Journal of Algebra **237,** 691–707 (2001)
doi:10.1006/jabr.2000.8604, available online at http://www.idealibrary.com on IDEAL®

A Presentation for the Unipotent Group over Rings with Identity

Daniel K. Biss

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
E-mail: daniel@math.mit.edu

and

Samit Dasgupta

Department of Mathematics, Harvard University, Cambridge, Massachusetts 02138
E-mail: dasgupta@post.harvard.edu

Communicated by Walter Feit

Received March 28, 2000

For a ring R with identity, define $\mathrm{Unip}_n(R)$ to be the group of upper-triangular matrices over R all of whose diagonal entries are 1. For $i = 1, 2, \ldots, n-1$, let S_i denote the matrix whose only nonzero off-diagonal entry is a 1 in the ith row and $(i+1)$st column. Then for any integer m (including $m = 0$), it is easy to see that the S_i generate $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$. Reiner gave relations among the S_i which he conjectured gave a presentation for $\mathrm{Unip}_n(\mathbf{Z}/2\mathbf{Z})$. This conjecture was proven by Biss [*Comm. Algebra* **26** (1998), 2971–2975] and an analogous conjecture was made for $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ in general. We prove this conjecture, as well as a generalization of the conjecture to unipotent groups over arbitrary rings.

1. INTRODUCTION AND PRELIMINARIES

For a ring R with identity, we define the unipotent group $\mathrm{Unip}_n(R)$ to be the group of $n \times n$ upper-triangular matrices over R with 1's along the diagonal. For $i = 1, 2, \ldots, n-1$, let S_i denote the matrix in $\mathrm{Unip}_n(R)$ whose only nonzero element above the diagonal is a 1 in the ith row and the $(i+1)$st column. It is easy to see that for any integer m, the set $\{S_1, S_2, \ldots, S_{n-1}\}$ generates $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$. In this paper, we prove the

0021-8693/01 $35.00

following theorem, conjectured by Biss [1], giving a presentation of $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ with these generators. Throughout, we use the convention $[a, b] = aba^{-1}b^{-1}$.

THEOREM 1. *The generators* $\{s_1, \dots, s_{n-1}\}$ *subject to the relations*

$$s_i^m = 1 \qquad \text{for } 1 \le i \le n-1, \tag{1}$$

$$[s_i, s_j] = 1 \qquad \text{for } i < j-1, \tag{2}$$

$$[s_i, [s_i, s_{i+1}]] = [s_{i+1}, [s_i, s_{i+1}]] = 1 \qquad \text{for } 1 \le i \le n-2, \tag{3}$$

$$[[s_i, s_{i+1}], [s_{i+1}, s_{i+2}]] = 1 \qquad \text{for } 1 \le i \le n-3 \tag{4}$$

give a presentation for the group $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ *under the map* $s_i \mapsto S_i$. *Furthermore, relation* (4) *is unnecessary for odd* m.

We also prove a generalization of this theorem for unipotent groups over arbitrary rings R with identity.

These results generalize theorems of Pavlov [8] and Levčuk [6]. In particular, [8, Theorem 1] for the case $R = \mathbf{Z}/p\mathbf{Z}$ with p prime and [6, Theorem 1] for the general case give presentations whose generators are all matrices with only one nonzero element off the diagonal. Also, [8, Theorem 2] gives a presentation in the case $R = \mathbf{Z}/p\mathbf{Z}$ with our generators. However, these presentations use many more relations than just (1)–(4) above; in particular, they include relations containing arbitrarily many of the s_i's. In contrast, our relations are "local" in the sense that, with the exception of relation (2), each one involves at most three consecutive s_i's.

Our results also generalize a paper by Biss [1] that uses different methods to prove Theorem 1 in the case $m = 2$. The work in [1] was in turn motivated by discussions with Reiner [9], who found a larger set of relations while investigating the restriction from $GL_n(\mathbf{Z}/2\mathbf{Z})$ to $\mathrm{Unip}_n(\mathbf{Z}/2\mathbf{Z})$ of the Steinberg representation (see [3, Example 2, p. 60]). Such a presentation may have applications in computing the cohomology of the unipotent group.

It is also possible to obtain presentations for $\mathrm{Unip}_n(R)$ from the fact that it is the unipotent subgroup of a Chevalley group (see, for example, [12]), but these presentations also have more generators and relations than the one we give. In fact, using results from group cohomology, we are able to prove that our presentation has a minimal number of both generators and relations for several rings, including $R = \mathbf{Z}/m\mathbf{Z}$ for m odd and $R = \mathbf{F}_q$ for odd prime powers q.

We now present the notation that will be used throughout the paper. Fix m and n, and let G denote the abstract group with generators $\{s_1, \dots, s_{n-1}\}$

and relations given in Eqs. (1)–(4) if m is even and (1)–(3) if m is odd. Let U denote the group $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$. For elements a_i of any group, we recursively define the generalized commutator $[a_1, \dots, a_k] = [[a_1, \dots, a_{k-1}], a_k]$. In the case $k = 1$, set $[a_1] = a_1$.

For $i < j$, define $\psi_{ij} = [s_i, s_{i+1}, \dots, s_j] \in G$ and the corresponding generalized commutators $\Psi_{ij} = [S_i, S_{i+1}, \dots, S_j] \in U$. If $i = j + 1$, then we let ψ_{ij} and Ψ_{ij} be the identity elements in their respective groups. We will not use the symbols ψ_{ij} or Ψ_{ij} for $i > j + 1$. Note that with these definitions, Ψ_{ij} is the matrix with 1's along the diagonal, a 1 in the $(i, j+1)$ position, and 0's elsewhere. As a final bit of notation, if $h \in H$ and K is a normal subgroup of H, we will write $\bar{h}$ for the image of h under the projection $H \to H/K$.

