MINIMAL FAITHFUL MODULES OVER ARTINIAN RINGS

George M. Bergman

Abstract: Let R be a left Artinian ring, and M a faithful left R-module such that no proper submodule or homomorphic image of M is faithful.

If R is local, and socle(R) is central in R, we show that length($M/J(R)M$) + length(socle(M)) $\leq$ length(socle(R)) + 1.

If R is a finite-dimensional algebra over an algebraically closed field, but not necessarily local or having central socle, we get an inequality similar to the above, with the length of socle(R) interpreted as its length as a bimodule, and the summand $+1$ replaced by the Euler characteristic of a graph determined by the bimodule structure of socle(R). The statement proved is slightly more general than this summary; we examine the question of whether much stronger generalizations are possible.

If a faithful module M over an Artinian ring is only assumed to have one of the above minimality properties – no faithful proper submodules, *or* no faithful proper homomorphic images – we find that the length of $M/J(R)M$ in the former case, and of socle(M) in the latter, is $\leq$ length(socle(R)). The proofs involve general lemmas on decompositions of modules.

2010 Mathematics Subject Classification: Primary: 13E10, 16G10, 16P20; Secondary: 16D60.

Key words: Faithful modules over Artinian rings, length of a module or bimodule, socle of a ring or module.

1. Background and motivation

This paper arose as a tangent to the unpublished note [**1**], which examines the question of which commutative Artinian rings R have the property that every faithful R-module M has length greater than or equal to that of R. (If R is a commutative algebra over a field k, it is known that this is true if R can be generated over k by 2 elements, but false for 4-generator algebras; it is an open problem whether it holds for 3-generator algebras. For more on this, see [**1**] and [**5**, Chapter 5].)

In studying that question, it is natural to focus on faithful modules M no proper factor-modules or submodules of which are faithful. I obtained some results showing that such M must have small "top" $M/J(R)M$ and "bottom" socle(M); Luchezar Avramov then pointed me to a 1972 paper of Tor Gulliksen, [**3**], which obtained a stronger result; I found, in turn, that Gulliksen's bound could be strengthened, and that the strengthened result could be applied to a wider class of rings (which, in particular, need not be commutative). This will be done in §2 below.

Some details: the relevant result of Gulliksen's paper is that if R is a commutative local Artinian ring, and M a faithful R-module no proper submodule or homomorphic image of which is faithful, then each of the semisimple R-modules $M/J(R)M$ and socle(M) has length less than or equal to that of socle(R), with at least one of these inequalities strict unless $M \cong R$. His proof is, in effect, a lemma in linear algebra, about bilinear maps $A \times B \to C$ of finite-dimensional vector spaces over a field, which have the property that every nonzero element of A acts nontrivially, but which lose this property both on restriction to any proper subspace of B, and on composition with any noninjective map out of C; though he only states it for the natural map socle$(R) \times M/J(R)M \to$ socle(M) of vector spaces over the field $R/J(R)$. In §2 we show that, in the general linear algebra setting, one has the stronger inequality $\dim(A) \geq \dim(B) + \dim(C) - 1$. We then, like Gulliksen, apply this result to maps socle$(R) \times M/J(R)M \to$ socle(M); here, rather than assuming the local ring R commutative, it is only necessary to assume socle(R) central in R.

Subsequent sections obtain inequalities of a similar nature for modules over not necessarily local Artinian rings R.

In §6, which is essentially independent of the rest of this note, we obtain results on lengths of modules satisfying only one of our minimality conditions.

2. Faithful modules over local Artin rings with central socles

Before formulating the promised linear algebra result, let us note that for finite-dimensional vector spaces B and C over a field k, to give a linear map $a\colon B \to C$ is equivalent to giving an element of $B^* \otimes_k C$. Hence a k-vector space A, given with a k-bilinear map $A \times B \to C$ such that every nonzero element of A induces a nonzero map $B \to C$ is equivalent to a subspace $A \subseteq B^* \otimes_k C$. This is a more symmetric situation; so we shall formulate the linear algebra result in that form, with B^* re-named B.

As noted in the preceding section, the bilinear maps $A \times B \to C$ of interest are those such that restriction to any proper subspace of B, or composition with the natural map into any proper homomorphic image of C, kills the action of some element of A. With B dualized as above, this becomes the condition that passing to a homomorphic image of either B or C has that effect. Now a minimal proper homomorphic image of B has the form B/kb for some $b \in B - \{0\}$, so the condition that passage to any such homomorphic image kills some element of A says that for each nonzero $b \in B$, the space A contains a nonzero element of the form $b \otimes c$. Likewise, the condition that passing to any proper homomorphic image C/kc of C kills some element of A means that for each nonzero $c \in C$, the space A contains a nonzero element of the form $b \otimes c$. This leads to the formulation of the next result.

My original proof required that the field k have cardinality at least $\max(\dim_k(B), \dim_k(C))$. For the present proof, I am indebted to Clément de Seguins Pazzis [**6**].

Proposition 1. *Let k be a field, and B and C nonzero finite-dimensional vector spaces over k. Suppose $A \subseteq B \otimes_k C$ is a subspace such that*

(1) $\qquad\qquad (\forall\, b \in B - \{0\})\, (\exists\, c \in C - \{0\})\ b \otimes c \in A,$

and

(2) $\qquad\qquad (\forall\, c \in C - \{0\})\, (\exists\, b \in B - \{0\})\ b \otimes c \in A.$

Then

(3) $\qquad\qquad \dim_k(A) \ \geq \ \dim_k(B) + \dim_k(C) - 1.$

Proof: (After de Seguins Pazzis [**6**].)

If $\dim_k(B)$ or $\dim_k(C)$ is 1, then (2), respectively (1), says that $A = B \otimes C$, and the desired result is immediate. So let $m, n > 1$, assume inductively that the result is known when B has dimension $m-1$ and C has dimension $n-1$, and suppose we are in a situation with $\dim_k(B) = m$ and $\dim_k(C) = n$. We consider two cases:

Case 1. There exists a linear functional $f: B \to k$ such that the induced map $f \otimes \mathrm{id}_C: B \otimes C \to k \otimes C \cong C$ carries $A \subseteq B \otimes C$ surjectively onto C.

Then taking any basis $\{c_1, \ldots, c_n\}$ of C, we can find elements $a_1, \ldots, a_n \in A$ whose images under $f \otimes \mathrm{id}_C$ are $c_1, \ldots, c_n$. On the other hand, letting $b_1, \ldots, b_{m-1}$ be any basis of $\ker(f) \subseteq B$, we can find, by (1), nonzero $c'_1, \ldots, c'_{m-1} \in C$ such that $b_1 \otimes c'_1, \ldots, b_{m-1} \otimes c'_{m-1} \in A$. We claim that $a_1, \ldots, a_n,\, b_1 \otimes c'_1, \ldots, b_{m-1} \otimes c'_{m-1} \in A$ are linearly independent.

Indeed, given any linear dependence relation among these elements, if we apply $f \otimes \mathrm{id}_C$ to it, this will annihilate the last $m-1$ terms, and looking at the remaining terms, we conclude from the linear independence of $c_1, \ldots, c_n$ that the coefficients of $a_1, \ldots, a_n$ in the relation must be zero. Given this fact, if for any $i \leq m-1$ we let $f_i \colon B \to k$ be a linear functional which takes b_i to 1 and all other b_j to 0, then application of $f_i \otimes \mathrm{id}_C$ to our relation shows that the coefficient of $b_i \otimes c_i'$ is also zero. So the indicated $m+n-1$ elements of A are linearly independent, establishing (3) in this case.

The negation of the condition of Case 1 says that for *every* linear functional $f \colon B \to k$, a certain conclusion holds. But to finish the proof, it will be enough to assume this for *some nonzero* f, as we do in

Case 2. There exists a nonzero linear functional $f \colon B \to k$ such that the induced map $f \otimes \mathrm{id}_C \colon B \otimes C \to k \otimes C \cong C$ carries $A \subseteq B \otimes C$ into a proper subspace $C_0 \subseteq C$.

Without loss of generality, we can take C_0 to have dimension $n-1$. Let $B_0 = \ker(f)$, which has dimension $m-1$. Then our assumption on f implies

(4)   For any element of A of the form $b \otimes c$, either $b \in B_0$ or $c \in C_0$.

Now let p_B be any retraction $B \to B_0$, let p_C be any retraction $C \to C_0$, and let A_0 be the image of A under $p_B \otimes p_C \colon B \otimes C \to B_0 \otimes C_0$. (It is not asserted that $A_0 \subseteq A$.) I claim that the analogs of (1) and (2) hold with A_0, B_0, C_0 in the roles of A, B, and C. Indeed, given nonzero $b \in B_0$, let $b' \in B$ be chosen which projects to b under p_B, but is *not* in B_0; and use (1) to find a nonzero $c \in C$ such that $b' \otimes c \in A$. Then by (4), $c \in C_0$, so c is fixed by p_C; so applying $p_B \otimes p_C$ to $b' \otimes c$, we get $b \otimes c \in A_0$; the analog of (1). The symmetric argument gives the analog of (2).

