi(A) \in S_N \), where \( \pi : \mathbb{R}^{N+M} \to \mathbb{R}^N \) is the projection onto the first \( N \) coordinates.

If \( S \) is a structure and, in addition,

5. \( S_1 \) consists of all finite union of open intervals and points,

then \( S \) is called an *o-minimal structure*.

Given a structure \( S \), we say that \( S \subseteq \mathbb{R}^N \) is a *definable set* if \( S \in S_N \).

Let \( U \subseteq \mathbb{R}^{N+M} \) and let \( \pi_1 \) and \( \pi_2 \) be the projection maps on the first \( N \) and on the last \( M \) coordinates, respectively. Now, for \( t_0 \in \pi_2(U) \), we define \( U_{t_0} = \{ x \in \mathbb{R}^N : (x, t_0) \in U \} = \pi_1(\pi_2^{-1}(t_0)) \) and call \( U \) a *family* of subsets of \( \mathbb{R}^N \), while
$U_0$ is called the fiber of $U$ above $t_0$. If $U$ is a definable set then we call it a definable family and one can see that the fibers $U_0$ are definable sets too. Let $S \subseteq \mathbb{R}^N$ and let $f : S \to \mathbb{R}^M$ be a function. We call $f$ a definable function if its graph \( \{(x, y) \in S \times \mathbb{R}^M : y = f(x)\} \) is a definable set. It is not hard to see that images and preimages of definable sets via definable functions are still definable.

There are many examples of o-minimal structures; see [van den Dries and Miller 1996]. In this article we are interested in the structure of globally subanalytic sets, usually denoted by $\mathbb{R}_{\text{an}}$. One of these is the fact that definable families have a uniform bound on the number of connected components of the fibers.

**Proposition 3.2** [van den Dries and Miller 1996, 4.4]. Let $U$ be a definable family. There exists a positive integer $\gamma$ such that each fiber of $U$ has at most $\gamma$ connected components.

Now, let $S \subseteq \mathbb{R}^N$ be a nonempty definable set and let $e$ be a nonnegative integer. The set of regular points of dimension $e$, denoted by $\text{reg}_e(S)$, is the set of points $x \in S$ such that there is an open neighborhood $I$ of $x$ for which $S \cap I$ is a $C^1$ (embedded) submanifold of $\mathbb{R}^N$ of dimension $e$. The dimension of $S$ is the maximum $e$ such that $S$ has a regular point of dimension $e$. Note that if $S$ has dimension $e$ then $S \setminus \text{reg}_e(S)$ has dimension $\leq e - 1$.

**Definition 3.3.** A definable block of dimension $e$ in $\mathbb{R}^N$ is a connected definable set $B$ of dimension $e$ contained in some semialgebraic set $A$ of dimension $e$, such that every point of $B$ is a regular point of dimension $e$ in $B$ and $A$. Dimension zero is allowed: a point is a definable block. Moreover, a definable block family is a definable family whose nonempty fibers are all definable blocks.

We now need to define the height of a rational point. The height used in [Pila 2011] is not the usual projective Weil height, but a coordinatewise affine height. If $a/b$ is a rational number written in lowest terms, then $H(a/b) = \max(|a|, |b|)$ and, for an $N$-tuple $(\alpha_1, \ldots, \alpha_N) \in \mathbb{Q}^N$, we set $H(\alpha_1, \ldots, \alpha_N) = \max H(\alpha_i)$. For a subset $Z$ of $\mathbb{R}^N$ and a positive real number $T$ we define

$$Z(\mathbb{Q}, T) = \{(\alpha_1, \ldots, \alpha_N) \in Z \cap \mathbb{Q}^N : H(\alpha_1, \ldots, \alpha_N) \leq T\}. \quad (3-1)$$

The following theorem is a special case of [Pila 2011, Theorem 3.6] (see also [Pila 2009]). Here, if $f$ and $g$ are real functions of $T$, the notation $f(T) \ll_{Z, \epsilon} g(T)$ means that there exists a constant $\gamma$, depending on $Z$ and $\epsilon$, such that $f(T) \leq \gamma g(T)$ for $T$ large enough.
Theorem 3.4 [Pila 2011]. Let \( Z \subseteq \mathbb{R}^N \times \mathbb{R}^M \) be a definable family, and let \( \epsilon > 0 \).
Then there exist a \( J = J(Z, \epsilon) \in \mathbb{N} \) and a collection of definable block families \( B^{(j)} \subseteq \mathbb{R}^N \times (\mathbb{R}^M \times \mathbb{R}^M_j) \), for \( j = 1, \ldots, J \), such that

1. each point in each fiber of \( B^{(j)} \) is regular of dimension \( e_j \);
2. for each \((t, u) \in \mathbb{R}^M \times \mathbb{R}^M_j\), the fiber \( B^{(j)}_{(t, u)} \) is contained in \( Z_t \);
3. for every \( t \in \pi_2(Z) \), the set \( Z_t(\mathbb{Q}, T) \) is contained in the union of \( \ll Z, \epsilon T^\epsilon \) definable blocks, each a fiber of one of the \( B^{(j)} \).

4. Points lying on rational linear varieties

Let \( n \geq 2 \) be an integer and let \( \ell_1, \ldots, \ell_n, f, g \) be holomorphic functions on a connected neighborhood \( I \) of some closed disc \( D \subseteq \mathbb{C} \). Suppose that

\[
\ell_1, \ldots, \ell_n \text{ are algebraically independent over } \mathbb{C}(f, g) \text{ on } D, \tag{4-1}
\]

and that \( f(\lambda) \) and \( g(\lambda) \) are \( \mathbb{R} \)-linearly independent for every \( \lambda \in D \).