We have a natural surjection $\pi : G \to U$ given by $\pi(s_i) = S_i$. Our goal is to prove that π is an isomorphism. We will outline a proof of this fact in the next section, fill in the details of the proof for m even in Section 3, and strengthen the result for m odd in Section 4. In Section 5, we generalize this presentation to arbitrary rings R with identity. That is, we give a set of generators of $\mathrm{Unip}_n(R)$ analogous to the S_i and give generalizations of relations (1)–(4). An additional relation that reflects the multiplicative structure of R is also necessary. We then show how to generalize the proof of the presentation of $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ to the group $\mathrm{Unip}_n(R)$. Although it would be possible to prove the result in full generality without first considering $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$, we feel that this exposition is preferable because it makes the notation less obscure for the technical heart of the proof and because the conjecture that motivated this work was stated for the ring $R = \mathbf{Z}/m\mathbf{Z}$. We conclude, in Section 6, by showing that our presentations are the best possible for many rings R.

2. OUTLINE OF THE PROOF

Consider the lower central series $G = G_0 \triangleright G_1 \triangleright G_2 \cdots$ defined by $G_{i+1} = [G, G_i]$. The map π induces surjections π_i: $G_i/G_{i+1} \to U_i/U_{i+1}$. We will show that these maps are in fact isomorphisms and that $G_{n-1} \cong U_{n-1} \cong 1$. A downward induction on i then shows that $G_i \cong U_i$ by applying the five lemma [7, Chap. VII] to the diagram

$$\begin{tikzcd}
1 \arrow[r] & G_{i+1} \arrow[r] \arrow[d, "\pi|_{G_{i+1}}"'] & G_i \arrow[r] \arrow[d, "\pi|_{G_i}"] & G_i/G_{i+1} \arrow[r] \arrow[d, "\pi_i"] & 1 \\
1 \arrow[r] & U_{i+1} \arrow[r] & U_i \arrow[r] & U_i/U_{i+1} \arrow[r] & 1
\end{tikzcd}$$

In particular, we will have that $\pi : G \to U$ is an isomorphism, as desired.

First of all, one easily checks that for $i = 0, \dots, n-1$, the quotient group U_i/U_{i+1} is isomorphic to $(\mathbf{Z}/m\mathbf{Z})^{n-i-1}$. Under this isomorphism, the element $\overline{\Psi}_{j,i+j} \in U_i/U_{i+1}$ corresponds to the jth standard basis vector of $(\mathbf{Z}/m\mathbf{Z})^{n-i-1}$. More explicitly, a matrix in U_i has 0's along the jth superdiagonal for $1 \le j \le i-1$, and its coset modulo U_{i+1} is determined by its values on the ith superdiagonal; the image of this coset in $(\mathbf{Z}/m\mathbf{Z})^{n-i-1}$ under our isomorphism is the vector containing the values along this ith superdiagonal.

To prove that π_i is an isomorphism, we construct inverse maps ϕ_i: $(\mathbf{Z}/m\mathbf{Z})^{n-i-1} \to G_i/G_{i+1}$ where ϕ_i sends the jth standard basis vector to $\overline{\psi}_{j,i+j}$. To demonstrate that ϕ_i is well defined, it suffices to prove that $\psi_{j,i+j}^m = 1$ since G_i/G_{i+1} is abelian.

To see this, we will show that s_{i+j} commutes with $\psi_{j,i+j}$, which along with Lemma 2 below gives the desired result, since then

$$\psi_{j,i+j}^m = \left[\psi_{j,i+j-1}, s_{i+j}\right]^m = \left[\psi_{j,i+j-1}, s_{i+j}^m\right] = 1.$$

Lemma 1. *If a, b, and c are elements of a group, and b commutes with $[a,c]$, then $[a,b][a,c] = [a,bc]$.*

Proof. We have

$$\begin{aligned}[a,b][a,c] &= aba^{-1}b^{-1}[a,c] = aba^{-1}[a,c]b^{-1}\\ &= abca^{-1}c^{-1}b^{-1} = [a,bc].\end{aligned}$$

∎

Lemma 2. *If a and b are elements of a group, and b commutes with $[a,b]$, then $[a,b^n] = [a,b]^n$ for any integer n.*

Finally, we will show that the $\overline{\psi}_{j,i+j}$ generate G_i/G_{i+1}, thereby showing that ϕ_i is surjective and hence indeed an inverse map to π_i. This will complete the proof that π is an isomorphism.

3. TECHNICAL DETAILS OF THE PROOF

In this section, we derive the relations that will be needed to complete the proof. To show that ϕ_i is well defined, we must verify that s_{i+j} commutes with $\psi_{j,i+j}$. To show that ϕ_i is surjective, we will need to prove that $\psi_{j,i+j} = [s_j, \psi_{j+1,i+j}]$. This will allow us to manipulate elements in the image of ϕ_i and show that they generate G_i/G_{i+1}. Since every relation is invariant under the operation of incrementing the indices, it suffices to verify these facts for $j = 1$. Indeed, we often implicitly use the fact that if an equation is satisfied in the symbols $s_1, \ldots, s_i$, then the analogous equation in $s_j, \ldots, s_{i+j-1}$ holds as well.

LEMMA 3. *For each $i \geq 2$, the following statements hold*:

1. $[s_{i-1}, s_i]$ *commutes with* $\psi_{1,i-1}$.
2. $s_1, \ldots, s_i$ *commute with* $\psi_{1,i}$.
3. $[s_1, \psi_{2,i-1}]$ *commutes with* $\psi_{2,i}$.
4. $\psi_{1,i}$ *commutes with* $\psi_{1,i-1}$.
5. $\psi_{1,i} = [s_1, \psi_{2,i}]$.

Proof. The proof proceeds by induction on i. The base cases $i = 2$ all follow directly from relation (3). We now carry out the inductive step, beginning with statement 1. By induction on statement 5, $\psi_{1,i-1} = [s_1, [s_2, \ldots, [s_{i-2}, s_{i-1}] \cdots]]$. However, $[s_{i-1}, s_i]$ commutes with $[s_{i-2}, s_{i-1}]$ as well as s_j for each $j < i - 2$, and statement 1 immediately follows.