Hence by our inductive assumption, $\dim(A_0) \geq (m-1)+(n-1)-1$, which is short by 2 of the desired lower bound on the dimension of A. Since A_0 is the image of A under $p_B \otimes p_C$, it will suffice to find two linearly independent elements of A which are in the kernel of that map. To do so, let b span $\ker(p_B)$ and c span $\ker(p_C)$, and use (1) and (2) to find nonzero elements $b \otimes c'$ and $b' \otimes c$ of A. From (4), one sees that these are linearly independent, completing the proof. $\square$

Dualizing B, to pass from this result to a statement about linear maps, as discussed earlier, we get

Corollary 2. *Suppose A, B, and C are finite-dimensional vector spaces over a field k, and $f \colon A \times B \to C$ a bilinear map, such that every nonzero element of A induces a nonzero map $B \to C$, but such that this property is lost on restriction to any proper subspace of B, and likewise on composition with the map to any proper homomorphic image of C. Then $\dim(A) \geq \dim(B) + \dim(C) - 1$.*

We shall now deduce the asserted generalization of Gulliksen's result. (Incidentally, Gulliksen does not explicitly say in [**3**] that his rings are commutative; but this is apparent from the techniques he uses, e.g., the duality called on at the top of [**3**, p. 79]; and is also evidenced by the fact that commutativity is one of the properties he verifies for the matrix example of [**3**, Theorem 2]. In the present note, rings are not assumed commutative unless this is stated.)

Recall that the *socle* of a left or right module M is the sum of its simple submodules. By the left and right socles of a ring R, we understand $\mathrm{socle}(_RR)$ and $\mathrm{socle}(R_R)$, the socles of R as a left and as a right module, each of which is a 2-sided ideal of R. On the other hand, by *the socle* of R, $\mathrm{socle}(R)$, we shall mean the sum of all its minimal nonzero 2-sided ideals, which, for R left or right Artinian, is the intersection of its right and left socles. (So, for example, in the algebra of $n \times n$ upper triangular matrices over a field, the left socle is the ideal of matrices with support in the top row, the right socle is the ideal of matrices with support in the last column, and $\mathrm{socle}(R)$ is the ideal of matrices with support in the upper right-hand corner.)

We denote the Jacobson radical of a ring R by $J(R)$.

Theorem 3 (cf. Gulliksen [**3**, Lemma 2]). *Suppose R is a left Artinian local ring such that $\mathrm{socle}(R)$ is central in R, and let M be any faithful left R-module such that no proper submodule or homomorphic image of M is faithful. Then*

$$(5) \quad \dim_{R/J(R)}(M/J(R)M) + \dim_{R/J(R)}(\mathrm{socle}(M))$$
$$\leq \dim_{R/J(R)}(\mathrm{socle}(R)) + 1.$$

Proof: $J(R)$ annihilates $M/J(R)M$ and $\mathrm{socle}(M)$ as left R-modules, and annihilates the (R, R)-bimodule $\mathrm{socle}(R)$ on both sides; so the first two become left vector spaces, and the latter a bimodule, over the division ring $R/J(R)$. The statement that $\mathrm{socle}(R)$ is central in R says that $R/J(R)$ has the same action on the two sides of $\mathrm{socle}(R)$, from which it immediately follows that the division ring $R/J(R)$ must be a field k. Writing $A = \mathrm{socle}(R)$, $B = M/J(R)M$, $C = \mathrm{socle}(M)$, the left module operation of R on M induces an action by which A carries B into C,

giving a k-bilinear map $A \times B \to C$. Because M is faithful, every nonzero element of A gives a nonzero map $B \to C$, while the minimality assumptions on M imply that this property of our bilinear map is lost when one passes to any proper subspace of B or proper homomorphic image of C. (Indeed, if B_0 is a proper k-subspace of B, equivalently, a proper R-submodule, then its inverse image in M is a proper submodule, hence non-faithful. The kernel of the action of R on this submodule must meet $\mathrm{socle}(R) = A$, and any $a \in A$ in that kernel acts trivially on B_0. The dual argument applies to proper homomorphic images C_0 of C.) That B and C are finite-dimensional is also not hard to see from these minimality conditions. (Some general results of which this finite dimensionality is a special case are developed in §6.)

Hence Corollary 2 applies, and gives the desired inequality.   □

Remarks. If R is a left Artinian local ring, then a *sufficient* condition for $\mathrm{socle}(R)$ to be central is that R be an algebra over a field k, and that the residue field of R be k itself. Such an R can be far from commutative; for instance, over any field we can take for R the ring of upper triangular $n \times n$ matrices over k with scalar main diagonal. On the other hand, if R contains a field k which is not central in R, but which again maps isomorphically onto $R/J(R)$, then $\mathrm{socle}(R)$ may or may not be central. For example, if we take a twisted polynomial ring $k[x;\theta]$ where θ is an automorphism of k of finite order $d > 1$, and look at its local factor-ring $k[x;\theta]/(x^{n+1})$ for some $n > 0$, then k is not central in R, but the socle, $k\,x^n$, is central if and only n is a multiple of d. Likewise, given a field k and automorphisms $\theta_2, \ldots, \theta_n$, we can generalize the triangular-matrices example by letting R be the ring of upper-triangular $n \times n$ matrices $((a_{ij}))$ over k satisfying $a_{ii} = \theta_i(a_{11})$ for $2 \leq i \leq n$. This will be Artinian and local with residue field isomorphic to k, and its socle, $k\,e_{1n}$, will be central if and only if $\theta_n = \mathrm{id}$.

Turning back to Proposition 1, one may ask what a minimal subspace $A \subseteq B \otimes C$ satisfying (1) and (2) can look like. Easy examples are the spaces of the form $B \otimes c + b \otimes C$ for arbitrary nonzero elements $b \in B$, $c \in C$. These spaces have dimension exactly $\dim(B) + \dim(C) - 1$, since $B \otimes c$ and $b \otimes C$ intersect in the one-dimensional subspace spanned by $b \otimes c$. In matrix notation, this example can be pictured as the vector space of $m \times n$ matrices with support in the union of the first row and the first column.

Using such matrix notation, one can describe further minimal families. For every positive integer $q \leq \min(m,n)$, I claim that the space A_q of $m \times n$ matrices with support in the union of the first q rows and the

first q columns, such that the first q entries along the main diagonal are all equal, satisfies (1) and (2). To see (1), note that given any nonzero column vector $b = (\beta_1, \ldots, \beta_m)^T$, we can find a nonzero row vector $c = (\gamma_1, \ldots, \gamma_q, 0, \ldots, 0)$ such that $\beta_1 \gamma_1 = \cdots = \beta_q \gamma_q$. Indeed, if at least one of $\beta_1, \ldots, \beta_q$ is zero, we can choose the γ's not all zero so that all products $\beta_i \gamma_i$ are zero, while if all of $\beta_1, \ldots, \beta_q$ are nonzero, we can choose the γ's so that all $\beta_i \gamma_i$ equal 1. In either case, the $m \times n$ matrix $b \otimes c$ will be nonzero and lie in A_q. Similarly, given a nonzero row vector c, we can find a nonzero column vector b such that $b \otimes c \in A_q$, proving (2).

It is not hard to see that for every column vector $b \in k^m$ at most one of whose first q entries is zero, and which has at least one nonzero entry after the first q entries, the above construction gives (up to scalars) the only c such that $b \otimes c \in A_q$. (The condition that at least one entry of b after the first q be nonzero guarantees that every c with $b \otimes c \in A_q$ must live in its first q entries, which we need to get uniqueness.) Likewise, if we are given c at most one of whose first q entries is zero, and having at least one later nonzero entry, there is up to scalars only one b such that $b \otimes c \in A_q$. Hence, any subspace of A_q that satisfies (1) and (2) must contain all the elements $b \otimes c$ of the two sorts just described. Combining these observations, one can deduce that if at least one of m, n is $> q$, and k has > 2 elements, then A_q indeed has no proper subspaces satisfying (1) and (2). (The condition that k has > 2 elements is used to get every row or column as a sum of rows or columns satisfying appropriate conditions on which entries are nonzero. For k the 2-element field and $q > 2$, there do in fact exist proper subspaces of A_q satisfying (1) and (2).) In the remaining case, $m = n = q$, one finds that the same argument works without the proviso in the first sentence of this paragraph about "at least one nonzero entry after the first q entries".

So if k has more than 2 elements, these constructions do give minimal examples. I don't know whether, conversely, every minimal subspace of a tensor product $B \otimes_k C$ satisfying (1) and (2) is, with respect to some bases of B and C, of one of these forms.

It is curious that the bound $\dim(A) \geq \dim(B) + \dim(C) - 1$ of Proposition 1 also appears in [**2**, Proposition 1.3, case $k = 1$], for subspaces $A \subseteq B \otimes C$ subject to a different condition involving rank-1 entities; namely, that A not be contained in the kernel of any tensor product $f \otimes g$ of linear functionals $f \in A^*$, $g \in B^*$ [**2**, Proposition-Definition 1.1(2), case $k = 1$]. But I cannot see a relation between the two results. Indeed, the result in [**2**] is proved only over an algebraically closed field, and it is noted that it fails without that condition; while the last clause of

[**2**, Proposition 1.3] implies that every minimal submodule A of the sort considered there has dimension exactly $\dim(B) + \dim(C) - 1$, though for our condition, we have seen the opposite. Finally, most of the minimal examples noted above for our condition do not satisfy the condition of [**2**], precisely because their matrix representations involve 0's in fixed locations. So if there is a relation between the two results, it must be a subtle one.

