THE STEINBERG REPRESENTATION IS IRREDUCIBLE

ANDREW PUTMAN AND ANDREW SNOWDEN

Abstract. We prove that the Steinberg representation of a connected reductive group over an infinite field is irreducible. For finite fields, this is a classical theorem of Steinberg and Curtis.

Contents

1. Introduction

Let $\mathbf{G}$ be a connected reductive group over a field k, e.g., $\mathbf{G} = \mathrm{GL}_n$. For another field $\mathbb{F}$, let $\mathrm{St}(\mathbf{G};\mathbb{F})$ be the Steinberg representation of the discrete group $\mathbf{G}(k)$ over $\mathbb{F}$. This representation plays a prominent role in the representation theory of $\mathbf{G}(k)$, and also has important connections to number theory and algebraic K-theory. When k is finite, $\mathrm{St}(\mathbf{G};\mathbb{F})$ is finite dimensional and Steinberg and Curtis showed that it is usually irreducible. When k is infinite, it is typically infinite dimensional. Our main theorem is that $\mathrm{St}(\mathbf{G};\mathbb{F})$ is always irreducible when k is infinite. As far as we are aware, this is new even for $\mathbf{G} = \mathrm{GL}_2$.

1.1. **Background.** Before explaining the contents of this paper in more detail, we recall the construction of the Steinberg representation, review some of its history, and discuss its connections to other topics.

1.1.1. *Tits building.* A tremendous amount of the structure of $\mathbf{G}(k)$ is encoded in its spherical Tits building $\mathcal{T}(\mathbf{G})$. This is a simplicial complex whose simplices are in bijection with the proper parabolic k-subgroups of $\mathbf{G}$. For proper parabolic k-subgroups P and P', the simplex corresponding to P is a face of the simplex corresponding to P' when $P' \subset P$. For example, if $\mathbf{G} = \mathrm{GL}_n$, then the proper parabolic k-subgroups are the stabilizers of nontrivial flags

$$0 \subsetneq V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_i \subsetneq k^n,$$

so $\mathcal{T}(\mathrm{GL}_n)$ is the simplicial complex whose i-simplices are such flags. The conjugation action of $\mathbf{G}(k)$ on itself permutes the parabolic k-subgroups and thus induces an action of $\mathbf{G}(k)$ on $\mathcal{T}(\mathbf{G})$. See [1, 33] for more about Tits buildings.

Date: July 1, 2021.
AP was supported in part by NSF grant DMS-1811210.
AS was supported in part by NSF grant DMS-1453893.

Remark 1.1. It is not obvious that the above description of $\mathcal{T}(\mathbf{G})$ specifies a simplicial complex. However, what we really care about is the homology of $\mathcal{T}(\mathbf{G})$, and for this it is enough to understand its barycentric subdivision, which is easy to describe completely: it is the simplicial complex whose i-simplices are decreasing chains

$$\mathbf{G} \supsetneq P_0 \supsetneq P_1 \supsetneq \cdots \supsetneq P_i \supsetneq 1$$

of proper parabolic k-subgroups. □

1.1.2. *Steinberg representation.* Let r be the semisimple k-rank of $\mathbf{G}$. By definition, $\mathcal{T}(\mathbf{G})$ is an $(r-1)$-dimensional simplicial complex. The Solomon–Tits theorem [28] says that in fact $\mathcal{T}(\mathbf{G})$ is homotopy equivalent to a wedge of $(r-1)$-dimensional spheres. For a field $\mathbb{F}$ (or, more generally, a commutative ring), the *Steinberg representation* of $\mathbf{G}(k)$ over $\mathbb{F}$, denoted $\mathrm{St}(\mathbf{G};\mathbb{F})$, is the unique nontrivial reduced homology group $\widetilde{\mathrm{H}}_{r-1}(\mathcal{T}(\mathbf{G});\mathbb{F})$. The action of $\mathbf{G}(k)$ on $\mathcal{T}(\mathbf{G})$ induces an action of $\mathbf{G}(k)$ on $\mathrm{St}(\mathbf{G};\mathbb{F})$, making it into a representation of $\mathbf{G}(k)$ over $\mathbb{F}$.

Remark 1.2. It might be the case that $\mathbf{G}$ is anisotropic, i.e., has no proper parabolic k-subgroups. This implies that $\mathcal{T}(\mathbf{G}) = \emptyset$ and that the semisimple k-rank of $\mathbf{G}$ is 0. Our convention then is that

$$\mathrm{St}(\mathbf{G};\mathbb{F}) = \widetilde{\mathrm{H}}_{-1}(\mathcal{T}(\mathbf{G});\mathbb{F}) = \widetilde{\mathrm{H}}_{-1}(\emptyset;\mathbb{F}) = \mathbb{F}$$

is the trivial representation. □

1.1.3. *Finite fields.* The representation $\mathrm{St}(\mathbf{G};\mathbb{F})$ was first studied for finite fields k. In this case, $\mathrm{St}(\mathbf{G};\mathbb{F})$ is a finite-dimensional representation of the finite group $\mathbf{G}(k)$ that is usually[1] irreducible. For instance, this holds if $\mathrm{char}(\mathbb{F}) = 0$ or if $\mathrm{char}(\mathbb{F}) = \mathrm{char}(k)$. Steinberg [29] initially proved this for $\mathbf{G} = \mathrm{GL}_n$, and then generalized it to many other finite groups [30, 31]. Curtis [11] proved the ultimate version for a finite group with a BN-pair. See [16, 32] for surveys of the fundamental role the Steinberg representation plays in the representation theory of finite groups of Lie type.

Remark 1.3. The above papers predate the definition of the Steinberg representation in terms of the Tits building, which first appeared in [28]. □

1.1.4. *Infinite fields.* For infinite fields k, the representation $\mathrm{St}(\mathbf{G};\mathbb{F})$ is usually[2] an infinite-dimensional representation of the infinite group $\mathbf{G}(k)$. In this context, it first appeared in work of Borel–Serre [5], who proved that the symmetric space associated to the Lie group $\mathbf{G}(\mathbb{R})$ has a $\mathbf{G}(k)$-equivariant bordification whose boundary is homotopy equivalent to $\mathcal{T}(\mathbf{G})$. They used this to show that $\mathrm{St}(\mathbf{G};\mathbb{F})$ is the "dualizing module" for arithmetic subgroups of $\mathbf{G}(k)$. This gave rise to a large literature using $\mathrm{St}(\mathbf{G};\mathbb{F})$ to study the cohomology of such arithmetic subgroups. Some representative papers include [2, 3, 7, 8, 9, 14, 17, 19, 21, 22, 23, 25].

A second important context for $\mathrm{St}(\mathbf{G};\mathbb{F})$ when k is infinite is algebraic K-theory, where Quillen [26] constructed a spectral sequence converging to the algebraic K-theory of a number ring $\mathcal{O}$ whose E^2-page involves the homology of $\mathrm{GL}_n(\mathcal{O})$ with coefficients in the Steinberg representation. His main application was to show that these K-groups are finitely generated. The Steinberg representation and related objects have since appeared in a variety of K-theoretic contexts.

[1]There are cases where it is reducible. For example, it is reducible if $\mathbf{G} = \mathrm{GL}_2$, the field k is finite of cardinality q, and $\mathbb{F}$ has finite characteristic ℓ with $\ell \mid q+1$.

[2]The only time when $\mathrm{St}(\mathbf{G};\mathbb{F})$ is finite-dimensional for infinite k is when $\mathbf{G}$ is anisotropic, in which case $\mathrm{St}(\mathbf{G};\mathbb{F}) = \mathbb{F}$ is the trivial representation.

1.2. **Main theorem.** In light of the importance of $\mathrm{St}(\mathbf{G};\mathbb{F})$, it is natural to wonder whether or not it is irreducible when k is infinite. As far as we are aware, this folklore question first appeared in writing in the first author's MathOverflow posting [24]. For GL_n, a partial answer was recently given by Galatius–Kupers–Randal-Williams [13], who showed that it is indecomposable[3] for all fields k. Our main theorem answers this question completely.

Theorem A. *Let* $\mathbf{G}$ *be a connected reductive group over an infinite field* k *and let* $\mathbb{F}$ *be an arbitrary field. Then the Steinberg representation* $\mathrm{St}(\mathbf{G};\mathbb{F})$ *is an irreducible* $\mathbf{G}(k)$*-representation.*

Remark 1.4. For a local field k, there is a variant of the Steinberg representation that takes into account the topology of k (see, e.g., [6]). This variant is irreducible, but it is different enough from the ordinary Steinberg representation that this does not seem to imply Theorem A when $\mathbf{G}$ is defined over a local field. □

Remark 1.5. There are a number of other representations that are similar to the Steinberg representation. For example, when k is finite and $\mathbb{F}$ has characteristic 0, there is one irreducible unipotent representation of $\mathrm{GL}_n(k)$ for each partition of n, with the Steinberg representation corresponding to the partition (1^n). Our methods should be able to prove irreducibility for these analogous representations, though we have not pursued this. □

1.3. **Necessity that k is infinite.** In most cases, our proof uses the fact that k is infinite in a serious way. Since $\mathrm{St}(\mathbf{G};\mathbb{F})$ is sometimes reducible when k is finite, this is inevitable. However, our work does give the most important special cases of irreducibility for finite k. Let k be a finite field with $\mathrm{char}(k) = p$.

- When $\mathrm{char}(\mathbb{F}) = p$, our proof works in complete generality and many aspects of it simplify.[4]
- The case where $\mathrm{char}(\mathbb{F}) = 0$ follows from the case where $\mathrm{char}(\mathbb{F}) = p$. To see this, observe that $\mathrm{St}(\mathbf{G};\mathbb{Z})$ is a finite-rank free abelian group with $\mathrm{St}(\mathbf{G};\overline{\mathbb{F}}_p) = \mathrm{St}(\mathbf{G};\mathbb{Z}) \otimes \overline{\mathbb{F}}_p$ and $\mathrm{St}(\mathbf{G};\mathbb{F}) = \mathrm{St}(\mathbf{G};\mathbb{Z}) \otimes \mathbb{F}$. We now quote the following standard result:[5] if G is a finite group, M is a $\mathbb{Z}[G]$-module whose underlying abelian group is finite-rank and free, and $M \otimes \overline{\mathbb{F}}_p$ is irreducible, then $M \otimes \mathbb{F}$ is irreducible for all fields $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) = 0$.

1.4. **Outline of proof of Theorem A.** Let the notation be as in Theorem A. Let $\mathbf{B}$ be a minimal parabolic k-subgroup of $\mathbf{G}$, let $\mathbf{U}$ be the unipotent radical of $\mathbf{B}$, and let $\mathbf{T}$ be a maximal k-split torus in $\mathbf{B}$. For example, if $\mathbf{G} = \mathrm{GL}_n$ then one can take $\mathbf{B}$ to be the Borel subgroup of upper triangular matrices, $\mathbf{U}$ to be group of upper triangular matrices with 1's on the diagonal, and $\mathbf{T}$ to be the group of diagonal matrices.

A strengthening of the Solomon–Tits theorem gives a linear isomorphism $\iota\colon \mathrm{St}(\mathbf{G};\mathbb{F}) \to \mathbb{F}[\mathbf{U}(k)]$. The map ι is equivariant for $\mathbf{U}(k)$ and $\mathbf{T}(k)$, which act by left multiplication and conjugation on the target, respectively. However, the action of a general element of $\mathbf{G}(k)$ on $\mathbb{F}[\mathbf{U}(k)]$ is opaque. The actions of $\mathbf{U}(k)$ and $\mathbf{T}(k)$ on $\mathbb{F}[\mathbf{U}(k)]$ preserve the augmentation ideal, i.e., the kernel of the augmentation $\epsilon\colon \mathbb{F}[\mathbf{U}(k)] \to \mathbb{F}$. We will prove the following:

[3] An indecomposable representation is one that cannot be decomposed as a nontrivial direct sum of two subrepresentations. When k is infinite, the Steinberg representation is typically infinite-dimensional and this is a weaker condition than being irreducible even when $\mathbb{F}$ has characteristic 0.

