Journal of Algebra 352 (2012) 232–267

Contents lists available at SciVerse ScienceDirect

Journal of Algebra

www.elsevier.com/locate/jalgebra

Metabelian Lie powers of the natural module for a general linear group

Karin Erdmann[a,*], L.G. Kovács[b]

[a] Mathematical Institute, 24-29 St. Giles, Oxford OX1 3LB, England, UK
[b] Mathematics, Australian National University, Canberra ACT 0200, Australia

ARTICLE INFO

Article history:
Received 30 June 2011
Available online 12 December 2011
Communicated by E.I. Khukhro

MSC:
17B01
20G05
08B15

Keywords:
Free metabelian Lie algebra
Infinite general linear group
Dual Weyl module
Submodule lattice

ABSTRACT

Consider a free metabelian Lie algebra M of finite rank r over an infinite field K of prime characteristic p. Given a free generating set, M acquires a grading; its group of graded automorphisms is the general linear group $GL_r(K)$, so each homogeneous component M^d is a finite dimensional $GL_r(K)$-module. The homogeneous component M^1 of degree 1 is the natural module, and the other M^d are the metabelian Lie powers of the title.

This paper investigates the submodule structure of the M^d. In the main result, a composition series is constructed in each M^d and the isomorphism types of the composition factors are identified both in terms of highest weights and in terms of Steinberg's twisted tensor product theorem; their dimensions are also given. It turns out that the composition factors are pairwise non-isomorphic, from which it follows that the submodule lattice is finite and distributive. By the Birkhoff representation theorem, any such lattice is explicitly recognizable from the poset of its join-irreducible elements. The poset relevant for M^d is then determined by exploiting a 1975 paper of Yu.A. Bakhturin on identical relations in metabelian Lie algebras.

1. Introduction

If p is a prime and G is a group acting on a finite group H of p-power order, the Frattini factor group of H (that is, the quotient of H modulo its Frattini subgroup) may be viewed as a G-module

* Corresponding author.
E-mail addresses: erdmann@maths.ox.ac.uk (K. Erdmann), lgkovacs@ozemail.com.au (L.G. Kovács).

0021-8693/$ – see front matter
doi:10.1016/j.jalgebra.2011.11.021

over the field of p elements, and the same can be said about the Frattini factor group of the last term of the lower central series of H. The close connection between these two modules has been exploited in many investigations. Similarly, consider a free Lie algebra over the same field and let G act on it by graded algebra automorphisms: then the homogeneous components of the algebra are G-modules and the action of G on the first of these determines the action on all the others. The connection between the two G-modules obtained from H can be explored in terms of the connection between the modules obtained from the free Lie algebra (see, for example, Chapter VIII in Huppert and Blackburn [16]). For applications with metabelian H, one can use the free metabelian Lie algebra instead. The 'largest' possible action arises when G is the group of all graded algebra automorphisms, that is, the general linear group whose natural module is the homogeneous component of degree 1. This was our original motivation for exploring the homogeneous components of free metabelian Lie algebras as modules for the relevant general linear group. These modules are the metabelian Lie powers referred to in the title.

In many respects, it is easier to deal with this matter over infinite fields (or over finite fields whose cardinality is at least as large as the degree of the homogeneous component that we are interested in). One may hope that the answers obtained here will eventually lead to useful conclusions also for the prime field case. Thus in this paper we take an infinite field K of prime characteristic p, a free metabelian Lie algebra M of finite rank r over K, and denote the degree d homogeneous component of M by M^d: our aim is to study the M^d as KG-modules where $G = GL_r(K)$. Another reason these modules are of interest is that, in the terminology of Green's lecture notes [14], they are the dual Weyl modules $D_{(d-1,1),K}$ (for a particularly direct proof, see [8, Lemma 4.2]).

The homogeneous components S^d of the analogous symmetric algebra are the dual Weyl modules $D_{(d),K}$, and their submodule structure has been completely described by Doty [10–12] and Krop [18, 19]; see also the exposition in Bryant [6]. Since this is not only the paradigm but also a prerequisite for our work here, we start with a sketch of their results. (We do not know of similarly conclusive results for any other dual Weyl modules in prime characteristic; of course, in characteristic 0, all such modules are simple.)

Perhaps the most important feature of S^d is that it is multiplicity-free (in the sense that the composition factors of any one composition series are pairwise non-isomorphic) and all its composition factors are absolutely simple. As was seen in [17, pp. 212–213], this has far-reaching consequences. The submodule lattice is finite and distributive, and the set of join-irreducible submodules is naturally bijective with the set of isomorphism types of the composition factors. The latter set has an obvious partial order matching the set-inclusion of the join-irreducibles (and a Hasse diagram of this poset is an example of what Alperin [2] called 'diagrams for modules'). The module has a basis such that each submodule is spanned by the basis elements it contains: so the submodule generated by any one basis element is join-irreducible, and each join-irreducible submodule is generated by some single basis element (which is usually not unique). Such a basis has a pre-order matching the set-inclusion of the submodules generated by the individual basis elements, and on this basis the matrices representing the elements of the algebra KG form the incidence algebra of this pre-order. Thus the relevant quotient of KG is completely described (see [17] and its review by Krop, MR 89k:20017). An equivalent description is in terms of the incidence matrix of the partial order on the set of isomorphism types of composition factors: in that, replace each zero entry by the zero matrix of a given size, and replace each non-zero entry by an arbitrary matrix of a given size. Here the rows and columns of the incidence matrix are indexed by (isomorphism types of) simple modules, and the size of a matrix replacing an entry is given by the dimensions of the simple modules indexing the position of that entry.

In the case of the S^d, a basis with this property is the obvious one, namely the set of monomials (when the symmetric algebra is viewed as a commutative polynomial algebra). Being a dual Weyl module, S^d has a unique simple submodule; less predictably, it also has a unique simple quotient (which is easy to identify by its highest weight), and its other composition factors may be identified as tensor products of 'twisted' versions of simple quotients of smaller symmetric powers. The partial order on their set is then explicitly described. We have not seen an explicit dimension formula in print, so to complete the picture we give one.

In this paper we aim to emulate much of this work. As a dual Weyl module, each M^d has a unique simple submodule; we show that it also has a unique simple quotient. We identify each composition factor of M^d either as a certain composition factor of S^d or as a tensor product of a twisted version of the simple quotient of a smaller metabelian power and a composition factor of some smaller symmetric power. The nature of the highest weights of these tensor factors make it possible to invoke Steinberg's theorem about the uniqueness of such 'twisted tensor factorizations' and conclude the multiplicity-free nature of M^d. We find the dimensions of the relevant simple modules. It is known that all composition factors of M^d are absolutely simple, so the deductions used in the case of S^d automatically apply: once we give the relevant partial order, the submodule structure of the module and the incidence algebra structure of the relevant quotients of KG will follow.

For the partial order, we adapt results from Bakhturin [4]. Certain elements of M^d with $r \geqslant d$ were singled out there as canonical words. It was proved that then each submodule of M^d is generated (as submodule) by the canonical words it contains; moreover, different canonical words generate different submodules, the submodule generated by any one canonical word is join-irreducible, and each join-irreducible submodule is generated by a canonical word. The relevant partial order on the set of canonical words was also determined. Once we match Bakhturin's canonical words to the composition factors we previously identified (see Theorem 7.1 and the subsequent paragraph), we have a description of the partial order we need, though some adjustments are needed when $r < d$.

For flavour, we give here the simplest case, namely that of a d not larger than r and not divisible by p. Consider all possible ways of writing d as a sum of powers of p, not caring about the order of summands. Call one such 'p-partition' a refinement of another if it breaks some parts into sums of smaller powers of p. The results of Doty and Krop mentioned above may be interpreted as saying that the set of all p-partitions of d partially ordered 'by refinement' is the poset relevant for the submodule structure of S^d. For M^d in this special case, all one has to change is to exclude the p-partitions which have only one part p^0.

In appealing to Bakhturin [4], we come to yet another reason why the submodules of M^d command attention. His paper was concerned with varieties (in the sense of equational classes, that is, classes defined by universally quantified equations) of metabelian Lie algebras. Given such a variety, $\mathfrak{V}$, an element of M is called an identity of $\mathfrak{V}$ if it lies in the kernel of every homomorphism $M \to A$ with $A \in \mathfrak{V}$. The set of all identities of $\mathfrak{V}$ in M is a fully invariant ideal; the intersection of that with M^d is a submodule, and each submodule arises in this way: thus questions about such varieties are intimately related to questions about submodules of the M^d.

While it would be useful to have an explicit basis for M^d with the property that each submodule is spanned by the basis elements it contains, here we make no attempt in that direction. Some other challenges are also left for the future. In the case of symmetric powers, one can avoid the assumption $r \geqslant d$ at the cost of a little extra complexity. In the case of metabelian Lie powers with $r < d$, there is a further price to pay: not only can we not find so good a basis, we cannot even write down explicit (submodule) generators for the join-irreducibles. Nevertheless, these M^d remain multiplicity-free and we can still identify the dimensions and isomorphism types of their composition factors and the relevant partial orders. In fact, some things become simpler when r is very small: for $r = 2$, see (5.1′) here; also, Parker [22, Section 3 and references].

In a different direction, it might be of some interest to note that the filtration of S^d in Lemma 4.5 and the filtration of M^d in Lemma 5.9 are p-good filtrations in the sense of Andersen [3]. For $r \leqslant 3$, Kühne-Hausmann [20] and Parker [22] constructed p-good filtrations for all (dual) Weyl modules, but we have not seen explicit p-good filtrations for $r > 3$.

2. Terminology and statement of results

2.1. Partitions and p-partitions

We need to start with some terminology. By a partition we mean a weakly decreasing sequence $\lambda = (\lambda_1, \lambda_2, \ldots)$ of non-negative integers with finite sum $|\lambda|$ (in notation, some or all of the trailing zeroes may be omitted); we also say that λ is a partition of $|\lambda|$. The sum of two partitions is formed according to the rule $(\lambda + \mu)_i = \lambda_i + \mu_i$, and a partition multiplied by a positive integer means a

sum with the relevant number of equal summands. An alternative notation takes the 'exponential' form: one writes $\lambda = (1^s, 2^t, \ldots)$ to indicate that precisely s of the λ_i are equal to 1, t of them are equal 2, and so on; the order of listing these 'powers' does not matter, and one usually omits powers with 'exponent' 0. This terminology is very convenient when we define, for each positive integer c, a partition σ^c of c: let

$$\sigma^c = \bigl((p-1)^a, b\bigr) \quad \text{where } a = \bigl\lfloor c/(p-1)\bigr\rfloor \text{ and } b = c - a(p-1)$$

(for a real number x, we denote by $\lfloor x\rfloor$ the unique integer such that $x - 1 < \lfloor x\rfloor \leqslant x$, and define $\lceil x\rceil$ similarly). We shall also need

$$\mu^c = \begin{cases} (c-1, 1) & \text{if } 2 \leqslant c < p, \\ (p-2, 1^2) & \text{if } c = p > 2, \\ (p, (p-1)^{a'}, b') & \text{if } c > p, \end{cases}$$

where $a' = \lfloor (c-p)/(p-1)\rfloor$ and $b' = (c-p) - a'(p-1)$. It will be convenient to have also $\sigma^0 = \mu^0 = (0)$, but we shall have no need for μ^1, and leave μ^2 also undefined when $p = 2$. Note that the number of (non-zero) parts of σ^c is $\lceil c/(p-1)\rceil$, while if $c > p$ then for μ^c this number is $\lceil (c-1)/(p-1)\rceil$.

If all positive λ_i are powers of p, we call λ a p-partition and use a variant of this exponential notation. To avoid confusion, for p-partitions we write $\alpha, \beta, \ldots$ instead of $\lambda, \mu, \ldots$, and use square brackets instead of round ones: by $[\alpha(0), \alpha(1), \ldots]$, we mean a sequence containing $\alpha(j)$ terms equal to p^j, so $|\alpha| = \sum_{j\geqslant 0} \alpha(j)p^j$. We shall also need the initial partial sums

$$\alpha_{<k} = \sum_{0\leqslant j<k} \alpha(j)p^j,$$

noting that $\alpha_{<k} \equiv |\alpha| \pmod{p^k}$; in particular,

$$\alpha(0) \equiv |\alpha| \pmod{p}.$$

Several more pieces of notation will be needed later. First, set

$$t_\alpha = \min\bigl\{ j \bigm| \alpha(j) > 0\bigr\},$$

and note that if $|\alpha|$ is prime to p then $t_\alpha = 0$. Second, for any α, define the p-partition α_+ of $(|\alpha| - \alpha(0))/p$ by

$$\alpha_+(j) = \alpha(j+1) \quad \text{for } j = 0, 1, \ldots.$$

Further, define the partitions σ^α and μ^α of $|\alpha|$ by

$$\sigma^\alpha = \sum_{j\geqslant 0} p^j \sigma^{\alpha(j)} = \sigma^{\alpha(0)} + p\sigma^{\alpha_+} \quad \text{and} \quad \mu^\alpha = p^{t_\alpha}\mu^{\alpha(t_\alpha)} + \sum_{j>t_\alpha} p^j\sigma^{\alpha(j)},$$

noting that if $\alpha(0) = 0$ then $t_{\alpha_+} = t_\alpha - 1$, $\alpha_+(t_{\alpha_+}) = \alpha(t_\alpha)$, and $\sigma^\alpha = p\sigma^{\alpha_+}$, $\mu^\alpha = p\mu^{\alpha_+}$.

Given a positive integer d, define a partial order $\alpha \preccurlyeq \beta$ on the set $\{\alpha \mid |\alpha| = d\}$ of all p-partitions of d by setting

$$\alpha \preccurlyeq \beta \quad \text{if and only if} \quad \alpha_{<k} \leqslant \beta_{<k} \quad \text{for } k = 1, 2, \ldots.$$

Heuristically, β is a refinement of α if it breaks up some parts of α into sums of smaller powers of p, and it is not hard to see that this is equivalent to $\alpha \preccurlyeq \beta$. Clearly, d has a unique finest p-partition, namely that with d parts, each equal to p^0. As agreed, we call this p-partition $[d, 0, \ldots]$ or simply $[d]$; beware of confusing it with the partition (d) which consists of a single part (and is not a p-partition unless d is a power of p). The digits of d in base p arithmetic describe the other extreme: the unique coarsest p-partition of d.

2.2. Symmetric powers

We paraphrase here the core results on submodules of symmetric powers, in a form suited for our present context.

Let S denote $K[x_1, \ldots, x_r]$, the polynomial algebra over K in the commuting variables $x_1, \ldots, x_r$. For each non-negative integer d, let S^d denote the space of homogeneous polynomials of degree d (so $S^0 = K$). With reference to the basis of S^1 formed by the variables, the matrix group $G = GL_r(K)$ has an obvious right action on S^1, turning this space into the natural module for G. The action extends to all of S (by algebra automorphisms), and so each S^d becomes a KG-module: this will be the form in which we think of the symmetric powers of the natural module. Given a partition λ of d, let x^λ denote the monomial $\prod_{i=1}^r x_i^{\lambda_i}$. (The symmetric group of all permutation matrices in G permutes the set of *all* monomials of degree d, and the x^λ form a convenient set of representatives for its orbits.)

As is well known,

$$\dim S^d = \binom{d+r-1}{r-1}. \tag{2.2.1}$$

The main result is that the inclusion-ordered poset of join-irreducible submodules of S^d is order-isomorphic to the refinement-ordered poset of those p-partitions α of d which satisfy

$$\alpha(j) \leqslant (p-1)r \quad \text{for } j = 0, 1, \ldots. \tag{2.2.2}$$

For such a p-partition α of d, denote by S^α the submodule generated by $\mathsf{x}^{\sigma^\alpha}$: these are the join-irreducible submodules of S^d, and $\alpha \mapsto S^\alpha$ is the order-isomorphism we have in mind. If $d < p$, then d has only one p-partition; for some d and r it happens that d has several p-partitions but only one of them satisfies (2.2.2): in all these cases, S^d is simple. In general, the unique simple quotient of S^α will be written as $\overline{S^\alpha}$, its highest weight is σ^α, and (the coset containing) $\mathsf{x}^{\sigma^\alpha}$ is a highest weight vector in it. Thus in the language of highest weights and twisted tensor products,

$$\overline{S^\alpha} = L(\sigma^\alpha) = \bigotimes_j L(\sigma^{\alpha(j)})^{F^j} \tag{2.2.3}$$

where F is the Frobenius functor. Accordingly,

$$\dim \overline{S^\alpha} = \dim L(\sigma^\alpha) = \prod_j \dim L(\sigma^{\alpha(j)})$$

and we shall see that the factors of this product are given by[1]

$$\dim L(\sigma^c) = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \binom{c-kp+r-1}{r-1}. \tag{2.2.4}$$

[1] The second author saw a similar formula in a 1994 message from Leonid Krop.

If $r \geqslant \sum_j \alpha(j)$, we can also consider the monomial

$$\mathsf{x}^\alpha = \prod_{i=1}^{\sum \alpha(j)} x_i^{p^{k(i)}}$$

where $k(i)$ is defined by the condition

$$\sum_{j<k(i)} \alpha(j) < i \leqslant \sum_{j \leqslant k(i)} \alpha(j).$$

(For example, if t $(= t_\alpha)$ is the smallest number such that $\alpha(t) > 0$ and t' is the next smallest number with $\alpha(t') > 0$, then $k(i)$ is t for $i = 1, \ldots, \alpha(t)$, it is t' for $i = \alpha(t) + 1, \ldots, \alpha(t) + \alpha(t')$, and so on. We hope the convention that α, β are always p-partitions will protect against confusing the x^α with the x^λ.) The submodule generated by x^α is S^α, and at times x^α is a more convenient generator for this submodule than $\mathsf{x}^{\sigma^\alpha}$, but of course it is not available when $r < \sum_j \alpha(j)$, and such α exist whenever $r < d$.