We now proceed to statement 2. Recall that $\psi_{1,i} = [\psi_{1,i-1}, s_i]$ by definition. Since s_j commutes with $\psi_{1,i-1}$ for $j \leq i - 1$ by induction and since $[s_j, s_i] = 1$ for $j \leq i - 2$, we need only verify that $\psi_{1,i}$ commutes with s_{i-1} and s_i. For now, we check only the case of s_{i-1}. We have

$$\begin{aligned} s_{i-1}\psi_{1,i} &= s_{i-1}\psi_{1,i-1}s_i\psi_{1,i-1}^{-1}s_i^{-1} \\ &= \psi_{1,i-1}s_{i-1}s_i\psi_{1,i-1}^{-1}s_i^{-1} \end{aligned}$$

by induction on statement 2. Writing $s_{i-1}s_i = [s_{i-1}, s_i]s_i s_{i-1}$ yields

$$\begin{aligned} &\psi_{1,i-1}[s_{i-1}, s_i]s_i s_{i-1}\psi_{1,i-1}^{-1}s_i^{-1} \\ &\quad = \psi_{1,i-1}[s_{i-1}, s_i]s_i\psi_{1,i-1}^{-1}s_{i-1}s_i^{-1} \\ &\quad = \psi_{1,i-1}[s_{i-1}, s_i]s_i\psi_{1,i-1}^{-1}\left[s_{i-1}, s_i^{-1}\right]s_i^{-1}s_{i-1}. \\ &\quad = \psi_{1,i-1}s_i[s_{i-1}, s_i]\psi_{1,i-1}^{-1}[s_{i-1}, s_i]^{-1}s_i^{-1}s_{i-1}. \end{aligned}$$

Using the just-proven statement 1, we see that this is equal to

$$\psi_{1,i-1}s_i\psi_{1,i-1}^{-1}s_i^{-1}s_{i-1} = \psi_{1,i}s_{i-1},$$

which is what we wanted. We postpone the demonstration that s_i and $\psi_{1,i}$ commute and instead move on to statement 3.

By induction on statement 5, $[s_1, \psi_{2,i-1}] = \psi_{1,i-1}$. Furthermore, $\psi_{2,i} = [s_2, [s_3, \ldots, [s_{i-1}, s_i] \cdots]]$. Now $\psi_{1,i-1}$ commutes with $s_2, \ldots, s_{i-2}$ by induction on statement 2 and also with $[s_{i-1}, s_i]$ by statement 1, which has been proven. This completes the proof of statement 3.

Statement 4 follows directly from the part of statement 2 which has already been proven, since all the terms in the commutator $\psi_{1,i-1}$ commute with $\psi_{1,i}$.

We can now prove statement 5. To simplify the notation, write $P = \psi_{2,i-1}$. Then by induction on statement 5, $\psi_{1,i} = [\psi_{1,i-1}, s_i] = [[s_1, P], s_i]$. We want to show that this is equal to $[s_1, [P, s_i]]$. The desired equality is then

$$\left(s_1 P s_1^{-1} P^{-1}\right) s_i \left(P s_1 P^{-1} s_1^{-1}\right) s_i^{-1} = s_1 \left(P s_i P^{-1} s_i^{-1}\right) s_1^{-1} \left(s_i P s_i^{-1} P^{-1}\right).$$

Cancelling the common terms on the left of the desired equality and noting that $s_i^{-1} s_1^{-1} s_i = s_1^{-1}$ on the right-hand side (since $i \geq 3$), we see that we need to show

$$s_1^{-1} P^{-1} s_i P s_1 P^{-1} s_1^{-1} s_i^{-1} = s_i P^{-1} s_1^{-1} P s_i^{-1} P^{-1}.$$

Multiplying on the left by s_i^{-1} and commuting this term past the s_1^{-1} on the left-hand side reduces the desired equality to

$$s_1^{-1} \left[s_i^{-1}, P^{-1}\right] s_1 P^{-1} s_1^{-1} s_i^{-1} = P^{-1} s_1^{-1} P s_i^{-1} P^{-1}.$$

Multiplying on the left by $s_1 P$ and on the right by s_i, we reduce what we need to

$$s_1 P s_1^{-1} \left[s_i^{-1}, P^{-1}\right] s_1 P^{-1} s_1^{-1} = \left[P, s_i^{-1}\right].$$

Now, the right-hand side is equal to $[P, s_i]^{-1}$ by induction on statement 2 and Lemma 2, and this commutes with P by induction on statement 4. Hence after conjugating by P on both sides, it remains to show

$$\left[P^{-1}, s_1\right] \left[s_i^{-1}, P^{-1}\right] \left[s_1, P^{-1}\right] = \left[P, s_i^{-1}\right].$$

Once again using induction and Lemma 2, this statement is equivalent to the fact that $[s_1, P]$ and $[P, s_i]^{-1}$ commute; this follows from statement 3, which has just been proven.

Only the last part of statement 2 remains to be shown: we must show that s_i commutes with $\psi_{1,i}$. However, we have just shown that $\psi_{1,i} = [s_1, \psi_{2,i}]$. Then since s_i commutes with s_1 (for $i \geq 3$) and with $\psi_{2,i}$ by induction on statement 2, our desired statement follows.

This completes the induction step and the proof of the lemma. ∎

As demonstrated in Section 2, statement 2 of Lemma 3 implies that ϕ_i is well defined. We now show how Lemma 3 can be used to show that ϕ_i is surjective.

