J. Inst. Math. Jussieu (2023), **22**(2), 961–983
doi:10.1017/S1474748021000426

THE p-ZASSENHAUS FILTRATION OF A FREE PROFINITE GROUP AND SHUFFLE RELATIONS

IDO EFRAT

Earl Katz Family Chair in Pure Mathematics, Department of Mathematics, Ben-Gurion University of the Negev, P.O. Box 653, Be'er-Sheva 8410501,Israel
(efrat@bgu.ac.il)

(Received 22 April 2020; revised 2 August 2021; accepted 4 August 2021; first published online 29 September 2021)

Abstract For a prime number p and a free profinite group S on the basis X, let $S_{(n,p)}$, $n=1,2,\ldots$, be the p-Zassenhaus filtration of S. For $p>n$, we give a word-combinatorial description of the cohomology group $H^2\left(S/S_{(n,p)},\mathbb{Z}/p\right)$ in terms of the shuffle algebra on X. We give a natural linear basis for this cohomology group, which is constructed by means of unitriangular representations arising from Lyndon words.

Key words and phrases: p-Zassenhaus filtration, modular dimension subgroups, Galois cohomology, shuffle algebra, shuffle relations, Massey products

2020 Mathematics Subject Classification: Primary 12G05
Secondary 68R15; 12F10; 12E30

1. Introduction

The purpose of this paper is to study the p-Zassenhaus filtration of a free profinite group S and its cohomology by means of the combinatorics of words. Here p is a fixed prime number, and we recall that the p-Zassenhaus filtration of a profinite group G is given by $G_{(n,p)}=\prod_{ip^j\geq n}\left(G^{(i)}\right)^{p^j}$, $n=1,2,\ldots$ – that is, $G_{(n,p)}$ is generated as a profinite group by all p^j-powers of elements of the ith term of the (profinite) lower central filtration $G^{(i)}$ of G for $ip^j\geq n$.

This filtration was introduced by Zassenhaus [39] for discrete groups (under the name *dimension subgroups modulo p*) as a tool to study free Lie algebras in characteristic p. It proved itself to be a powerful tool in a variety of group-theoretic and arithmetic problems: the Golod–Shafarevich solution to the class field tower problem ([20], [21, §7.7], [40], [13]), the structure of finitely generated pro-p groups of finite rank [5, Ch. 11], mild groups [24] and one-relator pro-p groups [15, §2.4], multiple residue symbols and their knot-theory analogues ([29], [30, Ch. 8], [37]), and more.

In the Galois-theory context, where $G=G_F$ is the absolute Galois group of a field F containing a root of unity of order p, it was shown in [12] that the quotient $G/G_{(3,p)}$

determines the full cohomology ring $H^*(G)=\bigoplus_{i\geq 0}H^i(G)$ with the cup product. Here and in the sequel we abbreviate $H^i(G)=H^i(G,\mathbb{Z}/p)$ for the profinite cohomology group of G with its trivial action on $\mathbb{Z}/p$. Moreover, $G/G_{(3,p)}$ is the smallest Galois group of F with this property (see also [3]).

In the present paper we focus on the cohomology group $H^2\left(G/G_{(n,p)}\right)$ for a profinite group G and $n\geq 2$. Its importance is that it controls the relator structure in the pro-p group $G/G_{(n,p)}$, whereas its generators are captured by the group $H^1(G/G_{(n,p)})$, which is well understood [31, §3.9].

Our main result gives, for a free profinite group S on a basis X, an explicit description of $H^2\left(S/S_{(n,p)}\right)$ in terms of the *combinatorics of words*. Namely, we consider X as an alphabet with a fixed total order, and let X^* be the monoid of words in X. For every $n\geq 0$, let $\mathbb{Z}\langle X\rangle_n$ be the free $\mathbb{Z}$-module generated by all words in X^* of length n. Let $\mathrm{Sh}(X)_{\mathrm{indec},n}$ be its quotient by the submodule generated by all *shuffle products* u ш v, where u,v are nonempty words in X^* with $|u|+|v|=n$. We recall that for words $u=(x_1\cdots x_r)$ and $v=(x_{r+1}\cdots x_{r+s})$ in X^*, one defines

$$u ш v=\sum_\sigma\left(x_{\sigma^{-1}(1)}\cdots x_{\sigma^{-1}(r+t)}\right)\in\mathbb{Z}\langle X\rangle,$$

where σ ranges over all permutations of $\{1,2,\dots,r+s\}$ such that $\sigma(1)<\cdots<\sigma(r)$ and $\sigma(r+1)<\cdots<\sigma(r+t)$. Thus $\mathrm{Sh}(X)_{\mathrm{indec},n}$ is the nth homogenous component of the indecomposable quotient of the *shuffle algebra* $\mathrm{Sh}(X)$ in the sense of [9, §5] (see §9). We prove the following word-combinatorial description of $H^2\left(S/S_{(n,p)}\right)$ for p sufficiently large:

Main Theorem. *Suppose that $n<p$. There is a canonical isomorphism of $\mathbb{F}_p$-linear spaces*

$$\left(\bigoplus_{x\in X}\mathbb{Z}/p\right)\oplus\left(\mathrm{Sh}(X)_{\mathrm{indec},n}\otimes(\mathbb{Z}/p)\right)\xrightarrow{\sim}H^2\left(S/S_{(n,p)}\right).$$

When $p\leq n$ we have a similar result, in the form of a canonical epimorphism.

More specifically, to any word w in X^* of length $1\leq|w|\leq n$ we associate a canonical cohomology element $\alpha_{w,n}\in H^2\left(S/S_{(n,p)}\right)$. Then the isomorphism in the Main Theorem is induced by the map $w\mapsto\alpha_{w,n}$, where w is either a single-letter word or a word of length n. In these cases, $\alpha_{w,n}$ turns out to be a *Bockstein element* or an element of an *n-fold Massey product*, respectively (see Examples 7.1–7.2 and the remarks below). The construction of $\alpha_{w,n}$ is based on a representation of $S/S_{(n,p)}$ in a group of *unitriangular* (i.e., unipotent upper-triangular) matrices, which we derive from the *Magnus map* – see §5 and §7 for details.

A main ingredient of the proof, of independent importance, is the construction of a canonical $\mathbb{F}_p$-linear basis of $H^2\left(S/S_{(n,p)}\right)$, which we call the *Lyndon basis*. Recall that a nonempty word w in X^* is called a *Lyndon word* if it is smaller in the alphabetic order (induced by the fixed total order on X) than all its nontrivial right factors (i.e., suffixes). The Lyndon basis then consists of all cohomology elements $\alpha_{w,n}$, where w is a Lyndon word of length $\lceil n/p^k \rceil$ for some $k \geq 0$. When $n \leq p$ the possible lengths are only 1 and n, leading to the two direct summands in the left-hand side of the Main Theorem.

We further use Lyndon words to give a canonical basis of the $\mathbb{F}_p$-linear space $S_{(n,p)}/\left(S_{(n,p)}\right)^p \left[S,S_{(n,p)}\right]$, and prove a duality (in a unitriangular sense) between the Lyndon basis of $H^2\left(S/S_{(n,p)}\right)$ and this latter basis (Corollary 8.2). In the smallest case, $n=2$, this recovers classical duality results between Bockstein elements/cup products and p-powers/commutators, respectively, proved by Labute in his classical work on Demuškin groups ([22, Prop. 8], [34], [31, Ch. III, §9], [34]). In the case $n=3$, it refines results by Vogel [37, §2].

The paper builds upon our earlier work [8], supplemented by [9], where we proved analogous results for the *lower p-central filtration*, defined inductively by $G^{(1,p)}=G$ and $G^{(n,p)}=\left(G^{(n-1,p)}\right)^p \left[G,G^{(n-1,p)}\right]$ for $n \geq 2$. In many respects, this filtration and the p-Zassenhaus filtration are the opposite extremes among the filtrations related to mod-p cohomology.

While we follow the general philosophy of [8] and [9], their methods fall short when applied to the p-Zassenhaus filtration. Therefore we modify these methods in several aspects: mainly, whereas in the lower p-central case one should consider words w of arbitrary lengths, in the case of Zassenhaus filtration we need to restrict to words of lengths $\lceil n/p^k \rceil$, $k \geq 0$, as above. These 'jumps' arise when we analyze the filtration for the group $\mathbb{U}_i\left(\mathbb{Z}/p^j\right)$ of unitriangular $(i+1)\times(i+1)$-matrices over $\mathbb{Z}/p^j$. They turn out to have crucial, and quite nonobvious, properties, which are in particular needed for handling the dual $S_{(n,p)}/\left(S_{(n,p)}\right)^p \left[S,S_{(n,p)}\right]$ of $H^2\left(S/S_{(n,p)}\right)$. Here commutator identities due to Shalev [36] also play a key role. By contrast, the corresponding quotient in the lower p-central case is $S^{(n,p)}/S^{(n+1,p)}$, which is considerably more tractable. In addition, the analysis for the lower p-central filtration in [8] is based on (mixed) Lie-algebra computations. In the case of Zassenhaus filtration we instead apply the theory of free *p-restricted* Lie algebra, following [25] and [15].

The correspondence in the Main Theorem demonstrates deep connections between the p-Zassenhaus filtration and its cohomology and the n-fold Massey product $H^1(G)^n \to H^2(G)$. In fact, it was shown in [7] that when S is a free profinite group, $S_{(n,p)}/S_{(n+1,p)}$ is dual to the subgroup of $H^2\left(S/S_{(n,p)}\right)$ generated by all such products. Moreover, the latter subgroup is the kernel of the inflation map $H^2\left(S/S_{(n,p)}\right) \to H^2\left(S/S_{(n+1,p)}\right)$. The size of $S_{(n,p)}/S_{(n+1,p)}$ was computed in [26]. The behavior of Massey products for absolute Galois groups $G=G_F$ has been the focus of extensive research in recent years, where the p-Zassenhaus filtration has played an important role (see, e.g., [10], [16], [17], [18], [27], [28] and the references therein).

2. Hall sets

Let X be a nonempty set, considered as an alphabet. Let again X^* be the free monoid on X. We consider its elements as *associative* words. It is equipped with the binary operation $(u,v) \mapsto uv$ of associative concatenation. Let $\mathcal{M}_X$ be the *free magma* on X (see [35, Part I, Ch. IV, §1], [8, §2]). Thus the elements of $\mathcal{M}_X$ are the nonempty *nonassociative* words in the alphabet X, and it is equipped with the binary operation $(u,v) \mapsto (uv)$

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

of nonassociative concatenation. There is a natural *foliage* (brackets-dropping) map $f\colon \mathcal{M}_X \to X^*$, which is the identity on X (considered as a subset of both $\mathcal{M}_X$ and X^*) and which commutes with the concatenation maps.

We fix a total order on X. It induces on X^* the *alphabetic* order $\leq_{\rm alp}$, which is also total. We denote the length of a word $w \in X^*$ by $|w|$.

Let $\mathcal{H}$ be a subset of words in $\mathcal{M}_X$ and $\leq$ any total order on $\mathcal{H}$. We say that $(\mathcal{H}, \leq)$ is a *Hall set in* $\mathcal{M}_X$ if the following conditions hold [33, §4.1]:

(1) $X \subseteq \mathcal{H}$ as ordered sets.