Turning back to Theorem 3, and restricting attention to commutative R, observe that if we wish to extend that result to a commutative not-necessarily-local Artinian ring R, then this will be a direct product of commutative local rings $e_1 R \times \cdots \times e_d R$, where the e_i are the minimal idempotents of R. By summing the inequalities (5) for these d rings, we get a similar inequality, with the final $+1$ replaced by $+d$. The dimensions in the inequalities we sum are with respect to different base fields $e_i R / J(e_i R)$; the most natural way to refer to these dimensions is as the *lengths* of those R-modules. So in this situation, the analog of (5) is

(6) $\mathrm{length}(M/J(R)M) + \mathrm{length}(\mathrm{socle}(M)) \leq \mathrm{length}(\mathrm{socle}(R)) + d.$

This points toward the form of the inequalities we will obtain in the next two sections, for not-necessarily-commutative Artinian R.

3. The general Artinian case – preliminary steps

The result we are now aiming for will again be an application of a statement about a bilinear map, but with semisimple left modules in the roles of B and C, and a bimodule in the role of A. In this case, I don't see how to turn statements about bilinear maps into statements about subobjects of tensor products, so we shall develop directly the "$A \times B \to C$" result analogous to Corollary 2. (However, there will be an obvious symmetry between what we do with B and with C; which suggests that I am missing some way that these can be unified.)

Throughout this and the next section, we will therefore assume that

(7) S and T are semisimple Artinian rings, $_S B$ and $_T C$ are left modules of finite lengths, $_T A_S$ is a bimodule, and $h \colon {_T A_S} \times {_S B} \to {_T C}$ a balanced bilinear map,

such that

(8) Every nonzero $a \in A$ induces a nonzero map of abelian groups $h(a, -) \colon B \to C$.

In our final application, S and T will be the same, namely $R/J(R)$; but keeping them distinct until then will make our manipulations clearer.

For $a \in A$, we shall think of $h(a, -) \colon B \to C$ as "the action of a", and thus speak of elements of A as annihilating certain elements of B, having certain elements of C in their images, etc.; thus, we will seldom mention h explicitly.

When we speak of the kernel or image of a *family* of maps, we shall mean the intersection of the kernels of those maps, respectively, the sum of their images. A "maximal submodule" of a module will mean a maximal *proper* submodule.

We now state the conditions corresponding to (1) and (2).

(9) For every maximal submodule $B_0 \subseteq B$, there exists a nonzero $a \in A$ which annihilates B_0; equivalently, such that the kernel of $a\,S$ is precisely B_0.

(10) For every simple submodule $C_0 \subseteq C$, there exists a nonzero $a \in A$ which carries B into C_0; equivalently, such that the image of $T a$ is precisely C_0.

The statements of equivalence follow by combining the fact that $\ker(a\,S)$ and $\operatorname{im}(T\,a)$ are submodules of ${}_S B$ and ${}_T C$ with the assumptions that B_0 is maximal and C_0 simple.

Let us now see what happens to conditions (9) and (10) when we write S and T as direct products of simple Artin rings, and decompose A, B, and C accordingly. Say the decompositions of the identity elements of S and T into minimal central idempotents are $1_S = e_1 + \cdots + e_m$ and $1_T = f_1 + \cdots + f_n$, so that $S \cong \prod_i e_i S$ and $T \cong \prod_j f_j T$ as rings. Then in the situation of (10), the simple submodule C_0 is necessarily contained in some summand $f_j C$, so the a that we get must satisfy $a = f_j a$. Moreover, if we take some e_i such that $a e_i \neq 0$, then the image of $a e_i$ must also generate the simple submodule C_0; so replacing a by $a e_i$ if necessary, we can assume $a \in f_j A e_i$.

Let us now fix some f_j. Then we can deduce from (10) and the above observations that every simple submodule of $f_j C$ is generated by the image of an element of $f_j A e_i$ for some $i \leq m$. It would be nice if we could reverse the order of quantifications, and say that there exists some $i \leq m$, such that every simple submodule of $f_j C$ is generated by the image of an element of $f_j A e_i$; but that is too much to hope for. However, it turns out that, under some weak assumptions, we can show that there exists an i such that *enough* of the simple submodules of $f_j C$ are images of elements of $f_j A e_i$ to suit our purposes. The key concept is introduced in the definition below. In that definition, think of X as either B or $e_i B$, of Y as $f_j C$, of W as $f_j A$, or $f_j A e_i$, and of V as $f_j T f_j$. The adjective "left" in that definition, and "right" in the one

that follows it, refer to whether we are thinking of fixing the left factor f_j in the product-symbol $f_j\,A\,e_i$ (as in the above discussion), or the right factor e_i.

Definition 4. If W is a set of homomorphisms from an abelian group X to a left module Y over a simple Artin ring V, and N is a positive real number, then we shall call W *left N-strong* if for every family of $\leq N$ proper submodules of Y, there exists $w \in W$ such that $Vw(X)$ is a simple submodule of Y not contained in any member of that family. We shall call W *left strong* if it is left N-strong for all positive real numbers N. We shall at times use "left ∞-strong" as a synonym for "left strong".

(We have taken N, when finite, to be a real number rather than an integer so that some later statements can be formulated more conveniently; e.g., so that we can say "N/d-strong" rather than "$\lfloor N/d \rfloor$-strong".)

We similarly define a variant of condition (9), namely

Definition 5. If W is a set of homomorphisms from a left module X over a simple Artin ring U to an abelian group Y, and N a positive real number, then we shall call W *right N-strong* if for every family of $\leq N$ nonzero submodules of X, there exists $w \in W$ such that $\ker(wU)$ is a maximal submodule of X, and does not contain any member of that family. We shall call W *right strong* if it is right N-strong for all positive real numbers N. We shall at times use "right ∞-strong" as a synonym for "right strong".

The observations following (9) and (10) will yield statements to the effect that for appropriate N, the sets $\bigcup_i f_j\,A\,e_i$ of maps $B \to f_j\,C$ are left N-strong, and that the sets $\bigcup_j f_j\,A\,e_i$ of maps $e_i\,B \to C$ are right N-strong. (We postpone the details for the moment.) The next lemma shows the virtue of the N-strong condition: it is inherited, in slightly weakened form, by at least one term of any such union.

Lemma 6. *Suppose, in the context of Definition 4, that W is the union $W_1 \cup \cdots \cup W_d$ of d subsets. Then if W is left N-strong, for N a positive real number or ∞, then at least one of the W_i is left N/d-strong. (So for $N = \infty$, this says that if W is left strong, so is one of the W_i.)*

Likewise, if in the context of Definition 5, $W = W_1 \cup \cdots \cup W_d$ is right N-strong, then at least one of the W_i is right N/d-strong.

Proof: Let us prove the statements for finite N in contrapositive form. In the case of the first assertion, if left N/d-strength fails for each of the W_i, then for each i we can find a family of $\leq N/d$ proper submodules of Y whose union contains all simple submodules generated by images

of members of W_i. Taking the union over i of these sets of submodules, we get a set of $\leq N$ proper submodules of Y whose union contains all such simple submodules, showing that W is not left N-strong. The dual argument works for the second statement. The cases with $N = \infty$ follow from the cases for finite N. $\qquad\square$

The next lemma gives the postponed argument which will allow us to obtain from conditions (9) and (10) statements about N-strength; i.e., which will show that "*all* simple submodules" as in (10) entails "a family of submodules not contained in the union of N proper submodules", for appropriate N.

We have spoken above of the *minimal central* idempotents of a semi-simple ring T; we shall now deal with *minimal* idempotents (where an idempotent e of a ring R is considered less than or equal to an idempotent f if $eRe \subseteq fRf$; equivalently, if $e = ef = fe$. By "minimal" we of course mean "minimal nonzero".) Note that if f is a minimal central idempotent of a semisimple Artinian ring T, then $V = fTf$ is a simple Artinian ring, that is, a matrix ring $\mathrm{Matr}_{n,n}(D)$ over some division ring D [**4**, Theorem 3.5]; and that by taking any minimal idempotent $f' \in \mathrm{Matr}_{n,n}(D)$, one can recover D up to isomorphism as $f'\mathrm{Matr}_{n,n}(D)\,f' = f'(f\,T f)f' = f'Tf'$.

Lemma 7. *Let V be a simple Artin ring, Y a left V-module, f' a minimal idempotent of V, and $N = \mathrm{card}(f'Vf')$.*

Then there exists no finite family of proper submodules $Y_1, \ldots, Y_n$ of Y with $n \leq N$ such that every simple submodule of Y is contained in some Y_i.

Likewise, there exists no finite family of nonzero submodules $Y_1, \ldots, Y_n$ of Y with $n \leq N$ such that every maximal submodule contains some Y_i.

Proof: Note that both conclusions are statements about the lattice of submodules of Y. Now letting $D = f'Vf'$, a division ring, we know that V is Morita equivalent to D, so the lattice of submodules of Y is isomorphic to that of the D-vector-space $f'Y$. Hence in our proof, we may assume without loss of generality that V is a division ring D, and Y a D-vector space.

In this situation, the simple submodules are just the cyclic submodules, so to prove the first statement, we have to show that given proper subspaces $Y_1, \ldots, Y_n$ of Y with $n \leq N$, there exists a $y \in Y$ which is not contained in any of the Y_i. Suppose inductively that for some $m < n$ we have found a $y \in Y$ which does not lie in any of $Y_1, \ldots, Y_m$. If y also does not lie in Y_{m+1}, we have our next inductive step. If, on the other hand,

$y \in Y_{m+1}$, take any $y' \notin Y_{m+1}$, and consider the N elements $y' + \alpha y$, as α runs over D. Clearly, none of these lie in Y_{m+1}, and it is easy to check that at most one can lie in each of $Y_1, \ldots, Y_m$. Since they constitute $N \geq n > m$ elements, at least one lies in none of these spaces, giving the inductive step.