It was noted on page 210 of [17] that the poset of the relevant p-partitions of d always has a unique largest element, there called β_1 but here simply β, defined inductively as follows. Take the largest number congruent (mod p) to d but no larger than d or $(p-1)r$, and call this $\beta(0)$. Once the $\beta(j)$ are defined for $j < k$, choose $\beta(k)$ as the largest number congruent to $(d - \beta_{<k})/p^k$ but no larger than $(d - \beta_{<k})/p^k$ or $(p-1)r$.

It follows that S^d is always join-irreducible, namely S^β, and its unique simple quotient is $\overline{S^\beta}$, that is, $L(\sigma^\beta)$. As long as $d \leqslant (p-1)r$, we have $\beta = [d]$ and so $S^d = S^{[d]}$; in this case the unique simple quotient of S^d is $\overline{S^{[d]}}$, that is, $L(\sigma^d)$. Thus all the tensor factors in (2.2.3) are twisted versions of simple quotients of symmetric powers.

However, when $d > (p-1)r$, the p-partition $[d]$ is irrelevant for our purposes and $S^{[d]}$ is undefined. Nevertheless, in (4.1), we shall find it convenient to define $\overline{S^d}$, which will turn out to be

$$\overline{S^d} = \begin{cases} \text{the unique simple quotient of } S^d & \text{if } d \leqslant (p-1)r, \\ 0 & \text{otherwise.} \end{cases}$$

In these terms, (2.2.3) can be re-phrased as

$$\overline{S^\alpha} = \bigotimes_j \left(\overline{S^{\alpha(j)}}\right)^{F^j}. \tag{2.2.3$'$}$$

2.3. Simple facts on metabelian Lie powers

Let M denote a free metabelian Lie algebra of rank r freely generated by $x_1, \ldots, x_r$ (with $r \geqslant 2$), and let M^d denote the degree d homogeneous component of M, regarded a right KG-module as previously indicated. As M^1 is the natural module $L(1)$ and M^2 is the exterior square $L(1,1)$, we may as well assume $d \geqslant 2$. As noted in (2.2) of [7] (but known at least since [9, Corollary 1]), then

$$\dim M^d = r\binom{d-1+r-1}{r-1} - \binom{d+r-1}{r-1} = (d-1)\binom{d+r-2}{r-2}. \tag{2.3.1}$$

Since M^d is the dual Weyl module corresponding to the partition $(d-1, 1)$, its socle is always simple and isomorphic to $L(d-1, 1)$. The left-normed Lie monomial $[x_1, x_2, x_1, \ldots, x_1]$ is a convenient highest weight vector for this submodule. This is all there is to be said when M^d is simple, so we list here the conditions under which that is the case.

The module M^d is simple for all p and r provided $d < p$ or $d = p+1$. If $r = 2$, then M^{kp^m+1} is simple whenever $0 < k < p$ and $m \geqslant 0$ (this includes the fact that M^p is simple when $r = 2$). Of course if $p = 2$ then M^p is simple for every r. Finally, if $p = 2$ and $r = 3$, then M^5 is also simple.

This exhausts the list of the simple M^d. In particular, if $p > 2$ and $r > 2$, then M^p is no longer simple: its socle $L(p-1, 1)$ is isomorphic to $\overline{S^p}$ (and we have the dimension of that), while its quotient over the socle is $L(p-2, 1, 1)$, and (the coset containing) $[x_1, x_2, x_3, x_1, \dots, x_1]$ may be chosen as highest weight vector for this quotient.

2.4. Coprime degrees

Suppose now that $d > p+1$ and $p \nmid d$, noting that the latter condition implies that $t_\alpha = 0$ and so $\mu^\alpha = \mu^{\alpha(0)} + p\sigma^{\alpha_+}$ for each p-partition α of d.

The main result for this case is very similar to that on symmetric powers: the inclusion-ordered poset of join-irreducible submodules of M^d is isomorphic to the refinement-ordered poset of those p-partitions α of d which satisfy

$$2 \leqslant \alpha(0) \leqslant (p-1)r + 1 \quad \text{and} \quad \alpha(j) \leqslant (p-1)r \quad \text{for } j = 1, 2, \dots, \tag{2.4.1}$$

and $\alpha \mapsto L(\mu^\alpha)$ is the corresponding isomorphism from the poset of these α to the relevant poset of the isomorphism types of the unique simple quotients of those join-irreducibles.

The dimensions of the composition factors come as special cases of a formula which holds without assuming coprimality: in general, $\mu^\alpha = p^t\mu^{\alpha(t)} + \sum_{j>t} p^j\sigma^{\alpha(j)}$ with $t = t_\alpha$, so we always have

$$L(\mu^\alpha) = L(\mu^{\alpha(t)})^{F^t} \otimes \bigotimes_{j>t} L(\sigma^{\alpha(j)})^{F^j} \quad \text{where } t = t_\alpha.$$

Consequently, $\dim L(\mu^\alpha)$ is the product of $\dim L(\mu^{\alpha(0)})$ and certain $\dim L(\sigma^{\alpha(j)})$. We have already seen how to calculate the latter; after Lemma 5.5, we shall see that

$$\dim L(\mu^c) = \begin{cases} r \text{ if } c = (p-1)r + 1 & \text{otherwise}, \\ r \dim L(\sigma^{c-1}) - \dim L(\sigma^c) & \text{if } p \nmid c, \\ r \dim L(\sigma^{c-1}) - 2\dim L(\sigma^c) & \text{if } p \mid c. \end{cases} \tag{2.4.2}$$

In contrast to the case previously discussed, here we have not succeeded in writing down highest weight vectors for the join-irreducibles. Unless $d \leqslant r$, we cannot always provide (submodule) generators either, though that can be done for the join-irreducibles whose simple quotient is an $L(\mu^\alpha)$ with $\sum_j \alpha(j) \leqslant r$. Namely, for any such p-partition α of d, the join-irreducible with simple quotient $L(\mu^\alpha)$ is generated by the *Bakhturin monomial* a^α which is defined as the left-normed Lie product with leftmost factor x_1, next factor x_2, and altogether $p^{k(i)}$ factors x_i where $k(i)$ is given by the condition

$$\sum_{j<k(i)} \alpha(j) < i \leqslant \sum_{j\leqslant k(i)} \alpha(j).$$

2.5. Other degrees

In general, the result simplest to state is that each M^d is multiplicity-free and join-irreducible, regardless of how d and r compare.

As long as $r \geqslant d$, M^d has a composition factor $L(\mu^\alpha)$ for each p-partition α of d such that

$$\alpha(t_\alpha) \geqslant \begin{cases} 3 & \text{when } p = 2, \\ 2 & \text{when } p > 2 \end{cases} \tag{2.5.1}$$

and a composition factor $L(\sigma^\alpha)$ for each α such that

$$\alpha(t_\alpha) \equiv 0 \pmod{p}. \tag{2.5.2}$$

These are all the composition factors of M^d, but their partial order is not simply the refinement order of the α. For a full description, we have to refer to Theorem 6.1 (which gives the relevant partial order in terms of the canonical generators chosen by Bakhturin [4] for the join-irreducible submodules which have those simple modules as quotients) and to Theorem 7.1 (which explains how to translate between canonical generators and highest weights).

The restriction $r \geqslant d$ can be replaced by excluding the cases in which μ^α or σ^α has more than r parts. For the μ^α, this amounts to demanding

$$\begin{gathered} \alpha(t_\alpha) \leqslant (p-1)r + 1, \\ \alpha(t_\alpha) \neq p \quad \text{when } r = 2, \quad \text{and} \\ \alpha(j) \leqslant (p-1)r \quad \text{when } j \neq t_\alpha, \end{gathered} \tag{2.5.1$'$}$$

while for the σ^α one has to impose

$$\alpha(j) \leqslant (p-1)r \quad \text{for } j = 0, 1, \ldots. \tag{2.5.2$'$}$$

The partial order on the set of composition factors read off Theorem 6.1 remains valid after these exclusions, though we lose the ability to name submodule generators for some of the join-irreducibles.

3. Preparations

Except where otherwise stated, all modules considered will be right modules, though scalars from a field will be written to the left of vectors. We write all maps on the right and form their composites accordingly.

3.1. Snakes

It will be convenient to start by isolating from our arguments two very simple but recurring steps. The first of them resembles the Snake Lemma but is even easier, so we leave the proof to the reader.

Half-a-snake Lemma. *To each commutative diagram*

$$\begin{tikzcd} \cdot \arrow[r, "\varphi"] \arrow[d, "\zeta"'] & \cdot \arrow[d, "\eta"] \\ \cdot \arrow[r, "\chi"'] & \cdot \end{tikzcd}$$

in which ζ is surjective, there is also a commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & \operatorname{im}\varphi \arrow[r] \arrow[d, "\zeta'"'] & \operatorname{codom}\varphi \arrow[r] \arrow[d, "\eta"] & \operatorname{coker}\varphi \arrow[r] \arrow[d, "\eta'"] & 0 \\ 0 \arrow[r] & \operatorname{im}\chi \arrow[r] & \operatorname{codom}\chi \arrow[r] & \operatorname{coker}\chi \arrow[r] & 0 \end{tikzcd}$$

with exact rows and surjective ζ', and a short exact sequence

$$0 \to \operatorname{im}\varphi \cap \ker\eta \to \ker\eta \to \ker\eta' \to 0.$$

If η is surjective, then so is η'.

The second result is also very simple; it comes by one-and-a-half applications of the Snake Lemma.

Snake-and-a-half Lemma. *Given a commutative diagram*

$$\begin{tikzcd} {} \arrow[r, "\varphi"] \arrow[d, "\zeta"'] & {} \arrow[d, "\eta"] \\ {} \arrow[r, "\chi"'] & {} \end{tikzcd}$$

with φ and η injective and χ surjective, there is a surjective $\psi : \operatorname{coker}\zeta \to \operatorname{coker}\eta$ and a short exact sequence $0 \to \ker\chi \to \ker\psi \to \operatorname{coker}\varphi \to 0$.

Proof. From the data one can form a commutative diagram

$$\begin{tikzcd} {} \arrow[r, "\varphi"] \arrow[d, "\zeta"'] & {} \arrow[r] \arrow[d, "\eta"] & \operatorname{coker}\varphi \arrow[r] \arrow[d] & 0 \\ {} \arrow[r, "\chi"'] & {} \arrow[r] & 0 & \end{tikzcd}$$

with exact rows. This part of the hypothesis of the Snake Lemma is just what is needed for the part of the conclusion we asserted first: there is an exact sequence $\operatorname{coker}\zeta \to \operatorname{coker}\eta \to 0$. With the ψ so defined, we form another commutative diagram

$$\begin{tikzcd} & {} \arrow[r, "\zeta"] \arrow[d, "\varphi"'] & {} \arrow[r] \arrow[d, "\chi"] & \operatorname{coker}\zeta \arrow[r] \arrow[d, "\psi"] & 0 \\ 0 \arrow[r] & {} \arrow[r, "\eta"'] & {} \arrow[r] & \operatorname{coker}\eta & \end{tikzcd}$$

with exact rows, and here the full Snake Lemma applies: there is an exact sequence

$$\ker\varphi \to \ker\chi \to \ker\psi \to \operatorname{coker}\varphi \to \operatorname{coker}\chi \to \operatorname{coker}\psi.$$

Since by assumption $\ker\varphi = \operatorname{coker}\chi = 0$, the second assertion of the lemma is also proved. □

3.2. Lattices

A module is called multiplicity-free if it has finite composition series and the composition factors of any one composition series are pairwise non-isomorphic.

The set of all submodules of any module is a modular lattice with (set theoretic) intersection and (additive group) sum as the lattice operations: this is called *the* submodule lattice of the module in question. In the multiplicity-free case, mapping each join-irreducible submodule to the isomorphism type of its unique simple quotient is clearly a bijection from the set of join-irreducible submodules

to the set of isomorphism types of composition factors. When we speak of the poset of composition factors, we mean the latter set with the inclusion order of the former transferred along this bijection.

It is well known (and follows for example from [2, Section 2]) that the submodule lattice of a multiplicity-free module is always finite and distributive. Thus Birkhoff's representation theorem applies (see [1, pp. 30–35] for all the lattice theory we need) and shows that the full submodule lattice may be reconstructed from the (inclusion-ordered) poset of the join-irreducible submodules. Explicitly, it yields that the submodule lattice is isomorphic to a sublattice in the lattice of all subsets of that poset, namely to the sublattice formed by the poset ideals. One can readily verify that the partial order on the set of composition factors of M defined above is the same as the partial order defined in [2, Section 2] in terms of non-split extensions occurring as subquotients of M.

3.3. Lattice homomorphisms

Given a lattice homomorphism θ between two finite lattices, we need to understand the very simple connection between the posets of join-irreducibles in the domain and in the image of θ.

Easy examples show that the image of a join-irreducible need not be join-irreducible. In the opposite direction, if U is a join-irreducible element of the image of θ, not all the pre-images of U need to be join-irreducibles, but they form a closed interval in the domain and their intersection is always join-irreducible. It is not hard to prove that *mapping each join-irreducible U of the image to this smallest pre-image is an order isomorphism from the poset of join-irreducibles of the image to a sub-poset of the join-irreducibles of the domain.* (By sub-poset we mean a subset with the partial order in which elements are comparable exactly as in the whole poset.)

A proof may be sketched as follows. For any element X in a finite lattice, set $X{\downarrow} = \bigvee\{Z \mid Z < X\}$, noting that X is join-irreducible if and only if $X{\downarrow} < X$. For each element U in the codomain of a surjective lattice homomorphism θ, let us write $U\tau = \bigwedge\{Z \mid Z\theta \geqslant U\}$. Note that $U\tau\theta = U$ so τ is one-to-one, and that if $U \leqslant V$ then $U\tau \leqslant V\tau$. Equally obviously, $Z < U\tau$ implies $Z\theta \leqslant U{\downarrow}$, and using this it is easy to see that if U is join-irreducible then so is $U\tau$:

$$\big((U\tau){\downarrow}\big)\theta = \Big(\bigvee\{Z \mid Z < U\tau\}\Big)\theta = \bigvee\{Z\theta \mid Z < U\tau\} \leqslant \bigvee\{W \mid W \leqslant U{\downarrow}\} = U{\downarrow}$$

shows that if $U{\downarrow} < U$ then we cannot have $(U\tau){\downarrow} = U\tau$.

3.4. Homogeneous polynomial representations

Let K be an infinite field of characteristic p and E (or E_r where this precision is relevant) an r-dimensional vector space over K, with a distinguished basis $\{x_1, \dots, x_r\}$. With reference to this basis, we consider E a right module for the multiplicative semigroup $\Gamma = \Gamma_r$ of all $r \times r$ matrices over K, and also for the group $G = GL_r(K)$ of the nonsingular matrices. Under the usual 'diagonal' action, the d-fold tensor power $E^{\otimes d}$ is also a $K\Gamma$-module and KG-module, and representations corresponding to direct sums of subquotients of $E^{\otimes d}$ are called homogeneous polynomial representations of degree d. Here, our principal references are Green [14] and Martin [21]. It follows from the Centralizer Theorem [14, (2.6c)], [21, 2.1.3], that *the images of $K\Gamma$ and KG in $\mathrm{End}_K(E^{\otimes d})$ coincide, and that this common image is the K-span of the image of $GL_r(K_0)$ whenever K_0 is an infinite subfield of K.* In particular, it does not matter whether we consider the $E^{\otimes d}$ or their subquotients as $K\Gamma$-modules or KG-modules. Also, all composition factors of these modules are absolutely simple [21, 1.6.1]. The symmetric algebra is a quotient of the tensor algebra; the free Lie algebra is a sub-Lie-algebra, and the free metabelian Lie algebra is a quotient of that: so symmetric powers and metabelian Lie powers are subquotients of tensor powers, and hence all the modules we deal with correspond to homogeneous polynomial representations.

3.5. Simple modules and highest weights

Let $\Lambda(r, d)$ be the set of all r-term sequences of non-negative integers summing to d, that is, of the $\lambda = (\lambda_1, \dots, \lambda_r)$ such that $\sum \lambda_i = d$. For $\lambda, \mu \in \Lambda(r, d)$, write $\lambda \geqslant \mu$ if either $\lambda = \mu$ or the least j for

which $\lambda_j \neq \mu_j$ is such that $\lambda_j > \mu_j$. Note that the highest element in this order is $(d, 0, \ldots, 0)$, often written briefly as (d). Denote the subset consisting of the weakly descending sequences by $\Lambda^+(r, d)$; if $r \geqslant d$, then the lowest element of this subset is $\omega = (1, \ldots, 1, 0, \ldots, 0)$.