For some positive real \( T \), denote by \( D(T) \) the set of \( \lambda \in D \) such that

\[
\begin{align*}
\sum a_1 \ell_1(\lambda) + \cdots + a_n \ell_n(\lambda) &= a_{n+1} f(\lambda) + a_{n+2} g(\lambda), \\
\sum b_1 \ell_1(\lambda) + \cdots + b_n \ell_n(\lambda) &= b_{n+1} f(\lambda) + b_{n+2} g(\lambda),
\end{align*}
\tag{4-2}
\]

for some linearly independent vectors \((a_1, \ldots, a_n), (b_1, \ldots, b_n) \in (\mathbb{Z} \cap [-T, T])^n\) and some \( a_{n+1}, a_{n+2}, b_{n+1}, b_{n+2} \in \mathbb{Z} \).

The following proposition gives the desired upper bound mentioned in the introduction. We postpone its proof until the end of this section after developing some auxiliary tools.

Proposition 4.1. Under the above hypotheses, for every \( \epsilon > 0 \), we have \(|D(T)| \ll_{\epsilon} T^\epsilon\).

Define

\[
\Delta = f \bar{g} - \bar{f} g,
\]

which does not vanish on \( D \), since \( f(\lambda) \) and \( g(\lambda) \) are \( \mathbb{R} \)-linearly independent for every \( \lambda \in D \). Moreover, let

\[
u_j = -\frac{\ell_j \bar{g} - \bar{\ell}_j g}{\Delta}, \quad v_j = -\frac{\ell_j \bar{f} - \bar{\ell}_j f}{\Delta}.
\]

One can easily check that these are real-valued and, furthermore, that we have

\[
\ell_j = u_j f + v_j g.
\]

If we view \( D \) and \( I \) as subsets of \( \mathbb{R}^2 \), then \( u_j \) and \( v_j \) are real analytic functions on \( I \).

Define

\[
\Theta : D \to \mathbb{R}^{2n}, \quad \lambda \mapsto (u_1(\lambda), v_1(\lambda), \ldots, u_n(\lambda), v_n(\lambda)),
\]
and set $S = \Theta(D)$. This is a definable set in $\mathbb{R}^a_n$. In what follows, $(u_1, v_1, \ldots, u_n, v_n)$ will just indicate coordinates in $\mathbb{R}^{2n}$.

For $T > 0$, we call $S(T)$ the set of points of $S$ of coordinates $(u_1, v_1, \ldots, u_n, v_n)$ such that there exist linearly independent vectors $(a_1, \ldots, a_{n+2})$, $(b_1, \ldots, b_{n+2})$ in $\mathbb{Q}^{n+2}$ of height at most $T$ with

$$\begin{cases} a_1 u_1 + \cdots + a_n u_n = a_{n+1}, \\ a_1 v_1 + \cdots + a_n v_n = a_{n+2}, \\ b_1 u_1 + \cdots + b_n u_n = b_{n+1}, \\ b_1 v_1 + \cdots + b_n v_n = b_{n+2}. \end{cases} \tag{4-3}$$

**Lemma 4.2.** For every choice of $a_1, \ldots, a_{n+2}, b_1, \ldots, b_{n+2} \in \mathbb{R}$, not all zero, the subset of $S$ for which (4-3) holds is finite.

**Proof.** By contradiction suppose that the subset of $S$ of points satisfying (4-3) for some choice of coefficients is infinite. We can suppose that at least one $a_j$ is nonzero. This would imply that there exists an infinite set $E \subseteq D$ on which, for every $\lambda \in E$,

$$a_1 \ell_1(\lambda) + \cdots + a_n \ell_n(\lambda) = a_{n+1} f(\lambda) + a_{n+2} g(\lambda).$$

Since this relation holds on a set with an accumulation point, it must hold on all of $D$ (see [Lang 1985, Chapter III, Theorem 1.2(ii)]), contradicting hypothesis (4-1).

The following proposition is the main tool used to prove Proposition 4.1.

**Proposition 4.3.** For every $\epsilon > 0$, we have $|S(T)| \ll_{\epsilon} T^\epsilon$.

**Proof.** We are counting points of $S$ that lie on linear varieties of $\mathbb{R}^{2n}$ defined by systems of the form (4-3).

Let us consider the set $W \subset \mathbb{R}^{4n+4}$ defined as

$$W = \{(u_1, v_1, \ldots, u_n, v_n, a_1, \ldots, a_{n+2}, b_1, \ldots, b_{n+2}) \in S \times \mathbb{R}^{2n+4} :$$

$$\begin{cases} (4-3) \text{ holds and } (a_1, \ldots, a_{n+2}), (b_1, \ldots, b_{n+2}) \text{ are linearly independent} \end{cases}\},$$

which is a definable set. Denote by $\pi_1$ the projection on $S$ and by $\pi_2$ the projection on the last $2n + 4$ coordinates. Given a point $L$ of $\pi_2(W)$, we write $\tau(L)$ for the set of points of $S$ that lie on the affine subspace corresponding to $L$, i.e., $\tau(L) = \pi_1(\pi_2^{-1}(L))$. In other words, if we consider $W$ as a family of subsets of $\mathbb{R}^{2n}$, then $\tau(L)$ is just the fiber $W_L$. This is a definable subset of $S$ and it must be zero-dimensional by Lemma 4.2. By Proposition 3.2, there exists a positive integer $\gamma_1$ such that $|\tau(L)| \leq \gamma_1$, for every $L \in \pi_2(W)$. If $V \subseteq \pi_2(W)$, we write $\tau(V)$ for $\pi_1(\pi_2^{-1}(V))$.

