

The Ascending Chain Condition

E. L. Lady

The goal is to find a condition on an R -module M which will ensure the following property:

(★) **Every surjective endomorphism of M is an automorphism.**

To start with, we should wonder how the above property could ever fail. If $\varphi: M \rightarrow M$ and φ is a surjection, then φ will be an automorphism if and only if φ is monic. Now if $K = \text{Ker } \varphi$, then $\varphi(M) \approx M/K$ and so if φ is surjective, then

$$M/K \approx M.$$

It seems a little strange that this could ever happen with $K \neq 0$, and in fact, a preliminary search fails to turn up any obvious examples. (It turns out that there's a good reason for this.)

Before continuing, consider the dual problem for a moment, viz. when can it happen that a monomorphism could fail to be surjective. If $\psi: M \rightarrow M$ is monic, then $\psi(M) \approx M$. If, now, ψ is not surjective, then $\psi(M)$ would be a proper submodule of M isomorphic to M itself. Although this may seem unlikely, a little thought brings up lots of examples. For instance, if we think of a ring R as a left module over itself, then a submodule L (i. e. a left ideal) will be isomorphic to R if and only if it is principal (**exercise**). Thus almost every ring has lots and lots of proper left ideals isomorphic to the whole ring. There are lots and lots of other examples of monic endomorphisms of modules which are not surjective.

So why does it seem, in comparison, so difficult to imagine a surjective endomorphism that's not monic?

The way I visualize this is that if K is a non-trivial submodule of M , then there is a collapsing involved in going from M to M/K . And it's hard to imagine how one can collapse a part of M and still wind up with M/K being isomorphic to M itself.

In fact, when you think about it, the situation seems more and more bizarre. If $M/K \approx M$, then M/K itself must have a submodule corresponding to K , and this

submodule must have the form K_2/K , where $K \subseteq K_2$. (In fact, if $\varphi : M \rightarrow M$ is the surjection with kernel K , then $K_2 = \varphi^{-1}(K)$.) And then

$$M/K_2 \approx \frac{M/K}{K_2/K} \approx \varphi(M)/\varphi(K_2) = M/K \approx M.$$

Continuing in this way inductively, we see that if $M/K \approx M$ for $K \neq 0$, then there would be a sequence of submodules

$$K \subset K_2 \subset K_3 \subset K_4 \subset \dots$$

continuing indefinitely, where each $K_i \neq K_{i-1}$ (WHY?) and $M/K_i \approx M$. In fact, K_i is the pre-image of K under the surjection $M \rightarrow M/K_{i-1}$.

What is happening then is that we are collapsing more and more out of the bottom (so to speak) of M , and yet the result is still isomorphic to all of M . It would seem, then, that somehow there must be an infinite springiness at the top of M to fill in what is getting collapsed. (Needless to say, this is a very unsatisfactory way for a mathematician to explain things. But this is the way I tend to think when I get lost.)

What has to be involved here is, of course, the “paradox of infinity.” One can think of the example of the Infinite Hotel (due to Hilbert, as I recall), where there are denumerably many rooms, numbered by the positive integers. When a new guest arrives, there is no problem, because each existing guest simply advances to the next room, leaving Room Number One available for the new guest.

The situation we’re thinking of now, though, is more like that where the guest in Room Number One checks out, and then all the other guests move down by one room, so that once again all the rooms are empty.

Now this is an admittedly very fanciful way of thinking about the example we seek. But the Infinite Hotel metaphor actually suggests a specific construction.

Take any module. To be specific, let’s take \mathbb{Z} , regarded as a \mathbb{Z} module. Now form the direct sum of a countable number of copies of \mathbb{Z} , indexed by the natural numbers:

$$M = \bigoplus_1^{\infty} \mathbb{Z}.$$

Think of the elements of M as being “infinity-tuples,” i. e. having the form (x_1, x_2, \dots) (with almost all coordinates being zero, so that we get the direct sum rather than the

direct product). Now define an endomorphism φ by letting φ shift the coordinates of every element of \bigoplus_1^∞ one space “to the left,” i. e.

$$\varphi(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots).$$

(In other words, φ is, roughly speaking, like having Guest Number 1 check out of the hotel and all the other guests move down a room.) Then it is easy to see that for any $m \in M$ there exists (at least one) $m' \in M$ such that $m = \varphi(m')$. Therefore φ is surjective. But φ is not monic, since $\text{Ker } \varphi \neq 0$. In fact, $\text{Ker } \varphi$ consists of all elements of M which “live completely in the first coordinate.”