To get the second statement, recall that maximal subspaces of Y are kernels of elements of the dual space $Y^* = \mathrm{Hom}(Y, {}_D D)$, a right D-vector-space. Hence we can apply to that space the left-right dual of the first statement, and get the desired result. □

The next corollary applies the two preceding results to the modules and map of (7). Note that the statement refers to both *minimal* idempotents and *minimal central* idempotents.

Corollary 8. *Suppose $h\colon {}_T A_S \times {}_S B \to {}_T C$ is as in (7), and satisfies (8), (9), and (10).*

Let N_T be the minimum of the cardinalities of the division rings fTf as f ranges over the minimal idempotents of T, if that minimum is finite, or the symbol ∞ if all those division rings are infinite. Let d_T be the maximum, as f ranges over the minimal central *idempotents of T, of the number of minimal central idempotents e of S such that $fAe \neq 0$. Then for each minimal idempotent f of T, there is at least one minimal idempotent e of S such that fAe is left N_T/d_T-strong as a set of maps $eB \to fC$.*

Likewise, let N_S be the minimum of the cardinalities of the division rings eSe as e ranges over the minimal idempotents of S if this is finite, or the symbol ∞ if that minimum is infinite, and d_S the maximum, as e ranges over the minimal central *idempotents of S, of the number of minimal central idempotents f of T such that $fAe \neq 0$. Then for each minimal idempotent e of S, there is at least one minimal idempotent f of T such that fAe is right N_S/d_S-strong as a set of maps $eB \to fC$.*

Proof: In the situation of the first assertion, we see from condition (10) that for each minimal central idempotent f of T, the set fA of morphisms $fa\colon B \to fC$ ($a \in A$) has the property that for every minimal fTf-submodule $C_0 \subseteq fC$, there is at least one nonzero $fa \in fA$ which takes B into C_0. Now the minimal central idempotents $e \in S$ sum to 1, so there is at least one such $e \in S$ such that $fae \neq 0$. This map clearly still carries B into C_0; so we conclude that for every minimal $C_0 \subseteq fC$, some nonzero element of $\bigcup_e (fAe - \{0\})$ has image in C_0, where the union is over the minimal central idempotents e of S. By Lemma 7,

this shows that if we write D for the division ring (unique up to isomorphism) given by $f'Tf'$ where f' is any minimal idempotent of fTf, then $\bigcup_e(fAe - \{0\})$ is left $\text{card}(D)$-strong; hence left N_T-strong.

Now by definition of d_T, there are at most d_T minimal central idempotents e such that $fAe \neq 0$; so $\bigcup_e(fAe - \{0\})$ involves at most d_T nonempty sets $fAe - \{0\}$. Hence by Lemma 6, at least one of these sets is left N_T/d_T-strong, as claimed.

The second assertion is proved in the analogous fashion. $\qquad\square$

The above corollary will enable us to prove our desired generalizations of (3) and (5) *unless* one or more of the division rings eSe and fTf are finite fields of small cardinality. (If such a field occurs, N_T/d_T or N_S/d_S may be too small for our arguments to work.) It is curious that a similar condition in my original proof of (3) and (5) was eliminated by de Seguins Pazzis's argument; but we shall see by example, in §5, that the corresponding condition in the present situation cannot be dropped.

While these cardinality conditions are not very restrictive, and are needed for the result to hold, we come now to an embarrassingly restrictive condition, needed for the proofs of the results of the next section, though I have no example showing that those results fail without it. The condition is awkward to state in maximum generality. A fairly natural special case is the hypothesis that in (7),

(11) The semisimple Artinian rings S and T are finite-dimensional algebras over a common algebraically closed field k, and the induced actions of k on the two sides of the bimodule $_TA_S$ are the same.

Actually, we need the assumption that k is algebraically closed only to make the simple factors of S and T full matrix algebras over k; so we can instead put that assumption on the table, as the condition

(12) For some field k, each of the semisimple Artinian rings S and T is a direct product of full matrix algebras over k, and the induced actions of k on the two sides of the bimodule $_TA_S$ are the same.

A condition that is still more general (as we shall show in the next lemma), and will suffice for our purposes, is

(13) If $a \in A$ is a nonzero element whose image aB is contained in a simple submodule of C, then there exists nonzero $a' \in TaS$ whose kernel contains a maximal submodule of B; and likewise if $a \in A$ is a nonzero element whose kernel contains a maximal submodule of B, then there exists nonzero $a' \in TaS$ whose image is contained in a simple submodule of C.

In fact, we can make do with the following still weaker (though still wordier) condition.

(14) For each minimal central idempotent $e \in S$ and minimal central idempotent $f \in T$ such that $f A e \neq \{0\}$, it is *either* true that for every $a \in f A e$ such that the induced map $e B \to f C$ has image in a simple submodule of $f C$, some nonzero $a' \in T a S$ has kernel containing a maximal submodule of B, *or* that for every $a \in f A e$ such that the induced map $e B \to f C$ has kernel containing a maximal submodule of B, some nonzero $a' \in T a S$ has image in a simple submodule of $f C$. (But which of these is true may vary with the pair (f, e).)

Let us note the implications among these conditions.

Lemma 9. *For $h \colon {}_T A_S \times {}_S B \to {}_T C$ a bilinear map as in* (7) *which satisfies* (8), *one has the implications* (11) $\Longrightarrow$ (12) $\Longrightarrow$ (13) $\Longrightarrow$ (14).

Proof: (11) $\Longrightarrow$ (12) follows from the standard description of the structures of semisimple Artin rings, and (13) $\Longrightarrow$ (14) is clear. Let us prove (12) $\Longrightarrow$ (13).

Suppose as in (13) that $a \neq 0$, and $a B$ is contained in a simple submodule $C_0 \subseteq C$. Since the identity elements of S and T are sums of minimal central idempotents, we can find such idempotents $e \in S$ and $f \in T$ such that $f a e \neq 0$; and we will have $f a e B \subseteq f a B \subseteq f C_0 \subseteq C_0$. Hence, replacing a by some $a' = f a e$ if necessary, we may assume without loss of generality that $a \in f A e$ for such a pair of idempotents.

Now by assumption, eSe and fTf have the forms $\mathrm{Matr}_{m,m}(k)$ and $\mathrm{Matr}_{n,n}(k)$ for some positive integers m and n. Identifying them with these matrix rings, it is easy to verify that there exist finite-dimensional k-vector-spaces B', C' such that we can identify eB and fC with the spaces B'^m and C'^n of column vectors of elements of B' and C', made into modules over $eSe = \mathrm{Matr}_{m,m}(k)$ and $fTf = \mathrm{Matr}_{n,n}(k)$ in the natural way; that fAe can then be identified with $\mathrm{Matr}_{n,m}(A')$ where A' is a k-vector-space of k-linear maps $B' \to C'$, and finally, that the simple submodule $C_0 \subseteq f C$ will have the form $C_0'^n$ for some 1-dimensional subspace $C_0' \subseteq C'$. (Explicitly, letting e' and f' denote minimal idempotents of S and T, say those given by the matrix units e_{11} of their matrix representations, we can take $B' = e'B$, $C' = f'C$, $C_0' = f'C_0$, and $A' = f'A e'$.)

Thus, our element $a \in f A e$ will be an $n \times m$ matrix of linear maps $e B' \to f C'$, each having range in C_0'. We can now choose $\bar{a} \in T a S - \{0\}$ to be nonzero and have all components in some 1-dimensional subspace $k a' \subseteq A'$. (E.g., we can let $\bar{a}$ be a product $e_{1,i} a e_{j,1}$ such that the (i, j) component of a is nonzero.) Since $a' \colon B' \to C'$ has range in the

1-dimensional subspace C_0' of C', it is a rank-1 k-linear map, hence has kernel $B_0' \subseteq B'$ of codimension 1. Hence $\bar{a}$, having all components in $k\,a'$, will have kernel containing the maximal proper submodule $(B_0')^m \subseteq B$, as required.

The second assertion of (13) is proved similarly. $\qquad \square$

4. Minimal faithful modules over left Artin rings

Recall that if G is a finite graph, its *Euler characteristic* $\chi(G)$ is the number of vertices of G minus the number of edges, an integer which may be positive, negative, or zero. Recall also that a *bipartite* graph is a graph whose vertex-set is given as the disjoint union of two specified sets, such that every edge connects a member of one set with a member of the other. We shall call those sets (nonstandardly) the *left* and *right* vertex-sets of G.

The hard work of this section comes right at the beginning: proving the following noncommutative analog of Proposition 1, or more precisely, of Corollary 2.