Now, let us set $\hat{W} = \pi_2(W) \subseteq \mathbb{R}^{2n+4}$. Recall the definition in (3-1) and note that $S(T) \subseteq \tau(\hat{W}(\mathbb{Q}, T))$. By Theorem 3.4, there is a finite number of definable
block families such that, for every $\epsilon_1$, the set $\hat{W}(\mathbb{Q}, T)$ is contained in the union of $\ll_{W, \epsilon_1} T^{\epsilon_2}$ definable blocks, each a fiber of one of these families. We have that $S(T) \subseteq \bigcup_B \tau(B)$, where the union is taken over the $\ll_{W, \epsilon_1} T^{\epsilon_2}$ definable blocks mentioned above.

Let us fix a definable block family $U$ with fibers $U_t \subseteq \hat{W}$. We claim that, for every $\epsilon_2$, each fiber $U_t$ of $U$ gives rise to $\ll_{U, \epsilon_2} T^{\epsilon_2}$ points on $S(T)$, i.e., that $|\tau(U_t) \cap S(T)| \ll_{U, \epsilon_2} T^{\epsilon_2}$ for every fiber $U_t$. Once we prove this, the claim of the proposition follows easily after fixing $\epsilon_1$ and $\epsilon_2$ with $\epsilon_1 \epsilon_2 = \epsilon$, e.g., $\epsilon_1 = \epsilon_2 = \sqrt{\epsilon}$.

We proceed by induction on the dimension $e$ of the fibers of $U$. By Lemma 4.2, the claim is true for $e = 0$.

Suppose now $e > 0$. We denote by $B_\eta(L)$ the Euclidean ball centered in $L$ of radius $\eta$ and define, for $m = 1, \ldots, \gamma_1$,

$$V^{(m)} = \{(L, t) \in U : \text{there exist an } \eta > 0 \text{ and } A_1, \ldots, A_m \in S \text{ such that, for all } L' \in B_\eta(L) \cap U_t, \text{ we have } \tau(L') = \{A_1, \ldots, A_m\}\}.$$ 

These are definable families and so is $V := \bigcup_{m=1}^{\gamma_1} V^{(m)}$, as it is a finite union of definable sets. Hence, by Proposition 3.2, there exists a $\gamma_2$ such that all fibers $V_t$ have at most $\gamma_2$ connected components. It is clear that, for each $L$ in the same connected component, $\tau(L)$ consists of the same set of not more than $\gamma_1$ points; therefore, each fiber $V_t$ of $V$ gives rise to at most $\gamma_1 \gamma_2$ points of $S(T)$, i.e., $|S(T) \cap \tau(V_t)| \leq \gamma_1 \gamma_2$.

Now we want to prove that all the fibers of $Z = U \setminus V$ have dimension $< e$. Suppose not and let $L$ be an $e$-regular point of a fiber $Z_t$. We fix a ball $B_\eta(L)$ such that $B_\eta(L) \cap U_t$ is connected and contained in $Z_t$. We set $\{A_1, \ldots, A_m\}$ equal to $\bigcap_{L' \in B_\eta(L) \cap Z_t} \tau(L')$, i.e., the set of points of $S$ that lie on all subspaces in $B_\eta(L) \cap Z_t$. By definition of $Z$, we know $\tau^{-1}(\{A_1, \ldots, A_m\}) \cap B_\eta(L) \cap Z_t$ must be of dimension $< e$; therefore, there exist an $L_0 \in B_\eta(L) \cap Z_t$ and an $\eta_0$ such that, for every $L' \in B_{\eta_0}(L_0) \cap Z_t$, we have $\tau(L') \supseteq \{A_1, \ldots, A_m\}$. Thus, we can define a function $f : B_{\eta_0}(L_0) \cap Z_t \to S$ that associates to $L'$ a point in $\tau(L')$ different from $A_1, \ldots, A_m$. This is a definable function and, taking $\eta_0$ small enough (and possibly choosing a different $L_0$), we can also suppose that it is differentiable [van den Dries and Miller 1996, C.2 Lemma].

Now, assume the derivative of $f$ is zero in all directions. Then $f$ is constant and there exists a point $A_{m+1} \in \tau(L')$ for all $L' \in B_{\eta_0}(L_0) \cap Z_t$. We repeat this procedure of finding a point, a ball and a function like above and continue until this function has nonzero derivative in some direction. This procedure must stop because otherwise we would have a point $L'$ with $|\tau(L')| > \gamma_1$, contradicting the above considerations.

We can therefore suppose that there are an $L_0 \in B_\eta(L) \cap Z_t$ and an $\eta_0$ such that $f$ is differentiable on $B_{\eta_0}(L_0) \cap Z_t$ and has nonzero derivative in some direction. Now, recall that $L_0$ is an $e$-regular point of $U_t$ and, by definition of definable block,
of a semialgebraic set that contains it. Therefore, $B_{n_0}(L_0) \cap U_i = B_{n_0}(L_0) \cap Z_i$ is semialgebraic. Thus, if we intersect it with a suitable linear variety, we get an algebraic curve segment $C$ in $B_{n_0}(L_0) \cap Z_i$, passing through $L_0$ in the direction for which the derivative of $f$ is nonzero. The function $f$ is nonconstant on $C$. Consider $C' = f(C) \times C$. By definition of $f$, we know $C'$ is a real-analytic curve segment in $W$. Moreover, let us define $D' = \Theta^{-1}(f(C))$. As $f$ is not constant on $C$, we know $D'$ is an infinite subset of $D$.