Given an element λ of $\Lambda(r, d)$ and a KG-module V, let V^λ be the set of the elements v of V such that $vt = (\prod_i t_{ii}^{\lambda_i})v$ for every diagonal matrix $t = (t_{ij})$ in G. The subspace V^λ is called a weight space of V; if $V^\lambda \neq 0$, then λ is called a weight of V, and the weights of V which lie in $\Lambda^+(r, d)$ are called its dominant weights. Each non-zero subquotient V of $E^{\otimes d}$ is the direct sum of its weight spaces; among its weights, there is a unique highest (in terms of the full order on $\Lambda(r, d)$ discussed above), and the highest weight is always a dominant weight. Every element of $\Lambda^+(r, d)$ occurs as the highest weight of some composition factor of $E^{\otimes d}$, and two composition factors of $E^{\otimes d}$ are isomorphic if and only if their highest weights are the same. Thus the set of isomorphism classes of composition factors of $E^{\otimes d}$ may be parametrized by $\Lambda^+(r, d)$: one writes $L(\lambda)$ for an arbitrary representative of the isomorphism class of simple modules with highest weight λ. By convention, $L(0)$ is the 1-dimensional trivial module, and $L(1) = E$.

One can also think of weights without specifying r or d: in that convention, a weight λ is just an infinite sequence $\lambda_1, \lambda_2, \ldots$ of non-negative integers almost all of which are zeros (and where the final zeros may be omitted in notation). To call a weight dominant still means that it is a weakly descending sequence.

Of course it is ambiguous to use $L(\lambda)$ as the common label for the simple $GL_r(K)$-modules with highest weight λ but unspecified r (as long as λ has no more than r positive parts). No confusion will come from this when r is fixed by the context, but one must be careful when changing r, as we are about to do.

3.6. *Change of rank*

Suppose $r < s$ so E_r is a subspace of E_s, and let ε denote the idempotent element of Γ_s which fixes the vectors in E_r and annihilates each of $x_{r+1}, \ldots, x_s$. One can identify $\varepsilon\Gamma_s\varepsilon$ with Γ_r in an obvious way, and consider a functor from the category of $K\Gamma_s$-modules V to $K\Gamma_r$-modules which takes V to $V\varepsilon$ and $K\Gamma_s$-module homomorphisms to the relevant restrictions: see the discussion in Section 6.2 and Section 6.5 of [14] (where our ε was called e). It is clear that this functor respects tensor products, and that it takes symmetric or metabelian Lie powers of E_s to such powers of E_r.

The corresponding map from the submodule lattice of V to the submodule lattice of $V\varepsilon$ is readily seen to be surjective (imitate the proof of [14, (6.2b)]), and then it is obviously a lattice homomorphism.

In particular, if V/W is simple, then $V\varepsilon/W\varepsilon$ is either 0 or simple: to be specific, [14, Theorems (6.5e) and (6.5f)] yield that $V\varepsilon/W\varepsilon$ is simple if and only if the highest weight of V/W has no more than r non-zero parts, and then the highest weight of $V\varepsilon/W\varepsilon$ is the same as that of V/W.

From Section 3.3 above we see that the poset of join-irreducible submodules of $V\varepsilon$ is order-isomorphic to the sub-poset of certain join-irreducible submodules of V. The present discussion and Remark 2 at the end of Section 6.2 of [14] show that the relevant join-irreducible submodules of V are precisely those whose simple quotients have highest weights with no more than r non-zero parts.

Thus *in order to prove that all metabelian Lie powers are multiplicity-free, and in order to specify the partial order on the set of the isomorphism types of their composition factors, no generality is lost by assuming that* $d < r$.

3.7. *The Frobenius functor*

The field endomorphism $\kappa \mapsto \kappa^p$ of K applied to the entries of matrices over K yields a group endomorphism of G, and composing that with the action of G on some KG-module V defines a new module on the same vector space V. (Note that only the action of G changes, the scalar action of K remains the same.) We denote this new module by V^F, and define the *Frobenius functor* F from the category of KG-modules to itself as the functor which takes V to V^F and leaves all homomorphisms unchanged. It follows from the paragraph above that if V is a sub-quotient of some $E^{\otimes d}$, then (as

subspaces) the submodules of V^F are precisely the submodules of V (because the image K_0 of $x \mapsto x^p$ is an infinite subfield of K). We shall also use that the Frobenius functor respects tensor products.

It is easy to see that $L(\lambda)^F = L(p\lambda)$ where $p\lambda$ is defined by $(p\lambda)_i = p\lambda_i$ for $i = 1, 2, \ldots$ (so in particular $L(0)^F = L(0)$).

3.8. *Steinberg's twisted tensor products*

A dominant weight λ is said to be p-restricted if $\lambda_i - \lambda_{i+1} \leqslant p - 1$ for $i = 1, 2, \ldots$. Note that the σ^c and μ^c defined in Section 2.1 are always p-restricted, and by convention so is (0), but if $\lambda \neq (0)$ then $p\lambda$ cannot be p-restricted.

It is easy to see that each dominant weight λ can be written in one and only one way as

$$\lambda = \lambda^0 + p\lambda^1$$

with λ^0, λ^1 dominant and the first of them p-restricted (hint: choose λ^1 so that $\lambda^1_i - \lambda^1_{i+1} = \lfloor(\lambda_i - \lambda_{i+1})/p\rfloor$ for $i = 1, 2, \ldots$). Steinberg's theorem says that in these terms

$$L(\lambda) \cong L(\lambda^0) \otimes L(\lambda^1)^F$$

and conversely, if λ^0 is p-restricted then the tensor product on the right hand side is $L(\lambda^0 + p\lambda^1)$. Repeated application yields a different variant, namely a unique decomposition

$$\lambda = \sum_j p^j \lambda^{(j)}, \qquad L(\lambda) \cong \bigotimes_j L(\lambda^{(j)})^{F^j}$$

with p-restricted $\lambda^{(j)}$. The uniqueness of these 'twisted tensor factorizations' of $L(\lambda)$ is critical for the present context: the way we prove that M^d is multiplicity-free is by describing a composition series whose composition factors have pairwise distinct twisted tensor factorizations.

3.9. *The Schur functor*

Assume $d \leqslant r$. Write Σ_d for the subgroup of G consisting of those permutation matrices which fix each x_i with $d < i \leqslant r$. Recall that ω stands for the smallest element $(1, \ldots, 1, 0, \ldots, 0)$ of $\Lambda^+(r, d)$. If V is a subquotient of $E^{\otimes d}$, then the weight space V^ω is setwise invariant under Σ_d. If $U \to V$ is a KG-homomorphism of subquotients of $E^{\otimes d}$, then it maps U^ω into V^ω, and its restriction $U^\omega \to V^\omega$ is a $K\Sigma_d$-homomorphism. These facts give us the Schur functor f from the category of direct sums of KG-subquotients of $E^{\otimes d}$ to the category of $K\Sigma_d$-modules.

We shall use that this functor is exact. It takes a simple module either to zero or to a simple module (see [14, (6.2g)]): namely, if λ is p-restricted then $fL(\lambda)$ is simple, otherwise $fL(\lambda) = 0$. In terms of twisted tensor factorizations, this means that $\bigotimes_{j \geqslant 0} L(\lambda^{(j)})^{F^j}$ is taken to zero unless each $\lambda^{(j)}$ with $j > 0$ is (0). In this way, the Schur functor gives a one-to-one correspondence between the (isomorphism types of) simple composition factors of $E^{\otimes d}$ with p-restricted highest weights and the simple $K\Sigma_d$-modules (see [14, (6.4b)]). Further, if V has only p-restricted dominant weights, then the Schur functor induces a lattice isomorphism between the submodule lattices of V and of fV.

We shall need just two applications of the Schur functor. The ω-weight space of the symmetric power S^d of E is the 1-dimensional space spanned by the monomial $\prod_{i=1}^d x_i$, and this fS^d is a trivial $K\Sigma_d$-module. As the first application, we conclude that in any composition series of S^d there is precisely one composition factor whose highest weight is p-restricted. In preparation for the second, put $u = \prod_{i=1}^d x_i$: it is easy to see that $\{x_i \otimes (\partial u/\partial x_i) \mid i = 1, \ldots, d\}$ is a basis of $f(E \otimes S^{d-1})$. Thus $f(E \otimes S^{d-1})$ is the natural permutation module for Σ_d, and the submodule structure of that is well known. Namely, if $p \nmid d$, then this module is the direct sum of two simple modules one of which is

trivial; if $p \mid d$ and $d > 2$ then it is uniserial of length three, with top and socle trivials and the middle nontrivial; while if $d = p = 2$ then it is uniserial of length two, with both composition factors trivial. Our investigation of M^d in Section 5 will be based on the fact that M^d can be viewed as the kernel of the surjective homomorphism $E \otimes S^{d-1} \to S^d$ which simply 'forgets the $\otimes$ sign'. The Schur functor takes this to the homomorphism which maps each $x_i \otimes (\partial u/\partial x_i)$ to u, so it takes M^d to a simple module when either $p \nmid d$ or $d = p = 2$, and to a uniserial of length two otherwise, with trivial socle and nontrivial top. The simplest part of the desired conclusion is that if $d < p$ or $d = p = 2$, then M^d is simple, as we claimed in Section 2.3. More generally, if $p \nmid d$, then in any composition series of M^d there is precisely one composition factor whose highest weight is p-restricted. (Remember: the assumption $d \leqslant r$ is still in force.) We shall draw some further conclusions once relevant terminology is developed in Section 5.

4. A composition series for symmetric powers

The core results on symmetric powers were collected in Section 2.2; for their proofs, the reader is referred to [6] and its list of references. The principal aim of this section is the construction of a particular composition series for S^d, to provide a basis for a similar construction in M^d. The dimension formula (2.2.4), for which we have no reference, will fall out as an early by-product.

Recall from Section 2.2 that S denotes $K[x_1, \dots, x_r]$, the polynomial algebra over K in the commuting variables $x_1, \dots, x_r$, with S^d the space of homogeneous polynomials of degree d. It will be convenient to interpret S^0 as the 1-dimensional trivial module, $L(0)$, and to identify S^1 with the natural module $E = L(1)$ introduced in Section 3.4.

If U and V are submodules of S, then so is UV (the span of all products uv with $u \in U$, $v \in V$), and there is a G-homomorphism from $U \otimes V$ with image UV defined by 'forget the $\otimes$ sign'. The subspace of S^{kp} spanned by the pth powers of the monomials of degree k is a submodule; for reasons which will be discussed later, we denote it by $(S^k)^{(p)}$. All specific G-homomorphisms considered in this section will be formed by restricting or composing or tensoring maps of this kind with each other or with inclusions $(S^k)^{(p)} \to S^{kp}$: we shall not spell out their identifications, and leave it to the reader to check that all relevant diagrams of such maps commute.

The maps $S^{d-p} \otimes E^{(p)} \to S^d$ with image $S^{d-p}E^{(p)}$ will be particularly important; we define the modules $\overline{S^d}$ as the corresponding cokernels, that is, by the exact sequence

$$S^{d-p} \otimes E^{(p)} \to S^d \to \overline{S^d} \to 0. \tag{4.1}$$

Our convention is that if $d - p < 0$ then $S^{d-p} = 0$ and $\overline{S^d} = S^d$, while of course if $d \geqslant p$ then $\overline{S^d} = S^d/S^{d-p}E^{(p)}$.

An obvious basis for $\overline{S^d}$ consists of the weight vectors which are known in this polynomial context as the monomials with total degree d and all partial degrees smaller than p (strictly, one should speak of the cosets modulo $S^{d-p}E^{(p)}$ that contain these monomials). In particular, all dominant weights of $\overline{S^d}$ are p-restricted, so if $r \geqslant d$ then the first application of the Schur functor in Section 3.9 shows that $\overline{S^d}$ has at most one composition factor, in other words, that if $\overline{S^d}$ is not 0, it is simple. If $r < d$, one can appeal to Section 3.6 with $s \geqslant d$: consider the variant of our $\overline{S^d}$ defined in terms of s variables instead of r; by the present discussion, that is a simple module; the functor discussed in Section 3.6 takes it to our $\overline{S^d}$, which is therefore either 0 or simple. It is obvious that the basis named is non-empty if and only if $0 \leqslant d \leqslant (p-1)r$, in which case the highest weight is the σ^d defined in Section 2.1: thus $\overline{S^d}$ is either 0 or $L(\sigma^d)$, as claimed at the end of Section 2.2.

Using the inclusion–exclusion principle to count the cardinality of the given basis, one obtains

$$\dim \overline{S^d} = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \dim S^{d-kp} \tag{4.2}$$

and hence the dimension formula (2.2.4). In particular, $\dim \overline{S^{(p-1)r}} = 1$, because in $S^{(p-1)r}$ there is only one monomial with all partial degrees smaller than p, namely $\prod_{i=1}^{r} x_i^{p-1}$.

Let $c \geqslant p$, and note that the image of the map $S^c \otimes (S^k)^{(p)} \to S^{c+kp}$ is $S^c(S^k)^{(p)}$, the span of the monomials $\prod x_i^{b(i)}$ such that $\sum b(i) = c + kp$ and $\sum \lfloor b(i)/p \rfloor \geqslant k$. Those with $\sum \lfloor b(i)/p \rfloor > k$ lie in $S^{c-p}(S^{k+1})^{(p)}$, and those with $\sum \lfloor b(i)/p \rfloor = k$ have unique factorizations as products of a pth power of a monomial from S^k and a monomial from S^c in which all partial degrees are smaller than p. The former span the image $S^{c-p}(S^{k+1})^{(p)}$ of $S^{c-p}E^{(p)} \otimes (S^k)^{(p)}$, while the set of the latter is bijective with a basis of $\overline{S^c} \otimes (S^k)^{(p)}$: hence the kernel of $S^c \otimes (S^k)^{(p)} \to S^{c+kp}$ lies in $S^{c-p}E^{(p)} \otimes (S^k)^{(p)}$, and therefore the quotient $S^c(S^k)^{(p)}/S^{c-p}(S^{k+1})^{(p)}$ of the images is isomorphic to $(S^c/S^{c-p}E^{(p)}) \otimes (S^k)^{(p)}$. More generally, if W is any submodule of S^k, then the kernel of the restriction to $S^c \otimes W^{(p)}$ lies in

$$\left(S^{c-p}E^{(p)} \otimes \left(S^k\right)^{(p)}\right) \cap \left(S^c \otimes W^{(p)}\right),$$

that is, in $S^{c-p}E^{(p)} \otimes W^{(p)}$, and from this we may conclude that

$$S^c W^{(p)}/S^{c-p}(EW)^{(p)} \cong \overline{S^c} \otimes W^{(p)}. \tag{4.3}$$

We record also the exact sequence version:

$$0 \to S^{c-p}(EW)^{(p)} \to S^c W^{(p)} \to \overline{S^c} \otimes W^{(p)} \to 0. \tag{4.4}$$

Because of our conventions, both versions remain valid even when $c < p$.

As a very special case of (4.3), we have that

$$S^{d-\lfloor d/p \rfloor p}\left(S^{\lfloor d/p \rfloor}\right)^{(p)} \cong \overline{S^{d-\lfloor d/p \rfloor p}} \otimes \left(S^{\lfloor d/p \rfloor}\right)^{(p)}. \tag{4.5}$$

A slightly less special case has the following useful paraphrase.

Lemma 4.6. *The module S^d has a filtration*

$$S^d \geqslant S^{d-p}E^{(p)} \geqslant \cdots \geqslant S^{d-kp}\left(S^k\right)^{(p)} \geqslant \cdots \geqslant S^{d-\lfloor d/p \rfloor p}\left(S^{\lfloor d/p \rfloor}\right)^{(p)} > 0$$

whose quotients are isomorphic to

$$\begin{gathered} \overline{S^d} \\ \vdots \\ \overline{S^{d-kp}} \otimes \left(S^k\right)^{(p)} \\ \vdots \\ \overline{S^{d-\lfloor d/p \rfloor p}} \otimes \left(S^{\lfloor d/p \rfloor}\right)^{(p)}. \qquad \square \end{gathered}$$

It should be noted that the first few terms of this filtration may coincide, giving filtrations quotients that vanish. This only happens when $d > (p-1)r$, in which case $S^d = \cdots = S^{d-cp}(S^c)^{(p)}$ with $c = \lceil (d-(p-1)r)/p \rceil$, but beyond this the filtration is strictly descending.

At this point we pause for some observations on the way the Frobenius functor acts on symmetric powers of the natural module: we shall use these in the sequel without reference. Since as commutative K-algebra $K[x_1, \ldots, x_r]$ is freely generated by $x_1, \ldots, x_r$, it has a (unique) K-algebra endomorphism which acts as $x_i \mapsto x_i^p$ on each of these free generators; let us write it as $(p): v \mapsto v^{(p)}$.

Note that $v^{(p)}$ is not v^p unless v is a linear combination of monomials with coefficients from the prime subfield of K, and that $\ker(p) = 0$. It is immediate to check that (p) intertwines the 'twisted' action of G on S, that is to say its action on S^F, with the natural action on S. Differently put: viewed as a map from S^F to S, (p) is a KG-homomorphism. More generally, if U is a submodule of S^k, the relevant restriction of (p) is a KG-isomorphism $U^F \to U^{(p)}$. In particular, (p) maps S^k onto the subspace we previously denoted by $(S^k)^{(p)}$ (this is the promised motivation for that notation), and by restriction yields a KG-isomorphism $(S^k)^F \to (S^k)^{(p)}$. We need one more fact: if U and V are submodules of S, then $U^{(p)}V^{(p)} = (UV)^{(p)}$ (because (p) is an algebra homomorphism) and there is a commutative diagram of surjective KG-homomorphisms

$$\begin{tikzcd} U^F \otimes V^F \arrow[r, "\varphi^F"] \arrow[d, "(p)\otimes(p)"'] & (UV)^F \arrow[d, "(p)"] \\ U^{(p)} \otimes V^{(p)} \arrow[r, "\chi"'] & (UV)^{(p)} \end{tikzcd}$$

in which φ and χ are the usual maps (which just forget the $\otimes$ sign) while the vertical maps are isomorphisms, and this yields an isomorphism $(\ker\varphi)^F \to \ker\chi$.