Proposition 10. *Suppose $h\colon {}_T A_S \times {}_S B \to {}_T C$ is a bilinear map as in (7), which satisfies (8), (9), (10), and (14). For notational convenience we shall assume S and T disjoint.*

Let N_S, d_S, N_T, and d_T be defined as in Corollary 8. Further, let l_S denote the maximum of the values $\mathrm{length}_{eSe}(e\,B)$ as e ranges over the minimal central idempotents $e \in S$, and l_T the maximum of $\mathrm{length}_{fTf}(f\,C)$ as f ranges over the minimal central idempotents $f \in T$; and assume that

$$(15) \qquad\qquad N_T \;\geq\; d_T\, l_S \quad and \quad N_S \;\geq\; d_S\, l_T.$$

Finally, let G be the bipartite graph whose right vertex-set is the set of minimal central idempotents $e \in S$ satisfying $e\,B \neq 0$ (equivalently, $A\,e \neq 0$), whose left vertex-set is the set of minimal central idempotents $f \in T$ satisfying $f\,C \neq 0$ (equivalently, $f\,A \neq 0$), and such that two such vertices e, f are connected by an edge (f,e) if and only if $f A e \neq \{0\}$.
Then

$$(16) \qquad \mathrm{length}({}_S B) + \mathrm{length}({}_T C) \leq \mathrm{length}({}_T A_S) + \chi(G).$$

Proof: The parenthetical equivalences in the definition of G follow from (8), (9), and (10). Combining Corollary 8 with our hypothesis (15), we find that

(17)   For each minimal idempotent f of T, there is at least one minimal idempotent e of S such that $f A e$ is left l_S-strong as a set of maps

$e\,B \to f\,C$; and for each minimal idempotent e of S, there is at least
one minimal idempotent f of T such that $f\,A\,e$ is right l_T-strong as
a set of maps $e\,B \to f\,C$.

We shall now perform a series of reductions and decompositions on
our system $_T A_S \times {}_S B \to {}_T C$, verifying at each stage that if the in-
equality corresponding to (16) holds for our simplified system(s), then
it also holds for the original system; and, finally, we shall establish that
inequality for the very simple sorts of system we end up with.

In preparation, let us harness (17) by choosing, arbitrarily, for each
minimal central idempotent $f \in T$, *one* minimal central idempotent $e \in S$
such that $f\,A\,e$ is left l_S-strong, and call (f, e) the *left-marked* edge of the
graph G associated with the vertex f; and similarly, for each minimal
central idempotent $e \in S$, let us choose a minimal central idempotent $f \in$
T such that $f\,A\,e$ is right l_T-strong, and call (f, e) the *right-marked* edge
of G associated with the vertex e. Some edges may be both right- and
left-marked (for their respective right and left vertices).

We begin our reductions by considering any edge $(f, e) \in G$ which is
neither right- nor left-marked, and seeing what happens if we drop the
summand $f\,A\,e$ from A; i.e., replace A with $(1-f)A + A(1-e)$; and thus
drop the edge (f, e) from G, leaving the rest of our system unchanged.
Because (f, e) is neither right nor left marked, condition (17) has not
been lost. (The constant d_S and/or d_T may have decreased by 1, but we
don't have to think about this, because our use of these constants was
only to obtain (17), which has been preserved.) The removal of $f\,A\,e$
has no effect on the left-hand side of (16), while on the right-hand side,
it decreases length($_T A_S$) by length($_T(f\,A\,e)_S$) ≥ 1, and increases $\chi(G)$
by 1. Hence if the new system satisfies (16), then the original system,
whose right-hand side is $\geq$ that of the new system, also did.

Hence by induction, the task of proving (16) is reduced to the case
where *every* edge of G is left and/or right marked.

Suppose, next, that there is some left vertex f of G such that the
only edge adjacent to f is its associated left-marked edge, say (f, e), and
such that this is *not* also right-marked, and consider what happens if we
remove both the vertex f and the edge (f, e); i.e., replace C by $(1-f)C$,
and A by $(1-f)A$.

Clearly, the *remaining* vertices and edges continue to witness condi-
tion (17). Our new system also has the same Euler characteristic as the
old one, since just one vertex and one edge have been removed from the
graph.

To see how (16) is affected, let $d = \operatorname{length}(_T f\, C)$. I claim that we can find elements $a_1, \ldots, a_d \in f A\, e$ such that the submodules $T\, a_i\, B \subseteq f\, C$ are simple and their sum is direct. Indeed, assuming we have constructed $a_1, \ldots, a_j$ with $j < d$, the submodule $\bigoplus_{i \leq j} T\, a_i\, B \subseteq f\, C$ will have length $\leq j < d = \operatorname{length}(_T f\, C)$, so the fact that $f A\, e$ is left 1-strong (a weak consequence of the fact that (f, e) is left marked, hence that $f A\, e$ is left l_S-strong) allows us to find a_{j+1} with $T\, a_{j+1}\, B$ simple and not contained in that proper submodule; and since it is simple, its sum with that submodule is direct. Once we have $a_1, \ldots, a_d$, we see that the sub-bimodules $\sum_{i \leq j} T\, a_i\, S \subseteq A$ $(j \leq d)$ form a chain of length d. Hence $\operatorname{length}(_T(f A\, e)_S) \geq d = \operatorname{length}(_T f\, C)$, from which we see that replacing A with $(1 - f)A$ decreases the right-hand side of (16) by at least as much as replacing C with $(1 - f)C$ decreases the left-hand side. So again, if the new system satisfies (16), so did the old one.

Similarly, if G has some right vertex e such that the only edge adjacent to e is its associated right-marked edge (f, e), and this edge is not also left-marked, we find that we can remove $f A\, e$ from A, and $e\, B$ from B, and that if the resulting system satisfies (16), so does our original system. The proof is the same, except that where above we used an increasing family of submodules $\bigoplus_{i \leq j} T\, a_i\, B \subseteq f\, C$, we now use a decreasing family of submodules $\bigcap_{i \leq j} \ker(a_i\, S) \subseteq e\, B$.

Repeating these two kinds of reductions until no more instances are possible, we are left with a system in which every vertex not only hosts its own marked edge, but also hosts the marked edge of *at least one* other vertex (which may or may not be the same as its own marked edge). By counting vertices, we immediately see that in the preceding sentence, "*at least one*" can be replaced by "*exactly one*". From this, it is easy to see that each connected component of G is now either a loop of even length > 2, or a graph having just two vertices, and a single edge which is marked for both of these.

It follows that our system $_T A_S \times _S B \to _T C$ decomposes into a direct sum of subsystems corresponding to those connected components, and that (16) will be the sum of the corresponding inequalities for those components. Hence, it will suffice to prove (16) in the two cases where G is a loop, and where G has just a single edge. The l_S and l_T for each such system are $\leq$ the l_S and l_T for our original system, so (17) will hold for these systems, because it held for the original system.

If G is a loop, then each edge is marked for only one vertex, and the proof is quick: The Euler characteristic $\chi(G)$ is zero, while for each vertex, the bimodule corresponding to its marked edge has at least

the length of the module corresponding to that vertex, by the same "$\sum_{i \leq j} T a_i S$" arguments used in the preceding reduction. Summing these inequalities over the vertices, we have (16).

We are left with the case where G has just one right vertex, e, one left vertex, f, and the single edge (f, e). Thus, $B = eB$ and $C = fC$. Let

(18) $$m = \text{length}(_S B), \quad n = \text{length}(_T C).$$

Note that the fact that (f, e) is both left- and right-marked tells us that A is both left l_S-strong and right l_T-strong.

It is now that we will use our hypothesis (14). Let us begin by assuming the second of the two alternatives it offers, which in this situation says that whenever an $a \in A$ has kernel containing a maximal submodule of B, then some nonzero element of TaS has image in a simple submodule of C. Under this assumption, we begin by constructing, for m as in (18), elements $a_1, \dots, a_m \in A$ such that

(19) $T a_1 S, \dots, T a_m S$ have for kernels maximal submodules of B, none of which contains the intersection of the kernels of the others, and each $T a_i S$ has for image a simple submodule of C.

To see that we can do this, suppose inductively that we have constructed $i < m$ elements $a_1, \dots, a_i$ with these properties. Since $i < m = \text{length}(_S B)$, the intersection of the kernels of $T a_1 S, \dots, T a_i S$ is a nonzero submodule of B, hence since A, being right l_T-strong, is in particular right 1-strong, we can find $a \in A$ such that the kernel of $a S$ is a maximal submodule $B_0 \subseteq B$ and does not contain that intersection. Now any $a' \in T a_i S - \{0\}$ will still annihilate B_0; hence, if it is nonzero, $T a' S$ must have exactly that annihilator. By our assumption from (14), we can find such a nonzero a' which has image in a simple submodule of C. Taking for a_{i+1} this element, we have our desired inductive step; for the intersection of the kernels of $T a_1 S, \dots, T a_{i+1} S$ has co-length $i + 1$ in B, hence none can contain the intersection of the kernels of the others. Thus, we get (19).

I claim next that we can find $n - 1$ more elements, $a'_1, \dots, a'_{n-1} \in A$, where each a'_j satisfies

(20) $T a'_j B$ is a simple submodule of C, and for each $i = 1, \dots, m$, we have $T a'_j B \nsubseteq T a_i B + T a'_1 B + \dots + T a'_{j-1} B$.

To see this, suppose inductively that $a'_1, \dots, a'_{j-1}$ have been chosen. For each $i \leq m$, both $T a_i B$ and each of $T a'_1 B, \dots, T a'_{j-1} B$ are simple submodules of C, and there are $1 + (j - 1) = j < n$ of these, so their sum is a proper submodule. As i ranges from 1 to m we get $m = l_S$

such proper submodules; so as A is left l_S-strong, our a'_j can be chosen to satisfy (20).