Now, on $D'$ the coordinate functions $a_1, \ldots, a_{n+2}, b_1, \ldots, b_{n+2}$ satisfy $2n + 3$ independent algebraic relations with coefficients in $\mathbb{C}$ and, combining the relations of (4-3), we have also

$$
\begin{align*}
    & a_1 \ell_1 + \cdots + a_n \ell_n = a_{n+1} f + a_{n+2} g, \\
    & b_1 \ell_1 + \cdots + b_n \ell_n = b_{n+1} f + b_{n+2} g.
\end{align*}
$$

Each of these two relations is independent of the previous $2n + 3$ relations, and they are independent of each other because $(a_1, \ldots, a_{n+2})$ and $(b_1, \ldots, b_{n+2})$ are required to be linearly independent. Therefore, as the $3n + 4$ functions $a_1, \ldots, a_{n+2}, b_1, \ldots, b_{n+2}, \ell_1, \ldots, \ell_n$ satisfy $2n + 5$ independent algebraic relations with coefficients in $\mathbb{C}[f, g]$ on the infinite set $D'$, they continue to do so on $I$. Therefore, if $F := \mathbb{C}(f, g)$,

$$
\text{tr \deg}_F F(\ell_1, \ldots, \ell_n) < n.
$$

This contradicts hypothesis (4-1).

We have just proved that there cannot be any $e$-regular point on any fiber of $Z$. We apply Pila’s result (Theorem 3.4) again on $Z$. There is a finite number of definable block families such that, for each $\epsilon_3$ and for each fiber $Z_t$, the set $Z_t(\mathbb{Q}, T)$ is contained in the union of $\ll_{Z, Z_t} T^{\epsilon_3}$ definable blocks, each a fiber of one of these families. The fibers of these families must have dimension $< e$; therefore, our inductive hypothesis implies that if $U'$ is one of them, then, for every $\epsilon_4 > 0$, we have $|\tau(U'_t) \cap S(T)| \ll_{U'_t, \epsilon_4} T^{\epsilon_4}$ for every fiber $U'_t$ of $U'$. This means that, after choosing $\epsilon_3 = \epsilon_4 = \sqrt{e_2}$, for each fiber $Z_t$, we have $|\tau(Z_t) \cap S(T)| \ll_{Z, \epsilon_2} T^{\epsilon_2}$. Now recall that we have $U_t = V_t \cup Z_t$ and that $V_t$ gives rise to at most $\gamma_1 \gamma_2$ points of $S(T)$. This proves our claim and the proposition.

\[\square\]

**Remark.** We would like to point out that this last proposition can be deduced from recent work of Habegger and Pila [2014, Corollary 7.2].

**Proof of Proposition 4.1.** Since $f$ and $g$ are linearly independent, if $\lambda \in D$ satisfies (4-2) then (4-3) holds for $\Theta(\lambda)$. Now, since $D$ is a compact subset of $\mathbb{R}^2$, each $\ell_j(D)$ is bounded and, therefore, if $\ell_1(\lambda), \ldots, \ell_n(\lambda), f(\lambda), g(\lambda)$ satisfy (4-2), then $|a_{n+1}|, |a_{n+2}|, |b_{n+1}|, |b_{n+2}|$ are bounded in terms of $|a_1|, \ldots, |a_n|, |b_1|, \ldots, |b_n|$ and thus of $T$. Therefore, $\Theta(\lambda) \in S(\gamma_3 T)$ for some $\gamma_3$ independent of $T$. Now, using Proposition 3.2 and Lemma 4.2, we see that there exists a $\gamma_4$ such that, for
any choice of \(a_1, \ldots, a_{n+2}, b_1, \ldots, b_{n+2}\), there are at most \(\gamma_4\) elements \(\lambda\) in \(D\) such that \(\ell_1(\lambda), \ldots, \ell_n(\lambda), f(\lambda), g(\lambda)\) satisfy (4-2). Thus \(|D(T)| \ll |S(\gamma_3 T)|\) and the claim follows from Proposition 4.3.

\[\Box\]

5. Periods and elliptic logarithms

In this section we introduce the functions to which we will apply Proposition 4.1. We follow [Masser and Zannier 2012].

It is well-known that there is an analytic isomorphism between \(E_\lambda(\mathbb{C})\) and \(\mathbb{C}/L_\lambda\), where \(L_\lambda\) is a rank-2 lattice in \(\mathbb{C}\). Consider the hypergeometric function

\[F(t) = F\left(\frac{1}{2}, \frac{1}{2}, 1; t\right) = \sum_{m=0}^{\infty} \frac{(2m)!^2}{2^{4m} m!^4} t^m,\]

and let

\[f(t) = \pi F(t) \quad \text{and} \quad g(t) = \pi i F(1 - t).\]  \tag{5-1}

Define

\[\Lambda = \{t \in \mathbb{C} : |t| < 1, |1 - t| < 1\}.

The functions \(f\) and \(g\) are well-defined and analytic in \(\Lambda\), as functions of \(t\). Moreover, they are well-defined as functions of \(c\) in \(\lambda^{-1}(\Lambda) \subset \mathcal{C}(\mathbb{C})\).

By [Husemöller 1987, Chapter 9, (6.1) Theorem, p. 179], \(f(\lambda)\) and \(g(\lambda)\) are basis elements of the period lattice \(L_\lambda\) of \(E_\lambda\) with respect to \(dX/(2Y)\). Therefore, if \(\exp_\lambda\) is the associated exponential map from \(\mathbb{C}\) to \(E_\lambda(\mathbb{C})\), we have

\[\exp_\lambda(f(\lambda)) = \exp_\lambda(g(\lambda)) = O,

where \(O\) denotes the origin in \(E_\lambda\). Let \(P_j = (x_j, y_j)\), where \(x_j, y_j\) are coordinate functions in \(\mathcal{C}(\mathbb{C})\). We can suppose, for every \(j\), that \(x_j \neq 0, 1, \lambda\) identically; otherwise the corresponding \(P_j\) would be identically 2-torsion, contradicting the hypothesis of Theorem 1.1.

Now, we want to define suitable functions \(z_j(c)\) such that \(\exp_{\lambda(c)}(z_j(c)) = P_j(c)\); in other words, we want \(z_j\) to be the elliptic logarithm of \(P_j\).

