8. Linear Transformations and module homomorphisms

Let M and N be modules over a commutative ring R (same ring for both!):

Definition 32. *A function $L\colon M \longrightarrow N$ is a* homomorphism *if*

1. $L(\mathbf{v} +_M \mathbf{w}) = L(\mathbf{v}) +_N L(\mathbf{w})$ *for all* $\mathbf{v}, \mathbf{w} \in M$, *and*
2. $L(\lambda \cdot_M \mathbf{v}) = \lambda \cdot_N L(\mathbf{v})$ *for all* $\mathbf{v} \in M$ *and numbers* $\lambda \in R$.

The set of all linear transformations from M to N will be denoted $\mathrm{Hom}_R(M, N)$

Note: When $R = \mathbb{F}$ is a field, these are called linear transformations of vector spaces and the set of all such transformations from the vector space V to the vector space W is denoted $\mathcal{L}_{\mathbb{F}}(V, W)$.

(Question 8.1) Suppose $L\colon M \longrightarrow N$ is a homomorphism of R-modules. Prove

1. $L(\mathbf{0}_M) = \mathbf{0}_N$
2. For any elements $\mathbf{v}, \mathbf{w} \in M$ and $\lambda \in R$, $L(\mathbf{v} +_M \lambda\mathbf{w}) = L(\mathbf{v}) +_N \lambda L(\mathbf{w})$
3. Suppose $[\lambda_1, \ldots, \lambda_k]$ is an ordered subset of R and $[\mathbf{v}_1, \ldots, \mathbf{v}_k]$ is an ordered subset of M . Prove by induction that

$$L(\lambda_1\mathbf{v}_1 + \cdots + \lambda_k\mathbf{v}_k) = \lambda_1 L(\mathbf{v}_1) + \cdots + \lambda_k L(\mathbf{v}_k)$$

In particular, the conditions in the definition force L to take linear combinations of $[\mathbf{v}_1, \ldots, \mathbf{v}_k]$ to linear combinations of $[L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)]$.

(Question 8.2) Use the results of the previous question to show that the following functions are NOT linear transformations:

1. $L\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by $L(x, y) = x^2 + 3y$,
2. $T\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ given by $T(x, y, z) = (2x + 3y, x + 4z - 2)$

(Question 8.3) Show that the following are R-module homomorphisms:

1. The identity map $I\colon M \longrightarrow M$ where $I(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in M$.
2. The zero map $\mathbf{0}\colon M \longrightarrow N$ where $\mathbf{0}(\mathbf{x}) = \mathbf{0}_N$ for all $\mathbf{x} \in M$.

bf Note: The zero map is the only map into or out of the trivial module $0 = \{\mathbf{0}_M\}$.

In the special case that $M = R^n$ we will denote the identity map by $\mathbf{I}_n$

(Question 8.4) Show that $L(\mathbf{A}) = \mathbf{A} - \mathbf{A}^T$ is a module homorphism $M_{n\times n}(\mathbb{F}) \longrightarrow M_{n\times n}(\mathbb{F})$.

We now show that the homomorphisms of R-modules generalize the functions that defined the systems of linear equations at the beginning of the class. Elements of $\mathrm{Hom}_R(M, N)$ still have properties similar to the systems of linear equations you are familiar with. Let $T \in \mathrm{Hom}_R(M, N)$ and $\mathbf{q} \in N$. In particular, if we try to solve the equation $T(\mathbf{x}) = \mathbf{q}$:

(Question 8.5) Suppose $\mathbf{u} \in M$ is any vector such that $T(\mathbf{u}) = \mathbf{0}_N$ and $\mathbf{p}$ is some (particular) solution to $T(\mathbf{x}) = q$. Show that $\mathbf{v} + \mathbf{u}$ is a solution to $T(\mathbf{x}) = q$.

Thus, given any solution to $T(\mathbf{x}) = \mathbf{q}$, we can obtain another solution by adding a vector $\mathbf{u}$ which solves the *homogeneous* equation $T(\mathbf{x}) = \mathbf{0}$. Furthermore,

(Question 8.6) Suppose $\mathbf{v}, \mathbf{w}$ are *distinct* solutions to $T(\mathbf{x}) = \mathbf{q}$. Show that $\mathbf{u} = \mathbf{v} - \mathbf{w}$ is a *non-zero* solution to the homogeneous equation $T(\mathbf{x}) = \mathbf{0}$.