Let us now show that as we add up successively the subbimodules $T a_1 S + \cdots + T a_m S + T a'_1 S + \cdots + T a'_{n-1} S$ of A, we get a chain of length $m + n - 1$. The first m steps are distinct by comparison of their kernels in B, in view of (19). If we had equality somewhere in the next $n - 1$ steps, this would mean that for some $j < n$ we would have

$$(21) \qquad T a'_j S \ \subseteq \ T a_1 S + \cdots + T a_m S + T a'_1 S + \cdots + T a'_{j-1} S.$$

To get a contradiction, let us, for $i = 1, \ldots, m$, write B_i for the intersection of the kernels of all the $T a_j S$ other than $T a_i S$. By (19), each B_i has co-length $m - 1$, hence is a simple submodule of B, and no B_i is contained in the sum of the others; so $\sum_i B_i = B$. Hence, some B_i is not in the kernel of $T a'_j S$; let us choose such a B_i and apply (21) to it. We get, on the left, $T a'_j B_i$, which by our choice of i is nonzero. Since it is contained in $T a'_j B$, which by (20) is simple, it must be equal thereto. On the right, by definition of B_i, we loose all of the first m terms other than the i-th. Replacing the remaining occurrences of B_i on the right by the larger module B, we get

$$(22) \qquad T a'_j B \subseteq T a_i B + T a'_1 B + \cdots + T a'_{j-1} B.$$

But this is one of the relations we chose a'_j to avoid in (20). This contradiction proves that our chain of submodules of A is strictly increasing, hence that A has length at least $m + n - 1 = \text{length}(_S B) + \text{length}(_T C) - \chi(G)$, as required.

If we are in the other case of (14), we operate dually, and first obtain n elements $a_1, \ldots, a_n$ of A which determine independent simple submodules of C, and each of which has kernel containing a maximal submodule of B, then choose $m - 1$ more elements $a'_1, \ldots, a'_{m-1}$, such that the submodule of B which each determines is maximal, and does not contain the intersection of the submodules determined by the proceeding members of that list, intersected with the submodule determined by any one of the a_i. $\qquad \square$

We can now get our main result. Since I don't see how to turn (13) or (14) into a condition on the ring R, I will only give the hypothesis corresponding to (12).

Condition (15) would put a finite lower bound on the size of k; but since this bound would involve both the structure of R and that of M, and we would like our restrictions on M to be stated in terms of the structure of R, I will achieve this simply by requiring k to be infinite.

The two rings S and T of the preceding development are the same ring $R/J(R)$ in the application below, so we shall distinguish the corresponding sorts of vertices of our graph with subscripts "left" and "right". We shall regard $\text{socle}(R)$ (the 2-sided socle of R, whose definition we recalled in the paragraph before Theorem 3) as a bimodule over $R/J(R)$; as such it will play the role of the A of Proposition 10.

Theorem 11. *Suppose k is an infinite field and R an overring of k, such that R is finite-dimensional as a left k-vector-space, k has central image in $R/J(R)$ and centralizes $\text{socle}(R)$, and $R/J(R)$ is a direct product of full matrix rings over k (this last condition being automatic if k is algebraically closed).*

Let G be the bipartite graph whose left vertex-set consists of symbols f_{left} for all minimal central idempotents $f \in R/J(R)$ such that $f\,\text{socle}(R) \neq \{0\}$, whose right vertex-set consists of symbols e_{right} for all minimal central idempotents $e \in R/J(R)$ such that $\text{socle}(R)\,e \neq \{0\}$, and whose edge-set is $\{(f_{\text{left}}, e_{\text{right}}) \mid f\,\text{socle}(R)\,e \neq \{0\}\}$.

Then for any faithful left R-module M such that no proper submodule or proper homomorphic image of M is faithful, we have

(23) $\text{length}(M/J(R)M) + \text{length}(\text{socle}(M)) \leq \text{length}(\text{socle}(R)) + \chi(G),$

where the two lengths on the left are as left R-modules, while the length on the right is as an (R, R)-bimodule.

Sketch of proof: Under the action of R on M, elements of $\text{socle}(R)$ annihilate $J(R)M$ and have image in $\text{socle}(M)$; so that action induces a balanced bilinear map of R-modules and bimodules, $\text{socle}(R) \times M/J(R)M \to \text{socle}(M)$. Since all four R-module structures involved (the three left module structures, and the right module structure of $\text{socle}(R)$) are annihilated by $J(R)$, that induced operation is a map of the form (7), with $R/J(R)$ in the roles of both S and T. Since M is faithful, this map satisfies (8), while as in the proof of Theorem 3, the minimality assumptions on M give (9) and (10). Since (12) implies (14), we can apply the preceding proposition, getting (23). $\qquad\qquad\square$

5. Examples, remarks, and questions

5.1. Counterexamples over small fields. To show that Proposition 10 can fail if condition (15) (the requirement that our division rings not be too small) is dropped, let k be the field of 2 elements, let $A = B = k \times k \times k$ and $C = k \times k$ (as abelian groups for the moment), and define $h\colon A \times B \to C$ by $h((\alpha_1, \alpha_2, \alpha_3), (\beta_1, \beta_2, \beta_3)) = \alpha_1\beta_1(1, 0) + \alpha_2\beta_2(0, 1) + \alpha_3\beta_3(1, 1)$. Letting $S = k \times k \times k$ and $T = k$,

and defining the module structures of $_T A_S$, $_S B$, and $_T C$ in the obvious ways, we see that h is indeed a balanced bilinear map, i.e., satisfies (7).

This map also clearly satisfies (12) and (8), and it is not hard to verify (9) and (10) as well. However, (15) fails, since $N_T = 2$, $d_T = 3$, $l_S = 1$. And in fact, the conclusion (16) fails: the left-hand side of that inequality is $3 + 2$, while the right-hand side is $3 + 1$. If we examine the steps of our proof of Proposition 10 in this case, we see that conditions (9) and (10) make A, regarded as a family of maps $B \to C$, left 2-strong, but of its three components $A\,e_i\colon e_i\, B \to C$, none is 1-strong. We can, as in the proof of that proposition, snip two leaves off G without changing the numerical relationship between the two sides of (16); but the remaining system, say $A\,e_1\colon e_1\, B \to C$, does not satisfy (16).

More generally, over any finite field k, say of q elements, one can get a similar example by taking $S = A = B = k^{q+1}$, $T = k$, and $C = k^2$, and letting the $q + 1$ components $A\,e_i\colon B \to C$ have for images the $q + 1$ one-dimensional subspaces of C. (The reader who has worked through the above $q = 2$ case should not find it hard to supply the details for this generalization.) Still more generally, if we take C to be d-dimensional ($d \geq 2$), we can make $S = A = B = k^{(q^d-1)/(q-1)}$, and let the natural basis of A act by maps having for images the $(q^d - 1)/(q - 1)$ one-dimensional subspaces of C.

We can adapt these constructions to get examples showing that for small k Theorem 11 likewise fails; but since in that case S and T must be the same, a bit of adjustment is needed. We can keep S as in those constructions, but let $T = S$, giving it actions on C and A under which all but one of its minimal idempotents annihilate those objects.

Let me describe in concrete terms an R and M based, in this way, on our initial example where k is the field of 2 elements. Our R will be the ring of all 4×4 matrices over this k with support in the union of the first row and the main diagonal, and whose $(1, 1)$ and $(2, 2)$ entries are equal. To obtain M, we start with the left R-submodule of $\mathrm{Matr}_{4,3}(k)$ spanned as a k-vector-space by $\{e_{11}, e_{12}, e_{13}, e_{21}, e_{32}, e_{43}\}$, and divide out by the subspace spanned by $e_{11} + e_{12} + e_{13}$. We find that $J(R)M = \mathrm{socle}(M) =$ the 2-dimensional space spanned by e_{11} and e_{12}; that M is a faithful R-module such that no proper submodule or homomorphic image of M is faithful, but that M does not satisfy (23), which for this case would say $3 + 2 \leq 3 + 1$.

5.2. An example not satisfying (14). I will give here an example of a system (7) arising as the map $\mathrm{socle}(R) \times M/J(M) \to \mathrm{socle}(M)$ for a faithful module M over an Artinian local ring R, which does not

satisfy (14), or our conclusion (16). In fact, it does not satisfy (9) or (10) either; but it may give some insight into how things can differ from the situation analyzed in the preceding sections.

Let K be the field $\mathbb{Q}(2^{1/3})$, let $F = K(\omega)$, where ω is a primitive cube root of unity, and let σ be the automorphism of F of order 3 which fixes ω, and takes $2^{1/3}$ to $\omega\,2^{1/3}$. Let $\mathrm{tr} = \mathrm{tr}_{F/K}\colon F \to K$ be the trace operation. (Thus, tr is K-linear, but σ is not.)

Let M be the $\mathbb{Q}$-vector-space F^2, and let us define the right shift operation $M \to M$,

$$(24) \qquad\qquad s\colon (a,b) \mapsto (0,a).$$

We shall understand tr, σ, and each element of F to act on M componentwise. (So if $a \in F$, the symbol a will also represent the operation of componentwise multiplication of elements of M by a, which does not in general commute with either tr or σ.)