(2) If $h = (h'h'') \in \mathcal{H} \setminus X$, then $h'' \in \mathcal{H}$ and $h < h''$.

(3) For $h = (h'h'') \in \mathcal{M}_X \setminus X$, one has $h \in \mathcal{H}$ if and only if
- $h', h'' \in \mathcal{H}$ and $h' < h''$, and
- either $h' \in X$ or $h' = (h_1h_2)$ with $h_2 \geq h''$.

In this case we say that $H = f(\mathcal{H})$ is a *Hall set in* X^*.

Every $w \in H$ can be written as $w = f(h)$ for a *unique* $h \in \mathcal{H}$ [33, Cor. 4.5]. If $w \in H \setminus X$, then we can uniquely write $h = (h'h'')$ with $h', h'' \in \mathcal{H}$ [33, p. 89]. Setting $w' = f(h'), w'' = f(h'') \in H$, we call $w = w'w''$ the *standard factorization* of w.

Example 2.1. The set of all Lyndon words in X^* (see the Introduction) is a Hall set with respect to $\leq_{\rm alp}$ [33, Th. 5.1].

The standard factorization of Lyndon words is explicitly given as follows:

Lemma 2.2. *Let $w, u, v \in X^*$ be nonempty words such that $w = uv$ and w is Lyndon. The following conditions are equivalent:*

(a) *$w = uv$ is the standard factorization of w in the set of Lyndon words.*

(b) *v is the $\leq_{alp}$-minimal nontrivial right factor of w which is Lyndon.*

(c) *v is the longest nontrivial right factor of w which is Lyndon.*

Proof. (a)⇔(b): This is shown in the proof of [33, Th. 5.1].

(b)⇒(c): Let v' by a nontrivial Lyndon right factor of w. By (b), $v \leq_{\rm alp} v'$. Since v' is Lyndon, v cannot be a nontrivial right factor of v'. Hence v' is a right factor of v, so $|v'| \leq |v|$.

(c)⇒(b): Let v' be a nontrivial Lyndon right factor of w. By (c), it is a right factor of v. Since v is Lyndon, $v \leq_{\rm alp} v'$. □

We order $\mathbb{Z}_{\geq 0} \times X^*$ lexicographically with respect to the usual order on $\mathbb{Z}_{\geq 0}$ and $\leq_{\rm alp}$. We then define a second total order $\preceq$ on X^* by setting

$$w_1 \preceq w_2 \quad \Longleftrightarrow \quad (|w_1|, w_1) \leq (|w_2|, w_2) \tag{2.1}$$

with respect to the latter order on $\mathbb{Z}_{\geq 0} \times X^*$.

3. Lie algebras

Let R be a unital commutative ring. We write $R\langle X\rangle$ for the free associative R-algebra over the set X. We view its elements as polynomials in the set X of noncommuting variables

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

and with coefficients in R. Alternatively, it is the free R-module on the basis X^* with multiplication induced by concatenation. The algebra $R\langle X\rangle$ is graded with respect to total degree.

We write $R\langle\langle X\rangle\rangle$ for the R-algebra of formal power series in the set X of noncommuting variables and with coefficients in R.

Let k be a field. For an associative k-algebra A, let A_{Lie} be the Lie algebra on A with Lie bracket $[a,b]=ab-ba$.

We now assume that X is a nonempty totally ordered set, and fix a Hall set H in X^*.

Let $L(X)$ be the free Lie k-algebra on the set X. The universal enveloping algebra of $L(X)$ is $k\langle X\rangle$ [35, Part I, Ch. IV, Th. 4.2].

Let L be a Lie k-algebra containing X. Define a map $P_L=P_L^H\colon H\to L$ by $P_L(x)=x$ for $x\in X$, and $P_L(w)=[P_L(u),P_L(v)]$, if $w=uv$ is the standard factorization of w, as in §2. This construction is functorial in L in the natural sense.

Proposition 3.1.

(a) *When $L=L(X)$, the images $P_L(w)$, where $w\in H$, form a k-linear basis of $L(X)$.*

(b) *Let L be a Lie k-algebra containing X. Then the image of P_L k-linearly spans the Lie k-subalgebra of L generated by X.*

Proof.

(a) See [33, Th. 4.9(i)].

(b) This follows from (a), the universal property of $L(X)$, and the functoriality of P_L. □

By Proposition 3.1(a) and the Poincaré–Birkhoff–Witt theorem [35, Part I, Ch. III, §4], the products $\prod_{i=1}^m P_{L(X)}(w_i)$, with $w_1\geq_{\mathrm{alp}}\cdots\geq_{\mathrm{alp}} w_m$ in H, form a k-linear basis of the universal enveloping algebra $k\langle X\rangle$ of $L(X)$.

Next assume that $\operatorname{char} k=p>0$. A *restricted Lie k-algebra* L is a Lie k-algebra with an additional unary operation $a\mapsto a^{[p]}$ for which there is an associative k-algebra A and a Lie k-algebra monomorphism $\theta\colon L\to A_{\mathrm{Lie}}$ such that $\theta\left(a^{[p]}\right)=\theta(a)^p$ for every $a\in L$ ([5, §12.1]; see also [19] for an alternative equivalent definition). A morphism of restricted Lie k-algebras is a morphism of Lie k-algebras which commutes with the $(\cdot)^{[p]}$-maps.

Every associative k-algebra A is endowed with the structure of a restricted Lie algebra $A_{\mathrm{res.Lie}}$, where we set $[a,b]=ab-ba$ and $a^{[p]}=a^p$. Every restricted Lie k-algebra L has a unique *restricted universal enveloping algebra* $\mathcal{U}_{\mathrm{res}}(L)$. This means that $\mathcal{U}_{\mathrm{res}}(L)$ is an associative k-algebra, and the functor $A\mapsto A_{\mathrm{res.Lie}}$ from the category of associative k-algebras to the category of restricted Lie k-algebras and the functor $L\mapsto\mathcal{U}_{\mathrm{res}}(L)$ from the category of restricted Lie k-algebras to the category of associative k-algebras are adjoint ([5, §12.1], [19, Ch. V, Th. 12]).

Given a restricted Lie k-algebra L containing X, we define a map $\widehat{P}_L=\widehat{P}_L^H\colon\mathbb{Z}_{\geq 0}\times H\to L$ by $\widehat{P}_L(j,w)=P_L(w)^{[p]^j}$, where $(\cdot)^{[p]^j}$ denotes applying j times the operation $(\cdot)^{[p]}$. In analogy with Proposition 3.1(b) we have the following:

Proposition 3.2. *The image of $\widehat{P}_L$ k-linearly spans the restricted Lie k-subalgebra of L generated by X.*

Proof. Let $\widehat{L}_0$ be the k-linear subspace of L spanned by $\mathrm{Im}\left(\widehat{P}_L\right)$. Let L_0 be the k-linear subspace of L spanned by $\mathrm{Im}(P_L)$. Clearly, $X \subseteq L_0 \subseteq \widehat{L}_0$. By Proposition 3.1(b), L_0 is the Lie k-subalgebra of L generated by X.

Since $\mathrm{char}\, k = p$, the binomial formula implies that the subspace $\widehat{L}_0$ is closed under $(\cdot)^{[p]}$.

If $w,u \in H$, then $[P_L(w),P_L(u)] \in L_0$. It follows from the k-bilinearity of the Lie bracket that for every $\alpha,\beta \in L_0$, also $[\alpha,\beta] \in L_0$. By induction on $m \geq 1$, the m-times iterated Lie brackets

$$[\alpha,_m,\beta] = [\alpha,[\alpha,[\dots[\alpha,\beta]\dots]]] \quad \text{and} \quad [\alpha,\beta,_m] = [\dots[[\alpha,\beta],\beta],\dots,\beta]$$

are also contained in L_0. Using the identities $\left[\alpha,\beta^{[p]}\right] = [\alpha,_p,\beta]$ and $\left[\alpha^{[p]},\beta\right] = [\alpha,\beta,_p]$ (see [5, p. 297]), we deduce that $\left[\alpha^{[p]^j},\beta^{[p]^r}\right] \in L_0$ for every $j,r \geq 0$. By the bilinearity again, $\widehat{L}_0$ is therefore closed under the Lie bracket.

Hence $\widehat{L}_0$ is the restricted Lie k-subalgebra of L generated by X. □

There is a *free restricted k-algebra* $\widehat{L}(X)$ on the generating set X, with the standard universal property. It is the restricted Lie k-subalgebra of $k\langle X\rangle_{\mathrm{res.Lie}}$ generated by X, and its restricted universal enveloping algebra is $k\langle X\rangle$ [15, Prop. 1.2.7]. We note that in the algebra $k\langle X\rangle$ one has

$$\widehat{P}_{\widehat{L}(X)}(j,w) = P_{L(X)}(w)^{p^j}$$

for every $j \geq 0$ and $w \in H$. The following analogue of Proposition 3.1(a) generalizes a result of Gärtner [15, Th. 1.2.11] (who considers a specific Hall family H):

Corollary 3.3. *The polynomials $\widehat{P}_{\widehat{L}(X)}(j,w)$, where $j \geq 0$ and $w \in H$, form a k-linear basis of $\widehat{L}(X)$.*

Proof. We consider $\widehat{L}(X)$ as a k-linear subspace of $k\langle X\rangle$. By Proposition 3.2, it is spanned by the powers $\widehat{P}_{\widehat{L}(X)}(j,w)$, where $j \geq 0$ and $w \in H$. As already observed, the products $\prod_{i=1}^m P_{L(X)}(w_i)$, with $w_1 \geq_{\mathrm{alp}} \cdots \geq_{\mathrm{alp}} w_m$ in H, form a k-linear basis of $k\langle X\rangle$. In particular, the powers $\widehat{P}_{\widehat{L}(X)}(j,w) = P_{L(X)}(w)^{p^j}$ are k-linearly independent. Hence they form a k-linear basis of $L(X)_{\mathrm{res}}$. □

We grade $L(X)$ and $\widehat{L}(X)$ by total degree, and write $L(X)_n$, $\widehat{L}(X)_n$ for their homogenous components of degree n.

Corollary 3.4. *Let n be a positive integer.*

(a) *The $P_{L(X)}(w)$, with $w \in H$ and $|w| = n$, form a k-linear basis of $L(X)_n$.*

(b) *The $\widehat{P}_{\widehat{L}(X)}(j,w)$, with $j \geq 0$ and $w \in H$ satisfying $n = |w|p^j$, form a k-linear basis of $\widehat{L}(X)_n$.*

Proof.

(a) This follows from Proposition 3.1(a), since $P_{L(X)}(w)$ has degree $|w|$ in $k\langle X\rangle$.

(b) This follows from Corollary 3.3, since $\widehat{P}_{\widehat{L}(X)}(j,w)$ has degree $|w|p^j$ in $k\langle X\rangle$. □

4. The p-Zassenhaus filtration

We fix as before a prime number p. For an integer $1 \le i \le n$, let $j_n(i) = \lceil \log_p(n/i) \rceil$ – that is, $j_n(i)$ is the least integer j such that $ip^j \ge n$.

Lemma 4.1. *The following conditions on $1 \le i \le n$ are equivalent:*

(a) $i'p^{j_n(i')} \ge ip^{j_n(i)}$ *for every* $1 \le i' \le i$.