[4] In particular, the second ingredient (Proposition 1.7) in the proof outline discussed in §1.4 is almost trivial in this case; see the very short §4 for details.

[5] Here is a quick proof. If $M \otimes \mathbb{F}$ is reducible for some $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) = 0$, then $M \otimes \overline{\mathbb{Q}}_p$ is reducible. Let $V \subset M \otimes \overline{\mathbb{Q}}_p$ be a nonzero proper subrepresentation. Letting $\overline{\mathbb{Z}}_p$ be the ring of integers in $\overline{\mathbb{Q}}_p$, the intersection $V \cap (M \otimes \overline{\mathbb{Z}}_p)$ is a nonzero proper direct summand of $M \otimes \overline{\mathbb{Z}}_p$, and hence maps to a nonzero proper subrepresentation of $M \otimes \overline{\mathbb{F}}_p$ under the reduction map $M \otimes \overline{\mathbb{Z}}_p \to M \otimes \overline{\mathbb{F}}_p$.

Proposition 1.6. *Let the notation be as above and let $x \in \mathrm{St}(\mathbf{G};\mathbb{F})$ be non-zero. Then there exists $g \in \mathbf{G}(k)$ such that $\iota(gx)$ is not in the augmentation ideal.*

Proposition 1.7. *Let the notation be as above and let $I \subset \mathbb{F}[\mathbf{U}(k)]$ be a left ideal that is stable under $\mathbf{T}(k)$ and not contained in the augmentation ideal. Then $I = \mathbb{F}[\mathbf{U}(k)]$.*

To deduce Theorem A, let $V \subset \mathrm{St}(\mathbf{G};\mathbb{F})$ be a non-zero subrepresentation. Then $\iota(V)$ is a $\mathbf{T}(k)$-stable left ideal of $\mathbb{F}[\mathbf{U}(k)]$ that by Proposition 1.6 is not contained in the augmentation ideal. Proposition 1.7 thus implies that $\iota(V) = \mathbb{F}[\mathbf{U}(k)]$, so $V = \mathrm{St}(\mathbf{G};\mathbb{F})$.

1.5. **Special case of Proposition 1.6.** To prove Proposition 1.6, we must relate the augmentation map $\epsilon\colon \mathbb{F}[\mathbf{U}(k)] \to \mathbb{F}$ to the structure of the Tits building $\mathcal{T}(\mathbf{G})$. Doing this in general requires introducing a lot of building-theoretic terminology (chambers, apartments, etc.). To give the basic idea, we will explain how this works for $\mathbf{G} = \mathrm{GL}_2$.

1.5.1. *Structure of building.* We must first construct ι. The parabolic subgroups of GL_2 are the stabilizers of lines in k^2. The Tits building $\mathcal{T}(\mathrm{GL}_2)$ can thus be identified with the discrete set $\mathbb{P}^1(k)$. Elements of $\mathrm{H}_0(\mathcal{T}(\mathrm{GL}_2);\mathbb{F})$ are formal $\mathbb{F}$-linear combinations of points of $\mathbb{P}^1(k)$. The Steinberg representation is the reduced homology:

$$\mathrm{St}(\mathrm{GL}_2;\mathbb{F}) = \widetilde{\mathrm{H}}_0(\mathcal{T}(\mathrm{GL}_2);\mathbb{F}) = \left\{ \sum_{i=0}^{n} c_i \ell_i \mid \ell_i \in \mathbb{P}^1(k),\ c_i \in \mathbb{F},\ \text{and } \sum_{i=0}^{n} c_i = 0 \right\}.$$

1.5.2. *Apartment classes.* The Steinberg representation is spanned by elements of the form $\ell - \ell'$ for distinct $\ell, \ell' \in \mathbb{P}^1(k)$, which are called *apartment classes*. These are not linearly independent. Using homogeneous coordinates on $\mathbb{P}^1(k)$, the apartment classes of the form $[1,0] - [\lambda, 1]$ are a basis for $\mathrm{St}(\mathrm{GL}_2;\mathbb{F})$. Let $\mathbf{U}$ be the unipotent subgroup of upper triangular 2×2 matrices with 1's on the diagonal. We thus have an isomorphism

$$\iota\colon\ \mathrm{St}(\mathrm{GL}_2;\mathbb{F}) \to \mathbb{F}[\mathbf{U}(k)], \qquad \iota\left([1,0] - [\lambda,1]\right) = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}.$$

1.5.3. *Making the augmentation nonzero.* Let $\epsilon\colon \mathbb{F}[\mathbf{U}(k)] \to \mathbb{F}$ be the augmentation. Consider a nonzero $x \in \mathrm{St}(\mathrm{GL}_2;\mathbb{F})$. We must find some $g \in \mathrm{GL}_2(k)$ with $\epsilon(\iota(gx)) \neq 0$. Write

$$x = \sum_{i=0}^{n} c_i \ell_i \quad \text{with } \ell_i \in \mathbb{P}^1(k),\ c_i \in \mathbb{F},\ \text{and } \sum_{i=0}^{n} c_i = 0$$

with the ℓ_i all distinct and the c_i all nonzero. Pick $g \in \mathrm{GL}_2(k)$ with $g\ell_0 = [1,0]$. For $1 \leq i \leq n$, write $g\ell_i = [\lambda_i, 1]$ with $\lambda_i \in k$. Since $\sum_{i=0}^{n} c_i = 0$, it follows that

$$gx = c_0[1,0] + \sum_{i=1}^{n} c_i[\lambda_i, 1] = \sum_{i=1}^{n} -c_i\left([1,0] - [\lambda_i,1]\right),$$

so

$$\epsilon(\iota(gx)) = \epsilon\left(\sum_{i=1}^{n} -c_i \begin{pmatrix} 1 & \lambda_i \\ 0 & 1 \end{pmatrix} \right) = \sum_{i=1}^{n} -c_i = c_0 \neq 0.$$

1.6. **Special case of Proposition 1.7.** We now explain our proof of Proposition 1.7 for $\mathbf{G} = \mathrm{GL}_2$ when $\mathrm{char}(k) = 0$. Here $\mathbf{U}$ is the unipotent group of upper triangular 2×2 matrices with 1's on the diagonal and $\mathbf{T}$ is the group of 2×2 diagonal matrices. Let $I \subset \mathbb{F}[\mathbf{U}(k)]$ be a left ideal that is stable under the conjugation action of $\mathbf{T}(k)$ and not contained in the augmentation ideal. We must prove that $I = \mathbb{F}[\mathbf{U}(k)]$.

1.6.1. *Torus action.* Since $I \not\subset \ker(\epsilon)$, we can find $x \in I$ with $\epsilon(x) = 1$. Write this as

$$x = \sum_{i=1}^{n} c_i \begin{pmatrix} 1 & \lambda_i \\ 0 & 1 \end{pmatrix} \in I \quad \text{with } \lambda_1, \ldots, \lambda_n \in k,\ c_1, \ldots, c_n \in \mathbb{F}, \text{ and } \sum_{i=1}^{n} c_i = 1.$$

Since $\mathrm{char}(k) = 0$, the torus $\mathbf{T}(k)$ contains matrices $\mathrm{diag}(d, 1)$ for all nonzero $d \in \mathbb{Z}$. We have

$$\begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & d\lambda \\ 0 & 1 \end{pmatrix} \quad \text{for all } \lambda \in k.$$

Letting $x_d \in I$ be the result of conjugating $x \in I$ by $\mathrm{diag}(d, 1)$, we thus have

$$x_d = \sum_{i=1}^{n} c_i \begin{pmatrix} 1 & d\lambda_i \\ 0 & 1 \end{pmatrix} \in I \quad \text{for all nonzero } d \in \mathbb{Z}. \tag{1.1}$$

1.6.2. *Laurent polynomials.* Define a ring homomorphism

$$\Psi\colon \mathbb{Z}[z_1^{\pm 1}, \ldots, z_n^{\pm 1}] \longrightarrow \mathbb{F}[\mathbf{U}(k)], \qquad \Psi(z_1^{d_1} \cdots z_n^{d_n}) = \begin{pmatrix} 1 & d_1\lambda_1 + \cdots + d_n\lambda_n \\ 0 & 1 \end{pmatrix}. \tag{1.2}$$

Use Ψ to make $\mathbb{F}[\mathbf{U}(k)]$ into a left module over $\mathbb{Z}[z_1^{\pm 1}, \ldots, z_n^{\pm 1}]$: for $f \in \mathbb{Z}[z_1^{\pm 1}, \ldots, z_n^{\pm 1}]$ and $x \in \mathbb{F}[\mathbf{U}(k)]$, define $f \cdot x = \Psi(f)x$. Letting $\mathrm{id} \in \mathbf{U}(k)$ be the identity matrix, we then have

$$x_d = \sum_{i=1}^{n} z_i^d \cdot (c_i \cdot \mathrm{id}) \in I \quad \text{for all nonzero } d \in \mathbb{Z}.$$

To prove the proposition, it is enough to show that $\mathrm{id} = \sum_{i=1}^{n} c_i \cdot \mathrm{id} \in I$.

1.6.3. *Modules over Laurent polynomials.* For this, apply the following lemma with

$$M = \mathbb{F}[\mathbf{U}(k)] \quad \text{and} \quad N = I \quad \text{and} \quad m_i = c_i \cdot \mathrm{id}.$$

Lemma 1.8. *Let $R = \mathbb{Z}[z_1^{\pm 1}, \ldots, z_n^{\pm 1}]$, let M be an R-module, let $N \subset M$ be a submodule, and let $m_1, \ldots, m_n \in M$. Assume that $z_1^d \cdot m_1 + \cdots + z_n^d \cdot m_n \in N$ for all $d \geq 1$. Then $m_1 + \cdots + m_n \in N$.*

Proof. Replacing M by M/N, we can assume that $N = 0$. Also, replacing M by the R-span of the m_i, we can assume that M is finitely generated. Let $\mathfrak{a}$ be a maximal ideal of R and let $r \geq 1$. Since $R/\mathfrak{a}^r$ is a finite ring and each z_i is a multiplicative unit in R, we can find some $d \geq 1$ such that $z_i^d \equiv 1 \pmod{\mathfrak{a}^r}$ for all $1 \leq i \leq d$. We then have

$$0 = z_1^d \cdot m_1 + \cdots + z_n^d \cdot m_n \equiv m_1 + \cdots + m_n \pmod{\mathfrak{a}^r},$$

so $m_1 + \cdots + m_n \in \mathfrak{a}^r M$. Since this holds for all r, we see that $m_1 + \cdots + m_n \in \bigcap_{r \geq 1} \mathfrak{a}^r M$, so by the Krull intersection theorem $m_1 + \cdots + m_r$ maps to 0 in the localization $M_{\mathfrak{a}}$. Since this holds for all maximal ideals $\mathfrak{a}$, it follows that $m_1 + \cdots + m_r = 0$, as required. □

Remark 1.9. The following special case of Lemma 1.8 might clarify its content. Consider $a_1, \ldots, a_n \in \mathbb{C}^\times$ and $b_1, \ldots, b_n \in \mathbb{C}$, and assume that $a_1^d b_1 + \cdots + a_n^d b_n = 0$ for all $d \geq 1$. Then Lemma 1.8 implies that $b_1 + \cdots + b_n = 0$. It is curious that we proved such a simple statement by reduction to finite characteristic[6]. When $\mathrm{char}(k) = 0$, our proof of Proposition 1.7 makes use of a similar reduction to finite characteristic. □

[6]Of course, this statement can be proven directly. However, the proof of the lemma shows that the conclusion still holds if the stated condition only holds for all d in a cofinal subset of $\mathbb{Z}$ (orded by divisibility). This stronger statement is not so easy to prove by hand.

