J. reine angew. Math. **788** (2022), 219–238
DOI 10.1515/crelle-2022-0028
Journal für die reine und angewandte Mathematik

Level structure, arithmetic representations, and noncommutative Siegel linearization

By *Borys Kadets* at Athens, GA and *Daniel Litt* at Athens, GA

Abstract. Let ℓ be a prime, k a finitely generated field of characteristic different from ℓ, and X a smooth geometrically connected curve over k. Say a semisimple representation of $\pi_1^{\text{ét}}(X_{\bar{k}})$ is arithmetic if it extends to a finite index subgroup of $\pi_1^{\text{ét}}(X)$. We show that there exists an effective constant $N = N(X, \ell)$ such that any semisimple arithmetic representation of $\pi_1^{\text{ét}}(X_{\bar{k}})$ into $\mathrm{GL}_n(\overline{\mathbb{Z}_\ell})$, which is trivial mod ℓ^N, is in fact trivial. This extends a previous result of the second author from characteristic zero to all characteristics. The proof relies on a new noncommutative version of Siegel's linearization theorem and the ℓ-adic form of Baker's theorem on linear forms in logarithms.

1. Introduction

The main goal of this note is to analyze representations of arithmetic fundamental groups, motivated by questions about level structure of Abelian varieties over function fields. Our main result (Theorem 1.1.2) implies that there is an absolute bound on the maximum N such that an Abelian scheme over a fixed curve over a field of characteristic prime to ℓ has full level ℓ^N-structure (Corollary 1.1.3). The contribution of this note is to show that this phenomenon, which was already known to hold in characteristic zero (see [9, Theorem 1.2] and [10, Theorem 1.1.13]) in fact holds in arbitrary characteristic. The proof in positive characteristic requires significant input from non-Archimedean dynamics and transcendence theory: in particular, a new noncommutative, non-Archimedean version of Siegel's linearization theorem (Theorem 3.2.1), which may be of independent interest, and the ℓ-adic version of Baker's theorem on linear forms in logarithms (due to Kunrui Yu [15]).

1.1. Main results. In preparation for the statement of the main theorem, we recall the definition of an arithmetic representation. The main interest in this notion stems from the fact

The corresponding author is Daniel Litt.

This material is partly based upon work supported by the NSF grant DMS-1928930 while the first author participated in a program hosted by the Mathematical Sciences Research Institute in Berkeley, California, during the Fall 2020 semester. The second author was supported by NSF grant DMS-2001196.

that representations "arising from geometry" are arithmetic. (See, e.g., [10] for a discussion of arithmeticity, its properties, and its consequences.)

Definition 1.1.1. Let k be a finitely generated field, and let X/k be a variety. Let k^{sep} be a separable closure of k, and let $\bar{x}$ be a geometric point of X. Then a continuous representation

$$\rho : \pi_1^{\text{ét}}(X_{k^{\text{sep}}}, \bar{x}) \to \mathrm{GL}_n(\overline{\mathbb{Q}_\ell})$$

is said to be *semisimple arithmetic* if

(i) ρ is semisimple, and

(ii) there exists a finite separable extension k'/k and a representation

$$\widetilde{\rho} : \pi_1^{\text{ét}}(X_{k'}, \bar{x}) \to \mathrm{GL}_n(\overline{\mathbb{Q}_\ell})$$

such that $\widetilde{\rho}|_{\pi_1(X_{k^{\text{sep}}}, \bar{x})}$ is conjugate to ρ.

Our main result about semisimple arithmetic representations is:

Theorem 1.1.2. *Let k be a finitely-generated field, and let X/k be a smooth curve. Let ℓ be a prime not equal to the characteristic of k, and let k^{sep} be a separable closure of k. Let $\bar{x}$ be a geometric point of X. There exists a positive constant $N = N(X, \ell)$ such that if*

$$\rho : \pi_1^{\acute{e}t}(X_{k^{\text{sep}}}, \bar{x}) \to \mathrm{GL}_n(\overline{\mathbb{Z}_\ell})$$

is a continuous representation such that

(i) *$\rho \otimes \overline{\mathbb{Q}_\ell}$ is semisimple arithmetic, and*

(ii) *ρ is trivial modulo ℓ^N,*

then ρ is trivial.

For a real number N we say that a representation is *trivial modulo ℓ^N* if it is trivial modulo the ideal

$$\{x \in \overline{\mathbb{Z}_\ell} \mid v_\ell(x) \geqslant N\}.$$

Theorem 1.1.2 immediately implies (by the Lang-Néron theorem, [3, Theorem 2.1]):

Corollary 1.1.3. *Let X, N, ℓ be as above, and let η be the function field of $X_{\bar{k}}$. Then for any integer $M > N$, and any Abelian scheme $A/X_{\bar{k}}$, the following holds: if the Abelian variety A_η has full ℓ^M-torsion (that is, $A_\eta[\ell^M](\eta) = A_\eta[\ell^M](\overline{\eta})$), then A_η is isogenous to an isotrivial abelian variety over η.*

Remark 1.1.4. In fact, the conclusion of Corollary 1.1.3 is equivalent to the claim that A_η is isogenous to $\mathrm{Tr}_{\eta/\bar{k}}(A_\eta)_\eta$. In characteristic zero, we may conclude that A_η is in fact isotrivial; this is not the case in positive characteristic. See, e.g., the well-known examples due to Moret-Bailly [11, 12].

Remark 1.1.5. The constant N in Theorem 1.1.2 and Corollary 1.1.3 is in principle explicit: it depends only on ℓ and on the natural Galois representation

$$\mathrm{Gal}(k^{\text{sep}}/k) \to \mathrm{GL}(H^1(X_{\bar{k}}, \mathbb{Z}_\ell)).$$

See Section 5 for related questions.

1.2. Relation to previous work. As far as we know, this is the first result of this form in positive characteristic. In characteristic zero, Theorem 1.1.2 was already known by work of the second author [9, Theorem 1.2]. In fact, more is known ([10, Theorem 1.1.13]): namely, that for fixed X, the constants $N(X, \ell)$ appearing in the statement of Theorem 1.1.2 may be taken to go to zero as $\ell \to \infty$. In positive characteristic we are unable to prove even that $N(X, \ell)$ may be bounded independent of ℓ, as the existing bounds arising from the ℓ-adic form of Baker's theorem on linear forms in logarithms *get worse* as $\ell \to \infty$ (see Question 5.3.1).

There is also related work of an analytic nature in characteristic zero; see, e.g., [1,2,7,13].

1.3. Outline of proof. We briefly sketch the idea of the proof of Theorem 1.1.2, which broadly follows the strategy of the works [9, 10], but replaces the use of Galois homotheties (arising from Bogomolov's ℓ-adic open image theorem) by the use of Frobenii, and replaces the naive estimates there with more sophisticated estimates arising from our form of Siegel's theorem.

A specialization argument reduces the theorem to the case where k is finite, X is affine, and $X(k)$ is nonempty. Let $x \in X(k)$ be a rational point and $\bar{x}$ an associated geometric point; let $\pi_1^\ell(X_{\bar{k}}, \bar{x})$ be the pro-ℓ completion of the geometric étale fundamental group of X. The pro-ℓ group ring $\mathbb{Z}_\ell\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ is (non-canonically) isomorphic to a ring of noncommutative power series over $\mathbb{Z}_\ell$. Letting K be a finite extension of $\mathbb{Q}_\ell$, the completion $K\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ of $\mathbb{Z}_\ell\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle \otimes K$ at the augmentation ideal is (non-canonically) isomorphic to a ring of noncommutative power series over K. For each positive real number $0 < r < 1$ we introduce a certain Banach sub-algebra $K^{\leqslant r}\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ of $K\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$, containing $\mathbb{Z}_\ell\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$, with the following property: any representation ρ of $\pi_1^\ell(X_{\bar{k}}, \bar{x})$ into $\mathrm{GL}_n(\overline{\mathbb{Z}_\ell})$, which is trivial modulo ℓ^N, extends canonically to a continuous homomorphism

$$\widetilde{\rho} : K^{\leqslant r'}\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle \to \mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}_\ell})$$

for every $r' > \ell^{-N}$.

Our main technical result is that for some finite extension K of $\mathbb{Q}_\ell$, the eigenvectors of the Frobenius action on the Banach algebra $K^{\leqslant r}\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ have dense span once r is sufficiently small, say for $0 < r \leqslant r_0$. This follows from a noncommutative variant of Siegel's linearization theorem (Theorem 3.2.1), whose proof is a version of Newton's method. The hypotheses of the theorem follow in our case from the Weil conjectures for curves, the ℓ-adic form of Baker's theorem on linear forms in logarithms (due to Kunrui Yu [15]), and a semisimplicity result proven in previous work of the second author [9, Theorem 2.20]. The introduction of these dynamical techniques is the main innovation of this work.

Now suppose we take N greater than $-\log r_0/\log \ell \in \mathbb{R}$, and ρ is as in Theorem 1.1.2. Then since $r_0 > \ell^{-N}$, we obtain a canonical continuous map

$$\widetilde{\rho} : K^{\leqslant r_0}\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle \to \mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}_\ell})$$

extending ρ. The homomorphism $\widetilde{\rho}$ is equivariant for the action of some power of Frobenius by arithmeticity of $\rho \otimes \mathbb{Q}_\ell$. Now the finite-dimensionality of $\mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}_\ell})$ implies that almost all eigenvectors of the Frobenius action on $K^{\leqslant r_0}\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ are sent to zero by $\widetilde{\rho}$; their density implies that some power of the augmentation ideal of $K^{\leqslant r_0}\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ and hence of $\mathbb{Z}_\ell\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x}\rangle\rangle$ is sent to zero. Thus the representation ρ was unipotent, and any unipotent semisimple representation is trivial, completing the proof.