(b) $i = \lceil n/p^k \rceil$ *for some* $k \ge 0$.

Proof. Set $i_k = \lceil n/p^k \rceil$. Thus $i_0 = n$, and the sequence i_k is weakly decreasing to 1. We may restrict ourselves to k such that $p^k \le n$. Then $(n/p^k) + 1 \le n/p^{k-1}$, so $n/p^k \le i_k < n/p^{k-1}$. Thus $j_n(i_k) = k$.

Since $n/p^k \le \lceil n/p^{k+1} \rceil p$, one has $i_k p^k \le i_{k+1} p^{k+1}$ – that is, the sequence $i_k p^{j_n(i_k)}$ is weakly increasing in the above range.

We also observe that if $i < i_{k-1}$, then $i < n/p^{k-1}$ – that is, $j_n(i) \ge k$.

(a) $\Rightarrow$(b): Since (b) certainly holds for $i = n$, we may assume that $i < n$, so there is k in the above range such that $i_k \le i < i_{k-1}$. By the previous observation, $j_n(i) \ge k$. We take in (a) $i' = i_k$ to obtain

$$i_k p^{j_n(i_k)} \ge ip^{j_n(i)} \ge i_k p^k = i_k p^{j_n(i_k)}.$$

Hence $i = i_k$.

(b) $\Rightarrow$(a): Suppose that $1 \le i' < i_k$. There exists l in the above range such that $i_l \le i' < i_{l-1}$. Necessarily, $l > k$, so $i_l p^l \ge i_k p^k$. As we have observed, $j_n(i') \ge l$. Hence

$$i'p^{j_n(i')} \ge i_l p^l \ge i_k p^k = i_k p^{j_n(i_k)}.$$ □

We define $J(n)$ to be the set of all $1 \le i \le n$ such that the equivalent conditions of Lemma 4.1 hold.

Remark 4.2.

(1) When $n \le p$, one has $J(n) = \{1,n\}$.

(2) Let $1 \ne i \in J(n)$ and take k such that $i = \lceil n/p^k \rceil$. By the first paragraph of the proof of Lemma 4.1, $j_n(i) = k$.

Now let G be a profinite group. Given closed subgroups K,K' of G and a positive integer m, we write $[K,K']$ (resp., K^m) for the closed subgroup of G generated by all commutators $[k,k'] = k^{-1}(k')^{-1}kk'$ (resp., powers k^m) with $k \in K$ and $k' \in K'$.

Recall that the (profinite) *lower central series* $G^{(i)}$, $i = 1,2,\dots,$ of G is defined inductively by $G^{(1)} = G$, $G^{(i+1)} = [G,G^{(i)}]$. As in the Introduction, we denote the *p-Zassenhaus filtration* of G by $G_{(n,p)}$, $n = 1,2,\dots$. Since $G^{(i)} \le G^{(n)}$ for $i > n$,

$$G_{(n,p)} = \prod_{ip^j \ge n} \left(G^{(i)}\right)^{p^j} = \prod_{i=1}^{n} \left(G^{(i)}\right)^{p^{j_n(i)}}.$$

The subgroups $G_{(n,p)}$ of G are characteristic, hence normal. We note that $G^{(n)} \le G_{(n,p)}$.

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

The Zassenhaus filtration can also be defined inductively by

$$G_{(1,p)} = G, \quad G_{(n,p)} = \left(G_{(\lceil n/p \rceil,p)}\right)^p \prod_{i+j=n} \left[G_{(i,p)}, G_{(j,p)}\right], \tag{4.1}$$

for $n \geq 2$. Indeed, this follows from a theorem of Lazard in the case of discrete groups ([5, Th. 11.2], [25, p. 209, Equation (3.14.5)]), and the profinite analog follows by a density argument. It follows from definition (4.1) that for $n \geq 2$,

$$G_{(np,p)} \leq \left(G_{(n,p)}\right)^p \left[G, G_{(n,p)}\right]. \tag{4.2}$$

Let $r \geq 0$. The following identity was proved in the discrete case by Shalev [36, Prop. 1.2]; the profinite analog follows again by a density argument:

$$\prod_{ip^j \geq n} \left(G^{(i+r+1)}\right)^{p^j} = \left[G, \prod_{ip^j \geq n} \left(G^{(i+r)}\right)^{p^j}\right].$$

In particular,

$$\prod_{ip^j \geq n} \left(G^{(i+1)}\right)^{p^j} = \left[G, G_{(n,p)}\right]. \tag{4.3}$$

Proposition 4.3. *Let $1 \leq i \leq n$ be an integer such that $i \notin J(n)$. Then*

$$\left(G^{(i)}\right)^{p^{j_n(i)}} \leq \left(G_{(n,p)}\right)^p \left[G, G_{(n,p)}\right].$$

Proof. As $i \notin J(n)$, there exists $1 \leq i' < i$ such that $ip^{j_n(i)} > i'p^{j_n(i')}$. We abbreviate $j = j_n(i)$ and $j' = j_n(i')$, so $ip^j, i'p^{j'} \geq n$.

If $j > j'$, then

$$\left(G^{(i)}\right)^{p^j} \leq \left(\left(G^{(i')}\right)^{p^{j'}}\right)^{p^{j-j'}} \leq \left(G_{(n,p)}\right)^p.$$

If $j \leq j'$, then the inequality $i > i'p^{j'-j}$ and equation (4.3) give

$$\left(G^{(i)}\right)^{p^j} \leq \left(G^{\left(i'p^{j'-j}+1\right)}\right)^{p^j} \leq \left[G, G_{(n,p)}\right]. \qquad \square$$

It follows from definition (4.1) that for every n the quotient $G_{(n,p)}/G_{(n+1,p)}$ is abelian of exponent dividing p. Consider the graded $\mathbb{F}_p$-module

$$\mathrm{gr}G = \bigoplus_{n \geq 0} G_{(n,p)}/G_{(n+1,p)}.$$

The commutator map and the p-power map induce on $\mathrm{gr}G$ the structure of a p-restricted Lie $\mathbb{F}_p$-algebra (see [5, §12.2], [15, Prop. 1.2.14]).

We now specialize to the case where S is a free profinite group on the basis X, in the sense of [14, §17.4]. It is the inverse limit of the free profinite groups on finite subsets of X [14, Lemma 17.4.9], so in our following results one may assume whenever convenient that X is actually finite, and use limit arguments for the general case.

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

By [15, Th. 1.3.8], there is a well-defined isomorphism $\mathrm{gr}S \xrightarrow{\sim} \widehat{L}(X)$ of graded restricted Lie algebras. Specifically, the coset of $x \in X$ in $\mathrm{gr}_1 S = S/S_{(2,p)}$ maps to x.

Let H be, as before, a fixed Hall set in X^*. For every word $w \in H$ we associate an element $\tau_w \in S$ as in [8]. Thus $\tau_{(x)} = x$ for $x \in X$, and for a word $w \in H$ of length $i > 1$ with standard factorization $w = uv$, where $u, v \in H$ (see §2), we set $\tau_w = [\tau_u, \tau_v]$. Then $\tau_w \in S^{(i)}$. Hence if $ip^j \geq n$, then $\tau_w^{p^j} \in \left(S^{(i)}\right)^{p^j} \leq S_{(n,p)}$.

Proposition 4.4. *Let $n \geq 1$. The cosets of the powers $\tau_w^{p^j}$, with $w \in H$ and $n = |w|p^j$, form an $\mathbb{F}_p$-linear basis of $S_{(n,p)}/S_{(n+1,p)}$.*

Proof. We use the terminology of §3 with the ground field $k = \mathbb{F}_p$. By induction on the structure of w, the isomorphism $\mathrm{gr}S \xrightarrow{\sim} \widehat{L}(X)$ of restricted Lie $\mathbb{F}_p$-algebras maps the coset of τ_w to $P_{L(X)}(w)$. Therefore it maps the coset of $\tau_w^{p^j}$ to $\widehat{P}_{\widehat{L}(X)}(j,w) = P_{L(X)}(w)^{p^j}$ considered as polynomials in $\mathbb{F}_p\langle X\rangle$. The assertion now follows from Corollary 3.4(b). $\square$

Remark 4.5. Vogel [37, Ch. I, §3] uses a specific Hall set H to give $\mathbb{F}_p$-linear bases of $S_{(n,p)}/S_{(n+1,p)}$ for $n = 2,3$, as well as generating sets for arbitrary n. Namely, for similarly defined basic commutators $c_w \in S^{(i)}$ of words $w \in H$ with $|w| = i$, the generating set consists of all $c_w^{p^j}$ with $n = ip^j$. Furthermore, according to [26, Cor. 3.12] the set of all such powers forms a basis of $S_{(n,p)}/S_{(n+1,p)}$, but the proof lacks details. I thank J. Mináč for a correspondence on the latter reference.

Let $n \geq 1$. For a word $w \in H$ of length $1 \leq i \leq n$ we abbreviate

$$\sigma_w = \tau_w^{p^{j_n(i)}}.$$

Thus $\sigma_w \in S_{(n,p)}$.

Theorem 4.6. *The cosets of σ_w, where $w \in H$ has length $i \in J(n)$, generate $S_{(n,p)}/\left(S_{(n,p)}\right)^p \left[S, S_{(n,p)}\right]$.*

Proof. Proposition 4.4 implies, by induction on $r \geq 1$, that $S_{(n,p)}/S_{(n+r,p)}$ is generated by the cosets of $\tau_w^{p^j}$, where $w \in H$ has length i, and $n \leq ip^j < n + r$. We apply this for $n + r = np$. By the inclusion (4.2), $S_{(np,p)} \leq \left(S_{(n,p)}\right)^p \left[S, S_{(n,p)}\right]$, and we deduce that $S_{(n,p)}/\left(S_{(n,p)}\right)^p \left[S, S_{(n,p)}\right]$ is generated by the cosets of $\tau_w^{p^j}$, where $w \in H$ has length i and $n \leq ip^j < np$. Moreover, it suffices to take such powers with $j = j_n(i)$, since otherwise $\tau_w^{p^j} \in \left(S_{(n,p)}\right)^p$. Finally, by Proposition 4.3, if $i \notin J(n)$ then the coset of $\sigma_w = \tau_w^{p^{j_n(i)}}$ is trivial. We are therefore left with the generators σ_w, as in the assertion. $\square$

5. The fundamental matrix

For a profinite ring R, let $R\langle\langle X\rangle\rangle^\times$ be the group of invertible elements in $R\langle\langle X\rangle\rangle$ (see §3). As before, let S be the free profinite group over the basis X. The *continuous Magnus homomorphism*

$$\Lambda = \Lambda_R \colon S \to R\langle\langle X\rangle\rangle^\times$$

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

is defined on the (profinite) generators $x \in X$ of S by $\Lambda(x) = 1 + x$ (see [7, §5] for details, and note that $1 + x$ is invertible by the geometric progression formula). For an arbitrary $\sigma \in S$ we write

$$\Lambda(\sigma) = \sum_{w \in X^*} \epsilon_{w,R}(\sigma) w,$$

with $\epsilon_{w,R}(\sigma) \in R$. The map $\epsilon_{w,R} \colon S \to R$ is continuous, and $\epsilon_{\emptyset,R}(\sigma) = 1$ for every σ (where $\emptyset$ denotes the empty word).