Remark 1.10. Our proof of Lemma 1.8 is a little abstract. It is an instructive exercise to prove it more concretely by exhibiting appropriate polynomial identities. For instance, the case $n = 3$ follows from the identity

$$\begin{aligned} z_1 z_2 z_3 (m_1 + m_2 + m_3) =& (z_1 z_2 + z_1 z_3 + z_2 z_3)(z_1 m_1 + z_2 m_2 + z_3 m_3) \\ &- (z_1 + z_2 + z_3)(z_1^2 m_1 + z_2^2 m_2 + z_3^2 m_3) \\ &+ (z_1^3 m_1 + z_2^3 m_2 + z_3^3 m_3). \end{aligned}$$

Since the right hand side lies in N, the left hand side does as well. As the z_i are units in R, we have $m_1 + m_2 + m_3 \in N$. $\square$

1.7. **Comments on general case of Proposition 1.7.** We close the introduction by saying a few words about the general case of Proposition 1.7. We actually prove a more general result that applies to arbitrary unipotent groups.

1.7.1. *Key identity.* To state this, we must abstract the necessary properties of the action of $\mathbf{T}$ on $\mathbf{U}$. The key identity that powered our proof of Proposition 1.7 when $\mathbf{G} = \mathrm{GL}_2$ and $\mathrm{char}(k) = 0$ is

$$\begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & d\lambda \\ 0 & 1 \end{pmatrix} \quad \text{for nonzero } d \in \mathbb{Z}.$$

We could also have used the matrices $\mathrm{diag}(1, d^{-1})$, for which the analogous formula is

$$\begin{pmatrix} 1 & 0 \\ 0 & d^{-1} \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & d^{-1} \end{pmatrix}^{-1} = \begin{pmatrix} 1 & d\lambda \\ 0 & 1 \end{pmatrix} \quad \text{for nonzero } d \in \mathbb{Z}.$$

The choice to use $d = d^1$ or d^{-1} reflects the fact that the weights of the actions of $\mathrm{diag}(*, 1)$ and $\mathrm{diag}(1, *)$ on the Lie algbera of $\mathbf{U}$ are 1 and -1, respectively.

1.7.2. *Positive actions.* When $\dim(\mathbf{U}) > 1$, there will be more than one such weight. If we try to imitate the above proof, it turns out that we will run into trouble if there are both positive and negative weights (roughly, we won't be able make a single "choice of d^1 or d^{-1}"). Composing a $\mathbb{G}_m$-action on $\mathbf{U}$ with the inversion involution on $\mathbb{G}_m$ changes the signs of the weights, so we might as well assume they are all positive as in the following:

Definition 1.11. An action of $\mathbb{G}_m$ on a smooth connected unipotent group $\mathbf{U}$ over a field k is said to be *positive* if the weights of the action of $\mathbb{G}_m$ on $\mathrm{Lie}(\mathbf{U})$ are positive. $\square$

We will prove the following.

Theorem B. *Let $\mathbf{U}$ be a smooth connected unipotent group over an infinite field k equipped with a positive action of $\mathbb{G}_m$ and let $\mathbb{F}$ be another field. Let $I \subset \mathbb{F}[\mathbf{U}(k)]$ be a left ideal that is stable under $\mathbb{G}_m$ and not contained in the augmentation ideal. Then $I = \mathbb{F}[\mathbf{U}(k)]$.*

We will also prove that if $\mathbf{U}$ is as in Proposition 1.7, then there is a 1-parameter subgroup $\mathbb{G}_m$ of $\mathbf{T}$ whose action is positive, so this includes Proposition 1.7 as a special case.

1.7.3. *Cases.* Most of this paper will be devoted to Theorem B. Its proof is quite different depending on the characteristics of k and $\mathbb{F}$:

(a) $\mathrm{char}(k) = 0$.
(b) $\mathrm{char}(k) = p$ is positive and $\mathrm{char}(\mathbb{F}) \neq \mathrm{char}(k)$.
(c) $\mathrm{char}(k) = p$ is positive and $\mathrm{char}(\mathbb{F}) = \mathrm{char}(k)$.

Case (c) turns out to be quite easy, and does not even require the positive action or for k to be infinite. For cases (a) and (b), we need to find appropriate generalizations of Lemma 1.8.

1.7.4. *Characteristic* 0. When $\mathrm{char}(k) = 0$, we use deep work of Philip Hall on representations of nilpotent groups to give a proof that in some sense is quite similar to the one we gave for Lemma 1.8, though by necessity the details are more abstract.

1.7.5. *Characteristic* p. When $\mathrm{char}(k) = p$ is positive, new ideas are needed even for $\mathbf{G} = \mathrm{GL}_2$ since the matrices $\mathrm{diag}(d, 1)$ we used there are not always invertible. Roughly speaking, we will use the positive action to "compress" the action of our group onto a small subgroup for which our representation is understandable. This subgroup must satisfy a lengthy sequence of hard-to-control polynomial conditions. Since k is infinite, we will be able to use the Chevalley–Warning theorem to ensure that no matter what those conditions are, they can always be satisfied.

1.8. **Outline.** We prove Proposition 1.6 in §2. Next, in §3 we give some background about unipotent groups and prove that Theorem B implies Proposition 1.7. Theorem B is then proved in §4–§6.

1.9. **Conventions.** To avoid cluttering our exposition, unless otherwise specified all subgroups, morphisms, quotients, etc., we discuss involving an algebraic group $\mathbf{G}$ defined over a field k are themselves defined over k; for instance, instead of saying that something is a parabolic k-subgroup of $\mathbf{G}$ we will just say that it is a parabolic subgroup of $\mathbf{G}$.

2. Buildings and the augmentation

In this section, we prove Proposition 1.6. Our proof uses the Borel–Tits structure theory for connected reductive groups, and all results we quote without proof or reference can be found in [4, Chapter V]. Let $\mathbf{G}$ be a connected reductive group over a field k, and let $\mathbb{F}$ be another field.

2.1. **Borel subgroups, unipotent radicals, and tori.** We start by introducing some key subgroups of $\mathbf{G}$.

- Let $\mathbf{B}$ be a minimal parabolic subgroup of $\mathbf{G}$. Since k is not assumed to be algebraically closed, $\mathbf{B}$ might not be a Borel subgroup, but in this more general context it serves as a suitable replacement.
- Let $\mathbf{T}$ be a maximal split torus contained in $\mathbf{B}$.
- Let $Z_{\mathbf{G}}(\mathbf{T})$ be the centralizer of $\mathbf{T}$. If $\mathbf{G}$ were a split group like GL_n, then $Z_{\mathbf{G}}(\mathbf{T})$ would be $\mathbf{T}$, but in general it can be larger. The group $Z_{\mathbf{G}}(\mathbf{T})$ is a Levi factor of $\mathbf{B}$, and in particular is contained in $\mathbf{B}$.
- Let $N_{\mathbf{G}}(\mathbf{T})$ be the normalizer of $\mathbf{T}$.
- Let $W = N_{\mathbf{G}}(\mathbf{T})/Z_{\mathbf{G}}(\mathbf{T})$ be the relative[7] Weyl group. This is a finite reflection group.
- Let $\mathbf{U}$ be the unipotent radical of $\mathbf{B}$.

With this notation, we have $\mathbf{B} = \mathbf{U} \rtimes Z_{\mathbf{G}}(\mathbf{T})$.

Remark 2.1. All choices of the pair of subgroups $(\mathbf{B}, \mathbf{T})$ are conjugate in $\mathbf{G}$. □

[7] The usual (or absolute) Weyl group is what one gets by working over an algebraic closure $\overline{k}$ and letting $\mathbf{T}$ be a maximal torus defined over $\overline{k}$. It is often different from the relative Weyl group.

2.2. **Tits building, chambers, and the Steinberg representation.** As described in [33], the Tits building $\mathcal{T}(\mathbf{G})$ is the Tits building associated to the group $\mathbf{G}(k)$ with the BN-pair $(\mathbf{B}(k), N_{\mathbf{G}}(\mathbf{T})(k))$. See [1, §6] for a textbook reference on the Tits building associated to a BN-pair[8]. We will not need to know the complete construction and structure of $\mathcal{T}(\mathbf{G})$, but only a few properties of it that we will try to isolate.

Letting r be the semisimple rank of $\mathbf{G}$, the building $\mathcal{T}(\mathbf{G})$ is an $(r-1)$-dimensional simplicial complex whose simplices are in bijection with the proper parabolic subgroups of $\mathbf{G}$. Let $\widetilde{\mathrm{C}}_{\bullet}(\mathcal{T}(\mathbf{G});\mathbb{F})$ be the reduced simplicial chain complex of $\mathcal{T}(\mathbf{G})$. The $(r-1)$-dimensional simplices of $\mathcal{T}(\mathbf{G})$ are in bijection with the minimal proper parabolic subgroups, and are called the *chambers*. Let $\mathfrak{P}_{\min}$ be the set of minimal proper parabolic subgroups of $\mathbf{G}$, so $\widetilde{\mathrm{C}}_{r-1}(\mathcal{T}(\mathbf{G});\mathbb{F}) \cong \mathbb{F}[\mathfrak{P}_{\min}]$. Since $\mathcal{T}(\mathbf{G})$ is an $(r-1)$-dimensional simplicial chain complex, we have $\widetilde{\mathrm{C}}_{r}(\mathcal{T}(\mathbf{G});\mathbb{F}) = 0$ and thus

$$\begin{aligned}\mathrm{St}(\mathbf{G};\mathbb{F}) &= \widetilde{\mathrm{H}}_{r-1}(\mathcal{T}(\mathbf{G});\mathbb{F})\\ &= \ker(\widetilde{\mathrm{C}}_{r-1}(\mathcal{T}(\mathbf{G});\mathbb{F}) \xrightarrow{\partial} \widetilde{\mathrm{C}}_{r-2}(\mathcal{T}(\mathbf{G});\mathbb{F}))\\ &= \ker(\mathbb{F}[\mathfrak{P}_{\min}] \xrightarrow{\partial} \widetilde{\mathrm{C}}_{r-2}(\mathcal{T}(\mathbf{G});\mathbb{F})).\end{aligned}$$

In particular, $\mathrm{St}(\mathbf{G};\mathbb{F})$ is a subrepresentation of $\mathbb{F}[\mathfrak{P}_{\min}]$.

2.3. **Apartments.** The homology group $\mathrm{St}(\mathbf{G};\mathbb{F})$ is spanned by the *apartment classes*. These are the homology classes of oriented subcomplexes of $\mathcal{T}(\mathbf{G})$ that are isomorphic to simplicial triangulations of an $(r-1)$-sphere. In fact, these subcomplexes are isomorphic to the Coxeter complex of the Weyl group W. One example of an apartment is as follows. Since W is a finite reflection group, each $w \in W$ has a sign $(-1)^w$. Since the group $Z_{\mathbf{G}}(\mathbf{T})$ is contained in $\mathbf{B}$, for $w \in W = N_{\mathbf{G}}(\mathbf{T})/Z_{\mathbf{G}}(\mathbf{T})$ the image $w \cdot \mathbf{B} = w\mathbf{B}w^{-1}$ makes sense and is an element of $\mathfrak{P}_{\min}$. We then have an apartment class

$$\mathcal{A}_0 = \sum_{w \in W} (-1)^w w \cdot \mathbf{B} \in \mathrm{St}(\mathbf{G};\mathbb{F}) \subset \mathbb{F}[\mathfrak{P}_{\min}].$$

The group $\mathbf{G}(k)$ acts transitively on the set of apartment classes.

2.4. **Basis for Steinberg.** The apartment classes are not linearly independent. One version of the Solomon–Tits Theorem (see [1, Theorem 4.73]) says that $\mathrm{St}(\mathbf{G};\mathbb{F})$ has for a basis the set of apartment classes that "contain $\mathbf{B}$" in the sense that as an element of $\mathbb{F}[\mathfrak{P}_{\min}]$ their $\mathbf{B}$-coefficient is 1. Letting $\mathfrak{A}_{\mathbf{B}}$ be the set of such apartment classes, we thus have an isomorphism $\mathbb{F}[\mathfrak{A}_{\mathbf{B}}] \cong \mathrm{St}(\mathbf{G};\mathbb{F})$.