2. Notation and preliminaries

2.1. Notation. Throughout X will be a smooth, geometrically connected curve over a field k, and ℓ will be a prime different from the characteristic of k. Fix an algebraic closure $\bar{k}$ of k and a geometric point $\bar{x}$ of X. Let K denote a finite extension of $\mathbb{Q}_\ell$. We denote by $\pi_1^\ell(X_{\bar{k}}, \bar{x})$ the pro-ℓ completion of $\pi_1^{\text{ét}}(X_{\bar{k}}, \bar{x})$. We denote by

$$\mathbb{Z}_\ell\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle := \varprojlim \mathbb{Z}_\ell[H]$$

the $\mathbb{Z}_\ell$-group ring of $\pi_1^\ell(X_{\bar{k}}, \bar{x})$; the inverse limit is taken over all finite quotients of $\pi_1^\ell(X_{\bar{k}}, \bar{x})$. Let $\mathscr{I} \subset \mathbb{Z}_\ell\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ be the augmentation ideal. Then we denote by

$$K\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle := \varprojlim_n (\mathbb{Z}_\ell\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle/\mathscr{I}^n \otimes K)$$

the completion of $\mathbb{Z}_\ell\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle \otimes K$ at the augmentation ideal. We abuse notation and denote the augmentation ideal of $K\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ by $\mathscr{I}$ as well.

The key ingredient of our argument is an analysis of the Frobenius action on the algebra $K\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ and related algebras. For X an affine curve, this ring is non-canonically isomorphic to a noncommutative power series ring in n variables $x_1, \dots, x_n$, where $x_i = \gamma_i - 1$ for $\{\gamma_1, \dots, \gamma_n\}$ a minimal set of topological generators of the free pro-ℓ group $\pi_1^\ell(X_{\bar{k}}, \bar{x})$. Dealing with coefficients of such noncommutative power series brings some notational difficulties. In this section we introduce some conventions that somewhat simplify the notation.

We write $\vec{x}$ as a shorthand for $x_1, \dots, x_n$, so that if X is affine, $K\langle\langle\pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle = K\langle\langle\vec{x}\rangle\rangle$, the non-commutative power series ring in $x_1, \dots, x_n$. If $I = \{i_1, \dots, i_m\}$ is a finite word in the alphabet $\{1, \dots, n\}$, we write $x^I := x_{i_1} \cdots x_{i_m}$. For an element $f \in K\langle\langle\vec{x}\rangle\rangle$ we write $f = \sum a_I x^I$ suppressing that summation is to be taken over all finite words in $\{1, \dots, n\}$. If I is a finite word, then $|I|$ denotes its length. The *weight* of the monomial x^I is $|I|$.

Definition 2.1.1. A power series $f \in K\langle\langle\vec{x}\rangle\rangle$ *converges on a disk of radius* $0 < r < 1$ if

$$\lim_{|I|\to\infty} |a_I| r^{|I|} = 0.$$

In this case we define the r-norm of f to be

$$\|f\|_r := \sup_I |a_I| r^{|I|}.$$

The set of convergent power series is denoted by $K^{\leq r}\langle\langle\vec{x}\rangle\rangle$.

We think of these convergent power series as functions on a closed noncommutative polydisk. From the point of view of this metaphor the Frobenius action is an automorphism of the polydisk fixing zero. We think of our main result as saying that this automorphism is conjugate to a linear map in a neighborhood of zero; the analogous result in holomorphic dynamics on a complex polydisk is Siegel's linearization theorem.

Proposition 2.1.2. *The pair* $K^{\leq r}\langle\langle\vec{x}\rangle\rangle, \|\cdot\|_r$ *is a Banach algebra.*

Proof. The fact that $\|\cdot\|_r$ is a non-Archimedean norm follows from the fact that $|\cdot|$ is a norm on K. For completeness, observe that a Cauchy sequence has to converge coefficient-wise, and the resulting limit is also the limit in the norm. □

We similarly define $K^{\leqslant r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ to be the subring of $K\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ consisting of those elements f that satisfy

$$\lim_{n\to\infty} \left(r^n \inf\{\ell^s \mid \ell^s \cdot f \in \mathscr{O}_K \otimes \mathbb{Z}_\ell\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle \bmod \mathscr{I}^{n+1}\}\right) = 0,$$

and define

$$\|f\|_r := \sup_n \left(r^n \inf\{\ell^s \mid \ell^s \cdot f \in \mathscr{O}_K \otimes \mathbb{Z}_\ell\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle \bmod \mathscr{I}^{n+1}\}\right).$$

Under a choice of isomorphism $K\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle \simeq K\langle\langle \vec{x}\rangle\rangle$ arising from a minimal set of topological generators of $\pi_1^\ell(X_{\bar{k}}, \bar{x})$ in the case X is affine, these two definitions are compatible, i.e., the isomorphism induces an isomorphism $K^{\leqslant r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle \simeq K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$. Moreover, these constructions are functorial: any endomorphism of $\pi_1^\ell(X_{\bar{k}}, \bar{x})$ induces an endomorphism of $K^{\leqslant r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ and hence of $K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$ when X is affine.

We denote by $\mathscr{I}_r$ the ideal of power series with zero constant term in $K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$; if the radius r is clear from context, we write simply $\mathscr{I}$. To emphasize the analogy with analysis, we write $f = g + O(x^n)$ to mean $f - g \in \mathscr{I}^n$.

2.2. Basic properties. We will study the dynamics of noncommutative power series on a polydisk. We use $\mathrm{End}^{\mathrm{op}}\, K\langle\langle \vec{x}\rangle\rangle$ to denote the monoid of n-tuples of power series

$$\vec{f} = (f_1, \ldots, f_n) \in K\langle\langle \vec{x}\rangle\rangle^n$$

with no constant term and with composition given by $\vec{f} \circ \vec{g} = (f_1(\vec{g}), \ldots, f_n(\vec{g}))$; elements of $\mathrm{End}^{\mathrm{op}}\, K\langle\langle \vec{x}\rangle\rangle$ define endomorphisms of $K\langle\langle \vec{x}\rangle\rangle$ that send x_i to f_i. Note that the order of composition on $\mathrm{End}^{\mathrm{op}}\, K\langle\langle \vec{x}\rangle\rangle$ is opposite of the natural composition on the endomorphisms, hence the notation. Similarly we use $\mathrm{End}^{\mathrm{op}}\, K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$ for the n-tuples $\vec{f} = (f_1, \ldots, f_n)$ of power series in $K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$ with no constant term and such that $\|f_i\|_r \leqslant r$ for all i. The norm of $\vec{f} \in \mathrm{End}^{\mathrm{op}}\, K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$ is defined by

$$\|\vec{f}\|_r := \max_i \|f_i\|_r.$$

We think of an element $\vec{f} \in \mathrm{End}^{\mathrm{op}}\, K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$ as a holomorphic map from the noncommutative polydisk $K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$ to itself, fixing the origin.

The algebras $K^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$ will be used to study representations trivial modulo ℓ^N with the help of the following lemma. In the lemma, the ℓ-adic valuation $v_\ell(A)$ of a matrix A is the minimal valuation of its entries, and the ℓ-adic norm is defined by $|A| = \ell^{-v_\ell(A)}$.

Lemma 2.2.1. *Suppose $\rho : \mathbb{Z}_\ell\langle\langle \vec{x}\rangle\rangle \to \mathrm{Mat}_{n\times n}(\overline{\mathbb{Z}_\ell})$ is a continuous representation such that $|\rho(x_i)| \leqslant C < 1$ for all i. Then for any $r > C$ the representation ρ extends to a continuous representation $\rho_r : \mathbb{Q}_\ell^{\leqslant r}\langle\langle \vec{x}\rangle\rangle \to \mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}_\ell})$.*

Proof. Suppose $f = \sum a_I x^I \in \mathbb{Q}_\ell^{\leqslant r}\langle\langle \vec{x}\rangle\rangle$ is a power series. Then

$$|a_I \rho(x)^I| \leqslant \|f\|_r r^{-|I|} C^{|I|}.$$

As $r^{-1}C < 1$, the series $f(\rho(\vec{x}))$ converges, and if $c := \max_i r^{-i} C^i$, then $|f(\rho(\vec{x}))| \leqslant c\|f\|_r$ by the ultrametric inequality. Therefore the representation $\rho_r : \mathbb{Q}_\ell^{\leqslant r}\langle\langle \vec{x}\rangle\rangle \to \mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}_\ell})$ given by $\rho_r : f \mapsto f(\rho(\vec{x}))$ is a well-defined continuous extension of ρ. □

Lemma 2.2.2. *The monomials of degree* $\geqslant m$ *have dense span in* $\mathscr{I}_r^m \subset K^{\leqslant r}\langle\!\langle \vec{x} \rangle\!\rangle$.

Proof. Let

$$f(\vec{x}) = \sum_{|I| \geqslant m} a_I x^I$$

be an element of $\mathscr{I}_r^m$. As $|a_I| r^I \to 0$ as $r \to \infty$, we have that, defining

$$f_M := \sum_{|I| \geqslant M} a_I x^I,$$

$f_M \to 0$ as $M \to \infty$. Hence $f = \lim_{M \to \infty} f - f_M$. Now we have written f as a limit of polynomials in the x_i, as desired. □

3. Siegel's linearization theorem

The goal of this section is to show that the Frobenius action on the convergent group ring $K^{\leqslant r}\langle\!\langle \pi_1^{\ell}(X, x) \rangle\!\rangle$ can be diagonalized once r is sufficiently small. To this end we prove a version of Siegel's linearization theorem for ℓ-adic noncommutative multivariate power series. Our argument is analogous to the one used in [6, Section 4], which itself is an ℓ-adic version of a classical argument of Rüssmann, as generalized by Zehnder [14, 16]. That said, we require these results without the "non-resonance" condition they impose. (In our application, this non-resonance condition is replaced by the use of the semisimplicity of the Frobenius action on the algebra $K\langle\!\langle \pi_1^{\ell}(X, x) \rangle\!\rangle$.) We start by establishing some simple lemmas on composition of power series.

3.1. Some lemmas. The following lemma establishes basic analytic properties of the set $\operatorname{End}^{\mathrm{op}} K^{\leqslant r}\langle\!\langle \vec{x} \rangle\!\rangle$.