Exercise. For a less contrived example, let M be the \mathbb{Z} -module \mathbb{Q}/\mathbb{Z} and let n be a fixed integer larger than 1. Define $\varphi: M \rightarrow M$ by setting $\varphi(x) = nx$. Show that φ is a surjection but is not monic.

Now that we have seen that there do indeed sometimes exist surjective endomorphisms which are not monic, we can see that it is worthwhile to consider the original question.

When will a module M have the property that all surjective endomorphisms are automorphisms?

The usual answer to this question is already inherent in the preceding discussion.

Definition. An R -module M is said to be **noetherian** or to satisfy the **ascending chain condition** if there does not exist any infinite strictly ascending chain

$$N_1 \subsetneq N_2 \subsetneq N_3 \subsetneq \dots$$

of submodules of M .

Exercise. Use the ideas in the preceding discussion to write a proof for the following theorem.

Theorem. A surjective endomorphism φ of a noetherian R -module M is an automorphism of M .

In my opinion, it is important to do the preceding exercise as stated. It doesn't lead to quite the most elegant proof, though.

SECOND PROOF OF THEOREM Let $\varphi: M \rightarrow M$ have the property that $\varphi(M) = M$ and let $K_i = \text{Ker}(\varphi^i)$. Then clearly $K_i \subseteq K_{i+1}$ (WHY?). If M is noetherian, then the chain $K_1 \subseteq K_2 \subseteq \dots$ cannot be strictly ascending forever, so eventually one must reach an integer n such that $K_{n+1} = K_n$.

First note that since $M = \varphi(M)$, we have in fact,

$$M = \varphi(M) = \varphi(\varphi(M)) = \dots = \varphi^n(M).$$

We will now prove that $\text{Ker } \varphi = 0$, so that φ is monic. In fact, let $x \in \text{Ker } \varphi$. Since $M = \varphi^n(M)$, there exists $m \in M$ such that $x = \varphi^n(m)$. Then $0 = \varphi(x) = \varphi^{n+1}(m)$, so $m \in \text{Ker}(\varphi^{n+1}) = K_{n+1} = K_n = \text{Ker}(\varphi^n)$. Thus $x = \varphi^n(m) = 0$. This shows that $\text{Ker } \varphi = 0$ so that φ is monic. \square

It can sometimes be worthwhile to see whether a result like this can be pushed a little further, i. e. whether the hypothesis can be weakened or the conclusion strengthened.

The crucial point in the second proof given above is where we have $x \in \text{Ker } \varphi$ and $x \in \varphi^n(M)$ and conclude that $x = 0$. Therefore, this proof actually yields the following result, which is stronger but seems of questionable value.

Mini-Result. If φ is an endomorphism of an R -module M and for some n , $\text{Ker } \varphi^n = \text{Ker } \varphi^{n+1}$, then $\varphi^n(M) \cap \text{Ker } \varphi = 0$.

If one ponders this little oddity for a few days, trying to find some value in it, it might occur to one that **if it were true** that $\varphi^n(M) + \text{Ker } \varphi$ were all of M , then since $\varphi^n(M) \cap \text{Ker } \varphi = 0$, we would have $M = \varphi^n(M) \oplus \text{Ker } \varphi$, a result that seems worthwhile.

But this is really stretching. It seems more in the realm of wishful thinking than productive exploration. However if one were to wonder under what circumstances it might be true that $M = \varphi^n(M) + \text{Ker } \varphi$, it might finally occur to one that since $\text{Ker } \varphi \subseteq \text{Ker } \varphi^2 \subseteq \dots$, one might stand a little better chance of proving that $M = \varphi^n(M) + \text{Ker}(\varphi^n)$, although this still seems like a long shot. We can record this observation for the record.