2.5. **Group-theoretic interpretation.** A standard property of BN-pairs is that the stabilizer of $\mathbf{B}(k)$ in $\mathbf{G}(k)$ acts transitively on $\mathfrak{A}_{\mathbf{B}}$. Since $\mathbf{B}$ is a parabolic subgroup, we have $N_{\mathbf{B}}(\mathbf{G}) = \mathbf{B}$, so the stabilizer of $\mathbf{B}(k)$ in $\mathbf{G}(k)$ is $\mathbf{B}(k)$. It follows that $\mathbf{B}(k)$ acts transitively on $\mathfrak{A}_{\mathbf{B}}$. Another standard property of BN-pairs is that the stabilizer of the apartment class $\mathcal{A}_0$ is $Z_{\mathbf{G}}(\mathbf{T})(k)$. Since $\mathbf{B}(k)$ is the semi-direct product of $\mathbf{U}(k)$ and $Z_{\mathbf{G}}(\mathbf{T})(k)$, the set map $\alpha\colon \mathbf{U}(k) \to \mathfrak{A}_{\mathbf{B}}$ defined by $\alpha(g) = g \cdot \mathcal{A}_0$ is a bijection.

The map α is $\mathbf{U}(k)$-equivariant with respect to the left action of $\mathbf{U}(k)$ on itself. It is also $Z_{\mathbf{G}}(\mathbf{T})(k)$-equivariant with respect to its conjugation action on $\mathbf{U}(k)$; indeed, for $h \in Z_{\mathbf{G}}(\mathbf{T})(k)$ and $g \in \mathbf{U}(k)$, we have

$$h \cdot \alpha(g) = hg \cdot \mathcal{A}_0 = hgh^{-1} \cdot \mathcal{A}_0 = \alpha(hgh^{-1}),$$

where in the second step we used that h stabilizes $\mathcal{A}_0$.

[8] Be warned that the standard notation in the theory of BN-pairs involves a group T, but this is *not* $\mathbf{T}(k)$. Instead, it is $Z_{\mathbf{G}}(\mathbf{T})(k)$. If k is algebraically closed, then $Z_{\mathbf{G}}(\mathbf{T})(k) = \mathbf{T}(k)$, but in general it is larger.

2.6. **Augmentation.** Let $\iota\colon \mathrm{St}(\mathbf{G};\mathbb{F}) \to \mathbb{F}[\mathbf{U}(k)]$ be the composition

$$\mathrm{St}(\mathbf{G};\mathbb{F}) \xrightarrow{\cong} \mathbb{F}[\mathfrak{A}_{\mathbf{B}}] \xrightarrow{\alpha^{-1}} \mathbb{F}[\mathbf{U}(k)],$$

so ι is a linear isomorphism. By the above, ι is equivariant for both $\mathbf{U}(k)$ and $Z_{\mathbf{G}}(\mathbf{T})(k)$; however, in what follows, we will only use the equivariance for $\mathbf{U}(k)$ and $\mathbf{T}(k)$. Let $\epsilon\colon \mathbb{F}[\mathbf{U}(k)] \to \mathbb{F}$ be the augmentation. The composition $\epsilon \circ \iota\colon \mathrm{St}(\mathbf{G};\mathbb{F}) \to \mathbb{F}$ is $\mathbf{B}(k)$-invariant but not $\mathbf{G}(k)$-invariant. It has the following simple interpretation:

Lemma 2.2. *Let the notation be as above, and let $x \in \mathrm{St}(\mathbf{G};\mathbb{F})$. Then $\epsilon(\iota(x)) \in \mathbb{F}$ is the $\mathbf{B}$-coefficient of x considered as an element of $\mathbb{F}[\mathfrak{P}_{\min}]$.*

Proof. It is enough to check this on the basis $\mathfrak{A}_{\mathbf{B}}$ for $\mathrm{St}(\mathbf{G};\mathbb{F})$, so consider $x \in \mathfrak{A}_{\mathbf{B}}$. By the definition of $\mathfrak{A}_{\mathbf{B}}$, the $\mathbf{B}$-coefficient of x is 1, so we must check that $\epsilon(\iota(x)) = 1$. By definition, $\iota(x) = \alpha^{-1}(x)$. Since this is an element of $\mathbf{U}(k)$, its image under ϵ is 1, as desired. ☐

2.7. **Proof of Proposition 1.6.** We finally turn to the proof of Proposition 1.6. Recall that this states that for all nonzero $x \in \mathrm{St}(\mathbf{G};\mathbb{F})$, there exists $g \in \mathbf{G}(k)$ such that $\epsilon(\iota(gx)) \neq 0$.

Proof of Proposition 1.6. Consider a nonzero $x \in \mathrm{St}(\mathbf{G};\mathbb{F})$. Regarding x as an element of $\mathbb{F}[\mathfrak{P}_{\min}]$, some coefficient must be nonzero. Since $\mathbf{G}(k)$ acts transitively on $\mathfrak{P}_{\min}$, there exists $g \in \mathbf{G}(k)$ such that gx has a nonzero $\mathbf{B}$-coefficient. By Lemma 2.2, we have $\epsilon(\iota(gx)) \neq 0$. ☐

3. Unipotent groups and positive actions

It remains to prove Theorem B and to show that Theorem B implies Proposition 1.7. The proof of Theorem B starts in §4. This section contains preliminaries about unipotent groups and positive actions. The final section (§3.6) proves that Theorem B implies Proposition 1.7. Throughout this section, we fix a field k. All algebraic groups we discuss are algebraic group schemes over k.

3.1. **Generalities about unipotent groups.** We will use the definition of unipotence in [12, IV.2.2.1]. This definition ensures that if $\overline{k}$ is an algebraic closure of k, then a smooth connected group $\mathbf{U}$ is unipotent if and only if its base change $\mathbf{U}_{\overline{k}}$ is [12, IV.2.2.6]. A smooth connected unipotent group $\mathbf{U}$ over k is *split* if there exists a central series

$$\mathbf{U} = \mathbf{U}_1 \rhd \mathbf{U}_2 \rhd \cdots \rhd \mathbf{U}_n \rhd \mathbf{U}_{n+1} = 1,$$

where the $\mathbf{U}_i$ are closed subgroups such that $\mathbf{U}_i/\mathbf{U}_{i+1} \cong \mathbb{G}_a$ for $1 \leq i \leq n$. If k is algebraically closed, then all smooth connected unipotent groups over k are split [12, IV.4.3.4, IV.4.3.14]. From this, we deduce the following:

Proposition 3.1. *Let $\mathbf{U}$ be an n-dimensional smooth connected unipotent group over a field k. Then $\mathbf{U}(k)$ is a nilpotent group. Moreover, if $\mathrm{char}(k) = p$ is positive, then all finitely generated subgroups of $\mathbf{U}(k)$ are finite p-groups of nilpotence class at most n and exponent at most p^n.*

Proof. Let $\overline{k}$ be an algebraic closure of k. Then $\mathbf{U}_{\overline{k}}$ is split, so $\mathbf{U}(\overline{k})$ has a central series of length n whose subquotients are isomorphic to $\overline{k}$. This implies the proposition if k is algebraically closed. The general case follows by considering the inclusion $\mathbf{U}(k) \hookrightarrow \mathbf{U}(\overline{k})$. ☐

3.2. **Splitting off unipotent subgroups.** Unipotent groups $\mathbf{U}$ can be studied inductively by identifying normal subgroups $\mathbf{U}'$ and then studying the unipotent groups $\mathbf{U}'$ and $\mathbf{U}/\mathbf{U}'$. The following is helpful for combining results about $\mathbf{U}'$ and $\mathbf{U}/\mathbf{U}'$ into results about $\mathbf{U}$:

Proposition 3.2. *Let $\mathbf{G}$ be a linear algebraic group and let $\mathbf{U} \lhd \mathbf{G}$ be a smooth connected split unipotent normal subgroup. There exists a subvariety $\mathbf{X}$ of $\mathbf{G}$ containing the identity such that the map $\mathbf{U} \times \mathbf{X} \to \mathbf{G}$ induced by the product on $\mathbf{G}$ is an isomorphism of varieties.*

Proof. Let $\mathbf{X} = \mathbf{G}/\mathbf{U}$ and let $\pi\colon \mathbf{G} \to \mathbf{X}$ be the quotient map. The map π gives $\mathbf{G}$ the structure of a $\mathbf{U}$-torsor over $\mathbf{X}$, and we can embed $\mathbf{X}$ into $\mathbf{G}$ as in the proposition precisely when that torsor is trivial. The result thus follows from two facts: the algebraic group $\mathbf{X}$ is affine (see [12, III.3.5.6]), and all such torsors over affine bases are trivial (see [12, IV.4.3.7]). ☐

3.3. **Positive actions.** Let $\mathbf{U}$ be a smooth connected unipotent group over k equipped with an action of $\mathbb{G}_m$. For $t \in \mathbb{G}_m(k)$ and $g \in \mathbb{U}(k)$, we will denote the action of t on g by ${}^t g$. Recall from the introduction that the action of $\mathbb{G}_m$ on $\mathbf{U}$ is defined to be positive if the weights of the action of $\mathbb{G}_m$ on $\mathrm{Lie}(\mathbf{U})$ are positive. If $\mathbb{G}_m$ acts on $\mathbb{G}_a$, then for $t \in \mathbb{G}_m(k) = k^\times$ and $x \in \mathbb{G}_a(k) = k$ we have ${}^t x = t^m x$ for some $m \in \mathbb{Z}$. The integer m is the *weight* of this action, and we say that $\mathbb{G}_m$ acts on $\mathbb{G}_a$ with *positive weight* if $m \geq 1$. With these definitions, we have the following:

Proposition 3.3. *Let $\mathbf{U}$ be a smooth connected unipotent group over a field k equipped with a positive $\mathbb{G}_m$-action. Then there exists a $\mathbb{G}_m$-stable central subgroup $\mathbf{A} \lhd \mathbf{U}$ with $\mathbf{A} \cong \mathbb{G}_a$ such that $\mathbb{G}_m$ acts on $\mathbf{A}$ with positive weight.*

Proof. First suppose that $\mathrm{char}(k) = 0$. The exponential map gives a $\mathbb{G}_m$-equivariant isomorphism of varieties $\mathrm{Lie}(\mathbf{U}) \to \mathbf{U}$. We can take $\mathbf{A}$ to be the image of a 1-dimensional weight space of $\mathrm{Lie}(\mathbf{U})$ contained in its center. See [12, IV.2.4] for more details.