Let $\mathbb{U}_i(R)$ be the profinite group of all unitriangular $(i+1) \times (i+1)$-matrices over R. Given a word $w = (x_1 \cdots x_i) \in X^*$ of length i, we define a continuous map $\rho_w \colon S \to \mathbb{U}_i(R)$ by

$$\rho_w(\sigma) = \left(\epsilon_{(x_k x_{k+1} \cdots x_{l-1}),R}(\sigma)\right)_{1 \le k \le l \le i+1}.$$

The fact that Λ is a homomorphism implies that ρ_w is a homomorphism of profinite groups [7, Lemma 7.5]. We call it the *Magnus representation* of S corresponding to w.

The subgroup $S^{(n)}$ of S is characterized in terms of the Magnus map as the set of all $\sigma \in S$ such that $\epsilon_{w,\mathbb{Z}_p}(\sigma) = 0$ for every word w of length $1 \le i < n$ [8, Prop. 4.1(a)]. The following result gives similar restrictions on the Magnus coefficients of elements of $S_{(n,p)}$. In the discrete case it was proved in [2, Example 4.6], where it was further shown that these restrictions in fact characterize $S_{(n,p)}$. While it is possible to derive the proposition from the discrete case using a density argument, we provide a direct proof.

Proposition 5.1. *If $\sigma \in S_{(n,p)}$, then $\epsilon_{w,\mathbb{Z}_p}(\sigma) \in p^{j_n(i)}\mathbb{Z}_p$ for every word $w \in X^*$ of length $i \ge 1$.*

Proof. Consider the subset

$$I = \sum_{i \ge 1} \sum_{|w|=i} p^{j_n(i)} \mathbb{Z}_p w = \sum_{1 \le i \le n} \sum_{|w|=i} p^{j_n(i)} \mathbb{Z}_p w + \sum_{i > n} \sum_{|w|=i} \mathbb{Z}_p w$$

of $\mathbb{Z}_p\langle\langle X \rangle\rangle$. It is an ideal in $\mathbb{Z}_p\langle\langle X \rangle\rangle$, and therefore $1 + I$ is closed under multiplication. Moreover, the identity $\alpha^{-1} = 1 - \alpha^{-1}(\alpha - 1)$ shows that $1 + I$ is in fact a subgroup of $\mathbb{Z}_p\langle\langle X \rangle\rangle^\times$.

As $S_{(n,p)} = \prod_{i=1}^n \left(S^{(i)}\right)^{p^{j_n(i)}}$, it therefore suffices to show that $\Lambda_{\mathbb{Z}_p}\left(\tau^{p^{j_n(i)}}\right) \in 1 + I$ for every $\tau \in S^{(i)}$ with $1 \le i \le n$. We abbreviate $j = j_n(i)$. Then $\Lambda_{\mathbb{Z}_p}(\tau) = 1 + \sum_{|w| \ge i} \epsilon_{w,\mathbb{Z}_p}(\tau) w$, by [8, Prop. 4.1(a)]. For every $1 \le l \le p^j$ such that $il \le n$, one has $p^{j_n(il)} \mid \binom{p^j}{l}$ [2, Example 3.9]. Hence

$$\begin{aligned}
\Lambda_{\mathbb{Z}_p}\left(\tau^{p^j}\right) &= \sum_{0 \le l \le p^j} \binom{p^j}{l} \left(\sum_{|w| \ge i} \epsilon_{w,\mathbb{Z}_p}(\tau) w\right)^l \subseteq 1 + \sum_{1 \le l \le p^j} \binom{p^j}{l} \left(\sum_{|w| \ge i} \mathbb{Z}_p w\right)^l \\
&\subseteq 1 + \sum_{1 \le l \le p^j, il \le n} p^{j_n(il)} \left(\sum_{|w| \ge i} \mathbb{Z}_p w\right)^l + \sum_{1 \le l \le p^j, il > n} \left(\sum_{|w| \ge i} \mathbb{Z}_p w\right)^l \\
&\subseteq 1 + I,
\end{aligned}$$

as desired. □

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

As before, let H be a Hall set in X^*.

Corollary 5.2. *Let w,w' be nonempty words in X^* of lengths $1 \le i,i' \le n$, respectively, with $w' \in H$. Then $\epsilon_{w,\mathbb{Z}_p}(\sigma_{w'}) \in p^{j_n(i)}\mathbb{Z}_p$.*

For an integer $1 \le i \le n$, let

$$\pi_i \colon \mathbb{Z}_p \to \mathbb{Z}/p^{j_n(i)+1}$$

be the natural epimorphism. For words w,w' of lengths i, i', respectively, with $w' \in H$, we define

$$\langle w,w' \rangle_n = \pi_i\left(\epsilon_{w,\mathbb{Z}_p}(\sigma_{w'})\right).$$

By Corollary 5.2, $\langle w,w' \rangle_n \in p^{j_n(i)}\mathbb{Z}_p/p^{j_n(i)+1}\mathbb{Z}_p$. We identify the latter group with $\mathbb{Z}/p$, and thus view $\langle w,w' \rangle_n$ as an element of $\mathbb{Z}/p$.

Consider the (possibly infinite) *transposed* matrix

$$\left[\langle w,w' \rangle_n\right]^T_{w,w'}$$

over $\mathbb{Z}/p$, where w,w' range over all words in H of lengths in $J(n)$, and indexed with respect to the total order $\preceq$ on X^* defined in §2. We call it the *fundamental matrix of level n of H.*

We now focus on the Hall set of Lyndon words (see the Introduction). We record the following fundamental *triangularity property* of H [33, Th. 5.1]: For every Lyndon word $w \in X^*$, one has

$$\Lambda_{\mathbb{Z}_p}(\tau_w) = 1 + w + \text{ a combination of words strictly larger than } w \text{ in } \preceq. \tag{5.1}$$

Proposition 5.3. *Let H be the Hall set of all Lyndon words in X^*. The fundamental matrix of H of level n is unitriangular (i.e., unipotent and upper-triangular).*

Proof. Let w be a Lyndon word of length $i \le n$. By (5.1),

$$\Lambda_{\mathbb{Z}_p}(\sigma_w) = (1 + w + \cdots)^{p^{j_n(i)}} = 1 + p^{j_n(i)}w + \cdots,$$

where the remaining terms are multiples of words strictly larger than w in $\preceq$. Therefore $\langle w,w \rangle_n = \pi_i\left(p^{j_n(i)}\right) = 1$ in $\mathbb{Z}/p$.

Furthermore, for Lyndon words $w \prec w'$ we get $\epsilon_{w,\mathbb{Z}_p}(\sigma_{w'}) = 0$, whence $\langle w,w' \rangle_n = 0$ (note that the empty word is not Lyndon).

Consequently, the matrix $[\langle w,w' \rangle_n]_{w,w'}$ is unipotent *lower*-triangular, and therefore its transpose is unitriangular. □

Example 5.4. Suppose that $n = 2$. Then $J(n) = \{1,2\}$.

The Lyndon words of length ≤ 2 are the words $w = (x)$ and $w = (xy)$, with $x,y \in X$, $x < y$. Then σ_w is $\tau_w^{p^{j_2(1)}} = x^p$ and $\tau_w^{p^{j_2(2)}} = [x,y]$, respectively. In [8, §10] it is shown that the value of $\langle w,w' \rangle$, where w,w' are Lyndon words of lengths ≤ 2, is 1 if $w = w'$ and is 0 otherwise. Thus the fundamental matrix of level 2 for the Lyndon words is the identity matrix.

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

Example 5.5. Suppose that $n=3$. Then $J(n)=\{1,3\}$ for $p\geq 3$ and $J(3)=\{1,2,3\}$ for $p=2$.

The Lyndon words w of length 3 are of the forms

$$(xxy),(xyy),(xyz),(xzy),$$

where $x,y,z\in X$ and $x<y<z$. For these words we have

$$\sigma_{(xxy)}=[x,[x,y]],\quad \sigma_{(xyy)}=[[x,y],y],\quad \sigma_{(xyz)}=[x,[y,z]],\quad \sigma_{(xzy)}=[[x,z],y],$$

respectively. We recall that $\langle w,w\rangle_3=1$ for every w, and $\langle w,w'\rangle_3=0$ when $w\prec w'$. It remains to compute $\langle w,w'\rangle_3$ when $w'\prec w$.

If $|w|,|w'|\leq 2$, then by Example 5.4, $\langle w,w'\rangle_3=0$. We may therefore assume that $|w'|\leq |w|=3$.

If w contains a letter which does not appear in w', then $\epsilon_{w,\mathbb{Z}_p}(\sigma_{w'})=0$, whence $\langle w,w'\rangle_3=0$. Thus we may assume that every letter in w appears in w'.

When $w=(xyy)$ and $w'=(xxy)$, where $x<y$, the proof of [8, Prop. 11.2] gives

$$\langle w,w'\rangle_3=\epsilon_{(xyy)}([x,[x,y]])=0.$$

Similarly, when $w=(xzy)$ and $w'=(xyz)$, where $x<y<z$, the proof of [8, Prop. 11.2] gives

$$\langle w,w'\rangle_3=\epsilon_{(xzy),\mathbb{Z}_p}([x,[y,z]])=-1.$$

This covers all possible cases when $p\geq 3$. When $p=2$ we also need to consider Lyndon words $w'=(xy)$ of length 2, where $x<y$. Then $w=(xxy)$ or $w=(xyy)$. An explicit computation gives

$$\Lambda_{\mathbb{Z}_2}([x,y])=1+xy-yx+xyx-yxy-x^2y+y^2x+\cdots,$$

where the remaining terms are of degree ≥ 4. The square of this series has no terms (xxy) and (xyy), so $\epsilon_{(xxy),\mathbb{Z}_2}\left([x,y]^2\right)=\epsilon_{(xyy),\mathbb{Z}_2}\left([x,y]^2\right)=0$. Therefore $\langle(xxy),(xy)\rangle_3=\langle(xyy),(xy)\rangle_3=0$.

Altogether, we have shown that

$$\langle w,w'\rangle_3=\begin{cases}1 & \text{if } w=w',\\ -1 & \text{if } w=(xzy),w'=(xyz), \text{ where } x,y,z\in X,\ x<y<z,\\ 0 & \text{otherwise.}\end{cases}$$

In particular, the fundamental matrix need not be the identity matrix.

6. Unitriangular matrices

Let $i\geq 1$ and $j\geq 0$ be integers and consider the ring $R=\mathbb{Z}/p^{j+1}$. In this section we study the p-Zassenhaus filtration of the group $\mathbb{U}=\mathbb{U}_i(R)$ of all unitriangular $(i+1)\times(i+1)$-matrices over R, and in particular characterize the values of i,j for which $\mathbb{U}_{(n,p)}\cong\mathbb{Z}/p$ (see §5 for the notation).

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

We denote the unit matrix in $\mathbb{U}$ by I, and write $E_{1,i+1}$ for the matrix which is 1 at entry $(1,i+1)$ and is 0 elsewhere. For $i' \geq 1$, the subgroup $\mathbb{U}^{(i')}$ of $\mathbb{U}$ consists of all matrices in $\mathbb{U}$ which are zero on the first $i'-1$ diagonals above the main diagonal [1, Th. 1.5(i)].