In view of $(S^k)^{(p)} \cong (S^k)^F$, if $k > p$ then Lemma 4.6 applied to S^k instead of S^d yields a filtration of $(S^k)^F$ with quotients $(\overline{S^{k-\ell p}})^F \otimes ((S^\ell)^{(p)})^F$. Inserting terms

$$S^{d-kp-p}\big(S^{k+1}\big)^{(p)} + S^{d-kp}\big(S^{k-\ell p}\big)^{(p)}\big(\big(S^\ell\big)^{(p)}\big)^{(p)}$$

with $\ell = 1, \dots, \lfloor k/p \rfloor$ into the filtration of S^d leads to a filtration with quotients

$$\overline{S^{d-kp}} \otimes \big(\overline{S^{k-\ell p}}\big)^F \otimes \big(S^\ell\big)^{F^2}.$$

Repeated moves like this eventually lead to a refined filtration of S^d whose quotients are of the form

$$\overline{S^\alpha} := \bigotimes_{j \geqslant 0} \big(\overline{S^{\alpha(j)}}\big)^{F^j}$$

where α ranges through the p-partitions of d that satisfy (2.2.2). By Steinberg's theorem, these twisted tensor products are simple, so the refined filtration is in fact a composition series. In this series, $\overline{S^\alpha}$ lies above $\overline{S^\beta}$ if and only if $\alpha(j) > \beta(j)$ for the smallest j at which α and β differ. Of course

$$\dim \overline{S^\alpha} = \prod_{j \geqslant 0} \dim \overline{S^{\alpha(j)}},$$

so $\dim \overline{S^\alpha}$ can be calculated using (4.2).

We shall use also the following.

Lemma 4.7. *The module S^d is simple if and only if either $d < p$ or $r = 2$ and $d = kp^m - 1$ with $0 < k < p$ and $m \geqslant 0$.*

Proof. We have seen that S^d is simple whenever $d < p$. Conversely, if S^d is simple, then all the nonzero terms of the filtration in Lemma 4.6 must be equal. This says nothing when there is only one such term, that is, when $d < p$, but otherwise it requires $\lfloor d/p \rfloor = \lceil (d-(p-1)r)/p \rceil$, which can only

hold when $r = 2$. It would not take long to complete the present argument, but for that case the result can be found in [23], immediately after Theorem (1.3). □

5. Composition factors for metabelian Lie powers

Let p and K be as above, let M be a free metabelian Lie algebra freely generated by $x_1, \dots, x_r$ (with $r \geqslant 2$), and let M^d denote the homogeneous component of degree d in M. We identify M^1 with the natural G-module E, extend the action of G to action on M by Lie algebra automorphisms, and regard the M^d as G-modules accordingly. Of course, then $M^2 \cong E^{\wedge 2}$. Our aim in this section is to identify the isomorphism types of the other composition factors of M^d, and so prove that M^d is multiplicity-free. The simple particular cases discussed in Section 2.3 will become apparent in the process.

As we did for S^d in the previous section, here we shall construct a filtration for M^d and show that the filtration quotients are simple and have pairwise different highest weights. Whenever convenient, we take advantage of the freedom obtained in Section 3.6 to assume that $r \geqslant d$.

In this section, our principal tool is a special case of an exact sequence which appeared as (2.5) in [15]; the special case ends in

$$\cdots \to E^{\wedge 3} \otimes S^{d-3} \to M^2 \otimes S^{d-2} \to E \otimes S^{d-1} \to S^d \to 0.$$

As was observed in Corollary 3.2 there, the tail yields

$$0 \to M^d \to E \otimes S^{d-1} \xrightarrow{\rho} S^d \to 0 \tag{5.1}$$

and so the beginning gives

$$\cdots \to E^{\wedge 3} \otimes S^{d-3} \to M^2 \otimes S^{d-2} \to M^d \to 0.$$

Of course if $r = 2$ then $E^{\wedge 3} = 0$ and $\dim M^2 = 1$, so we conclude that

if $r = 2$, then $M^d = M^2 \otimes S^{d-2}$
and the submodule lattice of M^d is isomorphic to that of S^{d-2}. (5.1′)

For the rest of this section, we take (5.1) as the characterization of M^d, never referring to M again and always assuming $d \geqslant 2$. The first result claimed in Section 2.3 is immediately at hand: in view of the simplest conclusion drawn from the second application of the Schur functor in Section 3.9, if $d < p$ or $d = p = 2$, then M^d is simple. (This is one of the points where we assume that $r \geqslant d$.) From [15, (2.2)] we also know that

$$\nu : S^d \to E \otimes S^{d-1}, \quad u \mapsto \sum_i x_i \otimes (\partial u / \partial x_i)$$

is a KG-homomorphism such that $\nu\rho$ is multiplication by d. It is easy to see that if $d \equiv 0$ (all congruences in this section will be mod p) then $\ker \nu$ is exactly $(S^{d/p})^{(p)}$, the last term of the filtration of S^d. Consequently,

if $d \not\equiv 0$, then $E \otimes S^{d-1} = S^d\nu \oplus M^d$ and $S^d\nu \cong S^d$, while
if $d \equiv 0$, then $S^d\nu$ is a submodule of M^d isomorphic to $S^d/(S^{d/p})^{(p)}$. (5.2)

Two particular cases of this will be especially useful:

$$\textit{if } d \equiv 1\textit{, then } E \otimes \left(S^{\lfloor d/p \rfloor}\right)^{(p)} \textit{ is } \left(E\left(S^{\lfloor d/p \rfloor}\right)^{(p)}\right)\nu \textit{ and hence avoids } M^d; \tag{5.2$'$}$$

$$\textit{if } d \equiv 0 \textit{ and } \overline{S^d} \neq 0\textit{, then } \overline{S^d} \textit{ is a composition factor of } M^d. \tag{5.2$''$}$$

Our next step is to prove the claims made in Section 2.3 about M^p. We have already covered this when $p = 2$, so assume for the moment that $p > 2$. Recalling that $\overline{S^p}$ is never 0 (because $r \geqslant 2$), we see that $S^p\nu$ is always a simple submodule of M^p. All but one of the dominant weights of $E \otimes S^{p-1}$ are p-restricted, the exception being (p); this has multiplicity 1, and is also a weight of S^p. It follows that M^p has only p-restricted dominant weights, and therefore the Schur functor maps its submodule lattice isomorphically: thus if $r \geqslant p$ then $M^p/S^p\nu$ is simple. To examine the dominant weights of $M^p/S^p\nu$, note first that $(p-2, 1^2)$ occurs more than once in M^p but only once in $S^p\nu$, so it must occur also in the quotient. Only two dominant weights of M^p are larger than $(p-2, 1^2)$, namely $(p-1, 1)$ and $(p-2, 2)$, but each of these has multiplicity 1 and occurs also in the submodule, so neither of them can occur in the quotient. This proves that the highest weight of the quotient is $(p-2, 1^2)$, a partition with precisely 3 non-zero parts. The rank changing functor of Section 3.6 now yields that $r \geqslant p$ can be relaxed to $r \geqslant 3$, but M^p is simple if $r = 2$. This settles the debts about M^p incurred in Section 2.3.

For the rest of this section, we need concern ourselves only with the case $d > p$. Then we have a commutative diagram of G-homomorphisms

$$\begin{tikzcd}
E \otimes S^{d-1-p} \otimes E^{(p)} \arrow[r, "\varphi"] \arrow[d, "\zeta"'] & E \otimes S^{d-1} \arrow[d, "\eta"] \\
S^{d-p} \otimes E^{(p)} \arrow[r, "\chi"'] & S^d
\end{tikzcd}$$

with surjective ζ and η. Apply the Half-a-snake Lemma with this input: the resulting commutative diagram with exact rows may be written as

$$\begin{tikzcd}
0 \arrow[r] & E \otimes S^{d-1-p}E^{(p)} \arrow[r] \arrow[d, "\zeta'"'] & E \otimes S^{d-1} \arrow[r] \arrow[d, "\eta"] & E \otimes \overline{S^{d-1}} \arrow[r] \arrow[d, "\eta'"] & 0 \\
0 \arrow[r] & S^{d-p}E^{(p)} \arrow[r] & S^d \arrow[r] & \overline{S^d} \arrow[r] & 0
\end{tikzcd}$$

while the resulting exact sequence is

$$0 \to M^d \cap \left(E \otimes S^{d-1-p}E^{(p)}\right) \to M^d \to N^d \to 0$$

where N^d is defined by the exact sequence

$$0 \to N^d \to E \otimes \overline{S^{d-1}} \xrightarrow{\eta'} \overline{S^d} \to 0. \tag{5.3}$$

We do not define N^d unless $d > p$.

It will be useful to note that $N^{p+1} = M^{p+1}$ (for, by (5.2$'$), if $d = p + 1$ then the left term of the exact sequence two lines above (5.3) is 0).

Of course if $d > (p-1)r+1$ then $N^d = 0$ (because then $\overline{S^{d-1}} = 0$); we also have that $N^{(p-1)r+1} \cong E \otimes \overline{S^{(p-1)r}}$ so $N^{(p-1)r+1}$ is simple of dimension r (because, as we have seen, $\dim \overline{S^{(p-1)r}} = 1$). In view of (5.3), if all we want to know is that $N^{(p-1)r+1} > 0$, it suffices to check that

$$x_1 \otimes x_1^{p-1} x_2^{p-1} \prod_{i=3}^{r} x_i^{p-1} - x_2 \otimes x_1^p x_2^{p-2} \prod_{i=3}^{r} x_i^{p-1}$$

lies in $M^{(p-1)r+1}$ but not in $E \otimes S^{(p-1)r-p} E^{(p)}$. One can show similarly that $N^d > 0$ whenever $p < d \leqslant (p-1)r+1$. For a general formula, (2.2.4) and (5.3) yield

$$\begin{aligned} \dim N^d &= r \dim \overline{S^{d-1}} - \dim \overline{S^d} \\ &= \sum_{k=0}^{r} (-1)^k \binom{r}{k} \left[r \binom{d-1-kp+r-1}{r-1} - \binom{d-kp+r-1}{r-1} \right] \\ &= \sum_{k=0}^{r} (-1)^k \binom{r}{k} (d-kp-1) \binom{d-kp+r-2}{r-2}. \end{aligned} \tag{5.4}$$

When $d \equiv 0$ and $p < d \leqslant (p-1)r$, one has $S^d \nu \nleqslant E \otimes (S^{d-p-1} E^{(p)})$ but $(S^{d-p} E^{(p)})\nu \leqslant E \otimes (S^{d-p-1} E^{(p)})$, so N^d has a simple submodule, U say, isomorphic to $\overline{S^d}$.

It is immediate from (5.3) and the restriction $d > p$ that all dominant weights of $E \otimes \overline{S^{d-1}}$ are p-restricted. As we remarked in Section 3.9, if $r \geqslant d$ then this implies that the Schur functor yields a lattice isomorphism from the lattice of KG-submodules of $E \otimes \overline{S^{d-1}}$ onto the lattice of $K\Sigma_d$-submodules of the ω-weight space of $E \otimes \overline{S^{d-1}}$, and from the rest of the discussion there we may therefore conclude that N^d is either simple or uniserial of length 2, depending on whether $d \equiv 0$. In fact, the highest weight of $E \otimes \overline{S^{d-1}}$ is μ^d, which is not a weight of $\overline{S^d}$, so μ^d must be the highest weight of N^d, and also the highest weight of N^d/U when $d \equiv 0$. The rank changing functor of Section 3.6 ensures that this remains the situation as long as r is at least as large as the number of non-zero parts in μ^d, and this condition comes to $d \leqslant (p-1)r+1$. We have proved the following.

Lemma 5.5. *If $d > (p-1)r+1$, then $N^d = 0$.*

If $p < d \leqslant (p-1)r+1$, then N^d has a simple quotient $L(\mu^d)$; in fact, $N^d \cong L(\mu^d)$ except when $p \mid d$ and $d \neq (p-1)r+1$, in which case N^d is uniserial of length 2, with socle $L(\sigma^d)$. $\square$

We noted after (5.3) that $N^{p+1} = M^{p+1}$, so this confirms the claim in Section 2.3 that M^{p+1} is always simple. We also noted that $\dim N^{(p-1)r+1} = r$, while (5.3) and Lemma 5.5 together yield that if $2 \leqslant d \leqslant (p-1)r+1$ then $\dim L(\mu^d)$ is either $r \dim \overline{S^{d-1}} - \dim \overline{S^d}$ or $r \dim \overline{S^{d-1}} - 2 \dim \overline{S^d}$, depending on whether $p \mid d$: this confirms (2.4.2).

We met N^d as the top quotient of the filtration of M^d given by

$$\begin{aligned} M^d \geqslant M^d \cap \big(E \otimes S^{d-p-1} E^{(p)}\big) \geqslant \cdots \geqslant M^d \cap \big(E \otimes S^{d-kp-1} \big(S^k\big)^{(p)}\big) \\ \geqslant \cdots \geqslant M^d \cap \big(E \otimes S^{d-\lfloor d/p \rfloor p-1} \big(S^{\lfloor d/p \rfloor}\big)^{(p)}\big) \geqslant 0. \end{aligned}$$

Most other quotients of this filtration can be identified similarly by showing that, as long as $d - kp > p$, there is an exact sequence

$$\begin{aligned} 0 &\to M^d \cap \big(E \otimes S^{d-(k+1)p-1} \big(S^{k+1}\big)^{(p)}\big) \\ &\to M^d \cap \big(E \otimes S^{d-kp-1} \big(S^k\big)^{(p)}\big) \to N^{d-kp} \otimes \big(S^k\big)^{(p)} \to 0. \end{aligned}$$

For application in Section 7, we prove here a more general result: if $c > p$, then for each submodule W of S^k there is an exact sequence

$$\begin{aligned} 0 &\to M^{c+kp} \cap \bigl(E \otimes S^{c-p-1}(EW)^{(p)}\bigr) \\ &\to M^{c+kp} \cap \bigl(E \otimes S^{c-1}W^{(p)}\bigr) \to N^c \otimes W^{(p)} \to 0. \end{aligned} \tag{5.6}$$

The input for the relevant application of the Half-a-snake Lemma is

$$\begin{tikzcd} E \otimes S^{c-p-1} \otimes (EW)^{(p)} \arrow[r, "\varphi"] \arrow[d, "\zeta"'] & E \otimes S^{c-1}W^{(p)} \arrow[d, "\eta"] \\ S^{c-p} \otimes (EW)^{(p)} \arrow[r, "\chi"'] & S^cW^{(p)} \end{tikzcd}$$

In proving (4.3), we argued that the homomorphism

$$S^{c-p} \otimes S^{(k+1)p} \to S^{c+kp}$$

maps $S^{c-p} \otimes (EW)^{(p)}$ into $S^cW^{(p)}$, and that the corresponding restriction χ has image $S^{c-p}(EW)^{(p)}$ and cokernel $\overline{S^c} \otimes W^{(p)}$. One obtains φ similarly. The obvious map ζ is surjective, and (5.6) is part of the conclusion of this application of the Half-a-snake Lemma.

A degenerate version of this argument shows that when $2 \leqslant c < p$,

$$M^{c+kp} \cap \bigl(E \otimes S^{c-1}W^{(p)}\bigr) \cong M^c \otimes W^{(p)}. \tag{5.6$'$}$$

When $d \not\equiv 0, 1$, this and (5.6), both with $W = S^k$, deal with all quotients of the filtration. When $d \equiv 1$, we are left with the last term of the filtration, but that is 0, as we have seen in (5.2$'$).