LEMMA 4. *For $k = 0, \dots, n-2$, the generalized commutator $[s_{n_0}, \dots, s_{n_k}]$ is the identity unless each initial subset of the indices $\{n_0, \dots, n_j\}$ for $j = 0, \dots, k$ is a set of consecutive integers. In this case, the commutator is a power of a commutator of the form $\psi_{i,i+k}$. More precisely,*

$$[s_{n_0}, \dots, s_{n_k}] = \begin{cases} [s_{i_k}, \dots, s_{i_k+k}]^{\pm 1}, & \textit{if for } j = 0, \dots, k, \\ & \exists i_j \textit{ with } \{n_0, \dots, n_j\} = \{i_j, \dots, i_j + j\}, \\ 1, & \textit{otherwise}. \end{cases}$$

Proof. We proceed by induction. The base case $k = 0$ is clear. For the inductive step, assume $k > 0$. We assume that there exist i_j for $j = 0, \dots, k-1$ such that $\{n_0, \dots, n_j\} = \{i_j, \dots, i_j + j\}$, since otherwise

$$[s_{n_0}, \dots, s_{n_k}] = \big[[s_{n_0}, \dots, s_{n_{k-1}}], s_{n_k}\big] = \big[1, s_{n_k}\big] = 1$$

by the induction hypothesis. Thus, we have

$$\begin{aligned} [s_{n_0}, \dots, s_{n_k}] &= \big[[s_{n_0}, \dots, s_{n_{k-1}}], s_{n_k}\big] \\ &= \big[[s_{i_{k-1}}, \dots, s_{i_{k-1}+k-1}]^{\pm 1}, s_{n_k}\big]. \end{aligned}$$

By statement 2 of Lemma 3, we see that if $n_k \in \{i_{k-1}, \dots, i_{k-1} + k - 1\}$, then the commutator above is the identity. Furthermore, if $n_k > i_{k-1} + k$ or $n_k < i_{k-1} - 1$, then s_{n_k} commutes with all the terms in $[s_{i_{k-1}}, \dots, s_{i_{k-1}+k-1}]$ by relation (2), and so the commutator above is once again the identity. Hence we are left with only the possibilities $n_k = i_{k-1} + k$ and $n_k = i_{k-1} - 1$. With $i_k = i_{k-1}$ or $i_k = i_{k-1} - 1$ in these two cases, respectively, we have $\{n_0, \dots, n_k\} = \{i_k, \dots, i_k + k\}$. In the first case we obtain

$$\begin{aligned} [s_{n_0}, \dots, s_{n_k}] &= \big[[s_{i_k}, \dots, s_{i_k+k-1}]^{\pm 1}, s_{i_k+k}\big] \\ &= [s_{i_k}, \dots, s_{i_k+k}]^{\pm 1} \end{aligned}$$

by Lemma 2 and statement 4 of Lemma 3. In the second case,

$$\begin{aligned} [s_{n_0}, \dots, s_{n_k}] &= \big[[s_{i_k+1}, \dots, s_{i_k+k}]^{\pm 1}, s_{i_k}\big] \\ &= \big[s_{i_k}, [s_{i_k+1}, \dots, s_{i_k+k}]^{\pm 1}\big]^{-1} \\ &= [s_{i_k}, \dots, s_{i_k+k}]^{\mp 1}. \end{aligned}$$

Here, the last equality uses Lemma 2 and statements 2 and 5 of Lemma 3. This proves the lemma. ∎

LEMMA 5. *The set* $\{\bar{\psi}_{j,i+j}\}_{j=1,\dots,n-i-1}$ *generates* G_i/G_{i+1}.

Proof. It is a general fact that if a set of elements $\{a_j\}$ generates a group H, then the cosets $\overline{[a_{n_0},\dots,a_{n_i}]}$ generate H_i/H_{i+1} [5, III.1.11]. However, the previous lemma shows that each generalized commutator $\overline{[s_{n_0},\dots,s_{n_i}]}$ is in the subgroup of G_i/G_{i+1} generated by the $\bar{\psi}_{j,i+j}$. Hence G_i/G_{i+1} is generated by the $\bar{\psi}_{j,i+j}$ as desired. ∎

The previous lemma shows that ϕ_i is surjective. Now to complete the proof that π is an isomorphism, it remains to prove that $G_{n-1} \cong U_{n-1} \cong 1$. It is known that if a set of elements $\{a_j\}$ generates a group H, then all conjugates of elements of the form $[a_{n_0},\dots,a_{n_i}]$ generate H_i [10, 5.1]. Hence the argument of Lemma 4 shows that G_{n-1} is trivial, since any generalized commutator involving n of the s_i is trivial. The surjection $G \to U$ induces a surjection $G_{n-1} \to U_{n-1}$, so U_{n-1} is trivial as well. This concludes the proof that π is an isomorphism and completes the proof of Theorem 1 in the case of m even.

4. ODD m

In this section, we show that when m is odd, the relation (4) actually follows from the other three relations. Let $m = 2k + 1$ and let G denote the group with generators s_1, s_2, and s_3 subject to relations (1)–(3). If we can show that the equation $[[s_1, s_2],[s_2, s_3]] = 1$ holds in G, then relation (4) will hold in general by the invariance of the relations under the operation of incrementing all the indices.

Since relation (4) is vacuous when $n = 3$, the previous section shows that the subgroup of G generated by s_1 and s_2 is isomorphic to $\mathrm{Unip}_3(\mathbf{Z}/m\mathbf{Z})$. Thus any equation involving only s_1 and s_2 can simply be checked in 3×3 matrices. Write A for the ring of rational numbers p/q in reduced form with denominator q relatively prime to m. For $a, b \in A$, we write $a \equiv b$ if $a - b$ has numerator divisible by m when written in lowest terms. (That is, A is the ring $\mathbf{Z}$ localized with respect to the integers relatively prime to m and $\equiv$ is congruence modulo the ideal mA.) Throughout this section, we will allow expressions of the form s_i^a, where $a \in A$: this will simply represent s_i^x where x is any integer such that $x \equiv a$. Note that this is well defined; indeed, $s_i^a = s_i^b$ if $a \equiv b$.

LEMMA 6. *Let* $a, b, c, x, y, z \in A$ *such that* $a + c$ *has numerator relatively prime to* m, $x \equiv \frac{bc}{a+c}$, $y \equiv a + c$, *and* $z \equiv \frac{ab}{a+c}$. *Then the equation* $s_2^a s_1^b s_2^c = s_1^x s_2^y s_1^z$ *holds in* G.