We record the following fact about binomial coefficients:

Lemma 6.1. *Let t, j' be positive integers such that $1 \leq t \leq p^{j'}$. The following conditions are equivalent:*

(a) $p^j \mid \binom{p^{j'}}{l}$, $l = 1,2,\ldots,t$.

(b) $p^j \mid \binom{p^{j'}}{l}$ *for* $l = p^{\lfloor \log_p t \rfloor}$.

(c) $j' \geq j + \lfloor \log_p t \rfloor$.

Proof. (a)⇒(b) is trivial. For (a) ⇒(c) and (b) ⇒(c) see [9, Prop. 2.2(c)] and its proof. □

Proposition 6.2. *Let $1 \leq i' \leq i$ and $j' \geq 0$.*

(a) *One has* $\left(\mathbb{U}^{(i')}\right)^{p^{j'}} = \{I\}$ *if and only if* $j' \geq j + 1 + \lfloor \log_p(i/i') \rfloor$.

(b) *One has* $\left(\mathbb{U}^{(i')}\right)^{p^{j'}} = I + p^j \mathbb{Z} E_{1,i+1}$ *if and only if* $j' = j + \log_p(i/i')$ *(in particular, i/i' is a p-power).*

(c) *One has* $\left(\mathbb{U}^{(i')}\right)^{p^{j'}} \leq I + p^j \mathbb{Z} E_{1,i+1}$ *if and only if* $j' \geq j + \log_p(i/i')$.

Proof. Let N be an $(i+1)\times(i+1)$-matrix over $\mathbb{Z}/p^{j+1}$ such that $I + N \in \mathbb{U}^{(i')}$. Then $N^l = 0$ for every integer l with $i/i' < l$. Hence

$$(I+N)^{p^{j'}} = \sum_{l=0}^{p^{j'}} \binom{p^{j'}}{l} N^l = \sum_{l=0}^{\min\left(p^{j'}, \lfloor i/i' \rfloor\right)} \binom{p^{j'}}{l} N^l.$$

Further, if $i' \mid i$, then $N^{i/i'} \in \mathbb{Z} E_{1,i+1}$.

In particular, let M be the $(i+1)\times(i+1)$-matrix over $\mathbb{Z}/p^{j+1}$ which is 1 on the (first) super-diagonal and is 0 elsewhere. Then the matrix $M^{i'l}$ is 1 on the $i'l$th diagonal above the main one and is 0 elsewhere. In particular, $I + M^{i'} \in \mathbb{U}^{(i')}$. By what we have justed noted,

$$\left(1 + M^{i'}\right)^{p^{j'}} = \sum_{l=0}^{\min\left(p^{j'}, \lfloor i/i' \rfloor\right)} \binom{p^{j'}}{l} M^{i'l}.$$

This matrix is $\binom{p^{j'}}{l}$ on the $i'l$th diagonals above the main one and is 0 elsewhere.

(a) By the previous observations, $\left(\mathbb{U}^{(i')}\right)^{p^{j'}} = \{I\}$ holds if and only if

$$p^{j+1} \mid \binom{p^{j'}}{l}, \quad l = 1,2,\ldots,\min\left(p^{j'}, \lfloor i/i' \rfloor\right).$$

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

In light of Lemma 6.1, this is equivalent to $j' \geq j+1+\min\left(j', \lfloor \log_p \lfloor i/i' \rfloor \rfloor\right)$, and it remains to note that $\lfloor \log_p \lfloor i/i' \rfloor \rfloor = \lfloor \log_p (i/i') \rfloor$.

(b) First assume that $i = i'$. Then $\mathbb{U}^{(i')} = I + \mathbb{Z}E_{1,i+1}$. Hence $\left(\mathbb{U}^{(i')}\right)^{p^{j'}} = I + \mathbb{Z}p^{j'}E_{1,i+1}$, and the equality $\left(\mathbb{U}^{(i')}\right)^{p^{j'}} = I + \mathbb{Z}p^j E_{1,i+1}$ means that $j' = j$, as desired.

Next we assume that $i > i'$. By the previous observations, $\left(\mathbb{U}^{(i')}\right)^{p^{j'}} = I + \mathbb{Z}p^j E_{1,i+1}$ holds if and only if the following conditions hold:

(i) i/i' is an integer $\leq p^{j'}$;

(ii) $p^{j+1} \mid \binom{p^{j'}}{l}$, $l = 1,2,\ldots,(i/i')-1$;

(iii) $p^j \mid \binom{p^{j'}}{i/i'}$, $p^{j+1} \nmid \binom{p^{j'}}{i/i'}$.

By Lemma 6.1 again, (i)–(iii) mean that i/i' is an integer $\leq p^{j'}$, and

$$\begin{aligned} j' &\geq j+1+\lfloor \log_p((i/i')-1) \rfloor \\ j' &\geq j+\lfloor \log_p(i/i') \rfloor \\ j' &< j+1+\lfloor \log_p(i/i')) \rfloor . \end{aligned}$$

This amounts to saying that $j' = j + \log_p(i/i')$.

(c) This follows from (a) and (b). □

The case $i' = 1$ of Proposition 6.2(a) was shown by Sawin (see [9, Prop. 2.3]).

The following corollary stands behind our definition of the sets $J(n)$. In the case $i = n$ it was proved by Mináč, Rogelstad and Tân [26, Cor. 3.7].

Corollary 6.3. *Suppose that $1 \leq i \leq n$ and $j = j_n(i)$. One has $\mathbb{U}_{(n,p)} = I + p^{j_n(i)}\mathbb{Z}E_{1,i+1}$ if and only if $i \in J(n)$.*

Proof. Recall that $\mathbb{U}_{(n,p)} = \prod_{i'=1}^n \left(\mathbb{U}^{(i')}\right)^{p^{j_n(i')}}$.

If $i' > i$, then $\mathbb{U}^{(i')} = \{I\}$, whence $\left(\mathbb{U}^{(i')}\right)^{p^{j_n(i')}} = \{I\}$.

Taking in Proposition 6.2(b), $i' = i$, and $j' = j = j_n(i)$, we obtain that $\left(\mathbb{U}^{(i)}\right)^{p^{j_n(i)}} = I + p^{j_n(i)}\mathbb{Z}E_{1,i+1}$.

Therefore, $\mathbb{U}_{(n,p)} = I + p^{j_n(i)}\mathbb{Z}E_{1,i+1}$ holds if and only if for every $1 \leq i' \leq i$ one has $\left(\mathbb{U}^{(i')}\right)^{p^{j_n(i')}} \leq I + p^{j_n(i)}\mathbb{Z}E_{1,i+1}$. By Proposition 6.2(c), this inclusion is equivalent to $i'p^{j_n(i')} \geq ip^{j_n(i)}$. □

Thus, for $i \in J(n)$ and $\mathbb{U} = \mathbb{U}_i\left(\mathbb{Z}/p^{j_n(i)+1}\right)$ there is a central extension

$$0 \to \mathbb{U}_{(n,p)}(\cong \mathbb{Z}/p) \to \mathbb{U} \to \overline{\mathbb{U}} := \mathbb{U}/\mathbb{U}_{(n,p)} \to 1, \tag{6.1}$$

where the isomorphism is the projection on the $(1,i+1)$-entry composed with the isomorphism $p^{j(i)}\mathbb{Z}/p^{j(i)+1}\mathbb{Z} \cong \mathbb{Z}/p$.

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

7. The cohomology elements $\alpha_{w,n}$

Let S be again a free profinite group on the basis X, and let $n \geq 2$. Consider the transgression homomorphism $\mathrm{trg}\colon H^1\left(S_{(n,p)}\right)^S \to H^2\left(S/S_{(n,p)}\right)$ (recall that the cohomology groups are with respect to the coefficient module $\mathbb{Z}/p$ with trivial action). It is the differential d_2^{01} in the Lyndon–Hochschild–Serre spectral sequence corresponding to the closed normal subgroup $S_{(n,p)}$ of S [31, Th. 2.4.3]. From the five-term sequence in profinite cohomology [31, Prop. 1.6.7] and the fact that S has cohomological dimension 1, it follows that trg is an isomorphism.

Now consider a word w of length $i \in J(n)$. Consider the ring $R_i = \mathbb{Z}/p^{j_n(i)+1}$, and set $\mathbb{U} = \mathbb{U}_i(R_i)$. As before, let $\overline{\mathbb{U}} = \mathbb{U}/\mathbb{U}_{(n,p)}$. By Corollary 6.3, the projection on the $(1,i+1)$-entry gives an isomorphism

$$\mathbb{U}_{(n,p)} \xrightarrow{\sim} p^{j_n(i)}\mathbb{Z}/p^{j_n(i)+1}\mathbb{Z}.$$

The Magnus representation $\rho = \rho_w\colon S \to \mathbb{U}$ induces continuous homomorphisms

$$\bar{\rho}_w\colon S/S_{(n,p)} \to \overline{\mathbb{U}}, \quad \rho_w^0 = \rho|_{S_{(n,p)}}\colon S_{(n,p)} \to \mathbb{U}_{(n,p)}.$$

Let $\bar{\rho}_w^*\colon H^2\left(\overline{\mathbb{U}}\right) \to H^2\left(S/S_{(n,p)}\right)$ be the pullback of $\bar{\rho}_w$.

Let $\gamma = \gamma_{n,R_i} \in H^2\left(\overline{\mathbb{U}}\right)$ correspond to the extension (6.1) under the Schreier correspondence [31, Th. 1.2.4]. We set

$$\alpha_{w,n} = \bar{\rho}_w^*(\gamma) \in H^2\left(S/S_{(n,p)}\right).$$

Example 7.1 *$\alpha_{w,n}$ for a word $w = (x)$ of length 1.* Let $j = j_n(1) = \lceil \log_p n \rceil$, so $\mathbb{U} = \mathbb{U}_1\left(\mathbb{Z}/p^{j+1}\right) \cong \mathbb{Z}/p^{j+1}$. As $1 \in J(n)$, we have $\mathbb{U}_{(n,p)} \cong \mathbb{Z}/p$, and the central extension (6.1) becomes

$$0 \to \mathbb{Z}/p \to \mathbb{Z}/p^{j+1} \to \mathbb{Z}/p^j \to 0. \tag{7.1}$$

We consider this extension as a sequence of trivial $S/S_{(n,p)}$-modules. The *Bockstein homomorphism*

$$\mathrm{Bock}_{p^j,S/S_{(n,p)}}\colon H^1\left(S/S_{(n,p)},\mathbb{Z}/p^j\right) \to H^2\left(S/S_{(n,p)}\right)$$

is the associated connecting homomorphism.