We give the set of solutions to the homogeneous equation a name:

Definition 33. *The* kernel[11] *of a R-module homorphism $T\colon M \longrightarrow N$ is the subset*

$$\mathrm{Ker}(T) = \left\{ \mathbf{u} \in M \,\middle|\, L(\mathbf{u}) = \mathbf{0}_N \right\}$$

(Question 8.7) Explain why the questions above imply: if $T(\mathbf{v}) = \mathbf{q}$ for some $\mathbf{v} \in M$ then the set of solutions to $T(\mathbf{x}) = \mathbf{q}$ equals

$$\mathbf{v} + \mathrm{Ker}(T) = \left\{ \mathbf{v} + \mathbf{u} \,\middle|\, \mathbf{u} \in \mathrm{Ker}(T) \right\}$$

This exercise implies that the set of solutions to $L(x) = \mathbf{q}$ for different q are all the same "size," since they will all be in bijection with $\mathrm{Ker}(L)$.

(Question 8.8) Let $T \in \mathrm{Hom}_R(M, N)$. Prove that $\mathrm{Ker}(T)$ is a subspace of M.

(Question 8.9) Use that $L(\mathbf{A}) = \mathbf{A} - \mathbf{A}^T$ is a linear transformation $M_{n\times n}(\mathbb{F}) \longrightarrow M_{n\times n}(\mathbb{F})$, to show that $\mathrm{Sym}(n)$ is a subspace of $M_{n\times n}(\mathbb{F})$. Can you find a linear transformation whose kernel is $\mathrm{Asym}(n)$?

(Question 8.10) Show that $f \longrightarrow \dfrac{d^n f}{dx^n}$ is a linear transformation $C^n(\mathbb{R}) \longrightarrow C^0(\mathbb{R})$. Use calculus to show that $\mathrm{Ker}(L)$ the set of polynomials with real coefficients of degree $\leq\ n-1$. Show that the addition and scalar multiplication is the same as from $\mathbb{R}_{n-1}[x]$.

Since $T \in \mathrm{Hom}_R(M, N)$ is a function, it has an image:

Definition 34. *The image of T is*

$$\mathrm{Im}(T) = \left\{ T(\mathbf{m}) \,\middle|\, \mathbf{m} \in M \right\}$$

(Question 8.11) Show that $Im(T)$ is an R-sub-module of N.

Recall that for any function $g\colon A \longrightarrow B$, where A and B are sets:

(1) g is *surjective* (or is a surjection, or is onto) if $g(x) = b$ has *at least* one solution for each $b \in B$. Thus g surjective if $\mathrm{Im}(g) = B$.
(2) g is *injective* (or is an injection, or is one-to-one) if $g(x) = b$ has *at most* one solution for each $b \in B$. More formally, if $g(a) = b$ and $a \neq a'$ then $g(a') \neq b$.
(3) g is *bijective* (or is a bijection, or is a correspondence) if g is simultaneously injective and surjective. That is, if $g(x) = b$ has exactly one solution for each $b \in B$.

[11]For vector spaces, kernels exist for matrices, and the set we define here is called the *null-space* of the linear transformation, and denoted $\mathrm{NS}(T)$. However, this adds different terminologies for essentially the same thing. In algebra similar constructions have now come to be known ubiquitously as kernels.

For $T \in \mathrm{Hom}_R(M,N)$ to be surjective, we only need $\mathrm{Im}(T) = N$. For T to be injective we need that $T(\mathbf{m}) = T(\mathbf{m}')$ implies $\mathbf{m} = \mathbf{m}'$.

(Question 8.12) Prove that T is injective if and only if $\mathrm{Ker}(T) = \{\mathbf{0}_M\}$

There is also a pictoral language for depicting many properties of homomorphisms. A diagram such as

$$M \xrightarrow{\quad T \quad} N \xrightarrow{\quad S \quad} O$$

where M, N and O are R-modules and T and S are R-module homomorphisms, is called a *sequence* of R-module homomorphisms. As with any functions, we can compose these homomorphisms since the domain of one is the co-domain of the other. That is, if $T \in \mathrm{Hom}_R(M,N)$ and $S \in \mathrm{Hom}_R(N,O)$ then there is a function

$$S \circ T \colon M \longrightarrow O$$

defined by $(S \circ T)(\mathbf{m}) = S\big(T(\mathbf{m})\big)$. The composition is thought of as following the path of the arrows in the diagram. Note that the arrows are applied in order from left to right, whereas the functions in a composition are applied in order from right to left.

(Question 8.13) Prove that $S \circ T \colon M \longrightarrow O$ is in $\mathrm{Hom}_R(M,O)$.

Definition 35. *A sequence containing*

$$\cdots \xrightarrow{\quad T \quad} N \xrightarrow{\quad S \quad} \cdots$$

is exact *at N if* $\mathrm{Ker}(S) = \mathrm{Im}(T)$. *The sequence is exact if it is exact at every module in the sequence.*

Example: The following is a sequence:

$$M \longrightarrow N_1 \longrightarrow N_2 \longrightarrow N_3 \longrightarrow O$$

It is exact if and only if it is exact at N_1, N_2 and N_3.