When $d \equiv 0$, the last term is obviously 0, and (5.6) deals with all but the second last term, namely $M^d \cap (E \otimes S^{p-1}(S^{(d/p)-1})^{(p)})$. We give a two-step filtration of this, with a view to refining the original filtration of M^d by adding a term at the lower end. For notational ease, what we show is that for each positive k there is an exact sequence

$$0 \to M^p \otimes \bigl(S^k\bigr)^{(p)} \to M^{(k+1)p} \cap \bigl(E \otimes S^{p-1}\bigl(S^k\bigr)^{(p)}\bigr) \to \bigl(M^{k+1}\bigr)^F \to 0. \tag{5.7}$$

For the first step in the proof, consider the commutative diagram with exact rows

$$\begin{tikzcd} 0 \arrow[r] & E^{(p)} \otimes (S^k)^{(p)} \arrow[r] \arrow[d] & S^p \otimes (S^k)^{(p)} \arrow[r] \arrow[d] & \overline{S^p} \otimes (S^k)^{(p)} \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & (S^{k+1})^{(p)} \arrow[r] & S^p(S^k)^{(p)} \arrow[r] & \overline{S^p} \otimes (S^k)^{(p)} \arrow[r] & 0 \end{tikzcd}$$

where the first row comes by setting $d = p$ in (4.1) and then tensoring the result with $(S^k)^{(p)}$, the second comes by setting $c = p$ and $W = S^k$ in (4.4), while the last vertical map is the relevant identity map. By (5.1) and the discussion of the Frobenius functor after Lemma 4.6, the kernel of the first vertical map is isomorphic to $(M^{k+1})^F$, and therefore (by the Snake Lemma) so is the kernel of the second: we get an exact sequence

$$0 \to \bigl(M^{k+1}\bigr)^F \to S^p \otimes \bigl(S^k\bigr)^{(p)} \to S^p\bigl(S^k\bigr)^{(p)} \to 0. \tag{5.8}$$

When W is a submodule of S^k, we also have a commutative diagram with exact rows

$$\begin{tikzcd}
0 \arrow[r] & M^p \otimes W^{(p)} \arrow[r] \arrow[d, "\varphi"'] & E \otimes S^{p-1} \otimes W^{(p)} \arrow[r] \arrow[d, "\chi"'] & S^p \otimes W^{(p)} \arrow[r] \arrow[d, "\psi"'] & 0 \\
0 \arrow[r] & M^{(k+1)p} \cap (E \otimes S^{p-1} W^{(p)}) \arrow[r] & E \otimes S^{p-1} W^{(p)} \arrow[r] & S^p W^{(p)} \arrow[r] & 0
\end{tikzcd}$$

in which χ is an isomorphism coming from (4.3), so φ is injective. By the Snake Lemma, $\operatorname{coker}\varphi \cong \ker\psi$. Since ψ is a restriction of one of the maps of (5.8), we know that $\ker\psi$ is isomorphic to a submodule of $(M^{k+1})^F$, and is equal to $(M^{k+1})^F$ when $W = S^k$. This has not only proved (5.7) but also the existence of an exact sequence

$$0 \to M^p \otimes W^{(p)} \to M^{(k+1)p} \cap \bigl(E \otimes S^{p-1} W^{(p)}\bigr) \to \bigl(M^{k+1}\bigr)^F \tag{5.7′}$$

for each W.

When $d \equiv 0$, after this refinement we have a filtration with top quotient N^d, one quotient $(M^{d/p})^F$, and all others quotients written as two-factor tensor products. Under the assumption $r \geqslant d$, we see from Lemma 5.5 that N^d and each of the left hand tensor factors are two-step uniserials. Accordingly, one can insert extra terms into the filtration, one between each pair of successive terms except the pair with quotient $(M^{d/p})^F$.

Using also the information provided by Lemma 5.5 for the case $d \not\equiv 0$, we conclude the following.

Lemma 5.9. *When $p < d \leqslant r$, the $M^d \cap (E \otimes S^{d-1-kp}(S^k)^{(p)})$ lead to a filtration of M^d such that if $d \not\equiv 0, 1 \pmod p$ then the filtration quotients are isomorphic to*

$$\begin{gathered}
L\bigl(\mu^d\bigr) \\
\vdots \\
L\bigl(\mu^{d-kp}\bigr) \otimes \bigl(S^k\bigr)^F \\
\vdots \\
L\bigl(\mu^{d-\lfloor d/p\rfloor p+p}\bigr) \otimes \bigl(S^{\lfloor d/p\rfloor-1}\bigr)^F \\
L\bigl(\mu^{d-\lfloor d/p\rfloor p}\bigr) \otimes \bigl(S^{\lfloor d/p\rfloor}\bigr)^F
\end{gathered}$$

while if $d \equiv 1 \pmod p$ they are

$$\begin{gathered}
L\bigl(\mu^d\bigr) \\
\vdots \\
L\bigl(\mu^{d-kp}\bigr) \otimes \bigl(S^k\bigr)^F \\
\vdots \\
L\bigl(\mu^{2p+1}\bigr) \otimes \bigl(S^{(d-2p-1)/p}\bigr)^F \\
L\bigl(\mu^{p+1}\bigr) \otimes \bigl(S^{(d-p-1)/p}\bigr)^F
\end{gathered}$$

and if $d \equiv 0 \pmod p$ *then we have*

$$\begin{gathered}
L(\mu^d)\\
L(\sigma^d)\\
\vdots\\
L(\mu^{d-kp}) \otimes (S^k)^F\\
L(\sigma^{d-kp}) \otimes (S^k)^F\\
\vdots\\
L(\mu^{2p}) \otimes (S^{(d/p)-2})^F\\
L(\sigma^{2p}) \otimes (S^{(d/p)-2})^F\\
(M^{d/p})^F\\
L(\mu^{p}) \otimes (S^{(d/p)-1})^F\\
L(\sigma^{p}) \otimes (S^{(d/p)-1})^F.
\end{gathered}$$

(*When* $p = 2$, *the case* $d \not\equiv 0, 1$ *does not arise and the second last filtration quotient* $L(\mu^p) \otimes (S^{(d/p)-1})^F$ *must be omitted.*) □

We note in passing that these filtrations are p-good in the sense of [3].

Think of the top filtration quotients as $L(\mu^d) \otimes (S^0)^F$ and $L(\sigma^d) \otimes (S^0)^F$, and think of $(M^{d/p})^F$ as $L(0) \otimes (M^{d/p})^F$, so that each filtration quotient is written as a two-factor tensor product. All left hand tensor factors are simple with pairwise distinct p-restricted highest weights, while the right hand tensor factors are twisted versions of smaller symmetric or metabelian powers. This property hands us the main qualitative result of the paper.

Theorem 5.10. *Each* M^d *is multiplicity-free.*

Proof. Taking advantage of the freedom obtained in Section 3.6, we assume that $d \leqslant r$, and argue by induction on d. We have seen in detail that M^d is multiplicity-free when $d \leqslant p + 1$, so we are assured of several initial steps. Once $d > p$, Lemma 5.9 is available. In the inductive step, the right hand tensor factors of those filtration quotients are all multiplicity-free: the symmetric powers because their weight spaces are 1-dimensional, and the one metabelian power by the obvious induction hypothesis. Thus Steinberg's theorem (discussed in Section 3.8) proves what we want.

Let us see in a little more detail just how this works. Given a non-zero filtration quotient $U \otimes V^F$ (with U simple with p-restricted highest weight and V multiplicity-free), take any composition series of V and use this to obtain a filtration of $U \otimes V^F$. The quotients of this filtration of $U \otimes V^F$ are the $U \otimes W^F$ as W ranges through the composition factors of V in the chosen composition series. By Steinberg's theorem, each $U \otimes W^F$ is simple, so the filtration so obtained for $U \otimes V^F$ is a composition series. Since V is multiplicity-free, that theorem also gives that no two of the $U \otimes W^F$ can be isomorphic. Composition series of the filtration quotients $U \otimes V^F$ combine to yield a composition series of M^d; quotients of this coming from different filtration quotients cannot be isomorphic, because the first tensor factors of their unique twisted tensor factorizations differ. This completes the proof. □

The quantitative version is also at hand.

Theorem 5.11. *If $r \geqslant d \geqslant 2$, the highest weights of the composition factors of M^d are the μ^α and the σ^α where α ranges through those p-partitions of d which satisfy*

$$\alpha(t_\alpha) \geqslant \begin{cases} 3 & \text{when } p = 2, \\ 2 & \text{when } p > 2 \end{cases} \tag{2.5.1}$$

in the case of the μ^α and

$$\alpha(t_\alpha) \equiv 0 \tag{2.5.2}$$

in the case of the σ^α.

In general, one must impose further restrictions as well, namely

$$\begin{aligned} &\alpha(t_\alpha) \leqslant (p-1)r + 1, \\ &\alpha(t_\alpha) \neq p \quad \text{when } r = 2, \quad \text{and} \\ &\alpha(j) \leqslant (p-1)r \quad \text{when } j \neq t_\alpha \end{aligned} \tag{2.5.1$'$}$$

in the case of the μ^α, and

$$\alpha(j) \leqslant (p-1)r \quad \text{for } j = 0, 1, \ldots \tag{2.5.2$'$}$$

in the case of the σ^α.

Proof. First we deal with the case $r \geqslant d$, using induction on d.

We have seen the composition factors when $d \leqslant p$ and can check that the claim holds under that assumption, so we have the initial step and can turn to $d > p$. For the inductive step, Lemma 5.9 is then available, and all we have to do is some straightforward book-keeping.

Let us start with a rough count. From the obvious fact that $\alpha(0) \equiv |\alpha|$ for all p-partitions, it follows that as α ranges through the p-partition of d, $\alpha(0)$ ranges through the $d - kp$ with $k = 0, \ldots, \lfloor d/p \rfloor$ while α_+ ranges through the p-partitions β of k. Accordingly, $\mu^\alpha = \mu^{\alpha(0)} + p\sigma^{\alpha_+}$ ranges through the $\mu^{d-kp} + p\sigma^\beta$, which are the highest weights of the composition factors of the $L(\mu^{d-pk}) \otimes (S^k)^F$. Similarly, $\sigma^\alpha = \sigma^{\alpha(0)} + p\sigma^{\alpha_+}$ ranges through the $\sigma^{d-kp} + p\sigma^\beta$, which are the highest weights of the composition factors of the $L(\sigma^{d-pk}) \otimes (S^k)^F$.

When $d \not\equiv 0, 1$, we must have $p > 2$ and $d - kp \geqslant 2$, so the rough count has done the job without any need for induction.

When $d \equiv 1$, the smallest value of $d - kp$ in the filtration provided in Lemma 5.9 is $p + 1$; in view of $t_\alpha = 0$ and $\alpha(0) \equiv d \equiv 1$, this is also the smallest value for $\alpha(t_\alpha)$ in the theorem: so again the rough count has done the job.

Similarly, one can see without induction that when $d \equiv 0$, the rough count has matched the μ^α and σ^α with $t_\alpha = 0$ to the composition factors of the filtration quotients other than $(M^{d/p})^F$ (though verifying the exceptional behaviour in case $p = 2$ takes some extra attention).

The inductive hypothesis will only be needed in the remaining case: matching the composition factors of $(M^{d/p})^F$ to the μ^α and σ^α with $t_\alpha > 0$, that is, with $\alpha(0) = 0$. As α ranges through these p-partitions of d, α_+ ranges through the p-partitions of d/p, and of course $\mu^\alpha = p\mu^{\alpha_+}$ and $\sigma^\alpha = p\sigma^{\alpha_+}$. By the inductive hypothesis, we have to select from the p-partitions of d/p those which satisfy the conditions imposed by the theorem; as here $\alpha(t_\alpha) = \alpha_+(t_{\alpha_+})$, an α_+ satisfies these conditions if an only if α does.

This completes the proof for the case $r \geqslant d$.

In view of Section 3.6, in the general case all that has to change is that one has to exclude the μ^α and the σ^α which have more than r (non-zero) parts. First, this requires excluding the μ^α such that $\mu^{\alpha(t_\alpha)}$ has too many parts, that is, those with $\alpha(t_\alpha) > (p-1)r + 1$, and if $r = 2$ then also those with $\alpha(t_\alpha) = p$. Second, one has to exclude μ^α if, for some $j \neq t_\alpha$, $\mu^{\alpha(j)}$ has too many parts, in

other words, if $\alpha(j) > (p-1)r$. Third, one has to exclude the σ^α such that some $\sigma^{\alpha(j)}$ has too many parts, which happens if $\alpha(j) > (p-1)r+1$. These are precisely the cases excluded by (2.5.1′) and (2.5.2′). □

In Section 2.3, we presented a list of the simple M^d, and claimed that it was complete. Most of that claim has been proved by now (the case $r=2$ in (5.1′) and Lemma 4.7); what is yet to be proved is the following.

Lemma 5.12. *Of the M^d with $d > p+1$ and $r > 2$, precisely one is simple, namely that with $d=5$, $p=2$, $r=3$.*

Proof. If such an M^d is simple, then all but the lowest filtration quotients discussed in Lemma 5.9 must vanish. If $d \not\equiv 0, 1$, then the second last filtration quotient has $L(p, d - \lfloor d/p \rfloor p) \otimes L(\lfloor d/p \rfloor - 1)^F$ as a non-zero composition factor. If $d \equiv 0$, then $(M^{d/p})^F \neq 0$. If $d \equiv 1$ and $p > 2$, then the second last filtration quotient has $L(p, p-1, 2) \otimes L((d-2p-1)/p)^F$ as a non-zero composition factor. If $d \equiv 1$, $p=2$ and $r > 3$, then the second last filtration quotient has $L(2,1,1,1) \otimes L((d-5)/2)^F$ as a non-zero composition factor. Finally, if $d \equiv 1$, $p=2$ and $r=3$, then Lemma 4.7 yields that $M^d = L(2,1) \otimes (S^{(d-3)/2})^F$: this is simple if and only if $d=5$. □

The main outstanding task is to describe, in terms of the isomorphism types so identified, the inclusion order of the corresponding join-irreducible submodules. For that, we need to adapt results from the 1975 paper [4] of Bakhturin. (While some of that paper was incorporated in his book [5], much of what we need here was not.) It is only with reference to this partial order that we shall be able to show that each M^d is join-irreducible.

6. From Bakhturin's results

Until the last paragraphs of this section, let M denote the free metabelian Lie algebra of countably infinite rank over an infinite field K of characteristic p. It is easy to see that $[M, M]$ contains all fully invariant ideals of M other than M itself. Bakhturin [4] selected certain canonical words (which he called canonical identities) in $[M, M]$, and in Lemma 3 of [4] showed that each fully invariant ideal contained in $[M, M]$ is generated by the canonical words it contains (and also by some finite subset of such words). Lemma 5 of [4] says in effect that if a canonical word lies in a sum of fully invariant ideals then it must lie in one of the summands. Consequently, sums and intersections of fully invariant ideals match unions and intersections of the sets of canonical words they contain, so the lattice of these ideals is distributive. It also follows that a fully invariant ideal is join-irreducible as element of this lattice if and only if it is generated by a single canonical word. Lemma 4 of [4] provides the key to deciding whether one canonical word lies in the fully invariant ideal generated by the other. This yields that different canonical words generate different fully invariant ideals, and yields the partial order on the set of canonical words that reflects how the join-irreducible fully invariant ideals they generate are compared by inclusion.

We had several problems with interpreting the intentions of [4], but leave discussion to Appendix A, concentrating here on a paraphrase of the conclusions that we intend to use.

The canonical words were described in terms of a free generating set $\{x_1, x_2, \ldots\}$ using (in disguise) p-partitions like those we saw in the context of symmetric powers, and the partial order again involved refinement of p-partitions. This time not all p-partitions were used, only those in which 'the smallest part is not alone': that is to say, if t_α is the smallest number such that $\alpha(t_\alpha) > 0$, then $\alpha(t_\alpha) \geqslant 2$. *This restriction will apply throughout the present section.* For each such p-partition α, define the *Bakhturin monomial* a^α as the (17) of [4], namely the left-normed Lie product with leftmost factor x_1, next factor x_2, and altogether $p^{k(i)}$ factors x_i where $k(i)$ is defined by the condition

$$\sum_{j<k(i)} \alpha(j) < i \leqslant \sum_{j \leqslant k(i)} \alpha(j).$$

There is no ambiguity here, because in a left-normed Lie product in a metabelian Lie algebra the order of the factors beyond the second place does not matter. Note that this monomial involves precisely the free generators x_i with $i \leqslant \sum_j \alpha(j)$, and that its total degree is given by $d = \sum_j \alpha(j)p^j$.

In the case $d \not\equiv 0$, the canonical words of degree d were simply the Bakhturin monomials a^α. Note that in this case one always has $t_\alpha = 0$ (because $\alpha(0) \equiv d \not\equiv 0$ so $\alpha(0) \neq 0$), confirming that the present definition agrees with what we gave in Section 2.4.

In the case $d \equiv 0$, there were other canonical words of degree d as well, namely the (18) of [4]: here we write these as b^α and call them *Bakhturin sums*. The p-partitions indexing these are subject to the condition $\alpha(t_\alpha) \equiv 0$. The definition is that b^α is a sum of $\alpha(t_\alpha) - 1$ monomials, one of which is the Bakhturin monomial a^α and the others are the monomials obtained from that by changing the order of its factors, keeping x_1 in first place but bringing into second place one of $x_3, x_4, \ldots, x_{\alpha(t_\alpha)}$. When $\alpha(t_\alpha) = p = 2$, this sum has only one summand, so $\mathsf{a}^\alpha = \mathsf{b}^\alpha$. Having two names for one word caused some confusion in [4], so from now on *whenever we write down* a^α *or* a^β, *we exclude* $\alpha(t_\alpha) = p = 2$ *or* $\beta(t_\beta) = p = 2$.

The hard technical result of [4] was the determination of the relevant partial order $\leqslant$ on the set of canonical words. From this, here we shall only need to know how canonical words with a common degree d compare. Recall from Section 2.1 that for p-partitions of a given d, refinement order is easy to test: β is a refinement of α (notation: $\alpha \preccurlyeq \beta$) if and only if $\alpha_{<k} \leqslant \beta_{<k}$ for $k = 1, 2, \ldots$ (as there, $\alpha_{<k}$ denotes $\sum_{j<k} \alpha(j)p^j$). Paraphrasing Bakhturin [4] in these terms, the partial order may be described as follows.