Lemma 3.1.1. *The following statements hold:*

(i) *The set* $\operatorname{End}^{\mathrm{op}} K^{\leqslant r}\langle\!\langle \vec{x} \rangle\!\rangle$ *is a complete abelian group under the norm* $\|\cdot\|_r$ *with respect to addition.*

(ii) *The group* $\operatorname{End}^{\mathrm{op}} K^{\leqslant r}\langle\!\langle \vec{x} \rangle\!\rangle \subset \operatorname{End}^{\mathrm{op}} K\langle\!\langle \vec{x} \rangle\!\rangle$ *is closed under composition.*

(iii) *Composition on* $\operatorname{End}^{\mathrm{op}} K^{\leqslant r}\langle\!\langle \vec{x} \rangle\!\rangle$ *is continuous and satisfies the inequalities*

$$\|\vec{f} \circ \vec{g} - \vec{f} \circ \vec{h}\|_r \leqslant \frac{1}{r} \|\vec{f}\|_r \|\vec{g} - \vec{h}\|_r,$$

$$\|\vec{g} \circ \vec{f} - \vec{h} \circ \vec{f}\|_r \leqslant \frac{1}{r} \|\vec{f}\|_r \|\vec{g} - \vec{h}\|_r$$

for all $f, g, h \in \operatorname{End}^{\mathrm{op}} K^{\leqslant r}\langle\!\langle \vec{x} \rangle\!\rangle$.

(iv) *For* $0 < r_1 < r_2 < 1$ *we have* $K^{\leqslant r_2}\langle\!\langle \vec{x} \rangle\!\rangle \subset K^{\leqslant r_1}\langle\!\langle \vec{x} \rangle\!\rangle \subset K\langle\!\langle \vec{x} \rangle\!\rangle$, *and the inclusion is continuous. Given* $\vec{f} \in K^{\leqslant r_2}\langle\!\langle \vec{x} \rangle\!\rangle$, *we have*

$$\|\vec{f}\|_{r_1} \leqslant \frac{r_1}{r_2} \|\vec{f}\|_{r_2}.$$

Proof. (i) This follows from the completeness of K; if $\{\vec{f}_j\}$ is a Cauchy sequence in $\operatorname{End}^{\mathrm{op}} K^{\leqslant r}\langle\!\langle \vec{x} \rangle\!\rangle$, then the coefficients of x^I in each component of $\vec{f}_j$ form a Cauchy sequence themselves.

(ii) Given $\vec{f}, \vec{g} \in \operatorname{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle\vec{x}\rangle\rangle$, we wish to show that $\|f_i(\vec{g})\|_r \leqslant r$ for all i. Writing $f_i = \sum a_I^i x^I$, $g_j = \sum b_{I,j} x^I$, the coefficient of x^I in $f_i(\vec{g})$ is an integer linear combination of $a_J b_{I_1,j_1} \cdots b_{I_{|J|},j_{|J|}}$, where $\sum_j |I_j| = |I|$. Hence

$$\|f_i(\vec{g})\|_r \leqslant \sup_{I,J,I_1,\dots,I_{|J|}} |a_J b_{I_1,j_1} \cdots b_{I_{|J|},j_{|J|}}| r^{|I|} \leqslant \sup_J |a_J| r^{|J|} \leqslant r$$

as desired.

(iii) We prove the first inequality; the second is the special case where we apply the first inequality to $\|(\vec{g}-\vec{h}) \circ \vec{f} - (\vec{g}-\vec{h}) \circ \vec{0}\|_r$. Setting

$$f_i = \sum a_I x^I, \quad g_j = \sum b_{I,j} x^I, \quad h_k = \sum c_{I,k} x^I,$$

we have that the coefficient of x^I in $f_i(\vec{g}) - f_i(\vec{h})$ is an integer linear combination of

$$a_J (b_{I_1,j_1} \cdots b_{I_{|J|},j_{|J|}} - c_{I_1,j_1} \cdots c_{I_{|J|},j_{|J|}})$$

with $\sum_s I_s = I$. Using the telescoping sum

$$\begin{aligned} &b_{I_1,j_1} \cdots b_{I_{|J|},j_{|J|}} - c_{I_1,j_1} \cdots c_{I_{|J|},j_{|J|}} \\ &\quad = \sum_{s=1}^{J} b_{I_1,j_1} \cdots b_{I_{s-1},j_{s-1}} (b_{I_s}, j_s - c_{I_s,j_s}) c_{I_{s+1},j_{s+1}} \cdots c_{I_{|J|},j_{|J|}}, \end{aligned}$$

we have

$$\begin{aligned} &|b_{I_1,j_1} \cdots b_{I_{|J|},j_{|J|}} - c_{I_1,j_1} \cdots c_{I_{|J|},j_{|J|}}| r^{|I|} \\ &\quad \leqslant \max_{s=1,\dots,|J|} |b_{I_1,j_1} \cdots b_{I_{s-1},j_{s-1}} (b_{I_s}, j_s - c_{I_s,j_s}) c_{I_{s+1},j_{s+1}} \cdots c_{I_{|J|},j_{|J|}}| r^{|I|} \\ &\quad \leqslant \max_{s=1,\dots,|J|} \left(\prod_{t=1}^{s-1} |b_{I_t,j_t}| r^{|I_t|} \right) \cdot |b_{I_s,j_s} - c_{I_s,j_s}| r^{|I_s|} \cdot \left(\prod_{t=s+1}^{|J|} |b_{I_t,j_t}| r^{|I_t|} \right) \\ &\quad \leqslant \left(\max_{s=1,\dots,|J|} |b_{I_s,j_s} - c_{I_s,j_s}| r^{|I_s|} \right) r^{|J|-1} \\ &\quad \leqslant \|\vec{g} - \vec{h}\|_r r^{|J|-1}. \end{aligned}$$

But then

$$\begin{aligned} |a_J (b_{I_1,j_1} \cdots b_{I_{|J|},j_{|J|}} - c_{I_1,j_1} \cdots c_{I_{|J|},j_{|J|}})| r^{|I|} &\leqslant |a_J| \cdot \|\vec{g} - \vec{h}\|_r r^{|J|-1} \\ &\leqslant \frac{1}{r} \|\vec{f}\|_r \|\vec{g} - \vec{h}\|_r \end{aligned}$$

as desired.

(iv) The inequality

$$\|\vec{f}\|_{r_1} \leqslant \frac{r_1}{r_2} \|\vec{f}\|_{r_2}$$

immediately implies that the inclusion $K^{\leqslant r_2}\langle\langle\vec{x}\rangle\rangle \subset K^{\leqslant r_1}\langle\langle\vec{x}\rangle\rangle$ is continuous, so it suffices to verify the inequality. Writing $f_i = \sum a_I x^I$, it is enough to show that

$$\sup_I |a_I| r_1^{|I|} \leqslant \frac{r_1}{r_2} \sup_I |a_I| r_2^{|I|},$$

where the supremum is taken over I with $|I| > 0$. But for each I with $|I| > 0$, we have

$$|a_I| r_1^{|I|} \leqslant |a_I| r_1 r_2^{|I|-1},$$

which gives the claim. □

As a special case of Lemma 3.1.1 (iii) above, we have:

Lemma 3.1.2. *For any $\vec{f}, \vec{\varepsilon} \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$ with $\|\vec{\varepsilon}\|_r < 1$, and for any diagonal matrix $A = \mathrm{diag}(\lambda_i) \in \mathrm{Mat}_{n\times n}(K)$, $|\lambda_i| \leqslant 1$ the following estimate holds:*

$$\|\vec{f}(A\vec{x} + \vec{\varepsilon}) - \vec{f}(A\vec{x})\|_r < \frac{1}{r}\|\vec{f}\|_r \|\vec{\varepsilon}\|_r.$$

We will require a criterion for invertibility of certain elements of $\mathrm{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$:

Lemma 3.1.3. *Suppose $\vec{\psi} \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$ is of the form $\vec{\psi} = \vec{x} + \widehat{\psi}$, $\widehat{\psi} = O(x^2)$ and $\|\widehat{\psi}\|_r < \varepsilon < r$. Then $\vec{\psi}$ admits a two-sided compositional inverse $\vec{g} = \vec{x} + \widehat{g}$ and $\|\widehat{g}\|_r < \varepsilon$.*

Proof. We first find the left inverse. Set $\vec{g}_L = \vec{x} - \widehat{\psi} + \widehat{\psi}^{\circ 2} - \widehat{\psi}^{\circ 3} + \cdots$. The sum converges as $\|\widehat{\psi}\|_r < \varepsilon < r$, and hence by Lemma 3.1.1 (iii), $\|\widehat{\psi}^{\circ n}\|_r < \frac{\varepsilon^n}{r^{n-1}}$, by induction on n. Thus $\|\widehat{\psi}^{\circ n}\|_r \to 0$ as $n \to \infty$. Moreover, $\widehat{g}_L := \vec{g}_L - \vec{x}$ satisfies $\|\widehat{g}_L\|_r < \varepsilon$ by the ultrametric inequality. Thus it suffices to show $\vec{g}_L$ is inverse to $\vec{\psi}$.

To see that $\vec{g}_L$ is left inverse to $\vec{\psi}$, we simply evaluate $\vec{g}_L \circ \vec{\psi}$; the sum telescopes.

The previous paragraphs shows that any tuple of power series $\vec{\psi}$ satisfying the conditions of the lemma admits a left inverse $\vec{g}_L$. Applying this to $\vec{g}_L$ shows that $\vec{g}_L$ admits a left inverse. Since $\vec{g}_L$ also has a right inverse, it follows that $\vec{\psi}$ is a two-sided inverse of $\vec{g}_L$, and $\vec{g}_L$ is a two-sided inverse of $\vec{\psi}$. □

Definition 3.1.4. An element $\vec{f} \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$ is *semisimple* if the operator given by $P \mapsto P \circ \vec{f}$ on $K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle / \mathscr{I}^m$ is semisimple for all m.

Remark 3.1.5. The property of being semisimple is preserved under conjugation. For all $r' < r$ an element $\vec{f} \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$ is semisimple as an element of $\mathrm{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$ if and only if it is semisimple as an element of $\mathrm{End}^{\mathrm{op}} K^{\leqslant r'}\langle\langle \vec{x} \rangle\rangle$.