Sequences and exactness provide

(Question 8.14) Show that

$$M \xrightarrow{\quad T \quad} N \longrightarrow 0$$

is exact at N if and only if T is surjective. (*Hint:* There is only one possible map from N to 0, the zero homomorphism!)

(Question 8.15) Show that

$$0 \longrightarrow M \xrightarrow{\quad T \quad} N$$

is exact at M if and only if T is injective.

Note: To simplify the following exposition, we will fix a commutative ring R, and use M, N, O and P to designate modules over R. L, T, and S will designate homomorphisms of R-modules.

We have seen that the composition of R-module homomorphisms is an R-module homomorphism: that is $(S,T) \longrightarrow S \circ T$ induces a map

$$\mathrm{Hom}_R(N,O) \times \mathrm{Hom}_R(M,N) \longrightarrow \mathrm{Hom}_R(M,O)$$

When $M = N = O$ the sets of homomorphisms are all equal, and this is an operation on $\mathrm{Hom}_R(M,M)$, such as addition or multiplication in a ring. In fact, this can be viewed as "multiplication" and the purpose of this section is to make this clear.

We give those homomorphisms which have the same domain and codomain endomorphisms:

Definition 36. *Elements of* $\mathrm{Hom}_R(M,M)$ *are called* endomorphisms *of* M*. We will denote also denote this set by* $\mathrm{End}_R(M)$*. An invertible endomorphism is called an* automorphism *and the set of automorphisms is denoted* $\mathrm{Aut}_R M$*.*

Vector Space Note: When $R = \mathbb{F}$ is a field, and $M = V$ is a vector space, endomorphisms are usually called *linear operators*.

We will derive properties of the endomorphisms from more general properties of the sets $\mathrm{Hom}_R(M,N)$.

We know that $\mathcal{F}(M,N)$ is an R-module since N is, and $\mathrm{Hom}_R(M,N) \subset \mathcal{F}(M,N)$.

(Question 8.16) Show that $\mathrm{Hom}_R(M,N)$ is an R sub-module of the module $\mathcal{F}(M,N)$.

(Question 8.17) Describe the additive inverse for L in $\mathrm{Hom}_R(M,N)$. What is the zero element in $\mathrm{Hom}_R(M,N)$?

In particular, we can add elements of $\mathrm{Hom}_R(M,N)$ and multiply by elements of R. Furthermore,the addition is commutative, associative, etc. since it is in $\mathcal{F}(M,N)$.

Thus, $\mathrm{Hom}_R(M,M)$ has an addition on it. Could composition play the role of a multiplication? The biggest stumbling block is usually associativity.

However, from the theory of functions, we know that composition is *always* associative: that is, if $T \in \mathrm{Hom}_R(M,N)$, $S \in \mathrm{Hom}_R(N,O)$ and $L \in \mathrm{Hom}_R(O,P)$ then $L \circ (S \circ T) = (L \circ S) \circ T$ since both evaluate on $\mathbf{x}$ to $L(S(T(\mathbf{x})))$.

Furthermore, there is an obvious "multiplicative" identity. Let $I_M \in \mathrm{Hom}_R(M,M)$, $I_N \colon N \longrightarrow N$ be the identify homomorphisms for M and N. If $T \in \mathrm{Hom}_R(M,N)$ then

$$I_N \circ T = T \circ I_M = T$$

Thus, in $\mathrm{Hom}_R(M,M)$, I_M will be the multiplicative identity for both left and right compositions.

Composition also distributes over the addition of homomorphisms:

(Question 8.18) Suppose $T, S \in \mathrm{Hom}(M,N)$, and $L \in \mathrm{Hom}(N,O)$. Show, by evaluating on $\mathbf{m} \in M$, that $L \circ (S+T) = L \circ S + L \circ T$ in $\mathrm{Hom}(M,O)$.

Similarly, if $L \in \mathrm{Hom}(P,M)$ $(S+T) \circ L = S \circ T + S \circ L$. We need to verify this identity separately because composition of functions isn't commutative.

Thus, if we think of composition as the "multiplication" on $\mathrm{End}_R(M)$, equipped with its addition as R-module homomorphisms, we have

Theorem 37. *The set* $\mathrm{End}_R(M)$ *is a non-commutative ring for the addition defined from* $\mathrm{Hom}_R(M,M)$ *and the multiplication* $(S,T) \longrightarrow S \circ T$. *The zero map* $\mathbf{0}$ *is the zero element for this ring, and the identity map on* M *is the multiplicative identity in this ring.*

This ring has an additional feature, however. Since M is an R-module elements in it can also be scaled by elements in R. However, this adds nothing since we can implement it using composition: since $\big((\lambda I_M) \circ T\big)(\mathbf{v}) = \lambda T(\mathbf{v})$ so $(\lambda I_M) \circ T = \lambda T$. That is, we can think of an element $\lambda \in R$ as the endomorphism λI_M, using the module structure on $\mathrm{Hom}_R(M,M)$. When an R-module has a multiplication which turns it into a (non-commutative or commutative) ring, and the scalar multiplication from R can be implemented using this multiplication, we call the module an *R-algebra*.