Now suppose that $\mathrm{char}(k) = p$ is positive. Let $\mathbf{U}_1$ be *cckp-kernel* of $\mathbf{U}$, i.e., the maximal smooth connected p-torsion central closed subgroup of $\mathbf{U}$; this exists and is non-trivial [10, §B.3]. The subgroup $\mathbf{U}_1$ is stable under automorphisms of $\mathbf{U}$, and is therefore $\mathbb{G}_m$-stable. Tits [10, Theorem B.4.3] proved that $\mathbf{U}_1 = \mathbf{U}_2 \times \mathbf{U}_3$ where $\mathbf{U}_2$ and $\mathbf{U}_3$ are closed smooth $\mathbb{G}_m$-stable subgroups of $\mathbf{U}_1$ such that $\mathbb{G}_m$ acts trivially on $\mathbf{U}_2$ and $\mathbf{U}_3$ is $\mathbb{G}_m$-equivariantly isomorphic to $\mathrm{Lie}(\mathbf{U}_3)$ (regarded as a group variety in the canonical manner). Since $\mathrm{Lie}(\mathbf{U})$ only has positive weights, we have $\mathrm{Lie}(\mathbf{U}_2) = 0$, so $\mathbf{U}_2 = 1$. Thus $\mathbf{U}_1 = \mathbf{U}_3$ is $\mathbb{G}_m$-equivariantly isomorphic to $\mathrm{Lie}(\mathbf{U}_1)$. We can now take $\mathbf{A}$ to be a subgroup of $\mathbf{U}_1$ corresponding to a weight space of $\mathrm{Lie}(\mathbf{U}_1)$ under this isomorphism. ☐

Remark 3.4. If $\mathrm{char}(k) = p$ then there are non-linear actions of $\mathbb{G}_m$ on vector spaces over k. For example, let $V = \mathbb{G}_a^2$, let σ be the linear action of $\mathbb{G}_m$ on V given by $\sigma(t)(x, y) = (tx, t^2 y)$, and let τ be the automorphism of V given by $\tau(x, y) = (x, y + x^p)$. Then conjugating σ by τ gives a non-linear action of $\mathbb{G}_m$ on V. This demonstrates one of the difficulties that Tits' theorem must handle. ☐

Remark 3.5. Applying Proposition 3.3 repeatedly, one can show that a smooth connected unipotent groups $\mathbf{U}$ over a field k equipped with a positive $\mathbb{G}_m$-action must be split. This implies in particular that as a variety, $\mathbf{U}$ is isomorphic to an affine space over k. ☐

3.4. **Characteristic 0.** The following proposition will be the key to understanding positive actions in characteristic 0:

Proposition 3.6. *Let $\mathbf{U}$ be an n-dimensional smooth connected unipotent group over a field k of characteristic 0 equipped with a positive $\mathbb{G}_m$-action. Then there exist $\mathbb{G}_m$-stable subgroups $\mathbf{G}_1, \ldots, \mathbf{G}_n$ of $\mathbf{U}$ with the following properties:*

- *The map* $\mathbf{G}_1 \times \cdots \times \mathbf{G}_n \to \mathbf{U}$ *arising from the product on* $\mathbf{U}$ *is an isomorphism of varieties.*
- *For* $1 \leq i \leq n$*, we have* $\mathbf{G}_i \cong \mathbb{G}_a$ *and* $\mathbb{G}_m$ *acts on* $\mathbf{G}_i$ *with positive weight.*

Proof. Let $\mathrm{Lie}(\mathbf{U}) = \oplus_{i=1}^n \mathfrak{u}_i$ be a decomposition into 1-dimensional weight spaces and let $\mathbf{G}_i$ be the image of $\mathfrak{u}_i$ under the exponential map (see [12, §IV.2.4.5]). □

3.5. **Extending positive actions.** Let $\overline{\mathbb{G}}_m = \mathrm{Spec}(k[x])$, which is an an algebraic monoid under multiplication. The following proposition shows that if a $\mathbb{G}_m$-action on a unipotent group is positive, then it can be extended to an action of $\overline{\mathbb{G}}_m$.

Proposition 3.7. *Let* $\mathbf{U}$ *be a smooth connected unipotent group over a field* k *equipped with a positive* $\mathbb{G}_m$*-action. Then the* $\mathbb{G}_m$*-action can be extended to an action of* $\overline{\mathbb{G}}_m$ *such that* ${}^0g = \mathrm{id}$ *for all* $g \in \mathbf{U}(k)$*.*

Proof. It is enough to prove this for the base change to an algebraic closure of k, so without loss of generality we can assume that k is algebraically closed. We say that a regular function $f\colon \mathbf{U}(k) \to k$ is a *weight function* of weight n if

$$f({}^tg) = t^n f(g) \quad \text{for } t \in \mathbb{G}_m(k) \text{ and } g \in \mathbf{U}(k).$$

The space of regular functions on $\mathbf{U}(k)$ can be decomposed into a direct sum of 1-dimensional subspaces spanned by weight functions, and the $\mathbb{G}_m$-action on $\mathbf{U}$ extends to $\overline{\mathbb{G}}_m$ if and only if there are no nonzero weight functions of negative weight.

By Proposition 3.3, there exists a $\mathbb{G}_m$-stable central subgroup $\mathbf{A} \lhd \mathbf{U}$ with $\mathbf{A} \cong \mathbb{G}_a$ such that $\mathbb{G}_m$ acts on $\mathbf{A}$ with positive weight. Let $\mathbf{U}' = \mathbf{U}/\mathbf{A}$, so that we have an exact sequence

$$1 \longrightarrow \mathbf{A} \longrightarrow \mathbf{U} \xrightarrow{\pi} \mathbf{U}' \longrightarrow 1.$$

There is a corresponding short exact sequence of Lie algebras, from which it follows that the induced $\mathbb{G}_m$-action on $\mathbf{U}'$ is positive. By induction on the dimension, the $\mathbb{G}_m$-action on $\mathbf{U}'$ extends to an action of $\overline{\mathbb{G}}_m$ such that ${}^0g = \mathrm{id}$ for all $g \in \mathbf{U}'(k)$.

Since the identity element of $\mathbf{U}'(k)$ is fixed by $\mathbb{G}_m$, the ideal of regular functions vanishing on it can be generated by weight functions $\overline{h}_1, \dots, \overline{h}_r$, necessarily of nonnegative weights. Let $h_i = \pi^*(\overline{h}_i)$. The h_i generate the ideal I of regular functions on $\mathbf{U}(k)$ vanishing on $\mathbf{A}(k)$.

Now, suppose f is a weight function on $\mathbf{U}(k)$ of negative weight. Then $f|_{\mathbf{A}(k)}$ is a weight function on $\mathbf{A}(k)$ of negative weight, so $f|_{\mathbf{A}(k)} = 0$. It follows that $f \in I$, so we can write $f = \sum_{i=1}^r g_i h_i$ for regular functions g_i; in fact, we can take the g_i to be weight functions such that $g_i h_i$ has the same weight as f. Since f has negative weight and h_i has non-negative weight, it follows that g_i must have negative weight. Hence, by the same argument, $g_i \in I$, so $f \in I^2$. Continuing in this manner, we find that

$$f \in \bigcap_{n \geq 1} I^n,$$

so $f = 0$ by Krull's intersection theorem. It follows that the $\mathbb{G}_m$-action on $\mathbf{U}$ extends to $\overline{\mathbb{G}}_m$.

Finally, if $g \in \mathbf{U}(k)$ then another application of induction implies that

$$\pi({}^0g) = {}^0\pi(g) = \mathrm{id},$$

so ${}^0g \in \mathbf{A}(k)$. Since $\mathbb{G}_m$ acts on $\mathbf{A}(k)$ with positive weight, this forces ${}^0g = \mathrm{id}$. □

Remark 3.8. The converse also holds: if $\mathbf{U}$ is a smooth connected unipotent group over a field k equipped with an action of $\mathbb{G}_m$ that extends to $\overline{\mathbb{G}}_m$ such that ${}^0g = \mathrm{id}$ for all $g \in \mathbf{U}(k)$, then the $\mathbb{G}_m$-action is positive. □

3.6. **Positive actions on unipotent radicals.** Our final result in this section shows that Theorem B implies Proposition 1.7.

Proposition 3.9. *Let $\mathbf{G}$ be a connected reductive group over a field k, let $\mathbf{B}$ be a minimal parabolic subgroup of $\mathbf{G}$, let $\mathbf{U}$ be the unipotent radical of $\mathbf{B}$, and let $\mathbf{T}$ be a maximal split torus of $\mathbf{B}$. Then there exists a one-parameter subgroup $\sigma\colon \mathbb{G}_m \to \mathbf{T}$ that acts positively on $\mathbf{U}$.*

Example 3.10. Suppose $\mathbf{G} = \mathrm{GL}_n$, the group $\mathbf{B}$ is the Borel subgroup of upper triangular matrices, $\mathbf{T}$ is the torus of diagonal matrices, and $\mathbf{U}$ is the unipotent subgroup of upper triangular matrices with 1's on the diagonal. We can then take $\sigma\colon \mathbb{G}_m \to \mathbf{T}$ to be the 1-parameter subgroup $\sigma(t) = \mathrm{diag}(t^n, t^{n-1}, \ldots, t^1)$. The key property of σ is that for $g \in \mathbf{U}(k)$ and $t \in \mathbb{G}_m(k)$, the matrix ${}^t g = \sigma(t) g\, \sigma(t)^{-1}$ is obtained from g by multiplying every entry above the diagonal by a *positive* power of t. □

Proof of Proposition 3.9. We start by recalling some basic facts about relative root systems (see [4, §21]). Let $X(\mathbf{T})$ be the group of characters $\chi\colon \mathbf{T} \to \mathbb{G}_m$ and let $Y(\mathbf{T})$ be the group of one-parameter subgroups $\gamma\colon \mathbb{G}_m \to \mathbf{T}$. For $\chi \in X(\mathbf{T})$ and $\gamma \in Y(\mathbf{T})$, the composition $\chi \circ \gamma\colon \mathbb{G}_m \to \mathbb{G}_m$ can be written in the form $\chi \circ \gamma(t) = t^n$ for some $n \in \mathbb{Z}$. Define $\langle \chi, \gamma \rangle = n$. This extends to a nondegenerate pairing between $X(\mathbf{T}) \otimes \mathbb{Q}$ and $Y(\mathbf{T}) \otimes \mathbb{Q}$. Let $\Phi_{\mathrm{rel}} \subset X(\mathbf{T})$ be the relative root system of $\mathbf{T}$ in $\mathbf{G}$. There is a choice of positive roots $\Phi^+_{\mathrm{rel}} \subset \Phi_{\mathrm{rel}}$ such that Φ^+_{rel} is precisely the set of weights of the action of $\mathbf{T}$ on the Lie algebra of $\mathbf{U}$. Let $\{\chi_1, \ldots, \chi_n\} \subset \Phi^+_{\mathrm{rel}}$ be the simple positive roots. These form a basis for $X(\mathbf{T}) \otimes \mathbb{Q}$, and every element of Φ^+_{rel} is a nonnegative integer linear combination of the χ_i. Using the nondegeneracy of our pairing, we can find some $\alpha \in Y(\mathbf{T}) \otimes \mathbb{Q}$ such that $\langle \chi_i, \alpha \rangle = 1$ for $1 \leq i \leq n$. Choosing $d \geq 1$ such that $d\alpha \in Y(\mathbf{T})$, we can then take $\sigma = d\alpha$. □

4. Ideals and positive actions I: equal characteristic p

We now turn to the proof of Theorem B. The following result implies this theorem in the special case where $\mathrm{char}(k) = \mathrm{char}(\mathbb{F}) = p$ is positive, but is more general since it does not require $\mathbf{U}$ to be equipped with a positive action or for k to be infinite:

Proposition 4.1. *Let $\mathbf{U}$ be a smooth connected unipotent group over a field k and let $\mathbb{F}$ be another field. Assume that $\mathrm{char}(k) = \mathrm{char}(\mathbb{F}) = p$ is positive. Let I be a left ideal in $\mathbb{F}[\mathbf{U}(k)]$ that does not lie in the augmentation ideal. Then $I = \mathbb{F}[\mathbf{U}(k)]$.*

Proof. Let $x \in I$ be an element that does does not belong to the augmentation ideal. Write $x = \sum_{i=1}^n c_i [g_i]$ with $c_i \in \mathbb{F}$ and $g_i \in \mathbf{U}(k)$. Let Γ be the subgroup of $\mathbf{U}(k)$ generated by the g_i. Proposition 3.1 implies that Γ is a finite p-group. Let $\epsilon_\Gamma\colon \mathbb{F}[\Gamma] \to \mathbb{F}$ be the augmentation. Then $\ker(\epsilon_\Gamma)$ is the Jacobson radical of $\mathbb{F}[\Gamma]$, i.e., the intersection of all maximal left ideals (see [18, Corollary 8.8]). Since $\ker(\epsilon_\Gamma)$ is itself a maximal left ideal, it follows that it is the unique maximal left ideal. The element $x \in I \cap \mathbb{F}[\Gamma]$ does not lie in $\ker(\epsilon_\Gamma)$, so $I \cap \mathbb{F}[\Gamma] = \mathbb{F}[\Gamma]$. In particular, $[\mathrm{id}] \in I$, so $I = \mathbb{F}[\mathbf{U}(k)]$. □

5. Ideals and positive actions II: characteristic 0

In this section, we prove Theorem B when $\mathrm{char}(k) = 0$. The proof is in §5.4 after three sections of preliminaries.