Theorem 6.1. *When $t_\alpha = t_\beta$, $\mathsf{a}^\alpha \leqslant \mathsf{b}^\beta$ can never hold, while each of the three statements $\mathsf{a}^\alpha \leqslant \mathsf{a}^\beta$, $\mathsf{b}^\alpha \leqslant \mathsf{a}^\beta$, $\mathsf{b}^\alpha \leqslant \mathsf{b}^\beta$ is equivalent to $\alpha \preccurlyeq \beta$.*

When $t_\alpha > t_\beta$, neither $\mathsf{a}^\alpha \leqslant \mathsf{b}^\beta$ nor $\mathsf{b}^\alpha \leqslant \mathsf{b}^\beta$ can ever hold, while each of $\mathsf{a}^\alpha \leqslant \mathsf{a}^\beta$ and $\mathsf{b}^\alpha \leqslant \mathsf{a}^\beta$ is equivalent to the conjunction of $\alpha \preccurlyeq \beta$ and $\beta_{<k} \geqslant 2p^k$ whenever $t_\alpha \geqslant k > t_\beta$.

When $t_\alpha < t_\beta$, neither $\mathsf{a}^\alpha \leqslant \mathsf{a}^\beta$ nor $\mathsf{a}^\alpha \leqslant \mathsf{b}^\beta$ can ever hold, while each of the two statements $\mathsf{b}^\alpha \leqslant \mathsf{a}^\beta$, $\mathsf{b}^\alpha \leqslant \mathsf{b}^\beta$ is equivalent to the conjunction of

$$\alpha(i) = \begin{cases} 0 & \text{if } i < t_\alpha, \\ p & \text{if } i = t_\alpha, \\ p-1 & \text{if } t_\alpha < i < t_\beta, \\ \beta(t_\beta) - 1 & \text{if } i = t_\beta \end{cases}$$

and

$$\alpha_{<k} \leqslant \beta_{<k} \quad \text{whenever } k > t_\beta.$$

How is this relevant in our present context? As usual, we regard M as graded in terms of the given free generating set $\{x_1, x_2, \ldots\}$. Number the terms $\gamma_d(M)$ of the lower central series of M so that $\gamma_d(M) = M^d \oplus \gamma_{d+1}(M)$, and write Γ for the multiplicative monoid of the K-endomorphisms of M^1. Each K-endomorphism of M^1 extends uniquely to a K-algebra endomorphism of M and that extension leaves each homogeneous component M^d setwise invariant, so each M^d is a $K\Gamma$-module. Each subspace V of $\gamma_d(M)$ which contains $\gamma_{d+1}(M)$ is an ideal, and it is fully invariant in M if and only if $V \cap M^d$ is a $K\Gamma$-submodule. Since we are interested in submodules of one M^d at a time, we may change first from M to $M/\gamma_{d+1}(M)$. Then we no longer need infinite rank: all relevant arguments remain valid if the rank r our free metabelian-and-nilpotent-of-class-d Lie algebra is finite but $r \geqslant d$. Finally, the only canonical words which remain of interest are those which correspond to p-partitions of this particular d. Of course we also need that, as we saw in Section 3.4, in this context $K\Gamma$-modules and KG-modules are the same.

The conclusion we want to take from here is this.

Theorem 6.2. *Suppose $r \geqslant d \geqslant 2$, and consider the canonical words a^α and b^α discussed above, with α ranging through the p-partitions of d that satisfy*

$$\alpha(t_\alpha) \geqslant \begin{cases} 3 & \text{when } p = 2, \\ 2 & \text{when } p > 2 \end{cases} \tag{2.5.1}$$

in the case of the a^α and

$$\alpha(t_\alpha) \equiv 0 \tag{2.5.2}$$

in the case of the b^α. The join-irreducible submodules of M^d are precisely the submodules which can be generated by a single canonical word from this collection, and the inclusion-order on the set of these submodules matches the partial order on the set of canonical words described in Theorem 6.1. □

Remarks. Before moving on, let us pause to domesticate some aspects of this partial order.

Write $d = p^m q$ with $p \nmid q$. In the coprime case $m = 0$, all p-partitions α of d have $t_\alpha = 0$, there are no b^α, and the partial order on the set of the a^α matches the refinement order of the relevant α. This is the simplest case, as highlighted in the Abstract and described in Section 2.4.

Consider next the other extreme, when $q = 1$ so d is a power of p. We have seen that the odd and even cases are different at $m = 1$: this distinction prevails for larger m as well. When $p > 2$, there are $2m$ kinds of canonical words: first, the a^α are divided into m kinds, according as $t_\alpha = 0, 1, \dots, m-1$, then the b^α fall into m classes as $t_\alpha = m-1, m-2, \dots, 0$. Note that t_α traverses the same interval twice, but in opposite order, and that in the middle we have two singletons, $\{\mathsf{a}^\alpha\}$ followed by $\{\mathsf{b}^\alpha\}$, involving the same α, namely that with $\alpha(m-1) = p$ and $\alpha(j) = 0$ whenever $j \neq m-1$. When $p = 2$, the first of these singletons is missing (because for this α we ruled out a^α), so we have only $2m-1$ kinds.

When $m > 0$ and $q > 1$, there are $2m+1$ kinds of canonical words: the a^α with $t_\alpha = 0, 1, \dots, m$, and the b^α with $t_\alpha = m-1, m-2, \dots, 0$.

The point to observe is that canonical words *of the same kind* compare exactly as the p-partitions indexing them do in refinement order, while canonical words of different kinds can never compare in the wrong direction: only the first listed kind can imply the other, and sometimes it does, sometimes it does not.

Another way of putting this is to say that M^d has a filtration described informally as follows. The small terms are the submodules generated by the b^α with t_α smaller than some bound. The large terms are submodules generated by all the b^α together with those a^α for which t_α is larger than some bound. For one or two terms in the middle, one may have to adjust the definition according to the case distinctions above. Each filtration quotient has for its composition factors the simple quotients of the join-irreducible submodules generated by the canonical words of one particular kind, partially ordered according to the refinement order of the p-partitions indexing them. This says nothing positive about how composition factors of different filtration quotients compare, but it does provide some information about how they do *not* compare. The filtrations seen in Lemma 5.9 allow a similar interpretation which tells a different part of the story. However, the two parts still do not add up to the whole: these observations form only a partial description of the partial order formally described in Theorem 6.1.

7. Matching canonical words to composition factors

Theorems 5.11 and 6.2 show that (the highest weights of) the composition factors of M^d and (the canonical words generating) the join-irreducible submodules of M^d are indexed by the same p-partitions of d. We still have to confirm that this correctly matches each join-irreducible submodule with its unique simple quotient.

Theorem 7.1. *Suppose $r \geqslant d \geqslant p$. If α is a p-partition of d which satisfies*

$$\alpha(t_\alpha) \geqslant \begin{cases} 3 & \textit{when } p = 2, \\ 2 & \textit{when } p > 2, \end{cases} \tag{2.5.1}$$

then the unique simple quotient of the submodule of M^d generated by the canonical word a^α has highest weight μ^α. If instead α satisfies

$$\alpha(t_\alpha) \equiv 0, \tag{2.5.2}$$

then the unique simple quotient of the submodule generated by b^α has highest weight σ^α.

The highest weights of the composition factors of M^d were listed in Theorem 5.11. Recall that the set of these weights is naturally bijective with the set of the join-irreducible submodules, and inherits (along this bijection) a partial order reflecting the inclusion order of the latter set. The last three theorems finally make this partial order explicit: the point of Theorem 7.1 is that it shows how to replace the canonical words in Theorem 6.1 by the highest weights listed in Theorem 5.11. As was shown in Section 3.6, the partial order does not change when the restrictions (2.5.1′), (2.5.2′) are imposed.

Proof of Theorem 7.1. In Section 5, we worked with a copy of M^d in $E \otimes S^{d-1}$ without having to say just what embedding we had in mind, but now we have to know what Bakhturin's canonical words look like as elements of this tensor product. We also have to go back and recall from Section 2.2 that S^α stands for the submodule of S^d generated by the monomial

$$\mathsf{x}^\alpha = \prod_{i=1}^{\sum \alpha(j)} x_i^{p^{k(i)}}$$

where $k(i)$ is defined by the condition

$$\sum_{j<k(i)} \alpha(j) < i \leqslant \sum_{j \leqslant k(i)} \alpha(j).$$

Taking the embedding $\varphi_d : M^d \to E \otimes S^{d-1}$ from [15, Theorem 3.1], we get

$$\mathsf{a}^\alpha \varphi_d = \left(x_1 \otimes x_1^{-1}\mathsf{x}^\alpha\right) - \left(x_2 \otimes x_2^{-1}\mathsf{x}^\alpha\right)$$

and if $\alpha(t_\alpha) \equiv 0$ then

$$\mathsf{b}^\alpha \varphi_d = \sum_{i=2}^{\alpha(t_\alpha)} \left[\left(x_1 \otimes x_1^{-1}\mathsf{x}^\alpha\right) - \left(x_i \otimes x_i^{-1}\mathsf{x}^\alpha\right)\right] = -\sum_{i=1}^{\alpha(t_\alpha)} \left(x_i \otimes x_i^{-1}\mathsf{x}^\alpha\right).$$

Recall ν from (5.2), and notice that if $t_\alpha = 0$ then $\mathsf{b}^\alpha \varphi_d = -\mathsf{x}^\alpha \nu$, so in this case the submodule generated by $\mathsf{b}^\alpha \varphi_d$ is $S^\alpha \nu$, with unique simple quotient $\overline{S^\alpha} = L(\sigma^\alpha)$ as claimed.

We now subdivide the case $t_\alpha = 0$ according to the value of $\alpha(0)$. As we did in (5.6), write $d = c + kp$ with $2 \leqslant c \leqslant d$ in every possible way and focus on one of these decompositions: for the moment, consider only the p-partitions α of d with $\alpha(0) = c$. From the set of these, $\alpha \mapsto \alpha_+$ is a bijection onto the set of all p-partitions of k; call that set P_k, say. Observe that $\mathsf{a}^\alpha \varphi_d$ lies in $E \otimes S^{c-1}(S^{\alpha_+})^{(p)}$ but not in $E \otimes S^{c-1-p}(ES^{\alpha_+})^{(p)}$, so a^α has a non-zero image in the quotient of the intersections of M^d with these submodules.

Suppose first that $c > p$: by (5.6) with $W = S^{\alpha_+}$, that quotient is then isomorphic to $N^c \otimes (S^{\alpha_+})^F$, so the unique simple quotient of the submodule generated by a^α is a composition factor of this tensor product. We know that the composition factors of S^{α_+} are the $\overline{S^\beta}$ with $\beta \in P_k$ subject to $\beta \preccurlyeq \alpha_+$, so the simple quotient is an $L(\mu^c + p\sigma^\beta)$ for such a β, or possibly an $L(\sigma^c + p\sigma^\beta)$ if $d \equiv 0$. In fact, this second possibility does not arise, because the $L(\sigma^c + p\sigma^\beta)$ with $\alpha(0) > 0$ have already been accounted for, as simple quotients of the submodules generated by the b^α. We therefore have a map from P_k into itself: each element is an α_+ (with a unique α with $\alpha(0) = c$), and its image is the β such that the simple quotient of the submodule generated by a^α is $L(\mu^c + p\sigma^\beta)$. This map is one-to-one, so it is a permutation (because P_k is finite); and it is weakly decreasing (with respect to the restriction of the partial order $\preccurlyeq$). It is an elementary fact that a finite poset can never have a weakly decreasing non-identity permutation, so we have proved that the simple quotient of the submodule generated by a^α is $L(\mu^c + p\sigma^{\alpha_+})$, as required.

When $c < p$ or $c = p$, one can use (5.6′) or (5.7′) in place of (5.6) and argue similarly.

What remains to do is to deal with the α such that $t_\alpha > 0$, that is, $\alpha(0) = 0$. Of course this case does not arise unless $d \equiv 0$. As we have seen, if $\alpha(0) = 0$ then $\alpha_+(t_{\alpha_+}) = \alpha(t_\alpha)$, $\mu^\alpha = p\mu^{\alpha_+}$ and $\sigma^\alpha = p\sigma^{\alpha_+}$, so what we have to show is that $L(\mu^{\alpha_+})^F$ or $L(\sigma^{\alpha_+})^F$ is a quotient of the submodule generated by a^α or b^α, respectively.

To this end, it will be convenient to write $k = (d/p) - 1$. The first thing is to note that $\mathsf{a}^\alpha \varphi_d \in E \otimes S^{p-1}(S^k)^{(p)}$. By (4.3) with $c = p - 1$ and $W = S^k$, we have an isomorphism

$$\chi : E \otimes S^{p-1}\big(S^k\big)^{(p)} \to E \otimes S^{p-1} \otimes \big(S^k\big)^{(p)},$$

and of course there is the obvious homomorphism

$$\psi : E \otimes S^{p-1} \otimes \big(S^k\big)^{(p)} \to S^p \otimes \big(S^k\big)^{(p)}:$$

applying one after the other, we get

$$\mathsf{a}^\alpha \varphi_d \chi \psi = \big(x_1^p \otimes \big(x_1^{-1} x^{\alpha_+}\big)^p\big) - \big(x_2^p \otimes \big(x_2^{-1} x^{\alpha_+}\big)^p\big).$$

On the other hand, for the embedding $\varphi_{d/p} : M^{d/p} \to E \otimes S^{(d/p)-1}$ we have

$$\mathsf{a}^{\alpha_+} \varphi_{d/p} = \big(x_1 \otimes x_1^{-1} x^{\alpha_+}\big) - \big(x_2 \otimes x_2^{-1} x^{\alpha_+}\big),$$

so

$$\big(\mathsf{a}^{\alpha_+} \varphi_{d/p}\big)^{(p)} = \big(x_1^p \otimes \big(x_1^{-1} x^{\alpha_+}\big)^p\big) - \big(x_2^p \otimes \big(x_2^{-1} x^{\alpha_+}\big)^p\big).$$

Of course the unique simple quotient of the submodule of $S^p \otimes S^{d-p}$ generated by $(\mathsf{a}^{\alpha_+}\varphi_{d/p})^{(p)}$ is $L(\mu^{\alpha_+})^F$. In view of $(\mathsf{a}^{\alpha_+}\varphi_{d/p})^{(p)} = \mathsf{a}^\alpha \varphi_d \chi \psi$, this submodule is a homomorphic image of the submodule of M^d generated by a^α, hence this latter submodule also has $L(\mu^{\alpha_+})^F$ as a quotient. This proves the claim for a^α and $L(\mu^\alpha)$, and the case of b^α and $L(\sigma^\alpha)$ is entirely similar. □

Remark. Let us return briefly to the filtration of M^d sketched very informally at the end of Section 6. We know that if $d \equiv 0$ then $S^d \nu \leqslant M^d \varphi_d$; since φ_d is one-to-one, there exists a $\nu_d : S^d \to M^d$ such that $\nu = \nu_d \varphi_d$. With this notation, the lowest term of that filtration may now be recognized as $S^d \nu_d$. Similarly, the next filtration quotient is isomorphic to $(S^{d/p} \nu_{d/p})^F$, and so on: the lower filtration quotients of M^d are all of the form $(S^{d/p^k} \nu_{d/p^k})^{F^k}$. These are also the quotients of the chain

$$S^d > \big(S^{d/p}\big)^{(p)} > \cdots > \big(S^{d/p^m}\big)^{(p)^m} = \big(S^q\big)^{(p)^m}$$

of submodules in S^d, except there they come in the opposite order. One may also note that the composition factors or S^d that do not appear in M^d are precisely those of $(S^q)^{(p)^m}$.

It would be interesting to find some similar recognition for the higher filtration quotients of the filtration of M^d under consideration.

We now turn to the last outstanding claim.

Theorem 7.2. *Each M^d is join-irreducible.*

We already know that M^d is join-irreducible when $d \leqslant p+1$, and by (5.1′) also when $r=2$; so assume $d \geqslant p+2$ and $r \geqslant 3$. Define a p-partition γ of d by imitating the greedy algorithm that yielded β in Section 2.2. Take the largest number congruent (mod p) to d but no larger than d or $(p-1)r+1$, and call this $\gamma(0)$, noting that $d \geqslant p+2$ and $r \geqslant 3$ imply $\gamma(0) \geqslant 3$. Once the $\gamma(j)$ are defined for $j<k$, choose $\gamma(k)$ as the largest number which is congruent to $(d-\gamma_{<k})/p^k$ but is no larger than $(d-\gamma_{<k})/p^k$ or $(p-1)r$. It is easy to verify that greedy works here as well: γ satisfies (2.5.1) and (2.5.1′), and *if a p-partition α of d satisfies* (2.5.1′)*, then $\alpha \preccurlyeq \gamma$.* [For a proof of the latter, consider a counterexample α. By the maximal choice of $\gamma(0)$, we have $\alpha_{<1} \leqslant \gamma_{<1}$. Since $\alpha \not\preccurlyeq \gamma$, one can choose k so that $\alpha_{<k} \leqslant \gamma_{<k}$ but $\alpha_{<k+1} > \gamma_{<k+1}$. Since $\alpha_{<k+1} \equiv d \equiv \gamma_{<k+1} \pmod{p^{k+1}}$, with this choice in fact $\alpha_{<k+1} \geqslant \gamma_{<k+1} + p^{k+1}$, and it follows that

$$\gamma_{<k} + \alpha(k)p^k \geqslant \alpha_{<k} + \alpha(k)p^k = \alpha_{<k+1} \geqslant \gamma_{<k+1} + p^{k+1} = \gamma_{<k} + \big(\gamma(k)+p\big)p^k,$$

whence $\gamma(k)+p \leqslant \alpha(k) \leqslant (p-1)r$. Similarly,

$$d \geqslant \alpha_{<k+1} \geqslant \gamma_{<k} + \big(\gamma(k)+p\big)p^k$$

whence $\gamma(k)+p \leqslant (d-\gamma_{<k})/p^k$. Recall that $\gamma(k)$ was chosen maximal in its congruence class with respect to two upper bounds, so $\gamma(k)+p$ must violate at least one of those. We have just seen that it violates neither, and this proves that no counterexample can exist.]