Mini-Proposition. If $\text{Ker}(\varphi^n) = \text{Ker}(\varphi^{n+1})$ and $M = \varphi^n(M) + \text{Ker}(\varphi^n)$, then $M = \varphi^n(M) \oplus \text{Ker}(\varphi^n)$.

PROOF: All this proposition is really saying is that $\varphi^n(M) \cap \text{Ker}(\varphi^n) = 0$. But this is clear from the reasoning we have already seen. In fact, if $x = \varphi^n(m)$ and $x \in \text{Ker}(\varphi^n)$, then $\varphi^{2n}(m) = 0$, so $m \in \text{Ker}(\varphi^{2n}) = \text{Ker}(\varphi^n)$, so $x = \varphi^n(m) = 0$. \square

Now this is still not a promising result. When I'm doing research and start getting a bunch of results like this that are purely stretching, wishful thinking, my inclination is to give up and start looking in other directions.

However in this particular case, that would be a mistake.

Suppose that we know not only that $\text{Ker}(\varphi^n) = \text{Ker}(\varphi^{n+1})$, but also that $\varphi^n(M) = \varphi^{n+1}(M)$. Then in fact we could prove that $M = \varphi^n(M) + \text{Ker}(\varphi^n)$, as stated below.

Proposition. Suppose that φ is an endomorphism of M such that for a certain $n \geq 1$, $\varphi^n(M) = \varphi^{n+1}(M)$ and $\text{Ker}(\varphi^n) = \text{Ker}(\varphi^{n+1})$. Then $M = \varphi^n(M) \oplus \text{Ker}(\varphi^n)$.

PROOF: We have already seen that $\varphi^n(M) \cap \text{Ker}(\varphi^n) = 0$. It remains to prove that $M = \varphi^n(M) + \text{Ker}(\varphi^n)$. So let $m \in M$. We need to find $y \in M$ and $k \in \text{Ker} \varphi^n$ such that $m = \varphi^n(y) + k$. Since in this case, $k = m - \varphi^n(y)$, another way of putting this is that we need to find $y \in M$ such that $m - \varphi^n(y) \in \text{Ker}(\varphi^n)$. In other words, we need to find $y \in M$ such that $\varphi^n(m - \varphi^n(y)) = \varphi^n(m) - \varphi^{2n}(y) = 0$.

A thought that comes to mind immediately is to choose $y = m$, but this turns out not to work. At this point, it is appropriate to remind oneself of the hypothesis of the theorem. We are given (among other things) that $\varphi^{n+1}(M) = \varphi^n(M)$ and it follows immediately that $\varphi^{2n}(M) = \varphi^n(M)$. And we are given $m \in M$ and looking for a y such that $\varphi^{2n}(y) = \varphi^n(m)$.

But the existence of such a y is precisely what is meant by the statement that $\varphi^{2n}(M) = \varphi^n(M)$. Therefore all the pieces of the proof are present, and it only remains to write them up in a more coherent form (**Exercise!**) that will convince the reader that we knew where we were going from the beginning. \square

Now why does this preceding proposition seem to have more apparent value than the two mini-results that precede it? (The reader, of course, is probably taking this on faith, just as the author has the benefit of hindsight in assessing the value of the proposition. But one should take time to think about such things.)

In trying to assess the value of a result, one needs to ask the following sorts of questions. (And even then, only time and further investigation will yield the true answer.)

- (1) *Does the conclusion of the theorem seem like something that would be useful?*
- (2) *Is the conclusion something inherently interesting?*
- (3) *Does the hypothesis give a condition that would be easy to recognize?*
- (4) *Is the hypothesis a condition that would frequently occur, or at least normally occur in certain important situations?*
- (5) *Does the conclusion of the theorem seem to say substantially more than the hypothesis does, or at least something substantially different?*

It's hard for a student to answer questions like this, of course, since a student doesn't know the territory and hasn't developed a good sense of what is surprising and what is commonplace. But let's consider the hypothesis of the proposition above.