Let \(\widehat{\mathcal{C}}\) be the subset of points \(c \in \mathcal{C}\) such that \(\lambda(c), x_j(c) \neq 0, 1, \infty\) and \(x_j(c) \neq \lambda(c)\) for every \(j = 1, \ldots, n\), and such that \(c\) is not a singular point or a point on which the differential of \(\lambda\) vanishes.

Note that, in this way, we exclude finitely many \(c \in \mathcal{C}\), and these are algebraic points of \(\mathcal{C}\). Moreover, on \(\widehat{\mathcal{C}}\), the coordinate function \(\lambda\) has everywhere a local inverse.

We now follow the construction of [Masser and Zannier 2012, p. 459]. Fix a point \(c_* \in \widehat{\mathcal{C}}\) and choose a path in the \(x_j\)-plane from \(x_j(c_*)\) to \(\infty\) and not passing through 0, 1 and \(\lambda(c_*)\). We also fix a determination of \(Y = \sqrt{X(X - 1)(X - \lambda(c_*)})\) that is equal to \(y_j(c_*)\) at \(X = x_j(c_*)\). Therefore, the path corresponds to a path on
the elliptic curve \( E_{\lambda(c_s)} \) from the point \( P_j(c_s) \) to the origin \( O \). Hence we can define \( z_j(c_s) \) as the integral

\[
z_j(c_s) = \int_{x_j(c_s)}^{\infty} \frac{dX}{2\sqrt{X(X-1)(X-\lambda(c_s))}}.
\]

We can extend it to a \( c \) close to \( c_s \) by

\[
z_j(c) = \int_{x_j(c_s)}^{\infty} \frac{dX}{2\sqrt{X(X-1)(X-\lambda(c))}} + \int_{x_j(c_s)}^{x_j(c)} \frac{dX}{2\sqrt{X(X-1)(X-\lambda(c))}}.
\]

In fact, in the first integral on the right we use the same path fixed before and the integrand is determined by continuity from the previously chosen determination of \( Y \). Hence, this term is an analytic function in \( \lambda(c) \). For the second term, we can take any local path from \( x_j(c_s) \) to \( x_j(c) \). We can extend the integrand as a double power series in \( \lambda(c) - \lambda(c_s) \) and in \( X - x_j(c_s) \); the result will be a double power series in \( \lambda(c) - \lambda(c_s) \) and \( x_j(c) - x_j(c_s) \). Notice that we have, at any rate, \( \exp\lambda(c)(z_j(c)) = P_j(c) \) for every \( j = 1, \ldots, n \).

In this way, fixing a \( c_s \in \lambda^{-1}(\Lambda) \cap \hat{\mathcal{C}} \), the functions \( z_1, \ldots, z_n \) are well-defined on a small neighborhood \( N_s \) of \( c_s \) on \( \mathcal{C} \). Moreover, if we take \( N_s \) small enough, we can see them as analytic functions of \( \lambda \) on \( \lambda^{-1}(N_s) \).

We will need the following transcendence result.

**Lemma 5.1.** The functions \( z_1, \ldots, z_n \) are algebraically independent over \( \mathbb{C}(f, g) \) on \( N_s \).

**Proof.** The functions \( z_1, \ldots, z_n, f, g \) are analytic functions of \( \lambda \), linearly independent over \( \mathbb{Z} \). Indeed, a relation \( a_1z_1 + \cdots + a_nz_n = a_{n+1}f + a_{n+2}g \), with integer coefficients, would map via \( \exp\lambda \) to a relation of the form (1-2) on \( N_s \), and therefore on all of \( \mathcal{C} \), which cannot hold by the hypothesis of the theorem. Moreover, if \( \varphi_{\lambda,j} \) is the Weierstrass \( \wp \)-function associated to \( L_\lambda \), the \( \varphi_{\lambda,j}(z_j) \) are algebraic functions of \( \lambda \) because \( \varphi_{\lambda,j}(z_j) = x_j - \frac{1}{3}(\lambda + 1) \) (see [Masser and Zannier 2010, (3.8), p. 1683]). Therefore, the hypotheses of [Bertrand 1990, Théorème 5, p. 136] are satisfied and we can apply it to get the claim.

We would like now to extend our functions \( f, g, z_1, \ldots, z_n \) on \( \hat{\mathcal{C}} \).

If \( c \in \hat{\mathcal{C}} \), one can continue \( f \) and \( g \) to a neighborhood \( N_c \) of \( c \). In fact, if we choose \( c \in \hat{\mathcal{C}} \) and a path from \( c_s \) to \( c \) lying in \( \hat{\mathcal{C}} \), we can easily continue \( f \) and \( g \) along the path using (5-1).

To continue \( z_j \) from a point \( c_s \) to a \( c \) in \( \hat{\mathcal{C}} \), it is sufficient to verify that if \( N_1 \) and \( N_2 \) are two open small subsets in \( \hat{\mathcal{C}} \), with \( N_1 \cap N_2 \) connected, and if \( z_j \) has analytic definitions \( z_j' \) on \( N_1 \) and \( z_j'' \) on \( N_2 \), then \( z_j \) has an analytic definition on the union \( N_1 \cup N_2 \). But we saw that \( \exp\lambda(z_j) = P_j \) for every \( j = 1, \ldots, n \); hence on \( N_1 \cup N_2 \) we have \( \exp\lambda(z_j') = \exp\lambda(z_j'') \). This means that there exist rational integers \( u, v \)
with \( z_j'' = z_j' + uf + vg \) on this intersection, and they must be constant there. Hence it is enough to change \( z_j'' \) to \( z_j'' - uf - vg \) on \( N_2 \).

Using the same path, it is clear that we can continue the function \((f, g, z_1, \ldots, z_n)\) from a small neighborhood of \(c_*\) to a small neighborhood \(N_c \subseteq \hat{\mathcal{C}}\) of \(c\), and that the obtained function \((f^c, g^c, z_1^c, \ldots, z_n^c)\) is analytic on \(N_c\). Moreover, the functions preserve algebraic independence, as the following lemma shows.