We now define two operations on M,

$$(25) \qquad \begin{aligned} x &= \mathrm{tr}\,\sigma\,s\colon (a,b) \mapsto (0, \mathrm{tr}(\sigma(a))), \\ y &= \sigma\,\mathrm{tr}\,s\colon (a,b) \mapsto (0, \sigma(\mathrm{tr}(a))). \end{aligned}$$

From the fact that $[F : K] = 2$, it is easy to see that $K\,\sigma(K) = F = K\,\sigma^{-1}(K)$. Combining this with the fact that tr and s commute with the action of elements of K, we see that

$$(26) \qquad \begin{aligned} K\,x\,K &= K\,\mathrm{tr}\,\sigma\,s\,K = \mathrm{tr}\,K\,\sigma\,s\,K \\ &= \mathrm{tr}\,\sigma\,(\sigma^{-1}(K))\,s\,K = \mathrm{tr}\,\sigma\,s\,(\sigma^{-1}(K)\,K) = x\,F, \\ K\,y\,K &= K\,\sigma\,\mathrm{tr}\,s\,K = K\,\sigma\,\mathrm{tr}\,K\,s \\ &= K\,\sigma\,K\,\mathrm{tr}\,s = (K\,\sigma(K))\,\sigma\,\mathrm{tr}\,s = F\,y. \end{aligned}$$

These calculation show in particular that the bimodule operations of the (K, K)-bimodules spanned by each of x and y contain the operations of a 1-dimensional F-module; so each of these bimodules is simple.

Now let R be the ring of $\mathbb{Q}$-vector-space endomorphisms of M generated by the actions of the elements of K, together with the two endomorphisms x and y. We see that

$$(27) \qquad \begin{aligned} R &= K + K\,x\,K + K\,y\,K = K + x\,F + F\,y, \quad \text{and} \\ J(R) &= \mathrm{socle}(R) = x\,F + F\,y. \end{aligned}$$

The subideals of $\mathrm{socle}(R)$ generated by x and by y are isomorphic to one another as bimodules over $R/J(R) \cong K$, namely, each is isomorphic to the (K, K)-bimodule $F\,\sigma = \sigma\,F$. But as systems of maps $M/J(R)M \to \mathrm{socle}(M)$ they behave differently: xF has range in $(0, K\,\mathrm{tr}(F)) = (0, K) \neq (0, F)$, but has trivial kernel; Fy, dually, has all

of $(0, F)$ as range, but, identifying $M/J(R)M$ with F, its kernel is the nontrivial K-subspace $\ker(\mathrm{tr})$.

5.3. An approach that probably doesn't go anywhere. Theorem 11 does not cover the case where R is a general finite-dimensional algebra R over an infinite field k, since in that situation, $R/J(R)$ need not be a direct product of matrix rings. But it is natural to ask: Can't we start with such an R and an R-module M, extend scalars to the algebraic closure of k, and apply Theorem 11 to the resulting structures?

The trouble is that the relevant properties of R and M may not carry over under this change of base field. I do not know whether we can expect the minimality conditions on M to carry over; but the set of minimal central idempotents of R can certainly grow under such an extension. So I do not see how anything can be achieved in this way.

5.4. Some ways our results *can* be strengthened. One step in our proof of Proposition 10 was noticeably wasteful. When we dropped all edges that were neither left nor right marked, we counted each of them as contributing "at least 1" to $\mathrm{length}({}_T A_S)$. But depending on what we know about the structure of A, we may be able to raise this estimate. A difficulty is that knowing the structure of A doesn't tell us *which* edges will be dropped, nor even exactly how many: the answers may depend on B and C.

However, we do know that dropping all unmarked edges must result in a graph in which there are at least as many vertices as edges (possibly more, if some edges are marked for both their vertices); so if the graph G determined by A has more edges than vertices, i.e., has $\chi(G) < 0$, then at least $-\chi(G)$ edges (f, e) must be dropped. Thus, if we have in front of us a list of the lengths of all the nonzero bimodules fAe (with repetitions shown), then we can say that when we delete $-\chi(G)$ such bimodules, the sum of their lengths must be at least the sum of the $-\chi(G)$ smallest elements on that list. Deleting $-\chi(G)$ edges will leave us a system whose graph has Euler characteristic 0, to which we can apply Proposition 10 as it stands. Consequently, we have

Corollary 12. *In the context of Proposition 10, let the family (set with multiplicity) of lengths of nonzero subbimodules fAe of A, listed in ascending order, be $d_1 \leq d_2 \leq \cdots$. Then if $\chi(G) < 0$, the term $+\chi(G)$ in (16) can be improved (decreased) to*

$$(28) \qquad\qquad - d_1 - d_2 - \cdots - d_{-\chi(G)}.$$

The same applies, mutatis mutandis, to the inequality (23) of Theorem 11.

One can do still a bit better, using the fact that we don't delete as unmarked *all* the edges adjacent to any vertex. Consequently, if the lengths of the bimodules corresponding to edges adjacent to a certain vertex all lie among the first $-\chi(G)$ terms of the above sequence $d_1 \leq d_2 \leq \cdots$, then in (28) we can skip the largest of these lengths, and instead throw in $d_{-\chi(G)+1}$, which may be larger, at the end of the list; and iterate this process for other vertices.

One can also strengthen Theorem 11 by weakening the assumption that R contains k, to say that k is the residue field of a local ring contained in R, which again becomes central in $R/J(R)$ and centralizes socle(R).

I have not put these observations into my formal statement of Theorem 11, because I feel that the more urgent task is to see whether one can generalize Proposition 10 to avoid or weaken the awkward restriction (14); and that if one can, some of these generalizations might be embraced by broader, more easily stated results.

5.5. Sketch of a more elaborate version of the N-strong condition. The reader may have noticed that in the proof of Proposition 10, the final case, where G has just one edge, both right and left marked, is roughly the situation of Proposition 1 (as reformulated in Corollary 2), but the proof is different from that of the earlier proposition. The reason is that the information we have available at that point in the proof of Proposition 10, that A is left l_S-strong and right l_T-strong, while useful in finding elements of A whose images lie in simple submodules of C *outside* of given submodules, and elements whose kernels have the dual property, does not provide a way of finding such elements whose images lie together *in* some proper submodule (and dually for kernels). But that was what we needed in Case 1 of the proof of Proposition 1, when we chose $b_1, \ldots, b_{m-1} \in \ker(f)$.

However, I think the concept of an N-strong map can be modified to make it compatible with the method of proof of Proposition 1. Let me sketch how.

Given W, X, and Y as in Definition 4, and adding the assumption that Y has finite length, our modified definition will involve a concept of W being left N-strong *relative to* a nonzero submodule $Y' \subseteq Y$. That condition will be defined by recursion on the length of Y'. If Y' has length 1, then W will be called left N-strong relative to Y' if and only if W has a nonzero element whose image is contained in Y'. If the concept has been defined relative to submodules of some length $r \geq 1$, then W will be said to be left N-strong relative to a submodule Y' of length $r+1$

if and only if for every submodule $Y'' \subseteq Y'$ of length $r - 1$, there exist $> N$ submodules of length r between Y'' and Y' relative to which W is left N-strong. We will simply say W is left N-strong (in our new sense) if it is left N-strong relative to its codomain Y.

It is now easy to verify that if, for every simple submodule of $Y_0 \subseteq Y$, there is a nonzero element of W with image in Y_0, and if the division ring over which V is a full matrix ring has cardinality $\geq N$, then W is left N-strong under our new definition. Moreover, an easy induction shows that if a union $W = W_1 \cup \cdots \cup W_d$ is left N-strong, then at least one of the W_i is left N/d-strong.

The definition of right N-strong would be modified analogously.

In the proof of Proposition 1, the hypotheses (1) and (2) could then be weakened to say that (up to the change of notation appropriate to a subspace $A \subseteq B \otimes_k C$ rather than a map $A \times B \to C$), A is both left and right 1-strong in our modified sense. I suspect that the same method could be adapted to the last part of the proof of Proposition 10, and would in fact allow us to weaken the hypothesis (15) by dropping the factors l_S and l_T.

I haven't worked out the details, because they do not get at the serious restrictions in our results. In particular, the suggested change in the end of the proof of Proposition 10 would not get rid of the need for condition (14).

Let us now look at what we wish we could do.

5.6. Some questions. Here is an innocent-sounding generalization of our first main result that we might ask for.

Question 13. Does Theorem 3 remain true if the assumption that socle(R) is central in R is deleted?

To prove such a result, we would want a version of Proposition 1 involving a division ring rather than a field. As in the development of our results on non-local rings, I don't see a convenient way of "symmetrizing" the general statement we need, so in the next question, I will ask, not for the analog of that proposition, but of its corollary. Moreover, although the result needed just to get a positive answer to Question 13 would have for B and C vector spaces over the same division ring, I expect it would be no more difficult to prove such a result without that restriction; so let us pose the question as follows.

Question 14. Let S and T be division rings, $_SB$ and $_TC$ be nonzero finite-dimensional vector spaces, and $_TA_S$ be a subbimodule of the (T, S)-bimodule of all additive group homomorphisms $_SB \to {}_TC$. Suppose moreover that

(29) For every proper subspace $B' \subsetneq B$, there is at least one $a \in A$ whose restriction to B' is zero,

and

(30) For every proper homomorphic image C/C_0 of C, there is at least one $a \in A$ whose composite with the factor map $C \to C/C_0$ is zero.

Then must it be true that

$$(31) \qquad \mathrm{length}(_T A_S) \geq \dim_S(B) + \dim_T(C) - 1$$

(where $\mathrm{length}(_T A_S)$ denotes the length of A as a bimodule)?