Proof. One simply checks the corresponding matrix equation in $\mathbf{Z}/m\mathbf{Z}$:

$$\begin{pmatrix} 1 & b & bc \\ & 1 & a+c \\ & & 1 \end{pmatrix} = \begin{pmatrix} 1 & x+z & xy \\ & 1 & y \\ & & 1 \end{pmatrix}.$$

∎

As noted earlier, the lemma remains true if the indices are incremented by 1. With the variables as above, we find that

$$\begin{aligned} s_3 s_2^a s_1^b s_2^c s_3 &= s_3 s_1^x s_2^y s_1^z s_3 \\ &= s_1^x s_3 s_2^y s_3 s_1^z \\ &= s_1^x s_2^{y/2} s_3^2 s_2^{y/2} s_1^z, \end{aligned} \tag{5}$$

where the first and last equalities above use Lemma 6. We will be interested in applying formula (5) in two specific cases:

$$s_3 s_2^3 s_1 s_2 s_3 = s_1^{1/4} s_2^2 s_3^2 s_2^2 s_1^{3/4} \tag{6}$$

and

$$\begin{aligned} s_2(s_3 s_2 s_1 s_2 s_3) &= \left(s_2 s_1^{1/2} s_2\right) s_3^2 s_2 s_1^{1/2} \\ &= \left(s_1^{1/4} s_2^2 s_1^{1/4}\right) s_3^2 s_2 s_1^{1/2} \\ &= s_1^{1/4} s_2^2 s_3^2 s_1^{1/4} s_2 s_1^{1/2}. \end{aligned} \tag{7}$$

Equations (6) and (7) imply that

$$s_2 s_3 s_2 s_1 s_2 s_3 = \left(s_3 s_2^3 s_1 s_2 s_3\right)\left(s_1^{-3/4} s_2^{-2} s_1^{1/4} s_2 s_1^{1/2}\right).$$

Now the second expression in parentheses can be simplified to $s_2^{-1}[s_1, s_2]$, as can be checked with a matrix computation. The resulting equation can be reduced to relation (4) as follows:

$$\begin{gathered} s_2 s_3 s_2 s_1 s_2 s_3 = s_3 s_2^3 s_1 s_2 s_3 s_2^{-1}[s_1, s_2], \\ \left[s_2^{-1}, s_3^{-1}\right] s_2 s_1 s_2 s_3 = s_2^2 s_1 s_2 s_3 s_2^{-1}[s_1, s_2], \\ [s_2, s_3] s_2^{-1} s_1 s_2 s_3 = s_1 s_2 s_3 s_2^{-1}[s_1, s_2], \\ [s_2, s_3]\left[s_2^{-1}, s_1\right] = (s_1 s_2 s_3)\left(s_2^{-1}[s_1, s_2] s_1^{-1} s_3^{-1}\right), \\ [s_2, s_3][s_1, s_2] = ([s_1, s_2] s_2 s_1 s_3)\left(s_1^{-1} s_2^{-1} s_3^{-1}\right) \\ = [s_1 s_2][s_2, s_3]. \end{gathered}$$

We have now collected enough facts to prove the desired result about $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$, which we restate below.

THEOREM 1 (Restated). *The group* $\mathrm{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ *has a presentation with generating set* $\{s_i\}_{1 \le i \le n-1}$ *and relations* (1)–(3), *and, if m is even,* (4).

5. THE MAIN THEOREM

The methods developed above can be used to produce a presentation of $\mathrm{Unip}_n(R)$ for any ring R with identity in terms of a presentation of R. Let T be a set of generators for the additive group of R containing the multiplicative identity 1. For notational convenience, we place an arbitrary strict order $\prec$ on T. Let K be a set of additive relations $\sum a_t \cdot t = 0$ on these generators which present the additive group of R. (In any sum indexed by the elements of T, we assume that all but finitely many of the coefficients are zero; similarly for products.)

For any $r \in R$, write $S_i(r)$ for the matrix in $\mathrm{Unip}_n(R)$ whose only nonzero entry above the diagonal is an r in the ith row and $(i+1)$st column. The elements $S_i(t)$ for $t \in T$ and $i = 1, \dots, n-1$ generate the group $\mathrm{Unip}_n(R)$.

THEOREM 2. *The generators* $\{s_i(t)\}$ *for* $i = 1, \dots, n-1$ *and* $t \in T$ *subject to the relations below give a presentation for the group* $\mathrm{Unip}_n(R)$ *under the map* $s_i(t) \mapsto S_i(t)$:

$$[s_i(t), s_i(u)] = 1 \qquad \text{for } t \prec u \in T \text{ and } i = 1, \dots, n-1; \tag{8}$$

$$\prod_{t \in T} s_i(t)^{a_t} = 1$$

$$\text{for } i = 1, \dots, n-1, \text{ for each relation } \sum a_t \cdot t = 0 \text{ in } K; \tag{9}$$

$$[s_i(t), s_j(u)] = 1 \qquad \text{for } 1 \le i < j-1 \le n-2 \text{ and } t, u \in T; \tag{10}$$

$$[s_i(1), [s_i(t), s_{i+1}(1)]] = [s_{i+1}(1), [s_i(t), s_{i+1}(1)]] = 1$$

$$\text{for } i = 1, \dots, n-2 \text{ and } t \in T; \tag{11}$$

$$[[s_i(t), s_{i+1}(1)], [s_{i+1}(u), s_{i+2}(1)]] = 1$$

$$\text{for } i = 1, \dots, n-3 \text{ and } t, u \in T; \tag{12}$$

and

$$[s_i(u)^{-1}, s_{i+1}(v)^{-1}] = \prod_{t \in T} [s_i(t), s_{i+1}(1)]^{c_t}$$

$$\text{for } i = 1, \dots, n-2 \text{ and } u, v \in T, v \ne 1. \tag{13}$$

Here, $\Sigma c_t \cdot t$ is an arbitrary expression of the product uv in terms of elements of T. The theorem holds without relation (12) *if R has odd characteristic.*

Relation (8) reflects the commutativity of the additive group of R. Relation (9) encodes the structure of the additive group of R and generalizes relation (1). Relations (10)–(12) generalize relations (2)–(4). Finally, relation (13) encodes the multiplicative structure of R. The presence of the inverses in the commutator on the left-hand side of relation (13) is unfortunate. We believe that the theorem remains true if both inverses are removed, but we are unable to prove this formulation.