Lemma 3.1.6. *Suppose $\vec{f} \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$ is semisimple, $\vec{f} = A\vec{x} + O(x^2)$, and A is a diagonal matrix with coefficients $\vec{\lambda} = (\lambda_1, \ldots, \lambda_n)$. If I is a word and j is an index such that $\vec{\lambda}^I = \lambda_j$, then the coefficient of x^I in f_j is zero.*

Proof. The operator $F \mapsto F \circ \vec{f}$ on $K\langle\langle \vec{x} \rangle\rangle / \mathscr{I}^{|I|+1}$ is upper triangular in the monomial basis. Therefore, since the diagonal entries corresponding to the coefficients of x^I and x_j are equal and the operator is semisimple, the corresponding off-diagonal coefficient a_I^j is zero. □

Definition 3.1.7. A tuple of numbers $\lambda_1, \ldots, \lambda_n \in \overline{\mathbb{Q}_\ell}$ is said to satisfy *ℓ-Siegel's condition with parameters $c, \mu > 0$* if the following condition holds: for any tuple of nonnegative integers $i_1, \ldots, i_n$ with $i_1 + \cdots + i_n = N$ and any index j such that $\lambda_1^{i_1}\lambda_2^{i_2}\cdots\lambda_n^{i_n} \neq \lambda_j$, the following inequality holds:

$$|\lambda_1^{i_1}\lambda_2^{i_2}\cdots\lambda_n^{i_n} - \lambda_j| \geqslant c\left(\frac{N}{2}\right)^{-\mu}.$$

Remark 3.1.8. Replacing λ_j with 1 and $\frac{N}{2}$ with N leads to an equivalent definition with different constants; the form given in Definition 3.1.7 is more convenient for our applications.

Siegel's condition holds for algebraic λ_i, as the following proposition shows.

Proposition 3.1.9 (Linear forms in logarithms, [15]). *Suppose $\lambda_1, \dots, \lambda_n \in \overline{\mathbb{Q}_\ell}$ are algebraic numbers. Then there exist constants $c, \mu > 0$ such that for any integers $i_1, \dots, i_n, j$ with $\lambda_1^{i_1} \cdots \lambda_n^{i_n} \neq \lambda_j$ the following inequality holds:*

$$|\lambda_1^{i_1} \cdots \lambda_n^{i_n} - \lambda_j| \geqslant c(|i_1| + \cdots + |i_n|)^{-\mu}.$$

Proof. After replacing the tuple $\lambda_1, \dots, \lambda_n$ with $\lambda_1, \dots, \lambda_n, \lambda_1^{-1}, \dots, \lambda_n^{-1}$ it suffices to prove the inequality for positive integers i_j. Also, by changing c to a different constant it suffices to show that $|\lambda_1^{i_1} \cdots \lambda_n^{i_n} - 1| \geqslant c(|i_1| + \cdots + |i_n|)^{-\mu}$. This estimate is proved in [15, Theorem 1']. □

Our goal is to conjugate a (semisimple vector of) power series $\vec{f} = A\vec{x} + O(x^2)$ to its linear part $A\vec{x}$. We do so iteratively, on every step conjugating f by some other power series so that the norm of $\vec{f} - A\vec{x}$ on a slightly smaller disk becomes much smaller than before. The key to the inductive step is Lemma 3.1.11, which itself uses the following simple estimate.

Lemma 3.1.10. *Suppose $\eta \in (0, 1)$ and $\mu > 0$ are real numbers. Then*

$$\sup_{i \in \mathbb{Z}_{\geqslant 0}} (1 - \eta)^i i^\mu \leqslant \left(\frac{\eta}{7\mu}\right)^{-\mu}.$$

Proof. For every positive x we have

$$e^{\frac{\eta x}{7\mu}} \geqslant \frac{\eta x}{7\mu}.$$

By convexity of log on $(1, 2)$ we have $\log(1 + \eta) \geqslant \frac{\eta}{7}$. Therefore

$$e^{\frac{\log(1+\eta)}{\mu} x} \geqslant e^{\frac{\eta x}{7\mu}} \geqslant \frac{\eta x}{7\mu}.$$

Since $\frac{1}{1+\eta} > 1 - \eta$ inverting both sides, we get

$$\frac{7\mu}{\eta x} \geqslant (1 - \eta)^{\frac{x}{\mu}},$$

or

$$x(1 - \eta)^{\frac{x}{\mu}} \leqslant \frac{7\mu}{\eta}.$$

Raising both sides to the power μ gives

$$x^\mu (1 - \eta)^x \leqslant \left(\frac{\eta}{7\mu}\right)^{-\mu}.$$ □

Lemma 3.1.11. *Given $\delta > 0$, suppose $\vec{f} \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$ is such that*

$$\vec{f} = A\vec{x} + \hat{f}(x), \quad \hat{f} = O(x^2) \quad \textit{and} \quad \|\hat{f}\|_r < \delta.$$

Suppose that $\vec{f}$ is semisimple, A is a diagonal matrix, and that the eigenvalues $\lambda_1, \dots, \lambda_n$ of A satisfy ℓ-Siegel's condition with parameters c, μ, $\mu > \frac{2}{7}$. Suppose also $|\lambda_i| \leqslant 1$. Suppose $\eta \in (0,1)$ satisfies $c^{-1}(7\mu)^{\mu}\eta^{-\mu}\delta < r$. Then there exists an invertible endomorphism $\vec{\psi} \in \operatorname{End}^{\mathrm{op}} K^{\leqslant r(1-\eta)}\langle\langle\vec{x}\rangle\rangle$, with $\vec{\psi} = \vec{x} + O(x^2)$, such that the following inequalities hold:

$$\|\vec{\psi} - \vec{x}\|_{r(1-\eta)} \leqslant c^{-1}\delta(1-\eta)\left(\frac{\eta}{7\mu}\right)^{-\mu} < r(1-\eta)$$

and

$$\|\vec{\psi}^{-1} \circ \vec{f} \circ \vec{\psi} - A\vec{x}\|_{r(1-\eta)} \leqslant c^{-1}(7\mu)^{\mu}\frac{\delta^2}{\eta^{\mu} r}.$$

Proof. Let $\widehat{\psi}$ be the solution to the equation $\widehat{\psi}(A\vec{x}) - A\widehat{\psi}(\vec{x}) = \widehat{f}(\vec{x})$ with $\widehat{\psi} = O(x^2)$. This equation can be solved as A is diagonal: if a_I^j is the coefficient of x^I in f_j, then the coefficient of x^I in ψ_j is $a_I^j/(\lambda^I - \lambda_j)$ if $\lambda^I - \lambda_j$ is nonzero, and it is zero otherwise. (Note that $a_I^j = 0$ whenever $\lambda^I = \lambda_j$ by Lemma 3.1.6.) To check that $\widehat{\psi}$ converges on the disk of radius $r(1-\eta)$, we use the Siegel condition:

$$\left|\frac{a_I^j}{\lambda^I - \lambda_j}\right| r^{|I|}(1-\eta)^{|I|} \leqslant \delta(1-\eta)^{|I|}|\lambda^{|I|} - \lambda_j|^{-1} \overset{(*)}{\leqslant} c^{-1}\delta(1-\eta)^{|I|}\left(\frac{|I|}{2}\right)^{\mu} \xrightarrow[|I|\to\infty]{} 0,$$

where we used Siegel's property for $(*)$. Similarly, we directly estimate the norm of $\widehat{\psi}$ using the Siegel condition as follows:

$$\begin{aligned}
\|\widehat{\psi}\|_{r(1-\eta)} &= \sup_{I,j}\left|\frac{a_I^j}{\lambda^I - \lambda_j}\right| r^{|I|}(1-\eta)^{|I|} \\
&\leqslant \delta \sup_{I,j}(1-\eta)^{|I|}|\lambda^I - \lambda_j|^{-1} \\
&\leqslant \delta \sup_i c^{-1}(1-\eta)^i\left(\frac{i}{2}\right)^{\mu} && \text{(by Siegel's property)} \\
&\leqslant c^{-1}\delta \max\bigl(\sup_{i\geqslant 2}(1-\eta)^i(i-1)^{\mu}, (1-\eta)2^{-\mu}\bigr) \\
&= c^{-1}\delta \max\bigl((1-\eta)\sup_{i\geqslant 1}(1-\eta)^i i^{\mu}, (1-\eta)2^{-\mu}\bigr) && \text{(Index shift)} \\
&\leqslant c^{-1}\delta(1-\eta)\left(\frac{\eta}{7\mu}\right)^{-\mu} && \text{(by Lemma 3.1.10).}
\end{aligned}$$

In the last inequality we have used the condition $\mu > \frac{2}{7}$ to resolve the maximum. Note that by our choice of η, we thus have $\|\widehat{\psi}\|_{r(1-\eta)} < r(1-\eta)$, and so $\widehat{\psi}$ is in $\operatorname{End}^{\mathrm{op}} K^{\leqslant r(1-\eta)}\langle\langle\vec{x}\rangle\rangle$.

We now show that the function $\vec{\psi} = \vec{x} + \widehat{\psi}$ satisfies the conditions of the lemma. Note that $\vec{\psi}$ is invertible by Lemma 3.1.3, again as $\|\widehat{\psi}\|_{r(1-\eta)} < r(1-\eta)$. We need to estimate the norm $\|\vec{\psi}^{-1} \circ \vec{f} \circ \vec{\psi} - A\vec{x}\|_{r(1-\eta)}$. Let $\vec{g}$ denote the function $\vec{\psi}^{-1} \circ \vec{f} \circ \vec{\psi}$, and write $\vec{g} = A\vec{x} + \widehat{g}$. We now use the functional equation for $\widehat{\psi}$ to derive an equation for $\widehat{g}$:

$$\begin{aligned}
\vec{\psi}(\vec{g}(\vec{x})) &= A\vec{\psi}(\vec{x}) + \widehat{f}(\vec{\psi}(\vec{x})), \\
\widehat{g}(\vec{x}) + \widehat{\psi}(A\vec{x} + \widehat{g}(\vec{x})) &= A\widehat{\psi}(\vec{x}) + \widehat{f}(\vec{x} + \widehat{\psi}(\vec{x})), \\
\widehat{g}(\vec{x}) + \widehat{\psi}(A\vec{x} + \widehat{g}(\vec{x})) &= \widehat{\psi}(A\vec{x}) - \widehat{f}(\vec{x}) + \widehat{f}(\vec{x} + \widehat{\psi}(\vec{x})), \\
\widehat{g}(\vec{x}) &= [\widehat{\psi}(A\vec{x}) - \widehat{\psi}(A\vec{x} + \widehat{g}(\vec{x}))] + [\widehat{f}(\vec{x} + \widehat{\psi}(\vec{x})) - \widehat{f}(\vec{x})].
\end{aligned}$$

The right-hand side of the last equation is visibly "small"; if $G = \|\hat{g}\|_{r(1-\eta)}$, then applying Lemma 3.1.2 and the ultrametric inequality we get

$$\begin{aligned} G &\leqslant \frac{1}{r(1-\eta)} \max\{\|\hat{\psi}\|_{r(1-\eta)} G, \|\hat{f}\|_{r(1-\eta)} \|\hat{\psi}\|_{r(1-\eta)}\} \\ &\leqslant \frac{1}{r(1-\eta)} \|\hat{f}\|_{r(1-\eta)} \|\hat{\psi}\|_{r(1-\eta)}. \end{aligned}$$

Here the last inequality holds because $G \leqslant \frac{1}{r(1-\eta)} \|\hat{\psi}\|_{r(1-\eta)} G$ is impossible, unless $G = 0$, since $\frac{1}{r(1-\eta)} \|\hat{\psi}\|_{r(1-\eta)} < 1$; and if $G = 0$, the inequality holds in any case. Using the estimates

$$\|\hat{\psi}\|_{r(1-\eta)} \leqslant c^{-1}(7\mu)^{\mu} \delta(1-\eta)\eta^{-\mu}$$

and $\|f\|_{r(1-\eta)} \leqslant \|f\|_r < \delta$, we get

$$G \leqslant \frac{c^{-1}(7\mu)^{\mu} \delta^2 \eta^{-\mu}}{r}. \qquad \square$$

To apply Lemma 3.1.11, we will need the following simple estimate.