Theorem 7.2 will follow once we prove our next lemma.

Lemma 7.3. *If $d \geqslant p+2$ and $r \geqslant 3$, then M^d is join-irreducible and its unique simple quotient is $L(\mu^\gamma)$ with the γ defined above.*

Proof. Call a canonical word relevant if it satisfies the conditions of Theorem 5.11. It will be sufficient to show that no a^α with $\alpha \neq \gamma$, and no b^α, can be maximal in the partial order of the relevant canonical words.

If $t_\alpha > 0$, define a p-partition α_- of d by

$$\alpha_-(j) = \begin{cases} 2p & \text{if } j = t_\alpha - 1, \\ \alpha(t_\alpha) - 2 & \text{if } j = t_\alpha, \\ \alpha(j) & \text{otherwise.} \end{cases}$$

It is easy to see that, as a^α is relevant, so is a^{α_-}. By Theorem 6.1, we have $\mathsf{a}^\alpha < \mathsf{a}^{\alpha_-}$: this proves that no a^α with $t_\alpha > 0$ can be maximal. On the other hand, if $t_\alpha = 0$ then $t_\alpha = t_\gamma$, so $\alpha \preccurlyeq \gamma$ implies that $\mathsf{a}^\alpha \leqslant \mathsf{a}^\gamma$: thus no a^α with $\alpha \neq \gamma$ can be maximal.

Because $r \geqslant 3$, if α satisfies (2.5.2′) then it also satisfies (2.5.1′). It follows that if b^α is relevant but a^α is not, then $p = \alpha(t_\alpha) = 2$: except possibly in this case, $\mathsf{b}^\alpha < \mathsf{a}^\alpha$, so b^α cannot be maximal.

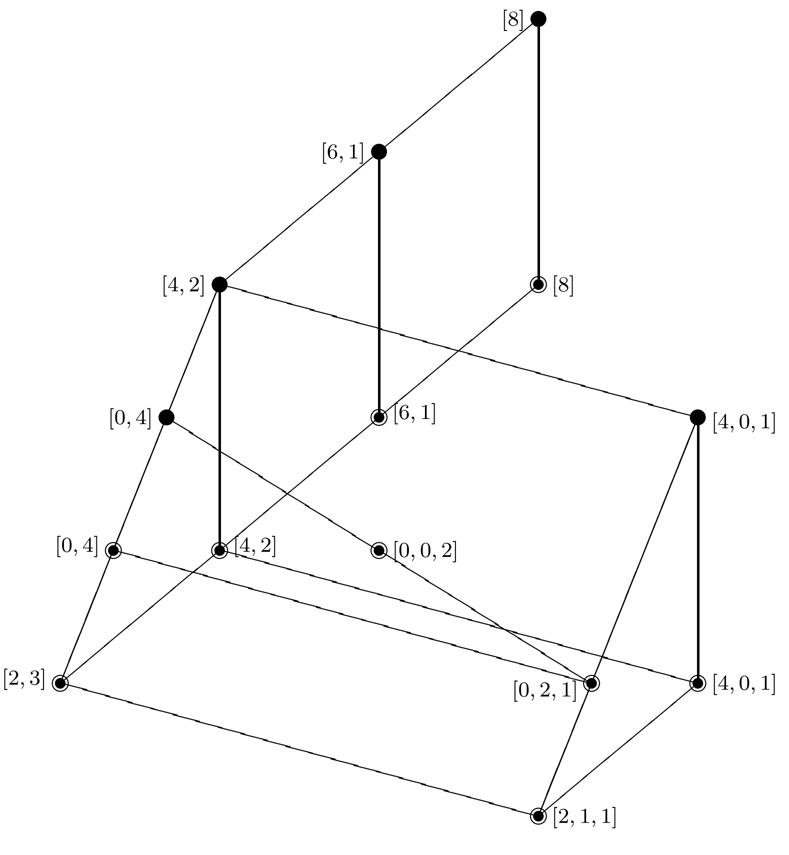
To examine this exceptional case more closely, put $t_\alpha = t$. If $d \neq 2^{t+1}$, let t' be the smallest number such that $t' > t$ and $\alpha(t') > 0$, and define α^- by

$$\alpha^-(j) = \begin{cases} 4 & \text{if } j = t, \\ 1 & \text{if } t < j < t', \\ \alpha(t') - 1 & \text{if } j = t', \\ \alpha(j) & \text{otherwise.} \end{cases}$$

In this case, b^{α^-} is relevant, $t_{\alpha^-} = t_\alpha$, and $\alpha \prec \alpha^-$: hence $\mathsf{b}^\alpha < \mathsf{b}^{\alpha^-}$, so again b^α is not maximal. In the remaining case $d = 2^{t+1}$, so $t > 0$ (because $d \geqslant p + 2$), and then $\mathsf{b}^\alpha < \mathsf{a}^{\alpha-}$ completes the proof. □

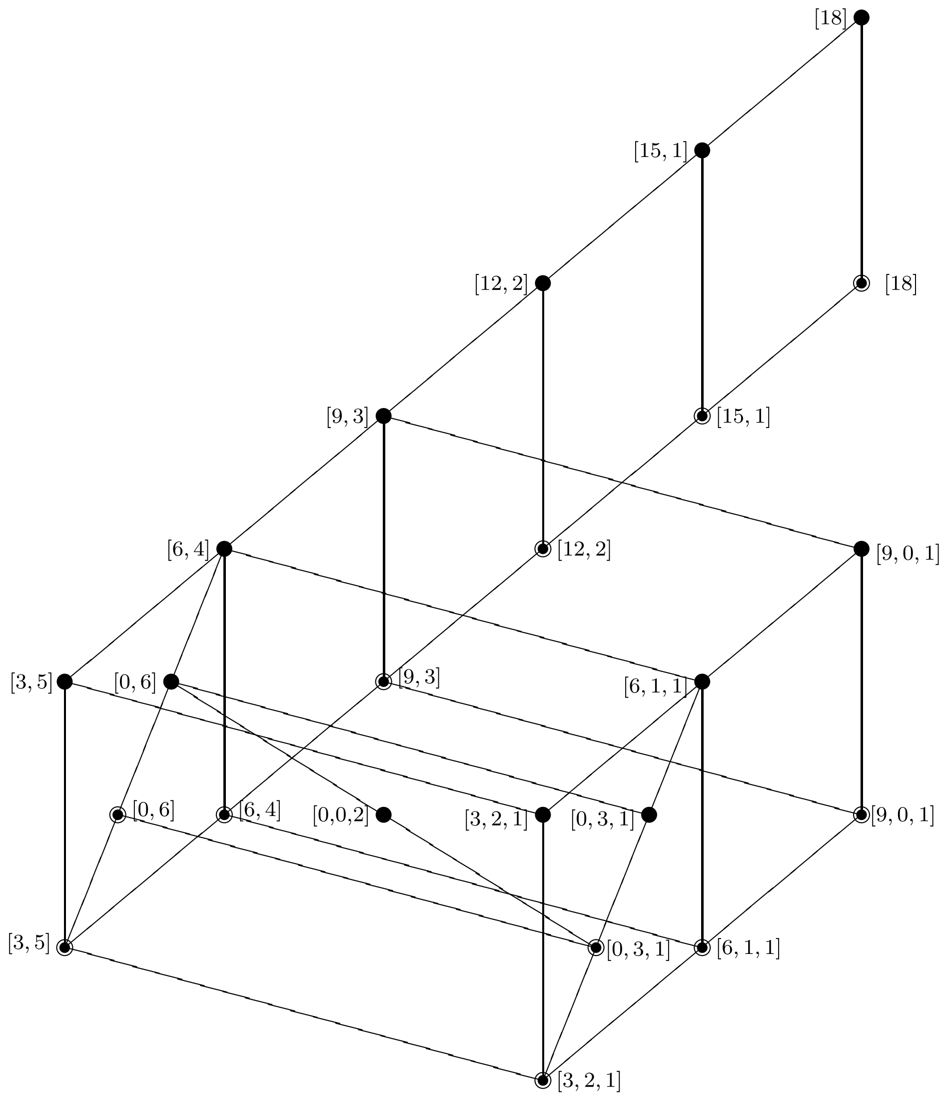
8. Examples

We start with the Alperin diagram of M^8 at $p = 2$ and $r \geqslant 8$. We think of it as the Hasse diagram of the inclusion-ordered poset of the join-irreducible submodules, so avoid horizontal edges and replace Alperin's arrows by the convention that all edges are directed downwards. Here some vertices are marked as solid dots and others as circled dots, and then labelled by p-partitions (that is, 2-partitions) of 8, omitting trailing zeros. For example, the solid dot with the label $[4, 2]$ stands for the join-irreducible whose simple quotient is $L(\mu^{[4,2]})$, while the circled dot with this label relates similarly to $L(\sigma^{[4.2]})$.



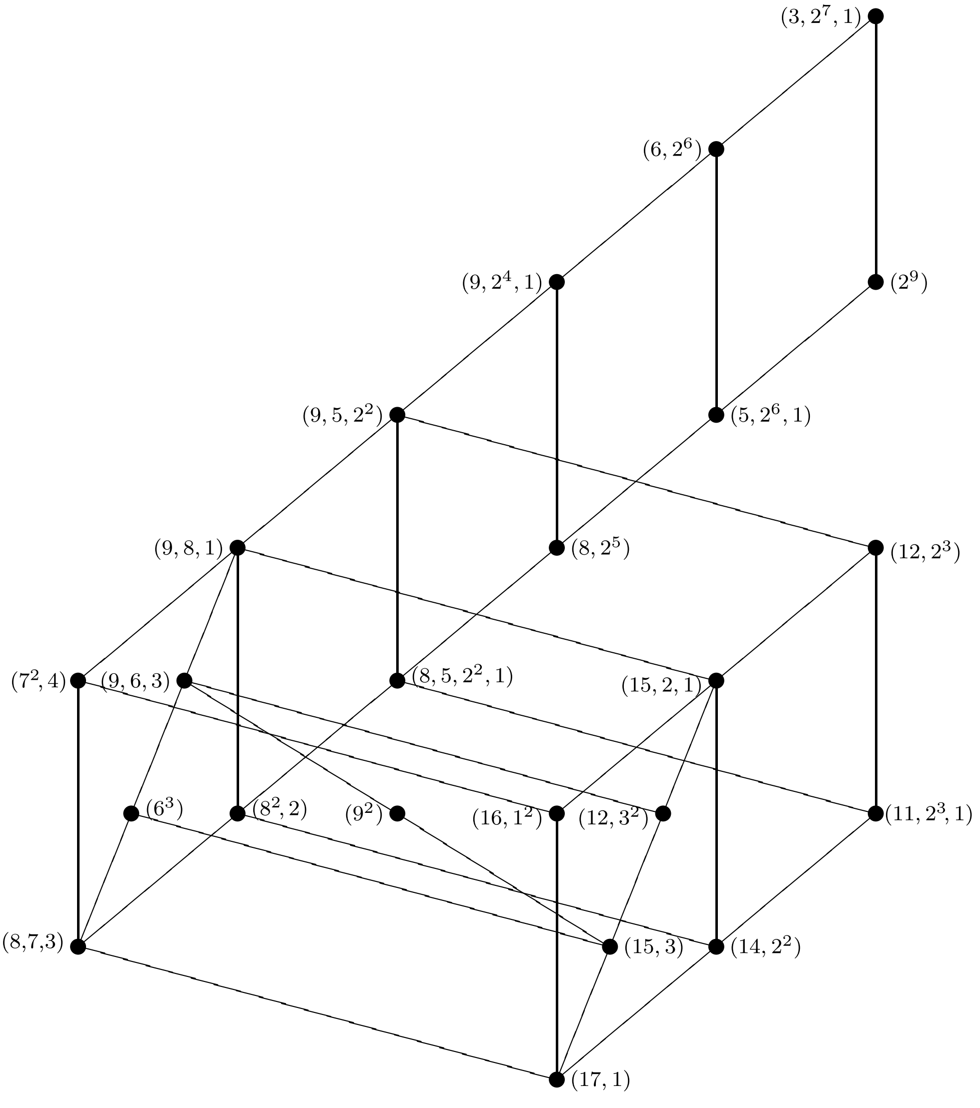
From this information one can calculate highest weights, such as $\mu^{[0,4]} = (4, 2^2)$ and $\sigma^{[0,4]} = (2^4)$, or $\sigma^{[2,3]} = (3^2, 2)$. Once all highest weights are known, one can re-draw the picture for smaller values of r, eliminating the nodes corresponding to weights with more than r parts. What about the edges? In general, one seems to have to test each pair of remaining nodes, putting an edge between the pair if and only if in the large diagram they are connected by one or more directed paths and no such path goes through any retained node. (For example, the first and last of the nodes looked at above are retained when $r = 3$, but the one in the middle is not, so while no edge joins the first and last at $r = 4$, there will be an edge between them at $r = 3$.)

Next we give two versions of the Alperin diagram of M^{18} at $p = 3$ and $r \geqslant 18$. In the first, we follow the previous conventions.



In the second, we label the nodes with the relevant highest weights, to make re-drawing for smaller values of r easier. It will be seen immediately that no such re-drawing is required as long

as $r \geqslant 9$. One notable feature is that, though the modules at $r = 4$ and $r = 5$ have different composition lengths, their simple quotients share the same highest weight, $(9, 5, 2^2)$.



Acknowledgments

Support from EPSRC Standard Research grant EP/G025487/1 is gratefully acknowledged. The authors also thank the Australian National University and the Mathematical Institute Oxford for support and hospitality during several visits.

Appendix A

In spite of a translation for which we are greatly indebted to Ralph Stöhr, we had problems with interpreting the intentions of [4]. In most cases, these difficulties can be overcome and one can arrive

at a satisfactory version of the proof offered there. Two challenges remain: the proof of Lemma 5 (of [4]) still does not convince us, and no explicit proof seems to have been offered for the crucially important 'only if' part of Lemma 4 of [4]. The aim of this appendix is to sketch a way for dealing with these challenges.

Instead of using the setting of [4], we do this in the version we are really interested in (see the discussion just after Theorem 6.1 above): when M is not free metabelian but free with respect to being metabelian and nilpotent of class d, so $\gamma_{d+1}(M)=0$ and $M^d=\gamma_d(M)$, whence each subspace of M^d is an ideal of M. We do not follow [4] in listing the free generators of M as $x_0, x_1, \ldots,$ but keep to $x_1, x_2, \ldots, x_r$ with $r \geqslant d$. On the other hand, we do switch from weight vectors and weight spaces (as used here up to now) to multihomogeneous elements and multihomogeneous components (as used in [4,5]). We also follow [4] by saying that an element of M^d implies another if the latter is contained in the submodule generated by the former.

Lemma 5 of [4] does not depend on Lemma 4 of [4] and is easier, so we take that one first, starting with a review of some of Bakhturin's arguments. From Lemma 1 of [4] we know that *each submodule of M^d is the direct sum of its intersections with the multihomogeneous components of M^d and is in fact generated by the intersections with the components all of whose positive partial degrees are powers of p*. These components are permuted by permutations of the free generators, and representatives of the orbits may be chosen as follows. To each p-partition α of d, consider the multihomogeneous component of M which involves precisely the generators x_i with $1 \leqslant i \leqslant \sum_j \alpha(j)$, the partial degree with respect to x_i being $p^{k(i)}$ where $k(i)$ is defined by $\sum_{j<k(i)} \alpha(j) < i \leqslant \sum_{j \leqslant k(i)} \alpha(j)$. Give this multihomogeneous component a name, M_α say, and conclude that each submodule is generated by its intersections with the M_α.

Within such a component, a left-normed monomial is determined by its two leftmost factors. For $i=2,\ldots,\sum_j \alpha(j)$, let u_i denote the left-normed monomial with first factor x_1 and second factor x_i: it is easy to see that these form a basis for the component in question. Call the span of $\{u_i \mid 2 \leqslant i \leqslant \alpha(t_\alpha)\}$ the initial span of M_α. Note that this span is zero unless $\alpha(t_\alpha) \geqslant 2$, so we may as well restrict attention to the α which satisfy this restriction. Then $\mathsf{a}^\alpha = u_2$, and if $\alpha(t_\alpha) \equiv 0$ then b^α is the sum of the u_i whose span we are discussing. In these terms, Lemma 2 of [4] asserts that *each submodule is generated by its intersections with these initial spans.*

Let H denote the subgroup of G consisting of the permutation matrices which fix each x_i with $i > \alpha(t_\alpha)$: this is isomorphic to the symmetric group of degree $\alpha(t_\alpha)$. Both M_α and its initial span are H-submodules, and the proof of Lemma 3 of [4] shows that *the initial span is simple as H-module when either $\alpha(t_\alpha) \not\equiv 0$ or $\alpha(t_\alpha)=p=2$, and otherwise it has just one non-zero proper H-submodule, namely the* 1*-dimensional subspace spanned by* b^α.