5.1. **Modules over nilpotent groups.** Recall that a group G is *abelian-by-nilpotent* if there is a normal abelian subgroup A of G such that G/A is nilpotent. Also, G is *residually finite* if it injects into its profinite completion, or equivalently if the intersection of all finite-index normal subgroups of G is trivial. Hall [15] proved the following. See [20, Theorem 4.3.1] for a textbook reference.

Theorem 5.1 (Hall). *All finitely generated abelian-by-nilpotent groups are residually finite.*

We will apply Theorem 5.1 in the form of the following corollary. If M is a module over a ring R, then an R-submodule M' of M is said to be *finite-index* if the cardinality of M/M' is finite. Just like for groups, we say that M is *residually finite* if the intersection of all finite-index submodules is 0.

Corollary 5.2. *Let G be a finitely generated nilpotent group and let M be a finitely generated $\mathbb{Z}[G]$-module. Then M is residually finite.*

Proof. Set $\Gamma = M \rtimes G$, so Γ is an abelian-by-nilpotent group. Since G is finitely generated and M is a finitely generated $\mathbb{Z}[G]$-module, the group Γ is finitely generated. Theorem 5.1 thus implies that Γ is residually finite, so the intersection of all elements of $\mathcal{F} = \{\Delta \mid \Delta \lhd \Gamma \text{ finite index}\}$ is trivial. This implies that the intersection of all elements of $\mathcal{F}' = \{\Delta \cap M \mid \Delta \in \mathcal{F}\}$ is 0. Each element of $\mathcal{F}'$ is a finite-index submodule of M, so we conclude that M is residually finite. □

5.2. **Finitely generated subgroups of unipotent groups.** For a group Γ and $d \geq 1$, let $\Gamma(d)$ be the subgroup of Γ generated by all d^{th} powers.

Proposition 5.3. *Let $\mathbf{U}$ be a smooth connected unipotent group over a field k of characteristic 0 equipped with a positive $\mathbb{G}_m$-action. Let S be a finite subset of $\mathbf{U}(k)$. There exists a finitely generated subgroup Γ of $\mathbf{U}(k)$ with $S \subset \Gamma$ such that ${}^{d}g \in \Gamma(d)$ for all $g \in \Gamma$ and $d \geq 1$.*

Proof. By Proposition 3.6, there exist $\mathbb{G}_m$-stable subgroups $\mathbf{G}_1, \ldots, \mathbf{G}_n$ of $\mathbf{U}$ such that the following hold:

- The map $\mathbf{G}_1 \times \cdots \times \mathbf{G}_n \to \mathbf{U}$ arising from the product on $\mathbf{U}$ is an isomorphism of varieties.
- For $1 \leq i \leq n$, we have $\mathbf{G}_i \cong \mathbb{G}_a$ and $\mathbb{G}_m$ acts on $\mathbf{G}_i$ with positive weight $m_i \geq 1$.

This second condition implies in particular that ${}^{d}x = x^{d^{m_i}}$ for all $x \in \mathbf{G}_i(k)$ and $d \geq 1$. For $1 \leq i \leq n$, choose a finite set $S_i \subset \mathbf{G}_i(k)$ such that every element of S is a product of elements of the S_i. Let Γ be the subgroup of $\mathbf{U}(k)$ generated by $S_1 \cup \cdots \cup S_n$. We thus have $S \subset \Gamma$. For $g \in \Gamma$, we can write

$$g = s_1 \cdots s_r \quad \text{with } s_j \in S_{i_j} \text{ for } 1 \leq j \leq r,$$

so for $d \geq 1$ we have

$${}^{d}g = {}^{d}s_1 \cdots {}^{d}s_r = (s_1)^{d^{m_{i_1}}} \cdots (s_r)^{d^{m_{i_r}}} \in \Gamma(d).$$ □

5.3. **Modules over unipotent groups in char 0.** The following is our generalization of Lemma 1.8 to the setting of unipotent groups over fields of characteristic 0.

Proposition 5.4. *Let $\mathbf{U}$ be a smooth connected unipotent group over a field k of characteristic 0 equipped with a positive $\mathbb{G}_m$-action. Let M be a $\mathbb{Z}[\mathbf{U}(k)]$-module and let $N \subset M$ be a submodule. For some $m_1, \ldots, m_n \in M$ and $g_1, \ldots, g_n \in \mathbf{U}(k)$, assume that*

$${}^{d}g_1 \cdot m_1 + \cdots + {}^{d}g_n \cdot m_n \in N \text{ for all } d \geq 1.$$

Then $m_1 + \cdots + m_n \in N$.

Proof. Replacing M by M/N, we can assume that $N = 0$. Proposition 5.3 says there exists a finitely generated subgroup Γ of $\mathbf{U}(k)$ containing $\{g_1, \ldots, g_n\}$ such that ${}^{d}g \in \Gamma(d)$ for all $g \in \Gamma$ and $d \geq 1$. Proposition 3.1 implies that Γ is a nilpotent group. Let M' be the $\mathbb{Z}[\Gamma]$-submodule of M generated by the m_i. Consider a finite-index submodule M'' of M'.

The group of automorphisms of the finite abelian group underlying M'/M'' is finite. Letting d be its exponent, the group $\Gamma(d)$ acts trivially on M'/M''. Thus

$$0 = {}^{d}g_1 \cdot m_1 + \cdots + {}^{d}g_n \cdot m_n \equiv m_1 + \cdots + m_n \pmod{M''},$$

so $m_1+\cdots+m_n \in M''$. Since M'' was an arbitrary finite-index submodule of M', Corollary 5.2 implies that $m_1 + \cdots + m_n = 0$, as required. ☐

5.4. **Conclusion.** The following is Theorem B in the special case where $\mathrm{char}(k) = 0$.

Theorem 5.5. *Let $\mathbf{U}$ be a smooth connected unipotent group over a field k of characteristic 0 equipped with a positive action of $\mathbb{G}_m$ and let $\mathbb{F}$ be another field. Let $I \subset \mathbb{F}[\mathbf{U}(k)]$ be a left ideal that is stable under $\mathbb{G}_m$ and not contained in the augmentation ideal. Then $I = \mathbb{F}[\mathbf{U}(k)]$.*

Proof. For $g \in \mathbf{U}(k)$, write $[g]$ for the associated element of $\mathbb{F}[\mathbf{U}(k)]$. Let $\epsilon\colon \mathbb{F}[\mathbf{U}(k)] \to \mathbb{F}$ be the augmentation. Since I is not contained in the augmentation ideal, there exists some $x \in I$ with $\epsilon(x) = 1$. Write this as

$$x = \sum_{i=1}^{n} c_i[g_i] \in I \quad \text{with } g_1, \ldots, g_n \in \mathbf{U}(k),\ c_1, \ldots, c_n \in \mathbb{F}, \text{ and } \sum_{i=1}^{n} c_i = 1.$$

Since I is stable under the action of $\mathbb{G}_m(k)$, for all $t \in \mathbb{G}_m(k)$ we have ${}^{t}x \in I$, so

$$\sum_{i=1}^{n} c_i[{}^{t}g_i] = \sum_{i=1}^{n} {}^{t}g_i \cdot c_i[1] \in I.$$

Applying Proposition 5.4 with $M = \mathbb{F}[\mathbf{U}(k)]$ and $N = I$, we deduce that $[1] = \sum_i c_i[1] \in I$, so $I = \mathbb{F}[\mathbb{U}(k)]$. ☐

6. Ideals and positive actions III: unequal characteristic p

In this section, we prove Theorem B when $\mathrm{char}(k) = p$ and $\mathrm{char}(\mathbb{F}) \neq p$. The proof is in §6.6 after five sections of preliminaries.

6.1. **Finitely generated subgroups.** If $\mathbf{U}$ is a smooth connected unipotent group over a field k of positive characteristic, then Proposition 3.1 says that all finitely generated subgroups of $\mathbf{U}(k)$ are finite. The following lemma says that if one bounds the size of a generating set, then only finitely many isomorphism classes of finite groups occur:

Proposition 6.1. *Let $\mathbf{U}$ be a smooth connected unipotent group over a field k of positive characteristic and let $m \geq 1$. Then there only exist finitely many isomorphism classes of subgroups of $\mathbf{U}(k)$ that are generated by m elements.*

Proof. Let $n = \dim(\mathbf{U})$ and $p = \mathrm{char}(k)$. Proposition 3.1 says that all finitely generated subgroups of $\mathbf{U}(k)$ are nilpotent of class at most n and have exponent at most p^n. A cheap way to proceed is to quote the Restricted Burnside Problem (proved by Zelmanov [34]), which says that there are only finitely many isomorphism classes of finite groups with m generators and exponent at most p^n. An easier approach[9] is as follows. Let $\mathcal{C}_d$ be set of isomorphism classes of finite groups of nilpotence class d that have m generators and exponent at most p^n. We will prove that $|\mathcal{C}_d| < \infty$ by induction on d.

The base case $d = 1$ is trivial, so assume that $d > 1$. Consider $G \in \mathcal{C}_d$. Let $\gamma(G)$ be the d^{th} term of the lower central series of G, so $G/\gamma(G) \in \mathcal{C}_{d-1}$. Since $|\mathcal{C}_{d-1}| < \infty$, it is enough to prove that there are finitely many possibilities for $\gamma(G)$. Taking d-fold iterated commutators, we get a surjective map of abelian groups $\wedge^d G^{\text{ab}} \to \gamma(G)$. Since G has m

[9] This is actually the first step in the restricted Burnside problem: much of the hard work in its proof is bounding the nilpotence class of finite p-groups in terms of their exponent and number of generators.

generators and exponent at most p^n, there are finitely many possibilities for G^{ab}, and thus also finitely many possibilities for $\gamma(G)$. $\square$

This has the following corollary. Let F_m denote the free group on generators $\{x_1, \ldots, x_m\}$. If G is a group, $g_1, \ldots, g_m \in G$ are elements, and $w \in F_m$, then let $w(g_1, \ldots, g_m) \in G$ be the image of w under the homomorphism $F_m \to G$ taking x_i to g_i for $1 \le i \le m$.

Corollary 6.2. *Let $\mathbf{U}$ be a smooth connected unipotent group over a field k of positive characteristic. For all $m \ge 1$, there exists a finite set $\mathcal{W}_m$ of elements of F_m such that for all $s_1, \ldots, s_m \in \mathbf{U}(k)$, the subgroup of $\mathbf{U}(k)$ generated by the s_i equals $\{w(s_1, \ldots, s_m) \mid w \in \mathcal{W}_m\}$.*

Proof. By Proposition 6.1, there are only finitely many possibilities for the isomorphism class of the subgroup generated by the s_i. For each of these groups and each choice of m-element generating for it, include words in $\mathcal{W}_m$ to express every element in terms of those generators. $\square$

6.2. **A-polynomials.** Let k be a field of positive characteristic p. A map $\lambda\colon k \to \mathbb{F}_p$ is an *additive map* if it is a homomorphism of additive groups. If X is a variety over k, then an *A-polynomial* on X is a function $f\colon X(k) \to \mathbb{F}_p$ of the form $\lambda \circ \phi$, where $\phi\colon X \to \mathbf{A}^1$ is a morphism of varieties over k and $\lambda\colon k \to \mathbb{F}_p$ is an additive map. Here the "A" stands for "additive". We will need to do a sort of algebraic geometry with A-polynomials, and the key result is as follows:

Proposition 6.3. *Let k be an infinite field of positive characteristic p, let $f_1, \ldots, f_r\colon \mathbb{A}^1(k) \to \mathbb{F}_p$ be A-polynomials such that $f_i(0) = 0$ for $1 \le i \le r$, and let $\mathfrak{a}$ be an infinite additive subgroup of k. Then there exists some nonzero $a \in \mathfrak{a}$ such that $f_i(a) = 0$ for all $1 \le i \le r$.*