Lemma 3.1.12. *Let $u \in (0, \frac{1}{2})$ and $\alpha > 1$ be real numbers. Then*

$$\prod_{n=0}^{\infty} \left(1 - \frac{u}{\alpha^n}\right) > e^{-\frac{\alpha}{\alpha-1}}.$$

Proof. For $x \in (0, \frac{1}{2})$ we have $\log(1-x) > -2x$. Taking the logarithm of the product and applying this inequality to every term gives

$$\sum_{n=0}^{\infty} \log\left(1 - \frac{u}{\alpha^n}\right) > -2u \sum_{n=0}^{\infty} \left(\frac{1}{\alpha^n}\right) = -2u \frac{\alpha}{\alpha-1} > -\frac{\alpha}{\alpha-1}. \qquad \square$$

3.2. The linearization theorem. We are now ready to prove the main result of this section.

Theorem 3.2.1 (noncommutative, non-archimedean Siegel linearization). *Suppose that $\vec{f} = A\vec{x} + \hat{f}(x)$ is an element of $\mathrm{End}^{\mathrm{op}}\, K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$. Suppose that $\vec{f}$ is semisimple, A is a diagonal matrix, and that the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A satisfy Siegel's condition with parameters c, μ. Suppose $|\lambda_i| \leqslant 1$. Then there exists a radius $r_\infty < r < 1$ and an invertible function $\vec{\psi} \in \mathrm{End}^{\mathrm{op}}\, K^{\leqslant r_\infty}\langle\langle \vec{x} \rangle\rangle$ such that $\vec{\psi}^{-1} \circ \vec{f} \circ \vec{\psi} = A\vec{x}$.*

Proof. The idea is to apply Lemma 3.1.11 iteratively, at every step conjugating $\vec{f}$ closer and closer to its linear part. At each step we are given a conjugate $\vec{f_n}$ of $\vec{f}$ on a disk of radius r_n with $\|\vec{f_n} - A\vec{x}\|_{r_n} = \delta_n < 1$, and we have to choose a suitable value of η_n to apply Lemma 3.1.11 and obtain a new conjugate $\vec{f}_{n+1}$ on the disk of radius $r_{n+1} := r_n(1-\eta_n)$. We make these choices so that the distance between $\vec{f_n}$ and its linear part tends to zero with n, and the radii of the relevant disks remains bounded away from zero. On each step we make the choice $\eta_n \approx \delta_n^{1/(\mu+1)}$, then the numerics of Lemma 3.1.11 show that

$$\delta_{n+1} \lessapprox \frac{\delta_n^2}{\eta_n^{\mu}} \approx \delta_n^{1+\frac{1}{\mu+1}},$$

so that the sequence δ_n converges to zero super-exponentially. As $\eta_n \approx \delta_n^{1/(\mu+1)}$, the sequence η_n will converge to zero super-exponentially as well, and so $r\prod_n(1-\eta_n) > 0$; by taking limits we will show that $\vec{f}$ is conjugate to $A\vec{x}$ on the disk of radius $r\prod_n(1-\eta_n)$.

After increasing μ if necessary we can assume $\mu > \frac{2}{7}$ (this is a condition necessary to apply Lemma 3.1.11). Before starting the iterative process, we need to (possibly) shrink the disk to make the norm of $\vec{f} - A\vec{x}$ small enough so that on every step of the iteration Lemma 3.1.11 can be applied.

Let δ denote $\|\vec{f} - A\vec{x}\|_r$. We choose a large positive constant B (to be determined later) and let $r_1 := \frac{r}{B}$. Then since $\vec{f} - A\vec{x}$ has no linear term, the norm $\delta_1 := \|\vec{f} - A\vec{x}\|_{r_1}$ satisfies $\delta_1 \leqslant \frac{\delta}{B^2}$. We now choose B large such that the following inequality holds:

$$c^{-1}(7\mu)^{\mu}\delta_1 3^{\mu} < \frac{1}{2} r_1 e^{-\frac{2^{1/(\mu+1)}}{2^{1/(\mu+1)}-1}}, \tag{3.2.2}$$

which is possible as δ_1 scales as $O(\frac{1}{B^2})$, while r_1 scales as $\sim \frac{1}{B}$.

We will produce a sequence of constants η_n, δ_n, r_n and elements $\vec{\psi}_n \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r_{n+1}}\langle\langle\vec{x}\rangle\rangle$, $\vec{f}_n \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r_n}\langle\langle\vec{x}\rangle\rangle$, with $\vec{\psi}_n$ invertible, and such that

$$\vec{f}_1 = \vec{f} \in \mathrm{End}^{\mathrm{op}} K^{\leqslant r_1}\langle\langle\vec{x}\rangle\rangle, \quad \delta_1 = \|\vec{f} - A\vec{x}\|_{r_1}, \quad \eta_1 = \frac{1}{3}.$$

We will have

$$\begin{aligned}
\vec{f}_{n+1} &= \vec{\psi}_n^{-1} \circ \vec{f}_n \circ \vec{\psi}_n, \\
r_n &= (1-\eta_{n-1}) r_{n-1}, \\
\delta_n &:= \|\vec{f}_n - A\vec{x}\|_{r_n} < \frac{\delta_{n-1}}{2}, \\
\|\vec{\psi}_n - \vec{x}\|_{r_{n+1}} &< c^{-1}\delta_n(1-\eta_n)(7\mu)^{\mu}\eta_n^{-\mu},
\end{aligned}$$

and

$$\eta_n^{\mu+1} = \frac{\delta_n}{3^{\mu+1}\delta_1} \in (0,1).$$

Indeed, iteratively apply Lemma 3.1.11 to $\vec{f}_n, r_n, \eta_n$. For the lemma to be applicable we need to check two conditions: $\eta_n \in (0,1)$ and $c^{-1}(7\mu)^{\mu}\eta_n^{-\mu}\delta_n < r_n$. The first condition follows since $\delta_n < \delta_1$ and so

$$\eta_n = \left(\frac{\delta_n}{3^{\mu+1}\delta_1}\right)^{\frac{1}{\mu+1}} < \frac{1}{3}.$$

For the second condition, we have

$$\begin{aligned}
\eta_n^{-\mu}\delta_n &= \left(\frac{\delta_n}{3^{\mu+1}\delta_1}\right)^{-\frac{\mu}{\mu+1}} \delta_n \\
&= \delta_n^{\frac{1}{\mu+1}} (3^{\mu+1}\delta_1)^{\frac{\mu}{\mu+1}} \\
&< \delta_1^{\frac{1}{\mu+1}} 3^{\mu} \delta_1^{\frac{\mu}{\mu+1}} \\
&= \delta_1 3^{\mu}.
\end{aligned}$$

Thus it is enough to show

$$c^{-1}(7\mu)^{\mu}\delta_1 3^{\mu} < r_n.$$

We estimate r_n from below using the estimate $\delta_n < \delta_1 2^{-(n-1)}$:

$$
\begin{aligned}
r_n &= r_1 \prod_{i=1}^{n-1} (1-\eta_i) \\
&= r_1 \prod_{i=1}^{n-1} \left(1 - \left(\frac{\delta_i}{3^{\mu+1}\delta_1}\right)^{\frac{1}{\mu+1}}\right) \\
&> r_1 \prod_{i=1}^{n-1} \left(1 - \frac{1}{3} 2^{-\frac{i-1}{\mu}+1}\right) \\
&= r_1 \prod_{i=0}^{n-1} \left(1 - \frac{1}{3} 2^{-\frac{i}{\mu}+1}\right) \\
&> r_1 \prod_{i=0}^{\infty} \left(1 - \frac{1}{3} 2^{-\frac{i}{\mu}+1}\right) \\
&> r_1 e^{-\frac{2^{1/(\mu+1)}}{2^{1/(\mu+1)}-1}} \quad \text{(by Lemma 3.1.12).}
\end{aligned}
\tag{3.2.3}
$$

Therefore using inequality (3.2.2) and the previous estimate, we get

$$c^{-1}(7\mu)^{\mu}\delta_1 3^{\mu} < \frac{1}{2} r_1 e^{-\frac{2^{1/(\mu+1)}}{2^{1/(\mu+1)}-1}} < r_n.$$

Thus, we can apply Lemma 3.1.11 with $\vec{f} = \vec{f}_n$, $r = r_n$ and $\eta = \eta_n$ to produce $\vec{\psi}_n$ and

$$\vec{f}_{n+1} = \vec{\psi}_n^{-1} \circ \vec{f}_n \circ \vec{\psi}_n$$

such that

$$\|\vec{\psi}_n - \vec{x}\|_{r_{n+1}} < c^{-1}\delta_n(1-\eta_n)(7\mu)^{\mu}\eta_n^{-\mu}$$

and

$$
\begin{aligned}
\delta_{n+1} &< c^{-1}(7\mu)^{\mu} \frac{\delta_n^2}{\eta_n^{\mu} r_n} \\
&= c^{-1}(7\mu)^{\mu} \delta_n^{\frac{1}{\mu+1}} \delta_1^{\frac{\mu}{\mu+1}} 3^{\mu} r_n^{-1} \delta_n \\
&< c^{-1}(7\mu)^{\mu} \delta_1 3^{\mu} r_n^{-1} \delta_n \\
&< c^{-1}(7\mu)^{\mu} \delta_1 3^{\mu} r_1^{-1} e^{\frac{2^{1/(\mu+1)}}{2^{1/(\mu+1)}-1}} \delta_n \quad \text{(by (3.2.3))} \\
&< \frac{1}{2}\delta_n \quad \text{(by (3.2.2)).}
\end{aligned}
$$

Thus the infinite sequence $\vec{f}_n$ can be constructed as claimed.