We now outline the proof of Theorem 2. Let G be the group generated by symbols $s_i(t)$ for $1 \le i \le n-1$ and $t \in T$ subject to the relations (8)–(13).

LEMMA 7. *The following statements hold in G for any $i = 1, \dots, n-2$ and any t, u, and v in T.*

1. *$s_{i+1}(u)$ commutes with $[s_i(t), s_{i+1}(1)]$.*
2. *$s_i(u)$ commutes with $[s_i(t), s_{i+1}(1)]$.*
3. *$[s_i(t), s_{i+1}(u)] = [s_i(t)^{-1}, s_{i+1}(u)^{-1}]$.*
4. *$s_i(t)$ and $s_{i+1}(t)$ commute with $[s_i(u), s_{i+1}(v)]$.*

Proof. For the first statement, we simply calculate

$$\begin{aligned} &s_{i+1}(u)\left[s_i(t), s_{i+1}(1)\right] \\ &\quad = s_i(t)s_{i+1}(u)\left[s_{i+1}(u)^{-1}, s_i(t)^{-1}\right]s_{i+1}(1)s_i(t)^{-1}s_{i+1}(1)^{-1} \\ &\quad = s_i(t)s_{i+1}(u)s_{i+1}(1)\left[s_{i+1}(u)^{-1}, s_i(t)^{-1}\right]s_i(t)^{-1}s_{i+1}(1)^{-1} \\ &\quad = s_i(t)s_{i+1}(1)s_{i+1}(u)\left[s_{i+1}(u)^{-1}, s_i(t)^{-1}\right]s_i(t)^{-1}s_{i+1}(1)^{-1} \\ &\quad = s_i(t)s_{i+1}(1)s_i(t)^{-1}s_{i+1}(u)^{-1}s_{i+1}(1)^{-1} \\ &\quad = \left[s_i(t), s_{i+1}(1)\right]s_{i+1}(u). \end{aligned} \tag{14}$$

The key step here is Eq. (14), which uses the relations (11) and (13). Now since both $s_{i+1}(t)$ and $s_i(1)$ commute with

$$\left[s_i(t), s_{i+1}(1)\right] = \left[s_i(1)^{-1}, s_{i+1}(t)^{-1}\right],$$

this commutator is equal to $[s_i(1), s_{i+1}(t)]$ by Lemma 2. Thus, as above, we have

$$\begin{aligned}
&[s_i(t), s_{i+1}(1)]s_i(u)^{-1}\\
&\quad = [s_i(1), s_{i+1}(t)]s_i(u)^{-1}\\
&\quad = s_i(1)s_{i+1}(t)s_i(1)^{-1}s_{i+1}(t)^{-1}s_i(u)^{-1}\\
&\quad = s_i(1)s_{i+1}(t)s_i(1)^{-1}\left[s_{i+1}(t)^{-1}, s_i(u)^{-1}\right]s_i(u)^{-1}s_{i+1}(t)^{-1}\\
&\quad = s_i(1)s_{i+1}(t)\left[s_{i+1}(t)^{-1}, s_i(u)^{-1}\right]s_i(1)^{-1}s_i(u)^{-1}s_{i+1}(t)^{-1}\\
&\quad = s_i(u)^{-1}[s_i(1), s_{i+1}(t)]\\
&\quad = s_i(u)^{-1}[s_i(t), s_{i+1}(1)].
\end{aligned}$$

This proves the second statement of the lemma. The third now follows from the first two and Lemma 2. Finally, the last statement of the lemma follows from the first three and relation (13). ▮

The rest of the proof of Theorem 2 follows the proof of Theorem 1. We have a natural surjection $\pi : G \to U = \mathrm{Unip}_n(R)$. To check that π is an isomorphism, we again pass to the successive quotients of the lower central series: it is easy to see that $U_i/U_{i+1} \cong R^{n-i-1}$, and we will construct maps $\phi_i : R^{n-i-1} \to G_i/G_{i+1}$ as before. For $1 \le l \le n-i-1$, let $e_l(t)$ denote the element of R^{n-i-1} whose only nonzero entry is a t in the lth coordinate. The $e_l(t)$ generate R^{n-i-1}, so to define ϕ_i we need only specify the values $\phi_i(e_l(t))$. We define

$$\phi_i(e_l(t)) = \overline{[s_l(1), \dots, s_{l+i-1}(1), s_{l+i}(t)]} \in G_i/G_{i+1}.$$

We must show that ϕ_i is well defined and surjective. To do this, we need the following analogue of Lemma 3.

LEMMA 8. *Fix $i \ge 2$ and $t_1, \dots, t_i, t'_1, \dots, t'_i \in T$. Then the following statements hold*:

1. $[s_{i-1}(t_{i-1}), s_i(t_i)]$ *commutes with* $[s_1(t'_1), \dots, s_{i-1}(t'_{i-1})]$.
2. $s_1(t_1), \dots, s_i(t_i)$ *commute with* $[s_1(t'_1), \dots, s_i(t'_i)]$.
3. $[s_1(t_1), [s_2(t_2), \dots, s_{i-1}(t_{i-1})]]$ *commutes with* $[s_2(t'_2), \dots, s_i(t'_i)]$.
4. $[s_1(t_1), \dots, s_i(t_i)]$ *commutes with* $[s_1(t'_1), \dots, s_{i-1}(t'_{i-1})]$.
5. $[s_1(t_1), \dots, s_i(t_i)] = [s_1(t_1), [s_2(t_2), \dots, s_i(t_i)]]$.