I have posed this question in a relatively easy-to-state form, but my hope is that if a positive answer can be proved, then one will be able to push the proof further, and get the same result with the hypotheses (29) and (30) weakened to say that A is left and right N-strong for appropriate N, in the modified sense sketched in §5.5 above (or something like it), and that this could be used to prove generalizations of Proposition 10 and Theorem 11.

Incidentally, it would be harmless to throw into our hoped-for variant of Proposition 10 the added assumption that A is semisimple as a left module, a right module, and a bimodule, since those conditions are true of the socle of a left Artin ring, the situation to which we apply that result in Theorem 11.

5.7. A waffle about wording. I would have preferred a more suggestive term for what I have called left and right N-strong families of maps. I was tempted to replace "strong" with "ubiquitous"; but this might suggest too much – a property like (9) and (10), rather than the weaker property actually defined. Another thought was "prevalent"; but this seemed a bit vague.

6. Getting minimal faithful modules from module decompositions

This section is essentially independent of the rest of this note, though the final assertion proved will complement the results proved above. Namely, Proposition 17(iii) below, though it has a weaker conclusion than Theorems 3 and 11, is applicable when the hypotheses of those theorems are not satisfied. Aside from that final result, the focus will be on modules with only one of the two minimality properties considered in preceding sections. (The one other exception to independence from the rest of this note is that at one point below, we will refer to an example from the preceding section.)

The two preliminary lemmas below may be of interest in their own right. The final statement of each describes how, under certain conditions, a module can be decomposed into "small pieces": in the first, as a sum of submodules N such that $N/J(R)N$ is simple; in the second, as a subdirect product of modules N with socle(N) simple.

Lemma 15. *Let R be a left Artinian ring, and M a left R-module (not necessarily Artinian). Then*

(i) *If L is a simple submodule of $M/J(R)M$, then M has a submodule N such that the inclusion $N \subseteq M$ induces an isomorphism $N/J(R)N \cong L \subseteq M/J(R)M$.*

Hence

(ii) *Given a decomposition of $M/J(R)M$ as a sum of simple modules $\sum_{i \in I} L_i$, one can write M as the sum of a family of submodules N_i $(i \in I)$, such that for each i, $N_i/J(R)N_i \cong L_i$, and L_i is the image of N_i in $M/J(R)M$.*

Proof: In the situation of (i), let x be any element of M whose image in $M/J(R)M$ is a nonzero member of L. Thus, the image of $R x$ in $M/J(R)M$ is L. Since R is left Artinian, so is $R x$, hence we can find a submodule $N \subseteq R x$ minimal for having L as its image in $M/J(R)M$. Now since $N/J(R)N$ is semisimple, it has a submodule L' which maps isomorphically to L in $M/J(R)M$. If L' were a proper submodule of $N/J(R)N$, then its inverse image in N would be a proper submodule of N which still mapped surjectively to L, contradicting the minimality of N. Hence $N/J(R)N = L' \cong L$, completing the proof of (i).

In the situation of (ii), choose for each L_i a submodule $N_i \subseteq M$ as in (i). Then we see that $M = J(R)M + \sum_i N_i$, hence since R is Artinian, $M = \sum_i N_i$ [**4**, Theorem 23.16, (1) $\Longrightarrow$ (2′)], as required. $\square$

The next result is of a dual sort, but the arguments can be carried out in a much more general context, so that the result we are aiming for (the final sentence) looks like an afterthought.

Lemma 16. *Let R be a ring and M a left R-module. Then*

(i) *If L is any submodule of M, then M has a homomorphic image M/N such that the composite map $L \hookrightarrow M \to M/N$ is an embedding, and the embedded image of L is essential in M/N (i.e., has nonzero intersection with every nonzero submodule of M/N).*

Hence

  (ii) *If E is an essential submodule of M, and $f\colon E \to \prod_I L_i$ a subdirect decomposition of E, then there exists a subdirect decomposition $g\colon M \to \prod_i M_i$ of M, such that each M_i is an overmodule of L_i in which L_i is essential, and f is the restriction of g to $E \subseteq M$.*

  In particular, every locally Artinian module can be written as a subdirect product of locally Artinian modules with simple socles.

Proof: In the situation of (i), let N be maximal among submodules of M having trivial intersection with L. The triviality of this intersection means that L embeds in M/N, while the maximality condition makes the image of L essential therein. (If it were not essential, M/N would have a nonzero submodule M' disjoint from the image of L, and the inverse image of M' in M would contradict the maximality of N.)

In the situation of (ii), for each $j \in I$ let K_j be the kernel of the composite $E \to \prod_I L_i \to L_j$. Applying statement (i) with M/K_j in the role of M, and $E/K_j \cong L_j$ in the role of L, we get an image M_j of M/K_j, and hence of M, in which L_j is embedded and is essential. Now since E is essential in M, every nonzero submodule $M' \subseteq M$ has nonzero intersection with E, and that intersection has nonzero projection to L_i for some i; so in particular, for that i, M' has nonzero image in M_i. Since this is true for every M', the map $M \to \prod_I M_i$ is one-to-one, and gives the desired subdirect decomposition.

To get the final assertion, note that the socle of a locally Artinian module is essential, and, being semisimple, can be written as a subdirect product (indeed, as a direct sum) of simple modules; so we can apply (ii) with $E = \mathrm{socle}(M)$ and the L_i simple. Each of the M_i in the resulting decomposition will have a simple essential submodule L_i, so that submodule must be its socle. $\qquad\Box$

We can now get the following result, showing that given a faithful module M over an Artinian ring, we can carve out of M a "small" faithful submodule, factor-module, or subfactor. Note that in the statement, the length of $\mathrm{socle}(R)$ as a bimodule may be less than its length as a left or right module. (For example, the full $n \times n$ matrix ring over a division ring is its own socle, and has length n as a right and as a left module, but length 1 as a bimodule. Similarly, in the R of §5.2, each of the direct summands $x\,F$ and $F\,y$ of $\mathrm{socle}(R)$ has length 2 as left and as right R-module, but length 1 as an R-bimodule; so $\mathrm{socle}(R)$ has length 4 on each side, but 2 as a bimodule.)

Proposition 17. *Let R be a left Artinian ring, let n be the length of* socle(R) *as a bimodule (equivalently, as a 2-sided ideal), and let M be a faithful R-module. Then*

(i) *M has a submodule M' which is again faithful over R, and satisfies* length$(M'/J(R)\,M') \le n$. *(In particular, M' is generated by $\le n$ elements.)*

(ii) *M has a homomorphic image M'' which is faithful over R, and satisfies* length(socle$(M'')) \le n$.

(iii) *M has a subfactor faithful over R satisfying both these inequalities.*

Proof: To get (i), note that since $R/J(R)$ is semisimple Artin, $M/J(R)M$ can be written as a direct sum of simple modules L_i $(i \in I)$ over that ring, and hence over R, so we can construct a generating family of submodules $N_i \subseteq M$ related to these as in Lemma 15(ii). Since $M = \sum_i N_i$ is faithful, and socle(R) has length n as a 2-sided ideal, the sum of some family of $\le n$ of these submodules, say

$$(32) \qquad\qquad M' = N_{i_1} + \cdots + N_{i_m} \quad \text{where } m \le n,$$

must have the property that M' is annihilated by no nonzero subideal of socle(R). (Details: one chooses the N_{i_j} recursively; as long as $N_{i_1} + \cdots + N_{i_j}$ is annihilated by a nonzero subideal $I \subseteq$ socle(R), one can choose an $N_{i_{j+1}}$ which fails to be annihilated by I. The annihilators in socle(R) of successive sums $M' = N_{i_1} + \cdots + N_{i_j}$ $(j = 0, 1, \dots)$ form a strictly decreasing chain, so this chain must terminate after $\le n$ steps.) Since an ideal of an Artinian ring having zero intersection with the socle is zero, M' has zero annihilator, i.e., is faithful. Since each N_i satisfies length$(N_i/J(R)N_i) = 1$, we have length$(M'/J(R)M') \le m \le n$.

Statement (ii) is proved in the analogous way from the final statement of Lemma 16, using images of M in products of finite subfamilies of the M_i in place of submodules of M generated by finite subfamilies of the N_i.

Statement (iii) follows by combining (i) and (ii). $\qquad\square$

References

[1] G. M. Bergman, Commuting matrices, and modules over Artinian local rings, unpublished note, `arXiv:1309.0053`.

[2] D. Eisenbud, Linear sections of determinantal varieties, *Amer. J. Math.* **110(3)** (1988), 541–575. DOI: `10.2307/2374622`.

[3] T. H. Gulliksen, On the length of faithful modules over Artinian local rings, *Math. Scand.* **31** (1972), 78–82.

[4] T. Y. Lam, *"A first course in noncommutative rings"*, Graduate Texts in Mathematics **131**, Springer-Verlag, New York, 1991. DOI: `10.1007/978-1-4684-0406-7`.

[5] K. C. O'Meara, J. Clark, and C. I. Vinsonhaler, *"Advanced topics in linear algebra"*, Weaving matrix problems through the Weyr form, Oxford University Press, Oxford, 2011.

[6] `http://mathoverflow.net/questions/141378/is-this-lemma-in-elementary-linear-algebra-new`.

University of California
Berkeley, CA 94720-3840
USA
E-mail address: `gbergman@math.berkeley.edu`