**Lemma 5.2.** The functions \(z_1^c, \ldots, z_n^c\) are algebraically independent over \(\mathbb{C}(f^c, g^c)\) on \(N_c\).

**Proof.** Any algebraic relation can be continued to a neighborhood \(N_*\) of some \(c_* \in \lambda^{-1}(\Lambda)\), contradicting Lemma 5.1.

Furthermore, the lattice \(L_1\) is still generated by \(f^c\) and \(g^c\) on \(N_c\); see Lemma 6.1 of [Masser and Zannier 2012] or Lemma 4.1 of [Masser and Zannier 2010].

Now fix \(c \in \mathcal{C}\) and \(N_c \subseteq \hat{\mathcal{C}}\). Since we are avoiding singular points and points on which the differential of \(\lambda\) vanishes, \(\lambda\) gives an analytic isomorphism \(\lambda : N_c \to \lambda(N_c)\). Therefore, we can view \(z_1^c, \ldots, z_n^c, f^c, g^c\) as analytic functions on \(\lambda(N_c)\).

### 6. Linear relations on a fixed curve

In this section we prove a general fact about linear relations on elliptic curves.

For a point \((\alpha_1, \ldots, \alpha_N) \in \overline{\mathbb{Q}}^N\), the absolute logarithmic Weil height is defined by

\[
h(\alpha_1, \ldots, \alpha_N) = \frac{1}{[\mathbb{Q}(\alpha_1, \ldots, \alpha_N) : \mathbb{Q}]} \sum_v \log \max \{1, |\alpha_1|_v, \ldots, |\alpha_N|_v\},
\]

where \(v\) runs over a suitably normalized set of valuations of \(\mathbb{Q}(\alpha_1, \ldots, \alpha_N)\).

Let \(\alpha\) be an algebraic number and consider the Legendre curve \(E = E_\alpha\) defined by the equation \(Y^2 = X(X - 1)(X - \alpha)\). Let \(P_1, \ldots, P_n\) be linearly dependent points on \(E\), defined over some finite extension \(K\) of \(\mathbb{Q}(\alpha)\) of degree \(\kappa = [K : \mathbb{Q}]\). Suppose that \(P_1, \ldots, P_n\) have Néron–Tate height \(\hat{h}\) at most \(q\) (for the definition of Néron–Tate height, see for example [Masser 1988, p. 255]). In case the \(P_1, \ldots, P_n\) are all torsion, i.e., \(\hat{h}(P_j) = 0\) for all \(j\), we set \(q = 1\). We define

\[
L(P_1, \ldots, P_n) = \{(a_1, \ldots, a_n) \in \mathbb{Z}^n : a_1P_1 + \cdots + anP_n = O\},
\]

a sublattice of \(\mathbb{Z}^n\) of some positive rank \(r\). We want to show that \(L(P_1, \ldots, P_n)\) has a set of generators with small max norm \(|a| = \max\{|a_1|, \ldots, |a_n|\}\).

**Lemma 6.1.** Under the above hypotheses, there are generators \(a_1, \ldots, a_r\) of \(L(P_1, \ldots, P_n)\) with

\[
|a_i| \leq \gamma_1 \kappa^{\gamma_2} (h(\alpha) + 1)^{2n} q^{\frac{1}{2}(n-1)},
\]

for some positive constants \(\gamma_1, \gamma_2\) depending only on \(n\).
Proof. The Weierstrass form \( \tilde{E} = \tilde{E}_\alpha \) of \( E = E_\alpha \) has equation
\[
\tilde{Y}^2 = 4\tilde{X}^3 - g_2\tilde{X} - g_3,
\]
where \( g_2 = \frac{4}{3}(\alpha^2 - \alpha + 1) \) and \( g_3 = \frac{4}{27}(\alpha^2 - 2)(\alpha + 1)(2\alpha - 1) \) (see [Masser and Zannier 2010, (3.7), p. 1683]). The isomorphism \( \phi \) from \( E \) to \( \tilde{E} \) is given by
\[
\tilde{X} = X - \frac{1}{3}(\alpha + 1), \quad \tilde{Y} = 2Y.
\]

Now, \( \tilde{E} \) is clearly defined over \( \mathbb{Q}(\alpha) \) and any linear relation \( a_1 P_1 + \cdots + a_n P_n = O \) on \( E \) carries on to \( \tilde{E} \) and vice versa. Moreover, the \( Q_i = \phi(P_i) \) will have coordinates in \( K \) and the same Néron–Tate height of the \( P_i \), and hence \( \hat{h}(Q_i) \leq q \).

First, suppose that at least one of the points has infinite order. By Theorem E of [Masser 1988], if \( Q_1, \ldots, Q_n \) are linearly dependent points on \( \tilde{E}(K) \) of Néron–Tate height at most \( q \geq \eta \), then \( L(Q_1, \ldots, Q_n) \) is generated by vectors with max norm at most
\[
L^n - 1 \omega \left( \frac{q}{\eta} \right)^{\frac{1}{2}(n-1)},
\]
where \( \omega = |\tilde{E}_{\text{tors}}(K)| \) and \( \eta = \inf \hat{h}(P) \), for \( P \in \tilde{E}(K) \setminus \tilde{E}_{\text{tors}}(K) \). We need to bound \( \omega \) and \( \eta \). The constants \( \gamma_3, \ldots, \gamma_9 \) are absolute constants.