The sequence δ_n converges to zero (at least) exponentially. Since $\eta_n = \frac{1}{3}(\frac{\delta_n}{\delta_1})^{\frac{1}{\mu+1}}$, the product $\prod_n (1-\eta_n)$ converges. Let

$$r_\infty := r_1 \prod_n (1-\eta_n).$$

By Lemma 3.1.1 (iv), we have $\|\vec{f}_n - A\vec{x}\|_{r_\infty} < \delta_n$, and so the sequence of conjugates $\vec{f}_n$ converges to the function $A\vec{x}$ on the disk of radius r_∞.

We now show that the limit of $\vec{f}_n$ is also a conjugate of $\vec{f}$. Let $\vec{\Psi}_n = \vec{\psi}_1 \circ \cdots \circ \vec{\psi}_n$. Then $\vec{\Psi}_n \in \operatorname{End}^{\mathrm{op}} K^{\leqslant r_\infty}\langle\langle \vec{x} \rangle\rangle$ is invertible, and $\vec{\Psi}_n^{-1} \circ \vec{f} \circ \vec{\Psi}_n = \vec{f}_n$. By construction,

$$\begin{aligned} \|\vec{\psi}_n - \vec{x}\|_{r_\infty} &< c^{-1}\delta_n(1-\eta_n)(7\mu)^{\mu}\eta_n^{-\mu} \\ &= 3^{\mu+1}c^{-1}(7\mu)^{\mu}(1-\eta_n)\delta_1\eta_n \\ &< 3^{\mu+1}c^{-1}(7\mu)^{\mu}\delta_1\eta_n \\ &= Q\eta_n, \end{aligned}$$

where $Q > 0$ does not depend on n. We have by Lemma 3.1.2

$$\begin{aligned} \|\vec{\Psi}_{n+1} - \vec{\Psi}_n\|_{r_\infty} &= \|\vec{\Psi}_n(\vec{\psi}_{n+1}(x)) - \vec{\Psi}_n(x)\|_{r_\infty} \\ &< \frac{1}{r_\infty}\|\vec{\Psi}_n\|_{r_\infty}\|\vec{\psi}_{n+1} - \vec{x}\|_{r_\infty} \\ &< \frac{Q}{r_\infty}\|\vec{\Psi}_n\|_{r_\infty}\eta_{n+1}. \end{aligned}$$

Since η_n converges to zero, it follows that the sequence $\vec{\Psi}_n$ is Cauchy, and thus has a limit $\vec{\Psi}$. Since $\vec{f}_n = \vec{\Psi}_n^{-1} \circ \vec{f} \circ \vec{\Psi}_n$, we have $\vec{\Psi}^{-1} \circ \vec{f} \circ \vec{\Psi} = A\vec{x}$, using continuity of composition (see Lemma 3.1.1 (iii)). □

Corollary 3.2.4. *Suppose $F \in \operatorname{End}^{\mathrm{op}} K^{\leqslant r}\langle\langle \vec{x} \rangle\rangle$ is a semisimple endomorphism. Suppose the eigenvalues λ_i of the action of F on $\mathscr{I}/\mathscr{I}^2$ are elements of K with $|\lambda_i| \leqslant 1$ that satisfy Siegel's condition with parameter μ. Then there exists a radius $r' < r$ and a collection of elements $y_1, \ldots, y_n \in K^{\leqslant r'}\langle\langle \vec{x} \rangle\rangle$ with the following two properties:*

(i) *If $\lambda_1, \ldots, \lambda_n \in \overline{\mathbb{Q}_\ell}$ are eigenvalues of F on $\mathscr{I}/\mathscr{I}^2$, then $Fy_i = \lambda_i y_i$.*

(ii) *For any integer m the monomials in y_i of degree at least m have dense span in $\mathscr{I}_{r'}^m$.*

Proof. After replacing r with a smaller radius $\tilde{r}$ we can do a linear change of variables $\vec{x}' = M\vec{x}$ to make the action of F on $\mathscr{I}_{\tilde{r}}/\mathscr{I}_{\tilde{r}}^2$ diagonal in the basis $\vec{x}'$. We can therefore assume that $F = A\vec{x}' + O(x^2)$, where A is a diagonal matrix $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. We can now apply Theorem 3.2.1: there exists a radius $r' < \tilde{r}$ and an element $\vec{\psi} \in \operatorname{End}^{\mathrm{op}} K^{\leqslant r'}\langle\langle \vec{x}' \rangle\rangle$ such that $\vec{\psi}^{-1} \circ F \circ \vec{\psi} = A\vec{x}'$. Let $y_i = \psi_i^{-1}(\vec{x}')$; then $Fy_i = y_i(F(\vec{x}')) = A\vec{y}$. The span of the monomials in the x_i' of degree m or larger is dense in $\mathscr{I}_{r'}^m$ by Lemma 2.2.2. Since $\vec{x}' = \vec{\psi}(\vec{y})$, any monomial in x_i' of degree m is a (convergent) sum of monomials in y_i of degree at least m. Hence monomials in y_i of degree m or larger have dense span in $\mathscr{I}_{r'}^m$. □

4. Main theorem

Having proven Theorem 3.2.1 and Corollary 3.2.4, we move on to the proof of Theorem 1.1.2, which is a straightforward application.

Proof of Theorem 1.1.2 (compare to [9, Proof of Theorem 1.2]). We first observe that it suffices to consider the case where X is affine and k is finite. Indeed, we may take X to be affine by deleting any closed point of X, which does not affect the hypotheses of the theorem. To see that we may reduce to the case where k is finite, choose a finitely-generated integral $\mathbb{Z}$-algebra R in which ℓ is invertible, a smooth proper R-curve $\overline{\mathscr{X}}$, a divisor D in $\overline{\mathscr{X}}$ étale over R, and an isomorphism $\mathrm{Frac}(R) \xrightarrow{\sim} k$ such that $(\overline{\mathscr{X}} \setminus D)_k$ is isomorphic to X. Now for any geometric point $\bar{p}$ lying over a closed point $p \in \mathrm{Spec}(R)$ with residue field $k(p)$ of characteristic prime to ℓ, the specialization map

$$\pi_1^{\ell}(X_{\bar{k}}) \to \pi_1^{\ell}((\overline{\mathscr{X}} \setminus D)_{\bar{p}})$$

is an isomorphism. Moreover, any semisimple arithmetic representation of $\pi_1^{\ell}(X_{\bar{k}})$ remains semisimple arithmetic when viewed as a representation of $\pi_1^{\ell}((\overline{\mathscr{X}} \setminus D)_{\bar{p}})$, by the argument of [10, Proof of Theorem 1.1.3, Step 2]. Thus it suffices to prove the theorem for $(\overline{\mathscr{X}} \setminus D)_p$, which is by construction a smooth affine curve over a finite field.

For the rest of the argument we assume k is finite of characteristic different from ℓ, and X/k is a smooth affine curve. Let $\bar{k}$ be an algebraic closure of k. We may, after replacing k with a finite extension, assume that X has a k-rational point x; we let $\bar{x}$ be the geometric point obtained from x via our choice of algebraic closure $\bar{k}$ of k. In this case $\pi_1^{\ell}(X_{\bar{k}}, \bar{x})$ is a free pro-ℓ group, and hence $\mathbb{Z}_\ell\langle\langle\pi_1^{\ell}(X_{\bar{k}}, \bar{x})\rangle\rangle$ is (non-canonically) isomorphic to a noncommutative power series ring over $\mathbb{Z}_\ell$; fix such an isomorphism. As $\bar{x}$ was obtained from a rational point of X, the absolute Galois group of k acts naturally on $\mathbb{Z}_\ell\langle\langle\pi_1^{\ell}(X_{\bar{k}}, \bar{x})\rangle\rangle$.

Let F be the Frobenius element in the absolute Galois group of k. Consider the action of F on $\mathbb{Q}_\ell\langle\langle\pi_1^{\ell}(X_{\bar{k}}, \bar{x})\rangle\rangle$. By [9, Theorem 2.20] the action of F is semisimple in the sense of Definition 3.1.4. By the Weil conjectures for curves, the eigenvalues $\lambda_1, \dots, \lambda_n$ of the action of F on $\mathscr{I}/\mathscr{I}^2 = H^1(X_{\bar{k}}, \overline{\mathbb{Q}_\ell})^\vee$ (see [9, Proposition 2.4]) are q-Weil numbers of weights -1 and -2. In particular, they are algebraic and therefore by Proposition 3.1.9 they satisfy Siegel's condition for some parameters c, μ. Let $K/\mathbb{Q}_\ell$ be a finite extension that contains all λ_i. By Theorem 3.2.1 and Corollary 3.2.4 there exists a radius r such that the ideal $\mathscr{I}_r^n \subset K^{\leqslant r}\langle\langle\pi_1^{\text{ét}}(X_{\overline{k}}, \bar{x})\rangle\rangle$ is (topologically) spanned by F-eigenvectors for all $n \geqslant 0$. These eigenvectors are monomials in the elements $y_1, \dots, y_n \in \mathscr{I}_r$ provided by Corollary 3.2.4, where the eigenvalue λ_i corresponding to y_i also appears as an eigenvalue of the Frobenius action on $H^1(X_{\bar{k}}, K)^\vee$, and hence is a q-Weil number of weight -1 or -2. In particular, $\mathscr{I}_r^n$ is topologically spanned by F-eigenvectors whose corresponding eigenvalues are q-Weil numbers of weight $\leqslant -n$.