We may identify $\rho_{(x)}\colon S \to \mathbb{U}$ with $\epsilon_{(x),\mathbb{Z}/p^{j+1}}\colon S \to \mathbb{Z}/p^{j+1}$, and $\bar{\rho}_{(x)}\colon S/S_{(n,p)} \to \overline{\mathbb{U}}$ with $\epsilon_{(x),\mathbb{Z}/p^j}\colon S/S_{(n,p)} \to \mathbb{Z}/p^j$, which are both continuous homomorphisms. Thus $\alpha_{(x),n}$ corresponds to the pullback of the extension (7.1) under $\epsilon_{(x),\mathbb{Z}/p^j}$. By [8, Remark 7.3],

$$\alpha_{(x),n} = \mathrm{Bock}_{p^j,S/S_{(n,p)}}\left(\epsilon_{(x),\mathbb{Z}/p^j}\right).$$

For the next Example, we first recall a few facts about Massey products. While these products are defined in the general context of differential graded algebras, in the special case of the n-fold Massey product $H^1(G,R)^n \to H^2(G,R)$ in profinite (or discrete) group cohomology it can be alternatively described in terms of unitriangular representations. This was discovered by Dwyer [6] in the discrete case, and we refer to [7, §8] for the profinite case, which is considered here. We assume as before that $n \geq 2$ and R is a finite commutative ring on which G acts trivially (see [38] for the case of a nontrivial action).

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

Specifically, let $\mathbb{U} = \mathbb{U}_n(R)$ and let $\overline{\mathbb{U}}$ be again the quotient of $\mathbb{U}$ by the central subgroup $I + RE_{1,n+1}\,(\cong R^+)$. The central extension

$$0 \to R^+ \to \mathbb{U} \to \overline{\mathbb{U}} \to 1 \tag{7.2}$$

of trivial G-modules corresponds to a cohomology element $\gamma_R \in H^2(G,R^+)$. Given $\psi_1,\dots,\psi_n \in H^1(G,R^+)$, we consider the continuous homomorphisms $\bar\rho\colon G \to \overline{\mathbb{U}}$ whose projection $\bar\rho_{k,k+1}\colon G \to R$ on the $(k,k+1)$-entry is ψ_k, for $k = 1,2,\dots,n$. As before, let $\bar\rho^*\colon H^2\left(\overline{\mathbb{U}},R^+\right) \to H^2(G,R^+)$ be the pullback of $\bar\rho$. Then $\bar\rho^*(\gamma_R)$ corresponds to the central extension

$$0 \to R^+ \to \mathbb{U} \times_{\overline{\mathbb{U}}} G \to G \to 1,$$

where the fiber product is with respect to the natural projection $\mathbb{U} \to \overline{\mathbb{U}}$ and to $\bar\rho$. The *n-fold Massey product* $\langle\psi_1,\dots,\psi_n\rangle$ is the subset of $H^2(G,R^+)$ consisting of all pullbacks $\bar\rho^*(\gamma_R)$ [7, Prop. 8.3]. Thus the n-fold Massey product $\langle\cdot,\dots,\cdot\rangle\colon H^1(G,R^+)^n \to H^2(G,R^+)$ is a *multivalued map*. In the special case $n = 2$, one has $\langle\psi_1,\psi_2\rangle = \{\psi_1 \cup \psi_2\}$.

Example 7.2 *$\alpha_{w,n}$ for a word w of length $n \ge 2$.* Since $j_n(n) = 0$ we have $R_n = \mathbb{Z}/p$, so $\mathbb{U} = \mathbb{U}_n(\mathbb{Z}/p)$. As $n \in J(n)$, Corollary 6.3 shows that $\mathbb{U}_{(n,p)} = I + \mathbb{Z}E_{1,n+1} \cong \mathbb{Z}/p$. Thus the extension (7.2) (for $R = \mathbb{Z}/p$) coincides with the extension (6.1) with $i = n$.

Now take a word $w = (x_1 \cdots x_n) \in X^*$ of length n. Let $\bar\rho = \bar\rho_w\colon S/S_{(n,p)} \to \overline{\mathbb{U}}$ and let $\bar\rho_{k,k+1}$ be homomorphisms as before. By its definition as the pullback of the extension (6.1), $\alpha_{w,n}$ is an element of the n-fold Massey product $\langle\rho_{12},\rho_{23},\dots,\rho_{n,n+1}\rangle$ in $H^2\left(S/S_{(n,p)}\right)$. Note that $\bar\rho_{k,k+1}$ is given by $\bar\rho_{k,k+1}(x_l) = \delta_{kl}$ for every $1 \le k,l \le n$.

8. The Lyndon bases

We continue with the setup of §7. Identifying $H^1\left(S_{(n,p)}\right) = \mathrm{Hom}\left(S_{(n,p)},\mathbb{Z}/p\right)$, we obtain a nondegenerate bilinear map

$$S_{(n,p)}/\left(S_{(n,p)}\right)^p\left[S,S_{(n,p)}\right] \times H^1\left(S_{(n,p)}\right)^S \to \mathbb{Z}/p, \quad (\bar\sigma,\varphi) \mapsto \varphi(\sigma)$$

[11, Cor. 2.2]. It gives rise to the bilinear *transgression pairing*

$$\begin{gathered}(\cdot,\cdot)_n\colon S_{(n,p)}/\left(S_{(n,p)}\right)^p\left[S,S_{(n,p)}\right] \times H^2\left(S/S_{(n,p)}\right) \to \mathbb{Z}/p,\\ (\bar\sigma,\alpha)_n = -\left(\mathrm{trg}^{-1}\alpha\right)(\sigma),\end{gathered} \tag{8.1}$$

where $\bar\sigma$ denotes the coset of $\sigma \in S_{(n,p)}$. It is therefore also nondegenerate.

By Proposition 5.1 and Corollary 6.3, for a word w of length $i \in J(n)$ there is a commutative diagram

$$\begin{array}{ccc} S_{(n,p)} & \xrightarrow{\epsilon_{w,\mathbb{Z}_p}} & p^{j_n(i)}\mathbb{Z}_p \\ {\scriptstyle\rho_w^0}\downarrow & & \downarrow{\scriptstyle\pi_i} \\ \mathbb{U}_{(n,p)} & \xrightarrow{\sim} & p^{j_n(i)}\mathbb{Z}/p^{j_n(i)+1}\mathbb{Z}, \end{array} \tag{8.2}$$

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

where, as before, $\pi_i \colon \mathbb{Z}_p \to \mathbb{Z}/p^{j_n(i)+1}$ is the natural projection, and the lower isomorphism is the projection on the $(1,i+1)$-entry. We deduce the following link between cohomology and the Magnus map. As before, we identify $p^{j_n(i)}\mathbb{Z}/p^{j_n(i)+1}\mathbb{Z}$ with $\mathbb{Z}/p$.

Proposition 8.1. *For $\sigma \in S_{(n,p)}$ and a word $w \in X^*$ of length $i \in J(n)$, one has $(\bar{\sigma}, \alpha_{w,n})_n = \pi_i\left(\epsilon_{w,\mathbb{Z}_p}(\sigma)\right)$.*

Proof. The central extension (6.1) gives rise to a transgression homomorphism $\mathrm{trg}\colon H^1\left(\mathbb{U}^{(n,p)}\right)^{\mathbb{U}} \to H^2\left(\overline{\mathbb{U}}\right)$. Let $\iota\colon \mathbb{U}_{(n,p)} \xrightarrow{\sim} \mathbb{Z}/p$ be the composition of the lower row in diagram (8.2) with the isomorphism $p^{j_n(i)}\mathbb{Z}/p^{j_n(i)+1}\mathbb{Z} \cong \mathbb{Z}/p$. By the results of [8, §7],

$$\gamma = -\,\mathrm{trg}(\iota).$$

The functoriality of transgression gives a commutative square

$$\begin{array}{ccc} H^1\left(\mathbb{U}_{(n,p)}\right)^{\mathbb{U}} & \xrightarrow{\mathrm{trg}} & H^2\left(\overline{\mathbb{U}}\right) \\ \Big\downarrow{\scriptstyle (\rho_w^0)^*} & & \Big\downarrow{\scriptstyle \bar{\rho}_w^*} \\ H^1\left(S_{(n,p)}\right)^{S} & \xrightarrow{\mathrm{trg}} & H^2\left(S/S_{(n,p)}\right). \end{array}$$

As $\sigma \in S_{(n,p)}$, this square and diagram (8.2) give

$$\begin{aligned} (\bar{\sigma}, \alpha_{w,n})_n &= (\bar{\sigma}, \bar{\rho}_w^*(\gamma))_n = -\,(\bar{\sigma}, \bar{\rho}_w^*(\mathrm{trg}(\iota)))_n = -\left(\bar{\sigma}, \mathrm{trg}\left(\left(\rho_w^0\right)^*(\iota)\right)\right)_n \\ &= \left(\left(\rho_w^0\right)^*(\iota)\right)(\sigma) = \iota\left(\rho_w^0(\sigma)\right) = \pi_i\left(\epsilon_{w,\mathbb{Z}_p}(\sigma)\right). \end{aligned}$$

□

Now consider words $w, w' \in X^*$ of lengths $i, i' \in J(n)$, respectively, with w' Lyndon. We deduce from Proposition 8.1 that

$$(\bar{\sigma}_{w'}, \alpha_{w,n})_n = \pi_i\left(\epsilon_{w,\mathbb{Z}_p}(\sigma_{w'})\right) = \langle w, w'\rangle_n.$$

We can therefore restate Proposition 5.3 cohomologically:

Corollary 8.2. *The transposed matrix $\left[(\bar{\sigma}_{w'}, \alpha_{w,n})_n\right]^T_{w,w'}$, where w, w' range over all Lyndon words in X^* of lengths i, i', respectively, in $J(n)$, and totally ordered by $\preceq$, coincides with the fundamental matrix of level n of the Lyndon words. In particular, it is unitriangular, whence invertible.*

Example 8.3. Let $n = 2$. Then $J(n) = \{1,2\}$. For every $x \in X$ let $\epsilon_x \in H^1\left(S/S_{(2,p)}\right)$ be the homomorphism induced by $\epsilon_{(x),\mathbb{Z}/p}$. It is 1 on the coset of x and is 0 on the coset of any $x' \in X$, $x' \neq x$.

For a one-letter word $w = (x)$ (which is always Lyndon) we have $\sigma_w = \tau_w^p = x^p$ and $\alpha_{w,2} = \mathrm{Bock}_{p,S/S_{(2,p)}}(\epsilon_x)$ (Example 7.1).

For a two-letter Lyndon word $w = (xy)$, $x < y$, the projections of the representation $\bar{\rho}_w$ on the (1,2)- and (2,3)-entries are $\bar{\rho}_{12} = \epsilon_x$ and $\bar{\rho}_{23} = \epsilon_y$. Thus $\sigma_w = \tau_w = [x,y]$, and $\alpha_{w,2} = \epsilon_x \cup \epsilon_y$ (Example 7.2).

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

Recall that the fundamental matrix for Lyndon words and for $n=2$ is the identity matrix (Example 5.4). Thus we recover the fundamental duality, discovered by Labute, between Bockstein elements/cup products and pth powers/commutators, respectively ([22, Prop. 8], [23, §2], [31, Ch. III, §9]).

We will need the following elementary fact in linear algebra [8, Lemma 8.4]:

Lemma 8.4. *Let R be a commutative ring and let $(\cdot,\cdot)\colon A\times B\to R$ be a nondegenerate bilinear map of R-modules. Let $(I,\le)$ be a finite totally ordered set, and for every $w\in I$ let $a_w\in A$, $b_w\in B$. Suppose that the matrix $\left[(a_w,b_{w'})\right]_{w,w'\in I}$ is invertible, and that a_w, $w\in I$, generate A. Then a_w, $w\in I$, is an R-linear basis of A, and b_w, $w\in I$, is an R-linear basis of B.*

We now deduce our first main result. Note that part (a) of the theorem strengthens Theorem 4.6 in the special case where H is the Hall set of Lyndon words.