For the first we use Théorème 1.2(i) of [David 1997]: choosing any archimedean \( v \) and noting that, by David’s definition, \( h_v(\tilde{E}) \geq \sqrt{3} \), one has
\[
\omega \leq \gamma_3(kh + k \log k),
\]
where \( h = \max\{1, h(j_\tilde{E})\} \). Now, \( j_{\tilde{E}} = 2^8(\alpha^2 - \alpha + 1)^3/(\alpha^2(\alpha - 1)^2) \) (see for instance [Husemöller 1987], p. 83). Therefore, \( h \leq \gamma_4(h(\alpha) + 1) \) and
\[
\omega \leq \gamma_5(h(\alpha) + 1)k^2. \tag{6-1}
\]

For the lower bound on \( \eta \), we use a result of Masser [1989a, Corollary 1]. In Masser’s bound a constant depending on \( \kappa \) appears in the denominator. However, going through the proof one can see that this constant is polynomial in \( \kappa \), as noted on [David 1997, p. 109]. Therefore,
\[
\eta \geq \frac{\gamma_6}{w\kappa^{\gamma_7}(w + \log k)^2} \geq \gamma_8k^{-(\gamma_7+5)}w^{-3},
\]
where \( w = \max\{1, h(g_2), h(g_3)\} \). As \( g_2 \) and \( g_3 \) are polynomials in \( \alpha \), we have \( w \leq \gamma_9(h(\alpha) + 1) \). Consequently, \( L(Q_1, \ldots, Q_n) \) will have generators of norms at most
\[
\gamma_1k^{\gamma_2}(h(\alpha) + 1)^2q^{\frac{1}{2}(n-1)},
\]
with \( \gamma_1, \gamma_2 \) depending only on \( n \).
If all the points are torsion points, it is clear that one can take $|a_i| \leq \omega$ and use (6-1).

\section{Bounded height}

In this section we see that the height of the points on the curve $C$ for which there is at least one dependence relation is bounded and a few consequences of this fact.

Let $k$ be a number field over which $C$ is defined. Suppose also that the finitely many points we excluded from $C$ to get $\hat{C}$, which are algebraic, are defined over $k$. Clearly, there are $f_1, \ldots, f_n \in k[T]$ such that $f_j(x_j, \lambda) = 0$ for every $j$, identically on the curve.

Let $C'$ be the set of points of $\hat{C}$ such that $P_1, \ldots, P_n$ satisfy two independent relations on the specialized curve and let $c_0 \in C'$. Since $C$ is defined over $\bar{Q}$, the $x_j(c_0)$ and $\lambda(c_0)$ must be algebraic, unless the $P_j$ are identically linearly dependent, which we excluded by hypothesis. Then by Silverman’s specialization theorem [1983] (see also of [Zannier 2012, Appendix C]) there exists a $\gamma_1 > 0$ such that

$$h(\lambda(c_0)) \leq \gamma_1. \quad (7-1)$$

We see now a few consequences of this bound. If $\delta > 0$ is a small real number, let us set

$$\Lambda_\delta = \{ t \in \mathbb{C} : |t| \leq 1/\delta, |t - \lambda(c)| \geq \delta \text{ for all } c \in C \setminus \hat{C} \}.$$

\textbf{Lemma 7.1.} There is a positive $\delta$ such that there are at least $\frac{1}{2}[k(\lambda(c_0)) : k]$ different $k$-embeddings $\sigma$ of $k(\lambda(c_0))$ in $\mathbb{C}$ such that $\sigma(\lambda(c_0))$ lies in $\Lambda_\delta$ for all $c_0 \in C'$.

\textbf{Proof.} See Lemma 8.2 of [Masser and Zannier 2012].

\textbf{Remark.} We would like to point out that, as suggested by the referee, it might be possible to avoid the restriction to a compact domain and the use of the previous lemma by exploiting the work of Peterzil and Starchenko [2004], who proved that it is possible to define the Weierstrass $\wp$ function globally in the structure $\mathbb{R}_{\text{an,exp}}$.

\textbf{Lemma 7.2.} There exist positive constants $\gamma_2, \gamma_3$ such that, for every $c_0 \in C'$ and every $j = 1, \ldots, n$, we have

$$\widehat{h}(P_j(c_0)) \leq \gamma_2,$$

and the $P_j(c_0)$ are defined over some number field $K \supseteq k(\lambda(c_0))$ with

$$[K : \mathbb{Q}] \leq \gamma_3[k(\lambda(c_0)) : k].$$

\textbf{Proof.} Recall that each $x_j(c_0)$ is a root of $f_j(X, \lambda(c_0))$. This already implies the second statement. Now, we have $h(P_j(c_0)) \leq \gamma_3(h(\lambda(c_0)) + 1)$ and, using the work of Zimmer [1976], we have $\widehat{h}(P_j(c_0)) \leq h(P_j(c_0)) + \gamma_3(h(\lambda(c_0)) + 1)$. The first claim now follows from (7-1).
We want to show that there are at most finitely many \( c \) such that \( P_1(c), \ldots, P_n(c) \) satisfy two linearly independent relations on \( E_{\lambda(c)} \). By Northcott’s theorem [1949] and (7-1), we only need to bound the degree \( d \) of \( \lambda(c) \) over \( k \).

Let \( c_0 \in C', \lambda_0 = \lambda(c_0) \) and \( d_0 = [k(\lambda(c_0)) : k] \). First, by Lemma 7.1, we can choose \( \delta \), independent of \( c_0 \), such that \( \lambda_0 \) has at least \( d_0/2 \) conjugates in \( \Lambda_\delta \). Now, since \( \Lambda_\delta \) is compact, it can be covered by \( \gamma_2 \) closed discs \( D_{c_1}, \ldots, D_{c_{\gamma_2}} \subseteq \hat{\Lambda} \), where \( D_{c_i} \) is centered in \( \lambda(c_i) \), for some \( c_i \in \hat{\Lambda} \).