Write an element u of M_α in terms of the basis described above, and suppose that some basis vector u_i beyond the initial span occurs in this expression with non-zero coefficient. One step in the proof of Lemma 2 of [4] is to show that in this case u implies u_i. Another step shows that u_i implies a sum v of p distinct basis vectors in the initial span of a certain M_β. We cannot have $v=\mathsf{b}^\beta$, because if that canonical word exists, it is a sum of $\beta(t_\beta)-1$ basis vectors, and this number is congruent to -1, not equal to p. By the information extracted above from the proof of Lemma 3 of [4], this means that u_i implies a^β. It is not hard to identify the relevant β, and to verify that a^α is an image of a^β under some endomorphism of M that maps the free generating set into itself. Thus u implies each element of the initial span of M_α, and we may conclude the following.

Lemma A1. *If the intersection of M_α with a submodule is neither* 0 *nor the one-dimensional subspace spanned by b^α, then this submodule must contain a^α as well.* □

We are now ready for Lemma 5 of [4], in the following paraphrase: *if a canonical word lies in a sum of submodules, then it must lie in one of the summands.*

Proof of Lemma 5 of [4]. Let V and W be G-submodules such that our canonical word lies in $M_\alpha \cap (V+W)$, and let M'_α denote the sum of all other multihomogeneous components of M^d. As $V = (M_\alpha \cap V) \oplus (M'_\alpha \cap V)$ and $W = (M_\alpha \cap W) \oplus (M'_\alpha \cap W)$, we have $M_\alpha \cap (V+W) = (M_\alpha \cap V) + (M_\alpha \cap W)$.

If one of $M_\alpha \cap V$ and $M_\alpha \cap W$ is 0, or if one of them contains our canonical word, then there is nothing to prove. Lemma A1 ensures that otherwise the two summands are equal to each other and therefore also equal to their sum. $\square$

Salvaging the 'only if' part of Lemma 4 will require a concept that was not mentioned in [4]. Given a submodule U of M^d and a multihomogeneous element w in M^d, say that w is *additive modulo U with respect to a certain variable* if, regarded as a function $M/U \to M/U$ of that variable while the other variables are replaced by constants, it is an additive homomorphism. It is *additive modulo U* if it is additive modulo U with respect to every variable. Suppose that is the case, and let B be any basis of M/U. In order to check whether $w = 0$ is an identity of M/U, it suffices to check whether w has value 0 whenever elements of B are substituted for its variables. (For, in an arbitrary substitution from M/U, the variables are replaced by linear combinations of elements of B; since w is additive modulo U, the value is a sum of values obtained on substituting scalar multiples of elements of B, and the multihomogeneous nature of w guarantees that the latter values are scalar multiples of the values we are checking.) Differently put, all values of w when substituting elements of M/U lie in the subspace of M/U spanned by the values at elements of B. Let W denote the span of the values of w at basic monomials of M. Like every subspace of M^d, $U + W$ is an ideal in M. The images of the basic monomials in M/U form a spanning set for M/U, and therefore some of them form a basis B. The values of w at elements of B all lie in $(U + W)/U$, so this subspace of M/U is in fact the span of all values of w on M/U; as such, it is a fully invariant ideal. It follows that $U + W$ is the fully invariant ideal closure of U and w in M. Finally, notice that any value of w at basic monomials not all of which have degree 1 will have degree at least $d + 1$ and so vanish: thus W is in fact spanned by the values of w at elements of $\{x_1, \ldots, x_r\}$. The conclusion we want is the following.

Lemma A2. *If a multihomogeneous element w of M^d is additive modulo a submodule U, then the submodule generated by U and w is spanned by U and the values of w at the free generators of M.* $\square$

(We were led to this by Theorem 1.4 in the roughly contemporaneous paper [13] of Drenski. Bakhturin's book [5] quoted that without linking it to the present context: see Theorem 4.2.6 in the original Russian version, Theorem 7 on p. 102 of the English edition.)

In the sequel, when we speak of values we always mean values at free generators.

The next thing is to note that if α is a p-partition of d and $\alpha(0) \geqslant 2$, then a^α is additive (modulo 0): additive in the x_i with $i = 1, \ldots, \alpha(0)$ because in these variables it is even linear, and additive in the others because in those it is symmetric, the partial degrees are powers of p, and the relevant multinomial coefficients are divisible by p. The same can be said of b^α whenever $0 < \alpha(0) \equiv 0$. It follows then that the submodule generated by a canonical word corresponding to an α with $t_\alpha = 0$ is spanned by the values of that word. All such values are multihomogeneous, so in fact each multihomogeneous component of that submodule is spanned by the values that fall into it. This makes it very much easier to decide whether some other canonical word lies in the submodule in question. For example, if a p-partition β is not refined by α, then no non-zero value of a^α or b^α can lie in M_β, and so neither a^β nor b^β can possibly lie in our submodule. In other cases, Lemma A1 can be of considerable help.

For each p-partition α of d with t_α positive, define the p-partition α' by setting $\alpha'(t_\alpha - 1) = p$, $\alpha'(t_\alpha) = \alpha(t_\alpha) - 1$, and otherwise $\alpha'(j) = \alpha(j)$. Of course, $t_{\alpha'}$ is $t_\alpha - 1$; if this is still positive, we can form $(\alpha')'$ or simply α'', and so on, the last of these 'derivatives' being $\alpha^{(t)}$ with $t = t_\alpha$.

The positive ('if') part of Lemma 4 of [4] is that the b^β, with β ranging through the derivatives of α, are consequences of a^α, and also of b^α. The negative ('only if') part is that a canonical word cannot lie in the submodule generated by a^α unless it is either a linear combination of values of a^α or a linear combination of values of one of the b^β with β a derivative of α; and similarly, a canonical word cannot lie in the submodule generated by b^α unless it is either a linear combination of values of b^α or a linear combination of values of one of the b^β with β a derivative of α. To put it concisely, *the submodule generated by $\mathsf{b}^{\alpha'}$ is spanned by the values of these b^β, and the submodule generated by a canonical word w in M_α is spanned by those values together with the values of w itself.*

In view of the third last paragraph and Lemma A2, the 'only if' part of Lemma 4 of [4] will follow by induction on t_α once we prove the following.

Lemma A3. *If α is a p-partition of d such that $t_\alpha > 0$ and $\alpha(t_\alpha) \geqslant 2$, then a^α is additive modulo the submodule generated by $\mathsf{b}^{\alpha'}$.*

We do not have to deal separately with the similar statement about b^α: when that word is defined, then so is a^α (here we do not have to worry about duplication at $\alpha(t_\alpha) = p = 2$), and if a^α is additive modulo a certain submodule, then so is b^α (because it is a sum of words obtained from a^α by permuting variables).

In the proof, we shall take for granted some congruences mod p, namely that

$$\binom{p^k}{j} \equiv 0 \quad \text{and} \quad \binom{p^k - 1}{j} \equiv (-1)^j \quad \text{whenever } 0 < j < p^k,$$
$$\binom{p^k - p^{k-1} - 1}{j-1} \equiv -\binom{p^k - p^{k-1} - 1}{j} \quad \text{when } p^{k-1} \nmid j, \quad \text{and}$$
$$\text{if also } 0 < j < p^k - p^{k-1}, \quad \text{then } \binom{p^k - p^{k-1} - 1}{j} \not\equiv 0.$$

Proof of Lemma A3. We can no longer avoid writing down some explicit Lie monomials in M. To make this easier, we take up several practices of [4]. The first of these is to use formal powers of variables: by $[x_1^a, x_2^b, \ldots]$ we mean a left-normed Lie product with first factor x_1, second factor x_2, altogether a of the factors x_1 and b of the factors x_2, and so on. This takes advantage of the fact that in the metabelian case the order of factors in beyond the first two does not matter, and leaves it to the reader to deduce from the context what other factors are represented by the dots at the end. The first significant example of this practice is that we write

$$\mathsf{a}^\alpha = \left[x_1^{p^{k_1}}, x_2^{p^{k_2}}, \ldots\right] \quad \text{with } k_1 = k_2 \leqslant \cdots.$$

Of course here $k_1 = k_2 = t_\alpha$, but to ease typographic congestion we shall in fact write it simply as k. The second significant example is

$$\mathsf{b}^{\alpha'} = \sum_{i=2}^{p} \left[y_1^{p^{k-1}}, y_i^{p^{k-1}}, x_1^{p^k}, \ldots\right].$$

This illustrates the practice that when we need variables not already involved in our expressions, rather than starting to use $x_r, x_{r-1}, \ldots$, we abuse notation by giving them more convenient names, like $u, v, y_1, \ldots, y_p$.

The first formula-manipulation we need is to take this expression of $\mathsf{b}^{\alpha'}$, first add to it 0 in the form of $p[x_1^{p^k}, y_1^{p^{k-1}}, \ldots]$, and then use the Jacobi identity

$$[y_1, y_i, x_1] + [x_1, y_1, y_i] = [x_1, y_i, y_1]$$

with $i = 2, \ldots, p$ to conclude that also

$$\mathsf{b}^{\alpha'} = \sum_{i=1}^{p} \left[x_1^{p^k}, y_i^{p^{k-1}}, \ldots\right].$$

Replacing $y_3, \ldots, y_p$ by y_2 now shows that $\mathsf{b}^{\alpha'}$ implies

$$\mathsf{w} := \left[x_1^{p^k}, y_2^{p^k - p^{k-1}}, y_1^{p^{k-1}}, \ldots\right]$$

(but beware, this w is not a canonical word).

Replacing x_2 in a^{α} by $u+v$, u, v in turn, form

$$f := \left[x_1^{p^k}, (u+v)^{p^k}, \ldots\right] - \left[x_1^{p^k}, u^{p^k}, \ldots\right] - \left[x_1^{p^k}, v^{p^k}, \ldots\right].$$

Our next task is to show that f is a consequence of $\mathsf{b}^{\alpha'}$. In preparation for this, we calculate the multihomogeneous components of f:

$$\begin{aligned}
f &= \sum_{i=0}^{p^k-2} \binom{p^k-1}{i} \left[x_1^{p^k}, u^{i+1}, v^{p^k-1-i}, \ldots\right] + \sum_{j=1}^{p^k-1} \binom{p^k-1}{j} \left[x_1^{p^k}, v^{p^k-j}, u^j, \ldots\right] \\
&= \sum_{j=1}^{p^k-1} \left(\binom{p^k-1}{j-1} \left[x_1^{p^k}, u^j, v^{p^k-j}, \ldots\right] + \binom{p^k-1}{j} \left[x_1^{p^k}, v^{p^k-j}, u^j, \ldots\right] \right) \\
&= \sum_{j=1}^{p^k-1} \left((-1)^{j-1} \left[x_1^{p^k}, u^j, v^{p^k-j}, \ldots\right] + (-1)^j \left[x_1^{p^k}, v^{p^k-j}, u^j, \ldots\right] \right) \\
&= \sum_{j=1}^{p^k-1} (-1)^j \left(\left[u^j, x_1^{p^k}, v^{p^k-j}, \ldots\right] + \left[x_1^{p^k}, v^{p^k-j}, u^j, \ldots\right] \right) \\
&= \sum_{j=1}^{p^k-1} (-1)^j \left[u^j, v^{p^k-j}, x_1^{p^k}, \ldots\right].
\end{aligned}$$

Thus it suffices to show that each $[u^j, v^{p^k-j}, x_1^{p^k}, \ldots]$ with $0 < j < p^k$ is a consequence of $\mathsf{b}^{\alpha'}$.

When j is a multiple of p^{k-1}, say, $j = ap^{k-1}$ with $0 < a < p$, this follows from the second form of $\mathsf{b}^{\alpha'}$: putting $y_1 = \cdots = y_a = u$ and $y_{a+1} = \cdots = y_p = v$ in that yields $a[x_1^{p^k}, u^j, \ldots] + (p-a)[x_1^{p^k}, v^j, \ldots]$ whence, by the Jacobi identity

$$[x_1, u, v] - [x_1, v, u] = -[u, v, x_1],$$

we get $a[u^j, v^{p^k-j}, x_1^{p^k}, \ldots]$ as required.

Suppose next that j is not a multiple of p^{k-1}. Swapping u and v if necessary, we may assume without loss of generality that $0 < j < p^k - p^{k-1}$. Put $y_2 = u + v$ in w to obtain $[y_1^{p^{k-1}}, (u+v)^{p^k-p^{k-1}}, \ldots]$ and calculate the multihomogeneous component of the result in which the degree of u is j: this is

$$\begin{aligned}
&\binom{p^k-p^{k-1}-1}{j-1} \left[y_1^{p^{k-1}}, u^j, v^{p^k-p^{k-1}-j}, \ldots\right] + \binom{p^k-p^{k-1}-1}{j} \left[y_1^{p^{k-1}}, v^{p^k-p^{k-1}-j}, u^j, \ldots\right] \\
&\quad = \binom{p^k-p^{k-1}-1}{j} \left(-\left[y_1^{p^{k-1}}, u^j, v^{p^k-p^{k-1}-j}, \ldots\right] - \left[v^{p^k-p^{k-1}-j}, y_1^{p^{k-1}}, u^j, \ldots\right] \right)
\end{aligned}$$

$$= \binom{p^k - p^{k-1} - 1}{j} \big[u^j, v^{p^k - p^{k-1} - j}, y_1^{p^{k-1}}, \ldots\big].$$

Putting $y_1 = v$ here and using that $\binom{p^k - p^{k-1} - 1}{j} \not\equiv 0$ completes the proof that all multihomogeneous components of f are consequences of $\mathsf{b}^{\alpha'}$, and therefore so is f itself.

We have proved that, modulo the submodule in question, a^α is additive with respect to x_2. Since swapping x_1 and x_2 only changes the sign of a^α, the same goes for x_1 as well, and of course a^α is additive with respect to all other variables. This completes the proof of Lemma A3. $\square$

References

[1] Martin Aigner, Combinatorial Theory, Springer-Verlag, Berlin, Heidelberg, New York, 1979.
[2] J.L. Alperin, Diagrams for modules, J. Pure Appl. Algebra 16 (1980) 111–119.
[3] Henning Haahr Andersen, p-filtrations and the Steinberg module, J. Algebra 244 (2001) 664–683.
[4] Yu.A. Bakhturin, On identities in metabelian Lie algebras, Tr. Semin. im. I. G. Petrovskogo 1 (1975) 45–56 (in Russian).
[5] Yu.A. Bakhturin, Identical Relations in Lie Algebras, Nauka, Moscow, 1985, (in Russian); English translation: VNU Science Press, Utrecht, 1987.
[6] R.M. Bryant, Groups acting on polynomial algebras, in: Finite and Locally Finite Groups, Istanbul, 1994, Kluwer, Dordrecht, 1995, pp. 327–346.
[7] R.M. Bryant, L.G. Kovács, Ralph Stöhr, Lie powers of modules for groups of prime order, Proc. Lond. Math. Soc. (3) 84 (2002) 343–374.
[8] R.M. Bryant, Ralph Stöhr, Lie powers in prime degree, Quart. J. Math. Oxford (2) 56 (2005) 473–489.
[9] Kuo-Tsai Chen, Integration in free groups, Ann. of Math. 54 (1951) 147–162.
[10] Stephen R. Doty, The submodule structure of certain Weyl modules for groups of type A_n, J. Algebra 95 (1985) 373–383.
[11] Stephen Doty, Submodules of the symmetric powers of the natural module for GL_n, in: Invariant Theory, Denton, TX, 1986, in: Contemp. Math., vol. 88, 1989, pp. 185–191.
[12] Stephen Doty, The symmetric algebra and representations of general linear groups, in: Proceedings of the Hyderabad Conference on Algebraic Groups, Hyderabad, 1989, Manoj Prakashan, Madras, 1991, pp. 123–150.
[13] V.S. Drenski, On identities in Lie algebras, Algebra Logika 13 (3) (1974) 265–290 (in Russian); English translation: Algebra Logic 13 (1974) 150–165.
[14] J.A. Green, Polynomial Representations of GL_n, Lecture Notes in Math., vol. 830, Springer-Verlag, Berlin, Heidelberg, New York, 1980; 2nd corrected and augmented ed., 2007.
[15] Torsten Hannebauer, Ralph Stöhr, Homology of groups with coefficients in free metabelian Lie powers and exterior powers of relation modules and applications to group theory, in: Proc. Second Internat. Group Theory Conf., Bressanone/Brixen, June 11–17, 1989, Rend. Circ. Mat. Palermo (2) Suppl. 23 (1990) 77–113.
[16] B. Huppert, N. Blackburn, Finite Groups II, Springer-Verlag, Berlin, Heidelberg, New York, 1982.
[17] L.G. Kovács, Some representations of special linear groups, in: The Arcata Conference on Representations of Finite Groups, Pt 2, in: Proc. Sympos. Pure Math., vol. 47, 1987, pp. 207–218.
[18] Leonid Krop, On the representations of the full matrix semigroup on homogeneous polynomials, J. Algebra 99 (1986) 370–421.
[19] Leonid Krop, On the representations of the full matrix semigroup on homogeneous polynomials. II, J. Algebra 102 (1986) 284–300.
[20] Kerstin Kühne-Hausmann, Zur Untermodulstruktur der Weylmoduln für SL_3, Bonner Math. Schriften, vol. 162, Univ. Bonn, Bonn, 1985.
[21] Stuart Martin, Schur Algebras and Representation Theory, Cambridge University Press, Cambridge, 1993.
[22] Alison E. Parker, The global dimension of Schur algebras for GL_2 and GL_3, J. Algebra 241 (2001) 340–378.
[23] P.W. Winter, On the modular representation theory of the two-dimensional special linear group over an algebraically closed field, J. Lond. Math. Soc. (2) 16 (1977) 237–252.