Theorem 8.5.

(a) *The $\mathbb{F}_p$-linear space $S_{(n,p)}/\left(S_{(n,p)}\right)^p\left[S,S_{(n,p)}\right]$ has a basis consisting of the cosets $\bar\sigma_w$ of σ_w, where w is a Lyndon word in X^* of length $i\in J(n)$.*

(b) *The $\mathbb{F}_p$-linear space $H^2\left(S/S_{(n,p)}\right)$ has a basis consisting of all $\alpha_{w,n}$, where w is a Lyndon word in X^* of length $i\in J(n)$.*

Proof. First assume that X is finite. By Theorem 4.6, the cosets in (a) generate $S_{(n,p)}/\left(S_{(n,p)}\right)^p\left[S,S_{(n,p)}\right]$. Furthermore, the bilinear map $(\cdot,\cdot)_n$ of (8.1) is nondegenerate, and the fundamental matrix $\left[(\bar\sigma_{w'},\alpha_{w,n})_n\right]_{w,w'}$ is invertible, by Corollary 8.2. Therefore Lemma 8.4 implies both assertions.

The case of general X follows from the finite case by a standard limit argument (see [31, Prop. 1.2.5]). □

When $2\le n\le p$ we have $J(n)=\{1,n\}$ (Remark 4.2(1)), $j_n(1)=1$, and $j_n(n)=0$. In view of Examples 7.1 and 7.2, we deduce the following:

Corollary 8.6. *Suppose that $2\le n\le p$.*

(a) *The $\mathbb{F}_p$-linear space $S_{(n,p)}/\left(S_{(n,p)}\right)^p\left[S,S_{(n,p)}\right]$ has a basis consisting of:*
   (i) *the cosets of x^p, $x\in X$, and*
   (ii) *the cosets of τ_w, where w is a Lyndon word in X^* of length n.*

(b) *The $\mathbb{F}_p$-linear space $H^2\left(S/S_{(n,p)}\right)$ has a basis consisting of:*
   (i) *the Bockstein elements $\mathrm{Bock}_{p,S/S_{(n,p)}}\left(\epsilon_{(x),\mathbb{Z}/p}\right)=\alpha_{(x),n}$, $x\in X$, and*
   (ii) *the n-fold Massey product elements $\alpha_{w,n}$, where w is a Lyndon word in X^* of length n.*

The number of words of a given length in a Hall set H can be expressed in terms of *Witt's necklace function*, defined for integers $i,m\ge1$ by

$$\varphi_i(m)=\frac{1}{i}\sum_{d\mid i}\mu(d)m^{i/d}.$$

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

Here μ is the Möbius function – that is, $\mu(d) = (-1)^k$ if d is a product of k distinct prime numbers, and $\mu(d) = 0$ otherwise. We also set $\varphi_i(\infty) = \infty$. Then the number of words of length i in H is $\varphi_i(|X|)$ [33, Cor. 4.14]. We deduce the following:

Corollary 8.7.

(a) *For every* $n \geq 2$, *one has*

$$\dim_{\mathbb{F}_p} H^2\left(S/S_{(n,p)}\right) = \sum_{i \in J(n)} \varphi_i(|X|).$$

(b) *If* $2 \leq n \leq p$, *then* $\dim_{\mathbb{F}_p} H^2\left(S/S_{(n,p)}\right) = |X| + \varphi_n(|X|)$.

9. Shuffle relations

Recall that the shuffle product $u ш v$ of words u, v was defined in the Introduction. It extends naturally to a bilinear, commutative, and associative product map $ш\colon \mathbb{Z}\langle X\rangle \times \mathbb{Z}\langle X\rangle \to \mathbb{Z}\langle X\rangle$. The *shuffle algebra* $\mathrm{Sh}(X)$ on X is the graded $\mathbb{Z}$-algebra whose underlying module is the free module on X^* (graded by the length of words), and its multiplication is ш.

We define the *infiltration product* $u \downarrow v$ of words $u = (x_1 \cdots x_r)$, $v = (x_{r+1} \cdots x_{r+t})$ in X^* as follows (see [4], [33, pp. 134–135]). Consider all maps $\sigma\colon \{1,2,\dots,r+t\} \to \{1,2,\dots,r+t\}$ with $\sigma(1) < \cdots < \sigma(r)$ and $\sigma(r+1) < \cdots < \sigma(r+t)$, and which satisfy the following weak form of injectivity: If $\sigma(i) = \sigma(j)$, then $x_i = x_j$. Let the image of σ consist of $l_1 < \cdots < l_{m(\sigma)}$. Then we set

$$u \downarrow v = \sum_{\sigma} \left(x_{\sigma^{-1}(l_1)} \cdots x_{\sigma^{-1}\left(l_{m(\sigma)}\right)} \right) \in \mathbb{Z}\langle X\rangle. \tag{9.1}$$

By our assumption, $x_{\sigma^{-1}(l_i)}$ does not depend on the choice of the preimages $\sigma^{-1}(l_i)$ of l_i. We also write $\mathrm{Infil}(u,v)$ for the set of all such words $\left(x_{\sigma^{-1}(l_1)} \cdots x_{\sigma^{-1}\left(l_{m(\sigma)}\right)}\right)$. Thus $u ш v$ is the part of $u \downarrow v$ of degree $r+t$ – that is, the partial sum corresponding to all such maps σ which in addition are bijective. The product $\downarrow$ on words extends by linearity to an associative and commutative bilinear map on $\mathbb{Z}\langle X\rangle$.

There is a well-defined $\mathbb{Z}_p$-bilinear map

$$(\cdot,\cdot)\colon \mathbb{Z}_p\langle\langle X\rangle\rangle \times \mathbb{Z}_p\langle X\rangle \to \mathbb{Z}_p, \quad (f,g) = \sum_{w \in X^*} f_w g_w,$$

where f_w, g_w are the coefficients of f, g, respectively, at w [33, p. 17].

The following connection between the Magnus representation and the infiltration product is proved in the discrete case in [4, Th. 3.6]. We refer to [37, Prop. 2.25] and [30, Prop. 8.16] for the profinite case. Here we view the infiltration and shuffle products as elements of $\mathbb{Z}\langle X\rangle \subseteq \mathbb{Z}_p\langle X\rangle$.

Proposition 9.1. *For every* $\emptyset \neq u,v \in X^*$ *and every* $\sigma \in S$, *one has*

$$\epsilon_{u,\mathbb{Z}_p}(\sigma)\epsilon_{v,\mathbb{Z}_p}(\sigma) = \left(\Lambda_{\mathbb{Z}_p}(\sigma), u \downarrow v\right).$$

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

Corollary 9.2. *Let u,v be nonempty words in X^* with $i=|u|+|v|\leq n$. For every $\sigma\in S_{(n,p)}$, one has* $\big(\Lambda_{\mathbb{Z}_p}(\sigma),u ⧢ v\big)\in p^{j_n(i-1)}\mathbb{Z}_p$.

Proof. Let w be a word of length $1\leq k\leq i-1$. Then $j_n(k)\geq j_n(i-1)$, so by Proposition 5.1, $\epsilon_{w,\mathbb{Z}_p}(\sigma)\in p^{j_n(k)}\mathbb{Z}_p\subseteq p^{j_n(i-1)}\mathbb{Z}_p$. In particular, this is the case for $w=u$, $w=v$, and for $w\in \mathrm{Infil}(u,v)$ of length smaller than i. Since $u ⧢ v$ is the part of $u\downarrow v$ consisting of summands of maximal length i, Proposition 9.1 implies that $\big(\Lambda_{\mathbb{Z}_p}(\sigma),u ⧢ v\big)\in p^{j_n(i-1)}\mathbb{Z}_p$. □

We obtain the following *shuffle relations*. Here X^i stands for the set of words in X^* of length i.

Theorem 9.3. *Let $\emptyset\neq u,v\in X^*$ with $i=|u|+|v|\in J(n)$. Then*

$$\sum_{w\in X^i}(u ⧢ v)_w\alpha_{w,n}=0.$$

Proof. As $2\leq i\in J(n)$, we have $(i-1)p^{j_n(i-1)}\geq ip^{j_n(i)}$, whence $j_n(i-1)>j_n(i)$.

We recall that $u ⧢ v$ is homogenous of degree i. For $\sigma\in S_{(n,p)}$, Corollary 9.2 gives

$$\begin{aligned}\sum_{w\in X^i}(u ⧢ v)_w\epsilon_{w,\mathbb{Z}_p}(\sigma)=\sum_{w\in X^*}(u ⧢ v)_w\epsilon_{w,\mathbb{Z}_p}(\sigma)&=(\Lambda_{\mathbb{Z}_p}(\sigma),u ⧢ v)\\ &\in p^{j_n(i-1)}\mathbb{Z}_p\subseteq p^{j_n(i)+1}\mathbb{Z}_p.\end{aligned}$$

Therefore, by Proposition 8.1,

$$\begin{aligned}\left(\bar\sigma,\sum_{w\in X^i}(u ⧢ v)_w\alpha_{w,n}\right)_n&=\sum_{w\in X^i}(u ⧢ v)_w\left(\bar\sigma,\alpha_{w,n}\right)_n=\sum_{w\in X^i}(u ⧢ v)_w\pi_i\left(\epsilon_{w,\mathbb{Z}_p}(\sigma)\right)\\ &=\pi_i\left(\sum_{w\in X^i}(u ⧢ v)_w\epsilon_{w,\mathbb{Z}_p}(\sigma)\right)=0.\end{aligned}$$

Now use the fact that $(\cdot,\cdot)_n$ is nondegenerate. □

Given a graded R-algebra $A=\bigoplus_{i\geq0}A_i$, we denote $A_+=\bigoplus_{i\geq1}A_i$. Let $\mathrm{WD}(A)$ be the R-submodule of A generated by all products aa', where $a,a'\in A_+$. We call $\mathrm{WD}(A)$ the submodule of *weakly decomposable elements* of A. It is also generated by all products aa', where $a,a'\in A_+$ are homogenous. Hence the quotient $A_{\mathrm{indec}}=A/\mathrm{WD}(A)$ has the structure of a graded R-module, which we call the *indecomposable quotient* of A.

Note that $\mathrm{WD}(A)_0=\mathrm{WD}(A)_1=\{0\}$, so the graded module morphism $A\to A_{\mathrm{indec}}$ is an isomorphism in degrees 0 and 1. For example, when $A=R\langle X\rangle$, one has $A_{\mathrm{indec},0}=R$, $A_{\mathrm{indec},1}$ is the free R-module on the basis X, and $A_{\mathrm{indec},i}=0$ for all $i\geq2$.

When $A=\mathrm{Sh}(X)$ is the shuffle algebra, we recover the module $\mathrm{Sh}(X)_{\mathrm{indec},n}$ as defined in the Introduction. The following key fact was proved in [9, Prop. 6.3]. It is based on a construction by Radford [32] and Perrin and Viennot of a basis of $\mathbb{Z}\langle X\rangle$, which arises from the decomposition of words in X^* into Lyndon words.