We can suppose that the closed disc \( D_{c_1} \) contains at least \( d_0/(2\gamma_2) \) conjugates \( \lambda_0 \). Now, each such conjugate comes from a \( c_0^* \in N_{c_1} \) and the corresponding points \( P_1(c_0^*), \ldots, P_n(c_0^*) \) satisfy the same linear relations. So there are linearly independent \( (a_1, \ldots, a_n), (b_1, \ldots, b_n) \) such that

\[
 a_1 P_1(c_0^*) + \cdots + a_n P_n(c_0^*) = b_1 P_1(c_0^*) + \cdots + b_n P_n(c_0^*) = O \quad (8-1)
\]
on \( E_{\lambda}(c_0^*) \).

By Lemma 7.2, \( \hat{h}(P_j(c_0^*)) \) is at most \( \gamma_3 \) and the points are defined over some finite extension of \( k(\lambda(c_0^*)) \) of degree at most \( \gamma_4 d_0 \). Therefore, applying Lemma 6.1 and recalling (7-1), we can suppose that the \( a_j \) and \( b_j \) are in absolute value less than or equal to \( \gamma_5 d_0^{\gamma_6} \).

Now, recall that, in Section 5, on \( \lambda(N_{c_1}) \supseteq D_{c_1} \), we defined \( f^{e_1}, g^{e_1} \) to be generators of the period lattice \( L_\lambda \) and the elliptic logarithms \( z_1^{e_1}, \ldots, z_n^{e_1} \) such that

\[
 \exp_\lambda(z_j^{e_1}(\lambda)) = P_j(\lambda) \quad (8-2)
\]on \( \lambda(N_{c_1}) \). We know that \( z_1^{e_1}, \ldots, z_n^{e_1}, f^{e_1}, g^{e_1} \) are holomorphic functions on a neighborhood of \( D_{c_1} \), with \( f^{e_1}(\lambda) \) and \( g^{e_1}(\lambda) \) linearly independent over \( \mathbb{R} \) for every \( \lambda \in D_{c_1} \), and, by Lemma 5.2, that \( z_1^{e_1}, \ldots, z_n^{e_1} \) are algebraically independent over \( \mathbb{C}(f^{e_1}, g^{e_1}) \) on \( D_{c_1} \). Therefore, the hypotheses of Proposition 4.1 are satisfied.

By (8-1) and (8-2), we have

\[
 a_1 z_1^{e_1}(\lambda_0^\sigma) + \cdots + a_n z_n^{e_1}(\lambda_0^\sigma) \equiv b_1 z_1^{e_1}(\lambda_0^\sigma) + \cdots + b_n z_n^{e_1}(\lambda_0^\sigma) \equiv 0 \pmod{L_{\lambda_0^\sigma}}.
\]

Therefore, there are \( a_{n+1}, a_{n+2}, b_{n+1}, b_{n+2} \in \mathbb{Z} \) such that

\[
 \begin{cases}
 a_1 z_1^{e_1}(\lambda_0^\sigma) + \cdots + a_n z_n^{e_1}(\lambda_0^\sigma) = a_{n+1} f^{e_1}(\lambda_0^\sigma) + a_{n+2} g^{e_1}(\lambda_0^\sigma), \\
 b_1 z_1^{e_1}(\lambda_0^\sigma) + \cdots + b_n z_n^{e_1}(\lambda_0^\sigma) = b_{n+1} f^{e_1}(\lambda_0^\sigma) + b_{n+2} g^{e_1}(\lambda_0^\sigma).
\end{cases}
\]

Thus all \( \lambda_0^\sigma \in D_{c_1} \) are in \( D_{c_1}(\gamma_5 d_0^{\gamma_6}) \) (recall the definition of \( D(T) \) just above Proposition 4.1).

By Proposition 4.1, we have \( |D_{c_1}(\gamma_5 d_0^{\gamma_6})| \ll d_0^{\gamma_6} \). But by our choice of \( D_{c_1} \) we have at least \( d_0/(2\gamma_2) \) points in \( D_{c_1}(\gamma_5 d_0^{\gamma_6}) \). Therefore, if we choose \( \epsilon < 1/\gamma_6 \) we have a contradiction when \( d_0 \) is large enough.
We just deduced that $d_0$ is bounded and, by \eqref{eq:1} and Northcott’s theorem, we have finiteness of the possible values of $\lambda(c_0)$, which proves Theorem \ref{thm:1}.

Acknowledgments

The authors would like to thank Umberto Zannier for his support, Daniel Bertrand, Philipp Habegger, Lars Kühne, Vincenzo Mantova and Harry Schmidt for useful discussions and the referee for very helpful comments that improved this article.

References


Linear relations in families of powers of elliptic curves


Communicated by Jonathan Pila
Received 2015-01-29 Revised 2015-10-01 Accepted 2015-11-27

fbarroero@gmail.com Classe di Scienze, Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy
laura.capuano1987@gmail.com Classe di Scienze, Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy
Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

**Originality.** Submission of a manuscript acknowledges that the manuscript is original and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

**Language.** Articles in ANT are usually in English, but articles written in other languages are welcome.

**Length** There is no a priori limit on the length of an ANT article, but ANT considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

**Required items.** A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

**Format.** Authors are encouraged to use \LaTeX{} but submissions in other varieties of \LaTeX{}, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

**References.** Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of Bib\LaTeX{} is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

**Figures.** Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

**White space.** Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

**Proofs.** Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.
Stable sets of primes in number fields
ALEXANDER IVANOV

Hopf–Galois structures arising from groups with unique subgroup of order $p$
TIMOTHY KOHL

On tensor factorizations of Hopf algebras
MARC KEILBERG and PETER SCHAUENBURG

Extension theorems for reductive group schemes
ADRIAN VASIU

Actions of some pointed Hopf algebras on path algebras of quivers
RYAN KINSER and CHELSEA WALTON

On the image of the Galois representation associated to a non-CM Hida family
JACLYN LANG

Linear relations in families of powers of elliptic curves
FABRIZIO BARROERO and LAURA CAPUANO