Proof. Lemma 7 shows that the $s_i(t)$ satisfy the formal properties used in showing Lemma 3, so the proof carries over to this case verbatim. ▮

Now, to see that ϕ_i is well defined, we use Lemmas 1 and 8 observe that for all relations $\sum a_t \cdot t = 0$ in our given presentation of the additive group of R, we have

$$\begin{aligned}\prod_{t\in T} \phi_i(e_l(t))^{a_t} &= \prod_{t\in T} [s_l(1),\dots,s_{l+i-1}(1),s_{l+i}(t)]^{a_t}\\ &= \prod_{t\in T} [[s_l(1),\dots,s_{l+i-1}(1)],s_{l+i}(t)]^{a_t}\\ &= \left[[s_l(1),\dots,s_{l+i-1}(1)],\prod_{t\in T} s_{l+i}(t)^{a_t}\right]\\ &= [[s_l(1),\dots,s_{l+i-1}(1)],1]\\ &= 1.\end{aligned}$$

To see that ϕ_i is surjective, let $t_1,\dots,t_{i+1} \in T$. Then we may express the product of the t_j as a linear combination of the additive generators,

$$\prod_{m=1}^{i+1} t_m = \sum_{t\in T} c_t t,$$

for some integers c_t. By repeated application of Eq. (13), we have

$$\begin{aligned}[s_l(t_1),\dots,s_{l+i}(t_{i+1})] &= \prod_{t\in T} [s_l(t)^{c_t},s_{l+1}(1),\dots,s_{l+i}(1)]\\ &= \prod_{t\in T} [s_l(1),\dots,s_{l+i-1}(1),s_{l+i}(t)]^{c_t}\\ &= \prod_{t\in T} \phi_i(e_l(t))^{c_t}\end{aligned}$$

so that $[s_l(t_1),\dots,s_{l+i}(t_{i+1})] \in \operatorname{im}\phi_i$. Then by the same reasoning used in the proof of Lemma 5, the map ϕ_i is surjective. The proof of Theorem 2 now directly parallels that of Theorem 1. In particular, the proofs given in Section 4 generalize to show that if R has odd characteristic, Eqs. (9)–(11) imply relation (12).

6. MINIMALITY AND THE SCHUR MULTIPLIER

Naturally, having found a presentation for the unipotent group, we would like to see that it is minimal—namely, that we cannot remove any of the relations or generators and still have a presentation. In many cases, we can do better: we show that for many rings, no presentation of the unipotent group has fewer generators or relations than ours. To prove this, we need to use some results from group cohomology and homology.

LEMMA 9. *Suppose the additive group of R is finitely generated, and that the generating set T is of minimal size. Then the generating set $\{s_i(t)\}$ for* $\mathrm{Unip}_n(R)$ *constructed in Theorem 2 is also of minimal size.*

Proof. This follows from the fact that $\mathrm{Unip}_n(R)/[\mathrm{Unip}_n(R), \mathrm{Unip}_n(R)] \cong R^{n-1}$, and so any generating set for $\mathrm{Unip}_n(R)$ must be at least as large as the minimal generating set for R^{n-1} (as an abelian group). However, it is easily seen from the classification of finitely generated abelian groups that a minimal generating set for R^{n-1} must have $n-1$ times as many elements as a minimal generating set for R. ▮

The relationship between group cohomology and presentations is given by the following lemma, proven in [11, I.4.3].

LEMMA 10. *Let p be a prime number and let G be a p group. Then the minimal number of relations needed to present G is equal to the $\mathbf{Z}/p\mathbf{Z}$ dimension of $H^2(G, \mathbf{Z}/p\mathbf{Z})$.*

Now, it is very easy to compute $H^2(G, \mathbf{Z}/p\mathbf{Z})$ given knowledge of $H_2(G, \mathbf{Z})$, which is known as the Schur multiplier of G, and so the size of the Schur multiplier has a great deal of control over the complexity of presentations of the group. This is unsurprising for the following reason. Suppose $G \cong F/K$ is a presentation of G, where F is a free group. Then it is known (see, for example, [2]) that the Schur multiplier is isomorphic to

$$\frac{K \cap [F, F]}{[F, K]}.$$

Because of this formula, as well as other applications, the Schur multiplier is a well-studied object and has been computed in many situations applicable to this setting. For $I = R = \mathbf{F}_q$, the following result is owing to Evens [4], and the general case follows from his arguments in a completely straightforward way.

LEMMA 11. *Let $q = p^k$ be an odd prime power, let R be any ring, and let $I \subset R$ be an ideal with q elements. Then since I is still an abelian group with a distributive multiplication, we can form the unipotent group* $\mathrm{Unip}_n(I)$. *For a minimal generating set T of I, the Schur multiplier* $H_2(\mathrm{Unip}_n(I), \mathbf{Z})$ *of*

$\mathrm{Unip}_n(I)$ *is a sum of cyclic p groups with generators*:

$$[s_i(t), s_i(u)] \qquad \textit{for } 1 \le i \le n-1, t \prec u \in T, \tag{15}$$

$$[s_i(t), s_j(u)] \qquad \textit{for } 1 \le i < j-1 \le n-2, t, u \in T, \tag{16}$$

$$[s_i(1), [s_i(1), s_{i+1}(t)]] \quad \textit{and} \quad [s_{i+1}(1), [s_i(1), s_{i+1}(t)]] \qquad \textit{for } 1 \le i \le n-2, t \in T, \tag{17}$$

and

$$[s_i(u), s_{i+1}(v)]^{-1} \prod_{t \in T} [s_i(t)^{c_t}, s_{i+1}(1)] \qquad \textit{for } 1 \le i \le n-2, u, v \in T, uv = \sum_{t \in T} c_t \cdot t. \tag{18}$$

Using this lemma and a few basic facts about group cohomology, it is not hard to obtain the minimality result we are aiming for.

THEOREM 3. *Let $q = p^k$ be an odd prime power and suppose R is a ring with q elements. Let T and K be minimal sets of additive generators and relations, respectively, for R. Then any presentation of* $\mathrm{Unip}_n(R)$ *has at least as many generators and relations as the presentation given by Theorem* 2.

Proof. Let $G = \mathrm{Unip}_n(R)$. The desired minimality property of our generating set is given by Lemma 9. By Lemma 10, the property for relations will follow if we can show that the rank of $H^2(G, \mathbf{Z}/p\mathbf{Z})$ is equal to the number of relations given in Theorem 2. Here, by the rank of an abelian p group M, we mean the size of the smallest generating set of M or, equivalently, the $\mathbf{Z}/p\mathbf{Z}$ rank of the vector space $M \otimes_{\mathbf{Z}} \mathbf{Z}/p\mathbf{Z}$.