Proof. Write $f_i = \lambda_i \circ \phi_i$ with $\phi_i \in k[z]$ and $\lambda_i\colon k \to \mathbb{F}_p$ an additive map. Let $d_i = \deg(\phi_i)$, and set $m = 1 + \sum_i d_i$. Regarding $\mathfrak{a}$ as an infinite-dimensional vector space over $\mathbb{F}_p$, we can choose $\mathbb{F}_p$-linearly independent elements $v_1, \ldots, v_m \in \mathfrak{a}$. For $1 \le i \le r$, define a function

$$h_i\colon \mathbb{F}_p^m \to \mathbb{F}_p, \qquad h_i(x_1, \ldots, x_m) = f_i(x_1 v_1 + \cdots + x_m v_m).$$

We claim that h_i is a polynomial of degree at most d_i. Indeed, ϕ_i is a sum of terms of the form cz^e with $c \in k$ and $e \le d_i$, so h_i is a sum of terms of the form

$$\begin{aligned}\lambda_i\left(c\,(x_1 v_1 + \cdots + x_m v_m)^e\right) &= \sum_{j_1 + \cdots + j_m = e} \lambda_i\left(\binom{e}{j_1, \ldots, j_m} x_1^{j_1} v_1^{j_1} \cdots x_m^{j_m} v_m^{j_m}\right) \\ &= \sum_{j_1 + \cdots + j_m = e} x_1^{j_1} \cdots x_m^{j_m} \lambda_i\left(\binom{e}{j_1, \ldots, j_m} v_1^{j_1} \cdots v_m^{j_m}\right).\end{aligned}$$

Here we are using the fact that $\lambda_i\colon k \to \mathbb{F}_p$ is an additive map, which implies that it is $\mathbb{F}_p$-linear. Since $f_i(0) = 0$, we also have $h_i(0, \ldots, 0) = 0$. The Chevalley–Warning theorem [27, Corollary I.2.2.1] thus implies that there exists some nonzero $(x_1, \ldots, x_m) \in \mathbb{F}_p^m$ such that $h_i(x_1, \ldots, x_m) = 0$ for all $1 \le i \le r$. The desired $a \in \mathfrak{a}$ is then $a = x_1 v_1 + \cdots + x_m v_m$. $\square$

6.3. **Subgroups satisfying A-polynomials.** Let $\mathbf{U}$ be a smooth connected unipotent group over an infinite field k of positive characteristic equipped with a positive action of $\mathbb{G}_m$. By Proposition 3.7, the $\mathbb{G}_m$ action on $\mathbf{U}$ extends to an action of $\overline{\mathbb{G}}_m$ satisfying ${}^0g = \mathrm{id}$ for all $g \in \mathbf{U}(k)$. For a subset S of $\mathbf{U}(k)$ and an additive subgroup $\mathfrak{a}$ of k, define $\mathbf{U}(S, \mathfrak{a})$ to be the subgroup of $\mathbf{U}(k)$ generated by $\{{}^a s \mid s \in S,\ a \in \mathfrak{a}\}$. The following demonstrates the flexibility of these subgroups:

Proposition 6.4. *Let* $\mathbf{U}$ *be a smooth connected unipotent group over an infinite field* k *of positive characteristic* p *equipped with a positive action of* $\mathbb{G}_m$*. Let* S *be a finite subset of* $\mathbf{U}(k)$*, let* $\mathfrak{a}$ *be an infinite additive subgroup of* k*, and let* $f\colon \mathbf{U}(k) \to \mathbb{F}_p$ *be an A-polynomial such that* $f(\mathrm{id}) = 0$*. Then there exists an infinite additive subgroup* $\mathfrak{b}$ *of* $\mathfrak{a}$ *such that* f *vanishes on* $\mathbf{U}(S, \mathfrak{b})$*.*

Proof. Say that an additive subgroup $\mathfrak{c}$ of k is *f-vanishing* if f vanishes on $\mathbf{U}(S, \mathfrak{c})$. Below we will construct a strictly increasing chain $\mathfrak{c}_1 \subsetneq \mathfrak{c}_2 \subsetneq \cdots$ of f-vanishing finite additive subgroups of $\mathfrak{a}$. Having done this, the union $\mathfrak{b}$ of the $\mathfrak{c}_i$ will be the desired f-vanishing infinite additive subgroup of $\mathfrak{a}$.

Start by setting $\mathfrak{c}_1 = 0$, so f vanishes on $\mathbf{U}(S, \mathfrak{c}_1) = \mathrm{id}$ by assumption. Assume now that we have constructed an f-vanishing finite additive subgroup $\mathfrak{c}_i$ of $\mathfrak{a}$. To construct an f-vanishing finite additive subgroup $\mathfrak{c}_{i+1}$ of $\mathfrak{a}$ with $\mathfrak{c}_i \subsetneq \mathfrak{c}_{i+1}$, it is enough to find some nonzero $d \in \mathfrak{a} \setminus \mathfrak{c}_i$ such that $\mathfrak{c}_i + \mathbb{F}_p d$ is f-vanishing. To simplify our notation, we will let $\mathfrak{c} = \mathfrak{c}_i$. Letting $\mathfrak{d}$ be an infinite additive subgroup of $\mathfrak{a}$ with $\mathfrak{c} \cap \mathfrak{d} = 0$, we will find the desired nonzero d in $\mathfrak{d}$.

For $u \in \mathfrak{c}$ and $c \in \mathbb{F}_p$ and $g \in S$, define a morphism of varieties

$$\gamma_{u,c,g}\colon \mathbb{A}^1 \to \mathbf{U}, \qquad \gamma_{u,c,g}(t) = {}^{u+ct}g.$$

Here have crucially used the fact (Proposition 3.7) that the $\mathbb{G}_m$ action extends to $\overline{\mathbb{G}}_m = \mathbb{A}^1$. Enumerate the finite set $\{\gamma_{u,c,g} \mid u \in \mathfrak{c},\, c \in \mathbb{F}_p,\, g \in S\}$ as $\{\gamma_1, \ldots, \gamma_N\}$. By definition, for $t \in k$ the group $\mathbf{U}(S, \mathfrak{c} + \mathbb{F}_p t)$ is the subgroup of $\mathbf{U}(k)$ generated by $\{\gamma_i(t) \mid 1 \le i \le N\}$. Let $\mathcal{W}_N \subset F_N$ be the finite set provided by Corollary 6.2, and for $w \in \mathcal{W}_N$ define a morphism of varieties

$$\phi_w\colon \mathbb{A}^1 \to \mathbf{U}, \qquad \phi_w(t) = w(\gamma_1(t), \ldots, \gamma_N(t)).$$

It follows that for $t \in k$ we have

$$\mathbf{U}(S, \mathfrak{c} + \mathbb{F}_p t) = \{\phi_w(t) \mid w \in \mathcal{W}_N\}. \tag{6.1}$$

Since $\gamma_{u,c,g}(0) = {}^u g$ for all $u \in \mathfrak{c}$ and $c \in \mathbb{F}_p$ and $g \in S$, we also have

$$\mathbf{U}(S, \mathfrak{c}) = \{\phi_w(0) \mid w \in \mathcal{W}_N\}. \tag{6.2}$$

For $w \in \mathcal{W}_N$, the function

$$f_w\colon \mathbb{A}^1(k) \to \mathbb{F}_p, \qquad f_w(t) = f(\phi_w(t))$$

is an A-polynomial. Since $\mathfrak{c}$ is f-vanishing, (6.2) implies that $f_w(0) = 0$. Proposition 6.3 thus implies that we can find some nonzero $d \in \mathfrak{d}$ such that $f_w(d) = 0$ for all $w \in \mathcal{W}_N$. By (6.1), this implies that $\mathfrak{c} + \mathbb{F}_p d$ is f-vanishing, as desired. ∎

6.4. **Replacement for Hall's theorem.** Our proof of Theorem B in characteristic 0 crucially relied Hall's theorem on abelian-by-nilpotent groups via Corollary 5.2. We now prove a result that will serve as a replacement for this in positive characteristic. For a group G and a $\mathbb{Z}[G]$-module M, let M_G denote the G-coinvariants of M, i.e., the largest quotient of M on which G acts trivially. We begin with a lemma.

Lemma 6.5. *Let* p *be a prime, let* $\Lambda = \mathbb{Z}[1/p]$*, and let* G *be an abelian group of exponent* p*. Let* M *be a* $\Lambda[G]$*-module and let* $m \in M$ *be nonzero. Then there exists a subgroup* $H \subset G$ *of index* 1 *or* p *such that the image of* m *in* M_H *is nonzero.*

Proof. Let $I \subset \Lambda[G]$ be the annihilator of m. Since $m \neq 0$, this is a proper ideal, so it is contained in a maximal ideal J. Let μ_p be the group of p^{th} roots of unity in the field $\Lambda[G]/J$. The map $\Lambda[G] \to \Lambda[G]/J$ induces a group homomorphism $G \to \mu_p$; let H be its kernel, which has index 1 or p in G. Since H acts trivially on $\Lambda[G]/J \neq 0$ and the map $\Lambda[G]/I \to \Lambda[G]/J$ is surjective, we deduce that $(\Lambda[G]/I)_H \neq 0$.

Since I is the annihilator of m, there is an injection $\Lambda[G]/I \to M$ whose image is the $\Lambda[G]$-span of m. This induces a map $(\Lambda[G]/I)_H \to M_H$ whose image is the $\Lambda[G/H]$-span of m in M_H. To prove that the image of m in M_H is nonzero, it is enough to show that the map $(\Lambda[G]/I)_H \to M_H$ is injective.

In fact, we claim that taking H-coinvariants is an exact functor on the category of $\Lambda[H]$-modules. First suppose H is finite. Then it is well-known that the group homology $\mathrm{H}_k(H; N)$ for $k > 0$ is annihilated by the order of H, which is a power of p. Since N is a Λ-module, multiplication by p is an isomorphism on N. It follows that $\mathrm{H}_k(H; N) = 0$ for $k > 0$. This implies the claim in this case, as group homology is the derived functor of coinvariants. We now treat the general case. Write $H = \bigcup_{i \in I} H_i$ where the H_i are finite subgroups of H. Then $M_H = \varinjlim M_{H_i}$. Since both direct limits and the formation of H_i-coinvariants are exact, the claim follows. $\square$

The following proposition is our substitute for Corollary 5.2 in positive characteristic:

Proposition 6.6. *Let $\mathbf{U}$ be a smooth connected unipotent group over an infinite field k of positive characteristic p equipped with a positive $\mathbb{G}_m$-action. Set $\Lambda = \mathbb{Z}[1/p]$. Let S be a finite subset of $\mathbf{U}(k)$, let $\mathfrak{a}$ be an infinite additive subgroup of k, let M be a $\Lambda[\mathbf{U}(S, \mathfrak{a})]$-module, and let $m \in M$ be nonzero. Then there exists an infinite additive subgroup $\mathfrak{c}$ of $\mathfrak{a}$ such that the image of m in $M_{\mathbf{U}(S,\mathfrak{c})}$ is nonzero.*

Proof. We will prove this by induction on $n = \dim(\mathbf{U})$. The case $n = 0$ being trivial, assume that $n > 0$ and that the lemma is true in smaller dimensions. Proposition 3.3 says there exists a $\mathbb{G}_m$-stable central subgroup $\mathbf{A} \lhd \mathbf{U}$ with $\mathbf{A} \cong \mathbb{G}_a$. The intersection

$$\mathbf{A}(k) \cap \mathbf{U}(S, \mathfrak{a}) \subset \mathbf{A}(k) = k$$

is an abelian group of exponent p. By Lemma 6.5, there is a subgroup $H \subset \mathbf{A}(k) \cap \mathbf{U}(S, \mathfrak{a})$ of index either 1 or p such that the image of m in M_H is nonzero. Choose an additive homomorphism $\lambda\colon \mathbf{A}(k) \to \mathbb{F}_p$ such that $\ker(\lambda) \cap \mathbf{U}(S, \mathfrak{a}) = H$. Proposition 3.2 says that there is a subvariety $\mathbf{X}$ of $\mathbf{U}$ containing the identity such that the multiplication map $\mathbf{A} \times \mathbf{X} \to \mathbf{U}$ is an isomorphism of varieties. Using this product structure, let $\pi\colon \mathbf{U} \to \mathbf{A}$ be the projection onto the first factor. Since $\mathbf{X}$ contains the identity, it follows that $\pi|_{\mathbf{A}}$ is the identity map. Define $f\colon \mathbf{U}(k) \to \mathbb{F}_p$ to be the composition

$$\mathbf{U}(k) \xrightarrow{\pi} \mathbf{A}(k) \xrightarrow{\lambda} \mathbb{F}_p.$$