Let $N(X, \ell)$ be an arbitrary real number strictly larger than $-\log(r)/\log(\ell) \in \mathbb{R}$. Now suppose ρ is a semisimple arithmetic representation trivial modulo $\ell^{N(X,\ell)}$. There exists a finite extension k'/k such that ρ extends to a representation $\rho' : \pi_1^{\text{ét}}(X_{k'}, \bar{x}) \to \mathrm{GL}_n(\overline{\mathbb{Z}_\ell})$ of the arithmetic fundamental group of $X_{k'}$. There exists an integer m such that F^m lifts to an element of $\pi_1^{\text{ét}}(X_{k'}, \bar{x})$. Let $A \in \mathrm{GL}_n(\overline{\mathbb{Z}_\ell})$ be the matrix $A := \rho'(F^m)$. Since the representation ρ is trivial modulo $\ell^{N(X,\ell)} > \frac{1}{r}$, it extends to a continuous representation

$$\widehat{\rho} : K^{\leqslant r}\langle\langle\pi_1^{\ell}(X_{\overline{k}}, \bar{x})\rangle\rangle \to \mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}_\ell})$$

by Lemma 2.2.1. Moreover, by arithmeticity,

$$\widehat{\rho}(F^m(g)) = A\widehat{\rho}(g)A^{-1} \quad \text{for every } g \in K^{\leqslant r}\langle\langle\pi_1^{\text{ét}}(X_{\overline{k}}, \bar{x})\rangle\rangle.$$

Let w' denote the most negative weight of an eigenvalue of the conjugation action of A on $\mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}_\ell})$, if any such exist, and 0 otherwise. Let $w = \max(-w', 0)$. Every monomial $Y \in \mathscr{I}_r^{w+1}$ in the y_i satisfies

$$A\widehat{\rho}(Y)A^{-1} = u\widehat{\rho}(Y)$$

for a q-Weil number u of weight less than $-w$. Since no such numbers are eigenvalues of the conjugation action of A on $\mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}_\ell})$, the image of every monomial in $\mathscr{I}_r^{w+1}$ under $\widehat{\rho}$ is zero. As such monomials topologically span $\mathscr{I}_r^{w+1}$ by Corollary 3.2.4, we have $\widehat{\rho}(\mathscr{I}_r^{w+1}) = 0$ and hence that

$$\rho(\mathscr{I}_r^{w+1} \cap \mathbb{Z}_\ell\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle) = 0.$$

But $\mathscr{I}_r^{w+1} \cap \mathbb{Z}_\ell\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ is $\mathscr{I}^{w+1}$, where $\mathscr{I}$ is the augmentation ideal of $\mathbb{Z}_\ell\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$. Therefore, ρ is unipotent. But a unipotent semisimple representation is trivial. □

Remark 4.0.1. In the course of the proof, we show that for X a smooth affine curve over a finite field k, and ℓ a prime different from the characteristic of k, there exists a finite extension K of $\mathbb{Q}_\ell$ and an $r > 0$ such that the Banach algebra $K^{\leq r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ is topologically spanned by Frobenius eigenvectors. In fact, the same statement for smooth proper curves follows immediately, as if X is smooth and proper and $y \in X$ is a closed point, the map

$$\pi_1^\ell((X \setminus y)_{\bar{k}}) \to \pi_1^\ell(X_{\bar{k}})$$

is surjective.

Proof of Corollary 1.1.3. As in the statement, we let X/k be a curve over a finitely-generated field, and $A/X_{\bar{k}}$ an Abelian scheme. Suppose that A_η had full ℓ^M-torsion. Then the natural geometric monodromy representation

$$\pi_1^{\text{ét}}(X_{\bar{k}}, \bar{x}) \to \mathrm{GL}(T_\ell(A_{\bar{x}}))$$

is trivial modulo ℓ^M for any geometric point $\bar{x}$ of $X_{\bar{k}}$. As this representation is semisimple arithmetic (as are all representations arising from geometry – semisimplicity follows from [4, 3.4.1 (iii)], and arithmeticity by spreading out), Theorem 1.1.2 implies that it is trivial, and in particular every ℓ-power torsion point of A_η is rational. Thus by the Lang–Néron theorem [3, Theorem 2.1], the natural map

$$\tau : \mathrm{Tr}_{\eta/\bar{k}}(A_\eta)_\eta \to A_\eta$$

had image containing all the ℓ-power torsion points of A_η. As the ℓ-power torsion is Zariski-dense, this implies that τ is surjective and hence an isogeny for dimension reasons, proving the statement. □

5. Remarks and extensions

5.1. A suggestive correspondence. Let X be a smooth proper curve over a finite field k and ℓ a prime different from the characteristic of k. Let $x \in X(k)$ be a rational point, and $\bar{x}$ the geometric point of X associated to x by a choice of algebraic closure of k. Let $N = N(X, \ell)$ be as in Theorem 1.1.2. One consequence of Theorem 3.2.1 is a Galois-equivariant description of the category of lisse $\overline{\mathbb{Q}_\ell}$-sheaves on $X_{\bar{k}}$ admitting lattices which are trivial mod ℓ^N in terms of linear algebra data. We view this as a (very weak) ℓ-adic analogue of non-abelian Hodge theory.

Definition 5.1.1. Let $\mathscr{H}_\ell(X)$ be the category whose objects consist of pairs

$$(V, \theta : V \to V \otimes H^1(X_{\bar{k}}, \overline{\mathbb{Q}_\ell})),$$

where V is a finite-dimensional $\overline{\mathbb{Q}_\ell}$-vector space and θ is a linear map. A morphism between (V, θ) and (V', θ') is a linear map $f : V \to V'$ so that the diagram

$$\begin{array}{ccc} V & \xrightarrow{\theta} & V \otimes H^1(X_{\bar{k}}, \overline{\mathbb{Q}_\ell}) \\ \downarrow f & & \downarrow f\otimes \mathrm{id} \\ V' & \xrightarrow{\theta'} & V' \otimes H^1(X_{\bar{k}}, \overline{\mathbb{Q}_\ell}) \end{array}$$

commutes.

Let $\mathrm{Sh}_{\ell,N}(X_{\bar{k}})$ be the full subcategory of the category of continuous representations

$$\rho : \pi_1^\ell(X_{\bar{k}}, \bar{x}) \to \mathrm{GL}(V),$$

where V is a finite-dimensional $\overline{\mathbb{Q}_\ell}$-vector space, such that there exists a $\overline{\mathbb{Z}_\ell}$-sublattice W of V, stable under the action of $\pi_1^\ell(X_{\bar{k}}, \bar{x})$, and such that $\pi_1^\ell(X_{\bar{k}}, \bar{x})$ acts on $W/(\ell^N)W$ trivially. We now construct a functor

$$H : \mathrm{Sh}_{\ell,N}(X_{\bar{k}}) \to \mathscr{H}_\ell(X).$$

Let K, r be as in the proof of Theorem 1.1.2 and Remark 4.0.1, so that Frobenius acts diagonalizably on $K^{\leq r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$. Letting $\mathscr{I}_r \subset K^{\leq r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ be the augmentation ideal, note that the natural map

$$\mathscr{I}_r \to \mathscr{I}_r/\mathscr{I}_r^2 \simeq H^1(X_{\bar{k}}, K)^\vee$$

admits a unique Frobenius-equivariant splitting, given by the span of the weight -1 Frobenius-eigenvectors. Thus the span of the weight -1 eigenvectors yields a copy of $H^1(X_{\bar{k}}, K)^\vee$ inside of $K^{\leq r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$. Now let V be an object of $\mathrm{Sh}_{\ell,N}(X_{\bar{k}})$. By Lemma 2.2.1, we have a natural action of $K^{\leq r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$ on V, and thus viewing $H^1(X_{\bar{k}}, K)^\vee$ as subspace of $K^{\leq r}\langle\langle \pi_1^\ell(X_{\bar{k}}, \bar{x})\rangle\rangle$, we obtain a natural map

$$H^1(X_{\bar{k}}, K)^\vee \otimes_K V \to V.$$

By adjointness we thus obtain an object

$$V \to V \otimes H^1(X_{\bar{k}}, \overline{\mathbb{Q}_\ell})$$

of $\mathscr{H}_\ell(X)$.

This construction is evidently functorial. One can verify from the definition that the functor H is fully faithful. Moreover, there is a natural Frobenius action on the set of isomorphism classes of objects of $\mathscr{H}_\ell(X)$ (via the action of Frobenius on $H^1(X_{\bar{k}}, \overline{\mathbb{Q}_\ell})$), and H induces a Frobenius-equivariant map from isomorphism classes of objects of $\mathrm{Sh}_{\ell,N}(X_{\bar{k}})$ to isomorphism classes of objects of $\mathscr{H}_\ell(X)$. We can interpret Theorem 1.1.2 as the full faithfulness of this functor, combined with the fact that any object of $\mathscr{H}_\ell(X)$, fixed up to isomorphism by the action of Frobenius, is *nilpotent*, in the sense that for $n \gg 0$, the composition

$$\theta^n : V \to V \otimes H^1(X_{\bar{k}}.\overline{\mathbb{Q}_\ell})^{\otimes n}$$

is zero.

Using the semisimplicity of the Frobenius action on $\mathbb{Q}_\ell\langle\langle \pi_1^{\text{ét}}(X_{\bar{k}}, \bar{x})\rangle\rangle$ (see [9, Theorem 2.20]), one may extend this correspondence to the case of non-proper X, though doing so seems to depend on some choices. It would of course be very interesting to find a variant of this construction for residually nontrivial representations.

5.2. Residually nontrivial representations. It is natural to ask if results similar to Theorem 1.1.2 hold for residually nontrivial arithmetic representations. Indeed, a version of Theorem 3.2.1 in the commutative setting (that is, a mild generalization of the main result of [6, Section 4], allowing "resonance"), with an identical proof, implies:

Theorem 5.2.1. *Let X be a smooth curve over a finite field k, $\bar{x}$ a geometric point of X, and*

$$\overline{\rho} : \pi_1^{\acute{e}t}(X, \bar{x}) \to \mathrm{GL}_n(\mathbb{F}_{\ell^r})$$

a representation which is absolutely irreducible when restricted to $\pi_1^{\acute{e}t}(X_{\bar{k}}, \bar{x})$, with ℓ different from the characteristic of k. Let $R_{\overline{\rho}}$ be the deformation ring of $\overline{\rho}|_{\pi_1^{\acute{e}t}(X_{\bar{k}}, \bar{x})}$, and let $U_{\overline{\rho}}$ be its rigid generic fiber. Let K be an ℓ-adic field with residue field $\mathbb{F}_{\ell^r}$, and let

$$\rho : \pi_1^{\acute{e}t}(X, \bar{x}) \to \mathrm{GL}_n(\mathscr{O}_K)$$

be a continuous lift of $\overline{\rho}$; let $[\rho] \in U_{\overline{\rho}}$ be the point corresponding to $\rho|_{\pi_1^{\acute{e}t}(X_{\bar{k}}, \bar{x})}$. If the action of Frobenius on $H^1(X_{\bar{k}}, \rho \otimes \rho^\vee)$ is semisimple, there exists an open neighborhood V of $[\rho]$ in $U_{\overline{\rho}}$ such that the Frobenius action on V is conjugate to a linear map.