By the universal coefficient theorem, we have

$$H^2(G, \mathbf{Z}/p\mathbf{Z}) \cong \mathrm{Hom}(H_2(G, \mathbf{Z}), \mathbf{Z}/p\mathbf{Z}) \oplus \mathrm{Ext}(H_1(G, \mathbf{Z}), \mathbf{Z}/p\mathbf{Z}).$$

Hence by Lemma 11 with $I = R$, it suffices to prove that the number of relations given in Theorem 2 exceeds the number of expressions given in Eqs. (15)–(18) by exactly the rank of $\mathrm{Ext}(H_1(G, \mathbf{Z}), \mathbf{Z}/p\mathbf{Z})$. By definition, $H_1(G, \mathbf{Z}) = G/[G, G]$, which, as we observed in the proof of Lemma 9, is isomorphic to a direct sum of $n-1$ copies of the additive group of R. Furthermore, this additive group is an abelian p group—denote its rank by r. Then $\mathrm{Ext}(H_1(G, \mathbf{Z}), \mathbf{Z}/p\mathbf{Z})$ has rank $r(n-1)$.

Now consider the number of relations given in Eqs. (8)–(11) and (13). Notice that since q is odd, we do not require the relations from Eq. (12). Equations (15) and (8) represent the same number of expressions, as do Eqs. (16) and (10), (17) and (11), and (18) and (13). This leaves only Eq. (9), but since the additive group of R has rank r, we know that there are precisely $r(n-1)$ relations in Eq. (9). This completes the proof. ▮

The rings R for which Theorem 3 holds do not include many of the rings for which the presentation of Theorem 2 is most appealing, for example, $\mathbf{Z}/m\mathbf{Z}$ for m not equal to an odd prime power. Fortunately, however, we can deduce from Theorem 3 similar results about other rings.

COROLLARY 1. *For any odd number m, any presentation of* $\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ *has at least as many generators* (*namely* $n-1$) *and at least as many relations* (*namely* $n-1+\binom{n-2}{2}+2(n-2)=\frac{(n-4)(n+3)}{2}$) *as the presentation given in Theorem* 2 (*corresponding to the generating set* $T=\{1\}$ *and the singleton relation set* $K=\{m\cdot 1=0\}$).

Proof. For $m=p^k$ an odd prime power, the result is simply a special case of Theorem 3. For the general case, notice that $\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ is nilpotent, and recall that for G a finite nilpotent group, we have

$$G \cong \prod_p G_p,$$

where $G_p < G$ is a Sylow p subgroup. Let $m=\prod_p p^{k_p}$ be the factorization of m into prime powers, and denote the ideal $\frac{m}{p^{k_p}}(\mathbf{Z}/m\mathbf{Z}) \subset \mathbf{Z}/m\mathbf{Z}$ by I_p. Then the subgroup $\operatorname{Unip}_n(I_p) \subset \operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ has $(p^{k_p})^{\binom{n}{2}}$ elements, so it is a Sylow p subgroup of $\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z})$. Thus, we have

$$\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z}) \cong \prod_p \operatorname{Unip}_n(I_p).$$

Denote by ϕ_p the projection map $\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z}) \to \operatorname{Unip}_n(I_p)$. Then clearly ϕ_p takes a presenting set of generators and relations for $\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ to a presentation of $\operatorname{Unip}_n(I_p)$. However, by Lemma 11 and an argument akin to the proof of Theorem 3, any presentation of any of the factors $\operatorname{Unip}_n(I_p)$ has at least as many generators and relations as the presentation of $\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ given in Theorem 2. Thus, any presentation of $\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ also has at least that many generators and relations. ▮

Of course, it would be very nice to have a similar result for m even (particularly since the result for $m=2$ alone would imply minimality of the representation of $\operatorname{Unip}_n(\mathbf{Z})$), but this does not seem possible without knowledge of the Schur multiplier of $\operatorname{Unip}_n(\mathbf{Z}/m\mathbf{Z})$ for m even. It appears that the Even's techniques would become hopelessly complicated in even characteristic, but perhaps there is other, more recent work which could shed light on this case.

ACKNOWLEDGMENTS

This work was done during the 1998 Research Experience for Undergraduates at the University of Minnesota, Duluth, directed by Joseph Gallian and sponsored by the National Science Foundation (grant number DMS-95-31373-001) and the National Security Agency (grant number MDA-904-98-1-0523). We thank Jacob Lurie for numerous enlightening discussions about this problem and extensive proofreading of the paper, and Dan Isaksen for pointing out the role played by lower central series.

REFERENCES

1. D. K. Biss, A presentation for the unipotent group over $\mathbf{F}_2$, *Comm. Algebra* **26** (1998), 2971–2975.
2. K. S. Brown, "Cohomology of Groups," Springer-Verlag, New York, 1982.
3. P. H. Edelman and V. Reiner, *H*-shellings and *h*-complexes, *Adv. Math.* **106** (1994), 36–64.
4. L. Evens, The Schur multiplier of a semi-direct product, *Illinois J. Math.* **16** (1972), 166–181.
5. B. Huppert, "Endliche Gruppen I," Springer-Verlag, Berlin, 1967.
6. V. M. Levčuk, Automorphisms of certain nilpotent matrix groups and rings, *Sov. Math. Dokl.* **16** (1975), 756–760.
7. W. S. Massey, "A Basic Course in Algebraic Topology," Springer-Verlag, New York, 1991.
8. P. P. Pavlov, Sylow *p*-subgroups of the full linear group over the prime field of characteristic *p*, *Izv. Akad. Nauk SSSR* **16** (1952), 437–458. MR14, 533. [Russian]
9. V. Reiner, personal communication.
10. D. J. S. Robinson, "A Course in the Theory of Groups," Springer-Verlag, New York, 1996.
11. J.-P. Serre, "Galois Cohomology," Springer-Verlag, New York, 1997.
12. R. Steinberg, "Lectures on Chevalley Groups," mimeographed notes, Yale, 1967.