Proposition 9.4. *Suppose that $1\leq n<p$. Then the images of the Lyndon words of length n span $\mathrm{Sh}(X)_{\mathrm{indec},n}\otimes(\mathbb{Z}/p)$ as an $\mathbb{F}_p$-linear space.*

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

In fact, in [9, Th. 7.3(b)] it is proved that these images form a linear basis of $\mathrm{Sh}(X)_{\mathrm{indec},n} \otimes (\mathbb{Z}/p)$, but we shall not use this stronger result.

Theorem 9.5. *Suppose that $n \geq 2$. The map $w \mapsto \alpha_{w,n}$ induces an epimorphism of $\mathbb{F}_p$-linear spaces*

$$\left(\bigoplus_{i \in J(n)} \mathrm{Sh}(X)_{\mathrm{indec},i} \right) \otimes (\mathbb{Z}/p) \to H^2\left(S/S_{(n,p)}\right).$$

Proof. For $i \in J(n)$, the map $X^i \to H^2\left(S/S_{(n,p)}\right)$, $w \mapsto \alpha_{w,n}$, extends by linearity to a $\mathbb{Z}$-module homomorphism

$$\Phi_i \colon \mathbb{Z}\langle X \rangle_i = \bigoplus_{w \in X^i} \mathbb{Z}w \to H^2\left(S/S_{(n,p)}\right), \quad f = \sum_{w \in X^i} f_w w \mapsto \sum_{w \in X^i} f_w \alpha_{w,n}.$$

By Theorem 9.3, $\Phi_i(u ш v) = 0$ for any nonempty words $u,v \in X^*$ with $i = |u| + |v|$. Consequently, Φ_i factors via $\mathrm{Sh}(X)_{\mathrm{indec},i}$, and induces an $\mathbb{F}_p$-linear map

$$\bar{\Phi}_i \colon \mathrm{Sh}(X)_{\mathrm{indec},i} \otimes (\mathbb{Z}/p) \to H^2\left(S/S_{(n,p)}\right),$$

where $\bar{\Phi}_i(\bar{w}) = \alpha_{w,n}$ for $w \in X^i$. Since the $\alpha_{w,n}$, where w ranges over all Lyndon words of an arbitrary length $i \in J(n)$, form an $\mathbb{F}_p$-linear basis of $H^2\left(S/S_{(n,p)}\right)$ (Theorem 8.5(b)), we obtain an epimorphism

$$\bigoplus_{i \in J(n)} \bar{\Phi}_i \colon \left(\bigoplus_{i \in J(n)} \mathrm{Sh}(X)_{\mathrm{indec},i} \right) \otimes (\mathbb{Z}/p) \to H^2\left(S/S_{(n,p)}\right). \qquad \square$$

We now obtain the Main Theorem from the Introduction:

Theorem 9.6. *Suppose that $2 \leq n < p$. Then there is an isomorphism of $\mathbb{F}_p$-linear spaces*

$$\left(\bigoplus_{x \in X} \mathbb{Z}/p \right) \oplus \left(\mathrm{Sh}(X)_{\mathrm{indec},n} \otimes (\mathbb{Z}/p) \right) \xrightarrow{\sim} H^2\left(S/S_{(n,p)}\right). \tag{9.2}$$

Specifically, this isomorphism maps a generator 1_x of the $\mathbb{Z}/p$-summand at $x \in X$ to $\mathrm{Bock}_{p,S/S_{(n,p)}}\left(\epsilon_{(x),\mathbb{Z}/p}\right)$, and maps the image $\bar{w}$ of a word $w \in X^*$ of length n to the n-fold Massey product element $\alpha_{w,n}$.

Proof. By Remark 4.2(1), $J(n) = \{1,n\}$. Therefore, Theorem 9.5 gives an *epimorphism* as in (9.2). The generators 1_x and the images $\bar{w}$ of words w of length n are mapped as specified, by Examples 7.1 and 7.2.

The generators 1_x, $x \in X$, clearly span $\bigoplus_{x \in X} \mathbb{Z}/p$, and by Proposition 9.4, the images $\bar{w}$ of the Lyndon words w in X^* of length n span $\mathrm{Sh}(X)_{\mathrm{indec},n} \otimes (\mathbb{Z}/p)$. Together they form a spanning set of the left-hand side of the epimorphism (9.2), which is mapped to a linear basis of the right-hand side (Corollary 8.6). It follows that this spanning set is a linear basis, and the map (9.2) is an isomorphism. $\square$

https://doi.org/10.1017/S1474748021000426 Published online by Cambridge University Press

Acknowledgments. I thank the referee for his/her comments and helpful suggestions. This research was supported by the Israel Science Foundation (grant 569/21).

Competing Interest. None.

References

[1] A. Bier and W. Hołubowski, A note on commutators in the group of infinite triangular matrices over a ring, *Linear Multilinear Algebra* **63**(11) (2015), 2301–2310.
[2] M. Chapman and I. Efrat, Filtrations of the free group arising from the lower central series, *J. Group Theory* **19** (2016), 405–433.
[3] S. K. Chebolu, I. Efrat and J. Mináč, Quotients of absolute Galois groups which determine the entire Galois cohomology, *Math. Ann.* **352** (2012), 205–221.
[4] K.-T. Chen, R. H. Fox and R. C. Lyndon, Free differential calculus. IV. The quotient groups of the lower central series, *Ann. Math.* **68** (1958), 81–95.
[5] J. D. Dixon, M. P. F. Du Sautoy, A. Mann and D. Segal, *Analytic Pro-p Groups* (Cambridge University Press, Cambridge, UK, 1999).
[6] W. G. Dwyer, Homology, Massey products and maps between groups, *J. Pure Appl. Algebra* **6** (1975), 177–190.
[7] I. Efrat, The Zassenhaus filtration, Massey products, and representations of profinite groups, *Adv. Math.* **263** (2014), 389–411.
[8] I. Efrat, The cohomology of canonical quotients of free groups and Lyndon words, *Documenta Math.* **22** (2017), 973–997.
[9] I. Efrat, The lower p-central series of a free profinite group and the shuffle algebra, *J. Pure Appl. Algebra* **224** (2020), 106260.
[10] I. Efrat and E. Matzri, Triple Massey products and absolute Galois groups, *J. Eur. Math Soc. (JEMS)* **19** (2017), 3629–3640.
[11] I. Efrat and J. Mináč, On the descending central sequence of absolute Galois groups, *Amer. J. Math.* **133** (2011), 1503–1532.
[12] I. Efrat and J. Mináč, Galois groups and cohomological functors, *Trans. Amer. Math. Soc.* **369** (2017), 2697–2720.
[13] M. Ershov, Golod-Shafarevich groups: A survey, *Internat. J. Algebra Comput.* **22** (2012), 1230001.
[14] M. D. Fried and M. Jarden, *Field Arithmetic*, 3rd ed. (Springer, Berlin, 2008).
[15] J. Gärtner, *Mild Pro-p-Groups with Trivial Cup-Product*, Dissertation, Universität Heidelberg, 2011.
[16] P. Guillot, J. Mináč, A. Topaz and O. Wittenberg, Four-fold Massey products in Galois cohomology, *Compos. Math.* **154** (2018), 1921–1959.
[17] Y. Harpaz and O. Wittenberg, The Massey vanishing conjecture for number fields, Preprint, 2019, https://arxiv.org/abs/1904.06512.
[18] M. Hopkins and K. Wickelgren, Splitting varieties for triple Massey products, *J. Pure Appl. Algebra* **219** (2015), 1304–1319.
[19] N. Jacobson, *Lie Groups* (Dover Publications, Inc., New York, 1962).
[20] H. Koch, Zum Satz von Golod-Schafarewitsch, *Math. Nachr.* **42** (1969), 321–333.
[21] H. Koch, *Galois Theory of p-Extensions* (Springer, Berlin, 2002).
[22] J. Labute, Demuškin groups of rank $\aleph_0$, *Bull. Soc. Math. France* **94** (1966), 211–244.
[23] J. Labute, Classification of Demuškin groups, *Canad. J. Math.* **19** (1967), 106–132.

[24] J. Labute, Mild pro-p groups and Galois groups of p-extensions of $\mathbb{Q}$, *J. Reine Angew. Math.* **596** (2006), 155–182.

[25] M. Lazard, Groupes analytiques p-adiques, *Publ. Math. Inst. Hautes Études Sci.* **26** (1965), 389–603.

[26] J. Mináč, M. Rogelstad and N. D. Tân, Dimensions of Zassenhaus filtration subquotients of some pro-p-groups, *Israel J. Math.* **212** (2016), 825–855.

[27] J. Mináč and N. D. Tân, The Kernel Unipotent Conjecture and the vanishing of Massey products for odd rigid fields (with an appendix by Efrat, I., Mináč, J. and Tân, N. D.), *Adv. Math.* **273** (2015), 242–270.

[28] J. Mináč and N. D. Tân, Triple Massey products vanish over all fields, *J. Lond. Math. Soc. (2)* **94** (2016), 909–932.

[29] M. Morishita, Milnor invariants and Massey products for prime numbers, *Compos. Math.* **140** (2004), 69–83.

[30] M. Morishita, *Knots and Primes*, Universitext (Springer, London, 2012).

[31] J. Neukirch, A. Schmidt and K. Wingberg, *Cohomology of Number Fields*, 2nd ed. (Springer, Heidelberg Berlin, 2008).

[32] D. E. Radford, A natural ring basis for the shuffle algebra and an application to group schemes, *J. Algebra* **58** (1979), 432–454.

[33] C. Reutenauer, *Free Lie Algebras*, *London Mathematical Society Monographs. New Series*, **7** (The Clarendon Press, Oxford University Press, New York, 1993).

[34] J.-P. Serre, Structure de certains pro-p-groupes (d'après Demuškin), *Séminaire Bourbaki* (1962/63), Exp. 252. Société Mathématique de France, 1964, eprint: http://www.numdam.org/item/SB_1962-1964__8__145_0/.

[35] J.-P. Serre, *Lie Algebras and Lie Groups* (Springer, Berlin Heidelberg, 1992).

[36] A. Shalev, Dimension subgroups, nilpotency indices, and the number of generators of ideals in p-algebras, *J. Algebra* **129** (1990), 412–438.

[37] D. Vogel, On the Galois group of 2-extensions with restricted ramification, *J. Reine Angew. Math.* **581** (2005), 117–150.

[38] K. Wickelgren, n-Nilpotent obstructions to π_1 sections of $\mathbb{P}^1 - \{0,1,\infty\}$ and Massey products, in *Galois-Teichmüller Theory and Arithmetic Geometry*, Advanced Studies in Pure Mathematics, **63**, pp. 579–600 (Mathematical Society of Japan, Tokyo, 2012).

[39] H. Zassenhaus, Ein Verfahren, jeder endlichen p-Gruppe einen Lie-Ring mit der Charakteristik p zuzuordnen, *Abh. Math. Sem. Univ. Hamburg* **13** (1939), 200–207.

[40] E. Zelmanov, On groups satisfying the Golod-Shafarevich condition, in. *New Horizons in Pro-p Groups*, Progress in Mathematics, **184**, pp. 223–232 (Birkhäuser Boston, Boston, MA, 2000).