The map f is an A-polynomial, so by Proposition 6.4 there exists an infinite additive subgroup $\mathfrak{b}$ of $\mathfrak{a}$ such that f vanishes on $\mathbf{U}(S, \mathfrak{b})$. This implies that $\mathbf{A}(k) \cap \mathbf{U}(S, \mathfrak{b}) \subset H$, so the image of m in $M_{\mathbf{A}(k) \cap \mathbf{U}(S,\mathfrak{b})}$ is nonzero. Let $\overline{\mathbf{U}} = \mathbf{U}/\mathbf{A}$ and let $\overline{S} \subset \overline{\mathbf{U}}(k)$ be the image of $S \subset \mathbf{U}(k)$. The action of $\mathbf{U}(S, \mathfrak{b})$ on $M_{\mathbf{A}(k) \cap \mathbf{U}(S,\mathfrak{b})}$ factors through $\overline{\mathbf{U}}(\overline{S}, \mathfrak{b})$. Using our inductive hypothesis, we can find an infinite additive subgroup $\mathfrak{c}$ of $\mathfrak{b}$ such that the image of m in

$$\left(M_{\mathbf{A}(k) \cap \mathbf{U}(S,\mathfrak{b})}\right)_{\overline{\mathbf{U}}(\overline{S},\mathfrak{c})}$$

is nonzero. This implies that the image of m in $M_{\mathbf{U}(S,\mathfrak{c})}$ is nonzero, as desired. $\square$

6.5. **Modules over unipotent groups in char p.** The following is our generalization of Lemma 1.8 to the setting of unipotent groups over fields of positive characteristic.

Proposition 6.7. *Let $\mathbf{U}$ be a smooth connected unipotent group over a field k of positive characteristic p equipped with a positive $\mathbb{G}_m$-action. Set $\Lambda = \mathbb{Z}[1/p]$. Let M be a $\Lambda[\mathbf{U}(k)]$-module and let $N \subset M$ be a submodule. For some $m_1, \ldots, m_n \in M$ and $g_1, \ldots, g_n \in \mathbf{U}(k)$, assume that*

$${}^{t}g_1 \cdot m_1 + \cdots + {}^{t}g_n \cdot m_n \in N \textit{ for all } t \in \mathbb{G}_m(k).$$

Then $m_1 + \cdots + m_n \in N$.

Proof. Replacing M by M/N, we can assume that $N = 0$. Let $S = \{g_1, \ldots, g_n\}$, let $\mathfrak{b}$ be a non-zero additive subgroup of k, and let $t \in \mathfrak{b}$ be non-zero. Letting $\equiv$ denote equality in the $\mathbf{U}(S, \mathfrak{b})$-coinvariants of M, since the elements ${}^t g_i \in \mathbf{U}(S, \mathfrak{b})$ act trivially on these coinvariants we have

$$0 = {}^t g_1 \cdot m_1 + \cdots + {}^t g_n \cdot m_n \equiv m_1 + \cdots + m_n.$$

We thus see that $m_1 + \cdots + m_n$ maps to 0 in $M_{\mathbf{U}(S,\mathfrak{b})}$ for all non-zero $\mathfrak{b}$. Proposition 6.6 (applied with $\mathfrak{a} = k$) implies that $m_1 + \cdots + m_n = 0$. □

6.6. **Conclusion.** The following is Theorem B in the special case where $\mathrm{char}(k)$ is positive and $\mathrm{char}(\mathbb{F}) \neq \mathrm{char}(k)$.

Theorem 6.8. *Let* $\mathbf{U}$ *be a smooth connected unipotent group over a field* k *of positive characteristic* p *equipped with a positive action of* $\mathbb{G}_m$ *and let* $\mathbb{F}$ *be another field with* $\mathrm{char}(\mathbb{F}) \neq p$. *Let* $I \subset \mathbb{F}[\mathbf{U}(k)]$ *be a left ideal that is stable under* $\mathbb{G}_m$ *and not contained in the augmentation ideal. Then* $I = \mathbb{F}[\mathbf{U}(k)]$.

Proof. The proof is nearly identical to that of Theorem 5.5. For $g \in \mathbf{U}(k)$, write $[g]$ for the associated element of $\mathbb{F}[\mathbf{U}(k)]$. Let $\epsilon\colon \mathbb{F}[\mathbf{U}(k)] \to \mathbb{F}$ be the augmentation. Since I is not contained in the augmentation ideal, there exists some $x \in I$ with $\epsilon(x) = 1$. Write this as

$$x = \sum_{i=1}^{n} c_i [g_i] \in I \quad \text{with } g_1, \ldots, g_n \in \mathbf{U}(k),\ c_1, \ldots, c_n \in \mathbb{F}, \text{ and } \sum_{i=1}^{n} c_i = 1.$$

Since I is stable under the action of $\mathbb{G}_m(k)$, for all $t \in \mathbb{G}_m(k)$ we have ${}^t x \in I$, so

$$\sum_{i=1}^{n} c_i [{}^t g_i] = \sum_{i=1}^{n} {}^t g_i \cdot c_i [1] \in I.$$

Since $\mathrm{char}(\mathbb{F}) \neq p$, the field $\mathbb{F}$ is an algebra over $\Lambda = \mathbb{Z}[1/p]$, so we can regard $\mathbb{F}[\mathbf{U}(k)]$ as a $\Lambda[\mathbf{U}(k)]$-module. Applying Proposition 5.4 with $M = \mathbb{F}[\mathbf{U}(k)]$ and $N = I$, we deduce that $[1] = \sum_i c_i [1] \in I$, so $I = \mathbb{F}[\mathbb{U}(k)]$. □

References

[1] P. Abramenko and K. S. Brown, *Buildings*, Graduate Texts in Mathematics, 248, Springer, New York, 2008.

[2] A. Ash, P. E. Gunnells and M. McConnell, Resolutions of the Steinberg module for $GL(n)$, J. Algebra 349 (2012), 380–390. arXiv:1106.5034

[3] A. Ash and L. Rudolph, The modular symbol and continued fractions in higher dimensions, Invent. Math. 55 (1979), no. 3, 241–250.

[4] A. Borel, *Linear algebraic groups*, second edition, Graduate Texts in Mathematics, 126, Springer-Verlag, New York, 1991.

[5] A. Borel and J.-P. Serre, Corners and arithmetic groups, Comment. Math. Helv. 48 (1973), 436–491.

[6] A. Borel and J.-P. Serre, Cohomologie d'immeubles et de groupes S-arithmétiques, Topology 15 (1976), no. 3, 211–232

[7] V. A. Bykovskiĭ, Generating elements of the annihilating ideal for modular symbols (Russian), Funct. Anal. Appl. 37 (2003), no. 4, 263–272; translated from Funktsional. Anal. i Prilozhen. 37 (2003), no. 4, 27–38, 95.

[8] T. Church, B. Farb and A. Putman, Integrality in the Steinberg module and the top-dimensional cohomology of $\mathrm{SL}_n\mathcal{O}_K$, Amer. J. Math. 141 (2019), no. 5, 1375–1419. arXiv:1501.01307

[9] T. Church and A. Putman, The codimension-one cohomology of $\mathrm{SL}_n\mathbb{Z}$, Geom. Topol. 21 (2017), no. 2, 999–1032. arXiv:1507.06306

[10] B. Conrad, O. Gabber, G. Prasad, *Pseudo-reductive groups.* Cambridge University Press, Cambridge, 2010.

[11] C. W. Curtis, The Steinberg character of a finite group with a $(B,\, N)$-pair, J. Algebra 4 (1966), 433–441.

[12] M. Demazure, P. Gabriel, *Groupes algébriques. Tome I.* North–Holland Publishing Company, Amsterdam, 1970.
[13] S. Galatius, A. Kupers, and O. Randal-Williams, E_∞-cells and general linear groups of infinite fields, preprint 2020. arXiv:2005.05620
[14] P. E. Gunnells, Symplectic modular symbols, Duke Math. J. 102 (2000), no. 2, 329–350. arXiv:math/9809059
[15] P. Hall, On the finiteness of certain soluble groups, Proc. London Math. Soc. (3) 9 (1959), 595–622.
[16] J. E. Humphreys, The Steinberg representation, Bull. Amer. Math. Soc. (N.S.) 16 (1987), no. 2, 247–263.
[17] A. Kupers, J. Miller, P. Patzt, and J. Wilson, On the generalized Bykovskii presentation of Steinberg modules, preprint 2020. arXiv:2006.10906
[18] T. Y. Lam, *A first course in noncommutative rings*, second edition, Graduate Texts in Mathematics, 131, Springer-Verlag, New York, 2001.
[19] R. Lee and R. H. Szczarba, On the homology and cohomology of congruence subgroups, Invent. Math. 33 (1976), no. 1, 15–53.
[20] J. C. Lennox and D. J. S. Robinson, *The theory of infinite soluble groups*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, Oxford, 2004.
[21] J. Miller, R. Nagpal and P. Patzt, Stability in the high-dimensional cohomology of congruence subgroups, Compos. Math. 156 (2020), no. 4, 822–861. arXiv:1806.11131
[22] J. Miller, P. Patzt and A. Putman, On the top-dimensional cohomology groups of congruence subgroups of SL(n, $\mathbb{Z}$), Geom. Topol. 25 (2021), no. 2, 999–1058. arXiv:1909.02661
[23] J. Miller, P. Patzt, J. Wilson, and D. Yakasi, Non-integrality of some Steinberg modules, J. Topol. 13 (2020), no. 2, 441–459. arXiv:1810.07683
[24] Andy Putman (https://mathoverflow.net/users/317/andy-putman), Is the Steinberg representation always irreducible?, URL (version: 2016-05-17): https://mathoverflow.net/q/228359
[25] A. Putman and D. Studenmund, The dualizing module and top-dimensional cohomology group of $\mathrm{GL}_n(\mathcal{O})$, to appear in Math. Z. arXiv:1909.01217
[26] D. Quillen, Finite generation of the groups K_i of rings of algebraic integers, in *Algebraic K-theory, I: Higher K-theories (Proc. Conf., Battelle Memorial Inst., Seattle, Wash., 1972)*, 179–198. Lecture Notes in Math., 341, Springer, Berlin.
[27] J.-P. Serre, *A course in arithmetic*, translated from the French, Springer-Verlag, New York, 1973.
[28] L. Solomon, The Steinberg character of a finite group with BN-pair, in *Theory of Finite Groups (Symposium, Harvard Univ., Cambridge, Mass., 1968)*, 213–221, Benjamin, New York.
[29] R. Steinberg, A geometric approach to the representations of the full linear group over a Galois field, Trans. Amer. Math. Soc. 71 (1951), 274–282.
[30] R. Steinberg, Prime power representations of finite linear groups, Canadian J. Math. 8 (1956), 580–591.
[31] R. Steinberg, Prime power representations of finite linear groups. II, Canadian J. Math. 9 (1957), 347–351.
[32] R. Steinberg, Comments on the Papers, in *Robert Steinberg*, Collected Works, 7, American Mathematical Society, Providence, RI, 1997.
[33] J. Tits, *Buildings of spherical type and finite BN-pairs*, Lecture Notes in Mathematics, Vol. 386, Springer-Verlag, Berlin, 1974.
[34] E. I. Zelmanov, On the restricted Burnside problem, in *Fields Medallists' lectures*, 623–632, World Sci. Ser. 20th Century Math., 5, World Sci. Publ., River Edge, NJ.

Department of Mathematics; University of Notre Dame; 255 Hurley Hall; Notre Dame, IN 46556

Email address: andyp@nd.edu

Department of Mathematics; University of Michigan; 2074 East Hall; 530 Church St; Ann Arbor MI 48109

Email address: asnowden@umich.edu