We know that $\text{Ker } \varphi^n = \text{Ker } \varphi^{n+1}$ will be true for some n if M satisfies the ascending chain condition. Analogously, it is easy to show that $\varphi^t(M) = \varphi^{t+1}(M)$ will be true for some t if M satisfied the opposite condition, a **descending chain condition**. Now, even assuming that both of these are true, there doesn't seem to be any good reason to suppose that $t = n$. But we can quickly remedy this glitch by simply taking the larger of t and n (EXPLAIN!). Thus the presence of both chain conditions would guarantee that for any endomorphism φ of M , there exists n such that $\text{Ker } \varphi^n = \text{Ker } \varphi^{n+1}$ and also $\varphi^n(M) = \varphi^{n+1}(M)$, and so the above theorem would apply.

Now would this be a good theorem or a contrived one? Well, it might seem contrived because at first thought, it seems like it will be pretty hard to decide whether a module satisfies both the ascending and descending chain conditions. After all, even for a small module, there are usually an infinite number of possible ascending and descending chains, and we can hardly examine them all to see whether they eventually stabilize.

So look at it this way: Are there any common situations where we can be sure that both chain conditions would hold?

Certainly for finite dimensional vector spaces one could not have any infinite strictly ascending or descending chains of subspaces (WHY?). So we get the following theorem in Linear Algebra:

Theorem. If V is a finite dimensional vector space and φ a linear operator on V , then $V = U \oplus W$, where U and W are subspaces of V such that $\varphi(U) \subseteq U$, $\varphi(W) \subseteq W$, and the restriction of φ to U is an automorphism of U , and the restriction to W is nilpotent.

PROOF: Let $U = \varphi^n(V)$ and $W = \text{Ker}(\varphi^n)$ for any sufficiently large n . (In fact, $n = \dim V$ will definitely be large enough.) And apply the above reasoning. The only assertion that is new is that φ restricts to an automorphism of $\varphi^n(V)$. In fact, the restriction of φ to an operator on $\varphi^n(V)$ is a surjection, since by choice of n , $\varphi(\varphi^n(V)) = \varphi^{n+1}(V) = \varphi^n(V)$. But it's known from linear algebra (or from preceding results in the notes here) that a surjective endomorphism of a finite-dimensional vector space is an automorphism. \square

Now vector spaces generally play the role of the trivial case for theorems in module theory. So if the above theorem turns out to be worthwhile even in this quasi-trivial case, this is a pretty clear indication that we've found something good.

In fact, the theorem given is a major result in linear algebra, a foreshadowing of the theory of canonical forms. To see this better, let us phrase it in terms of matrices.

If we choose a basis for the subspace U above and a basis for W , then together these form a basis for V . If we look at the matrix for φ in terms of this basis, the fact that $\varphi(U) \subseteq U$ and $\varphi(W) \subseteq W$ tells us that this matrix splits up into four blocks, where the upper right and lower left block are zero (EXPLAIN!). A little more thought shows that our theorem can be rephrased as follows.

Theorem. Let A be an $n \times n$ -matrix over any field or skewfield. Then A is similar to a matrix of the form

$$\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix},$$

where B is a non-singular square matrix and C a nilpotent square matrix.

PROOF: **Exercise.**

Now part of the thrust of the preceding development I have presented is that modules with finite length are very much like finite-dimensional vector spaces. So now that we have a theorem for vector spaces, we can try to generalize it to modules with finite length. And, in fact, this is easy, because clearly a module with finite length must satisfy both the ascending and descending chain conditions.

Therefore we now have what is in fact one of the most important basic results in the theory of modules with finite length.

Fitting's Lemma. If M is a module with finite length and φ is an endomorphism of M , then there exist submodules N and K of M such that $M = N \oplus K$ and $\varphi(N) \subseteq N$, $\varphi(K) \subseteq K$, and the restriction of φ to an endomorphism of N is an automorphism, and the restriction of φ to K is nilpotent.

On the face of it, at least, this seems fairly promising. But we can see even more value in it if we play "What if?" a little more.

What if a module M with finite length was of such a nature that it was not possible to write it as a direct sum in a non-trivial way?

A module like this is called **indecomposable**. In other words,

Definition. M is indecomposable if whenever $M = N \oplus P$, then either $N = 0$ or $P = 0$.