In dynamics, a neighborhood V as above is referred to as a *Siegel disk*. Note that the hypothesis on the semisimplicity of the Frobenius action on $H^1(X_{\bar{k}}, \rho \otimes \rho^\vee)$ would follow for all ρ from the Tate conjecture, by L. Lafforgue's work on the Langlands program: [8, Corollaire VII.8] implies that $\rho \otimes \rho^\vee$ "arises from geometry", whence the Tate conjecture would imply that the Frobenius action on its cohomology groups is semisimple.

As a corollary of Theorem 5.2.1, one obtains that for ρ as in the theorem statement, there exists a neighborhood of $[\rho]$ containing no Frobenius-periodic points aside from $[\rho]$ (that is, no arithmetic representations). This is proven unconditionally in [10, Theorem 1.1.3]. That said, it would in our view be quite interesting to understand Siegel disks in $U_{\bar{\rho}}$; for example, if $U_{\bar{\rho}}$ was covered by Siegel disks for iterates of Frobenius, the Hard Lefschetz theorem would follow for all lifts of $\bar{\rho}$, by the strategy of [5].

Sketch proof of Theorem 5.2.1. The proof of Theorem 3.2.1 works verbatim in the commutative setting, giving the following result. Let K be an ℓ-adic field, R a Tate algebra over K, and $F : R \to R$ a continuous endomorphism. Suppose F preserves a maximal ideal $\mathfrak{m}$ of R, and acts semisimply on the completion $\widehat{R}$ of R at $\mathfrak{m}$ (i.e., F acts semisimply on the finite-dimensional K-vector spaces $R/\mathfrak{m}^n$ for all n). Suppose moreover that the action of F on $\mathfrak{m}/\mathfrak{m}^2$ has eigenvalues satisfying ℓ-Siegel's condition with parameters c, μ for some $c, \mu > 0$. Then there exists an affinoid neighborhood of $[\mathfrak{m}] \in \mathrm{Sp}(R)$ on which F is conjugate to a linear map.

We now choose a Frobenius-stable open ball U containing $[\rho]$ in the rigid generic fiber of $R_{\bar{\rho}}$; as $R_{\bar{\rho}}$ is a power series ring over $W(k)$ by the absolute irreducibility of $\bar{\rho}$, we may choose U to be the spectrum of a Tate algebra R. Thus it is enough to check the hypotheses of the result of the previous paragraph, taking F to be the Frobenius automorphism of R and $\mathfrak{m}$ to be the maximal ideal corresponding to ρ. The semisimplicity hypothesis follows from the assumption of the semisimplicity of the Frobenius action on $H^1(X_{\bar{k}}, \rho \otimes \rho^\vee) = (\mathfrak{m}/\mathfrak{m}^2)^\vee$ by an argument identical to the proof of [10, Theorem 5.1.8]. And [8, Corollaire VII.8] implies that the eigenvalues of the Frobenius action on $\mathfrak{m}/\mathfrak{m}^2$ are Weil numbers, hence algebraic; thus they satisfy ℓ-Siegel's condition for some μ by Proposition 3.1.9. □

Question 5.2.2. Let $U_{\bar{\rho}}$ be as in the statement of Theorem 5.2.1. Is $U_{\bar{\rho}}$ covered by Siegel disks for iterates of Frobenius? That is, for each point ν of $U_{\bar{\rho}}$, does there exist a finite extension k' of k, a representation $\rho_\nu : \pi_1^{\text{ét}}(X_{k'}, \bar{x}) \to \mathrm{GL}_n(\mathscr{O}_L)$ lifting $\bar{\rho}$ (with L an ℓ-adic field with residue field $\mathbb{F}_{\ell^r}$), and a neighborhood V of $[\rho_\nu]$ containing ν, such that the action of the Frobenius of k' on V is conjugate to a linear map?

We view our proof of Theorem 1.1.2 as showing that the trivial representation is contained in a "noncommutative Siegel disk".

5.3. Other questions. Theorem 1.1.2 is unsatisfying in a number of ways. First, it is natural to ask if the constant $N(X, \ell)$ appearing in the statement of the theorem may be taken to tend to zero as $\ell \to \infty$; as remarked in the introduction, this is known in characteristic zero by [10, Theorem 1.1.13]. It is not clear to us if there is a plausible improvement of the ℓ-adic form of Baker's theorem on linear forms in logarithms which would, by our method, imply that one may take $N(X, \ell) \to 0$ as $\ell \to \infty$, but the following would imply that one may take $N(X, \ell)$ to be independent of ℓ:

Question 5.3.1. Let $\lambda_1, \dots, \lambda_n \in \overline{\mathbb{Q}}$ be a collection of algebraic numbers, stable under the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. For each prime ℓ, the ℓ-adic form of Baker's theorem yields constants c_ℓ, μ_ℓ such that the λ_i satisfy ℓ-Siegel's condition with parameters c_ℓ, μ_ℓ, as in Definition 3.1.7. Is it true that the constants c_ℓ, μ_ℓ may be taken to be independent of ℓ?

It is also natural to ask for an analogue of Theorem 1.1.2 with better uniformity in e.g. the genus or gonality of the curve X, or one which depends only on the function field $\bar{k}(X)$. For example, the geometric torsion conjecture [2] predicts that in the case that ρ arises from the ℓ-adic Tate module of a traceless Abelian scheme A, there is a bound on the torsion

$$(\rho \otimes \mathbb{Q}_\ell/\mathbb{Z}_\ell)^{\pi_1^\ell(X_{\bar{k}}, \bar{x})}$$

depending only on the gonality of X and the rank of ρ (here $(\rho \otimes \mathbb{Q}_\ell/\mathbb{Z}_\ell)^{\pi_1^\ell(X_{\bar{k}}, \bar{x})}$ is naturally isomorphic to the rational ℓ-power torsion of the generic fiber of A). A weaker question, plausibly approachable via the methods of this paper, is to ask for a bound on N depending only on the function field $k(X)$ and not on X itself.

Acknowledgement. We are grateful for useful conversations with Hélène Esnault, Moritz Kerz, Samit Dasgupta, and Simion Filip. We would also like to thank the referee for their many extremely helpful suggestions, and for catching and fixing an error in an earlier version of the paper.

References

[1] *B. Bakker* and *J. Tsimerman*, p-torsion monodromy representations of elliptic curves over geometric function fields, Ann. of Math. (2) **184** (2016), no. 3, 709–744.
[2] *B. Bakker* and *J. Tsimerman*, The geometric torsion conjecture for abelian varieties with real multiplication, J. Differential Geom. **109** (2018), no. 3, 379–409.
[3] *B. Conrad*, Chow's K/k-image and K/k-trace, and the Lang–Néron theorem, Enseign. Math. (2) **52** (2006), no. 1–2, 37–108.

[4] *P. Deligne*, La conjecture de Weil. II, Inst. Hautes Études Sci. Publ. Math. (1980), no. 52, 137–252.
[5] *H. Esnault* and *M. Kerz*, Étale cohomology of rank one ℓ-adic local systems in positive characteristic, Selecta Math. (N.S.) **27** (2021), no. 4, Paper No. 58.
[6] *M. Herman* and *J.-C. Yoccoz*, Generalizations of some theorems of small divisors to non-Archimedean fields, in: Geometric dynamics (Rio de Janeiro 1981), Lecture Notes in Math. **1007**, Springer, Berlin (1983), 408–447.
[7] *J.-M. Hwang* and *W.-K. To*, Uniform boundedness of level structures on abelian varieties over complex function fields, Math. Ann. **335** (2006), no. 2, 363–377.
[8] *L. Lafforgue*, Chtoucas de Drinfeld et correspondance de Langlands, Invent. Math. **147** (2002), no. 1, 1–241.
[9] *D. Litt*, Arithmetic representations of fundamental groups I, Invent. Math. **214** (2018), no. 2, 605–639.
[10] *D. Litt*, Arithmetic representations of fundamental groups, II: Finiteness, Duke Math. J. **170** (2021), no. 8, 1851–1897.
[11] *L. Moret-Bailly*, Familles de courbes et de variétés abéliennes sur $\mathbb{P}^1$. I. Descente des polarisations, in: Séminaire sur les pinceaux de courbes de genre au moins deux, Astériques **86**, Société Mathématique de France, Paris (1981), 109–124.
[12] *L. Moret-Bailly*, Familles de courbes et de variétés abéliennes sur $\mathbb{P}^1$. II. Exemples, in: Séminaire sur les pinceaux de courbes de genre au moins deux, Astériques **86**, Société Mathématique de France, Paris (1981), 125–140.
[13] *A. M. Nadel*, The nonexistence of certain level structures on abelian varieties over complex function fields, Ann. of Math. (2) **129** (1989), no. 1, 161–178.
[14] *H. Rüssmann*, Kleine Nenner. II. Bemerkungen zur Newtonschen Methode, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II (1972), 1–10.
[15] *K. R. Yu*, Linear forms in p-adic logarithms. III, Compositio Math. **91** (1994), no. 3, 241–276.
[16] *E. Zehnder*, A simple proof of a generalization of a theorem by C. L. Siegel, in: Geometry and topology (Rio de Janeiro 1976), Lecture Notes in Math. **597**, Springer, Berlin (1977), 855–866.

Borys Kadets, Department of Mathematics, University of Georgia,
200 D. W. Brooks Drive, 30602, Athens, GA, USA
e-mail: kadets.math@gmail.com

Daniel Litt, Department of Mathematics, University of Georgia,
200 D. W. Brooks Drive, 30602, Athens, GA, USA
https://orcid.org/0000-0003-2273-4630
e-mail: dlitt@uga.edu

Eingegangen 3. September 2021, in revidierter Fassung 25. März 2022