Now it is fairly clear that any module with finite length can be broken down into a finite direct sum of indecomposable summands (EXPLAIN!). (Usually this can be done in several different ways.) So indecomposable modules seem of legitimate interest.

Surely any module with length 1 (i. e. a simple module) is indecomposable (EXPLAIN!). But over most rings there also exist indecomposable modules which are not simple.

From the preceding theorem, we immediately get

Corollary. If M is an indecomposable module with finite length, then every endomorphism of M is either an automorphism or is nilpotent.

Finally, from this we can get a theorem which is both intrinsically interesting and useful.

Theorem. Let M be an indecomposable R -module with finite length. Then the set of nilpotent endomorphisms of M forms a two-sided ideal in $\text{End}_R M$. This ideal is maximal and is the only maximal two-side ideal in $\text{End}_R M$. Every element of $\text{End}_R M$ not belonging to the maximal ideal is invertible.

Before giving the proof, I would like to point out that this makes $\text{End}_R M$ a very exceptional type of ring. (These called, in the trade, *local* rings. We'll talk about them quite a bit later.) In general, most rings have quite a few different maximal (two-sided) ideals, often an infinite number. Furthermore, in most non-commutative rings (which includes almost all endomorphism rings), the set of nilpotent elements does not form an ideal, since the sum of two non-commuting nilpotent elements is not usually nilpotent.

Proof of Theorem. Let \mathcal{J} be the set of nilpotent endomorphisms of M . Since M is indecomposable with finite length, we have proved that all non-nilpotent endomorphisms are automorphisms of M , and this is the same as saying that they are invertible elements of $\text{End}_R M$ (EXPLAIN!).

Now we have seen previously that since M has finite length, if $\varphi \in \text{End}_R M$ and φ is either monic or surjective, then φ is an automorphism. Thus \mathcal{J} can be characterized as the set of endomorphisms of M which are not monic, and also as the set of endomorphisms which are not surjective. Now let $\varphi \in \mathcal{J}$ and $\psi \in \text{End}_R M$. Then φ is not monic, so $\psi\varphi$ cannot be monic (EXPLAIN!), and it follows that $\psi\varphi \in \mathcal{J}$. Likewise, φ is not surjective, so $\varphi\psi$ cannot be surjective (EXPLAIN!) and it follows the $\varphi\psi \in \mathcal{J}$.

Therefore to see that \mathcal{J} is a two-sided ideal, we need only prove that if $\varphi_1, \varphi_2 \in \mathcal{J}$ then $\varphi_1 + \varphi_2 \in \mathcal{J}$. This turns out to be the hard part. Suppose by way of contradiction that $\varphi_1 + \varphi_2 \notin \mathcal{J}$. Then, as shown previously, $\varphi_1 + \varphi_2$ must be invertible in $\text{End}_R M$, i.e. there exists θ such that $\theta(\varphi_1 + \varphi_2) = 1$ (where here, 1 denotes the identity map on M). Now as just shown, $\theta\varphi_1$ and $\theta\varphi_2$ belong to \mathcal{J} , hence they are nilpotent, so there exist integers r and s such that $(\theta\varphi_1)^r = 0 = (\theta\varphi_2)^s$. But since $\theta\varphi_2 = 1 - \theta\varphi_1$, these two maps $\theta\varphi_1$ and $\theta\varphi_2$ commute with each other, and one then sees from the Binomial Theorem that $1^{r+s} = (\theta\varphi_1 + \theta\varphi_2)^{r+s} = 0$, which is absurd. This contradiction shows that $\varphi_1 + \varphi_2 \in \mathcal{J}$, finishing the proof that \mathcal{J} is an ideal.

Finally, we see that \mathcal{J} is the unique maximal ideal of $\text{End}_R M$. In other words, \mathcal{J} contains every proper ideal in $\text{End}_R M$. In fact, if I is an ideal not contained in \mathcal{J} , then I contains at least one element not in \mathcal{J} , and we have seen that such an element is invertible in $\text{End}_R M$. But from this, it follows that I must be the whole ring (Exercise). \square