Example: The polynomial ring $R[x]$ is a ring as well as an R-module. Scaling $p(x)$ by $r \in R$ is the same as multiplying $p(x)$ by the constant polynomial $q(x) = r$ when using the multiplication we defined on polynomials. Thus $R[x]$ is and R-algebra.

Since $\mathrm{End}_R(M)$ is a non-commutative ring, we can try to find modules over it. One stands out:

$$M \text{ is a (left) module for the ring } \mathrm{End}_R(M) \text{ for its usual addition and the "scalar" multiplication } T\cdot\mathbf{v} \longrightarrow T(\mathbf{v}).$$

We already know that the addition on M satisfies the relevant properties. The identity in $\mathrm{End}_R(M)$, which is the identity homomorphism, by definition takes $\mathbf{v}$ To $\mathbf{v}$. The definition of composition shows that $(S \circ T)\cdot\mathbf{v} = \big(S \circ T\big)(\mathbf{v}) = S(T(\mathbf{v})) = S\cdot(T\cdot\mathbf{v})$, so all that remains to show is the distributivity properties:

(1) That $T\cdot(\mathbf{v}+\mathbf{w}) = T\cdot\mathbf{v} + T\cdot\mathbf{w}$ is just a restatement of the linearity properties of $T \in \mathrm{Hom}_R(M,M)$.
(2) That $(T+S)\cdot\mathbf{v} = T\cdot\mathbf{v} + S\cdot\mathbf{v}$ is just a restatement of the definition of addition in $\mathrm{Hom}_R(M,M)$

We think of the elements of $\mathrm{End}_R(M)$ as acting on the elements of M and moving them around.

That $\mathrm{End}_R(M,M)$ makes M into a (left) module just from already understood properties shows that it is fundamental to understanding M. In fact, in some ways, the subject should be the study of this ring and this module, rather than M as a module over R.

Or, restated for vector spaces, linear algebra, rather than studying vector spaces V over a field $\mathbb{F}$, should instead study vector spaces V as *modules* over the non-commutative *ring* $\mathrm{End}_{\mathbb{F}}(V)$.

Indeed, this additional module structure incorporated the original module structure since $\mathrm{End}_R(M,M)$ is an R-algebra.

However, considering this module structure as a whole is not easy. Instead, we will pick a single and us it to define an $R[x]$-module structure on M. As we will see, properties of this module reflect the properties of T in a manner which, as unbelievable as it may seem, is of direct use when applying linear algebra over fields to problems in the real world.

Let $T \in \mathrm{End}_R(M)$:

(1) We define the powers of T: $T^n \in \mathrm{End}_R(M)$, to be the result of composing T with itself n times: $T \circ T \circ \cdots \circ T$, with n T's in the composition. More formally, we define T^n via the

recursion $T^n = T \circ T^{n-1}$ and $T^0 = I$, the identity. For example, $T^3(\mathbf{v}) = T(T(T(\mathbf{v})))$.

(2) The *associativity* of composition guarantees that $T^k \circ T^l = T^{k+l}$. For example, we can factor $T^{\circ 5} = T \circ T \circ T \circ T \circ T$ as $T^{\circ 2} \circ T^{\circ 3} = \big(T \circ T\big) \circ \big(T \circ T \circ T\big)$ or as $T^{\circ 1} \circ T^{\circ 4} = T \circ \big(T \circ T \circ T \circ T\big)$ since this is just a matter of where we put the parentheses. Thus, when taking composition as our product, these really do behave like powers of T.

(3) For any polynomial, $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in R[x]$ there is an associated endomorphism

$$p(T) = a_n T^n + a_{n-1}T^{(n-1)} + \cdots + a_1 T + a_0 I_M \in \mathrm{End}_R(M)$$

which evaluates on $\mathbf{v} \in M$ by

$$\big(p(T)\big)(\mathbf{v}) = a_n T^n(\mathbf{v}) + a_{n-1}T^{(n-1)}(\mathbf{v}) + \cdots + a_1 T(\mathbf{v}) + a_0\mathbf{v}$$

where $a_0\mathbf{v}$ is just the R-module scaling from M being a module over R.

This action of $R[x]$ on M makes M into an $R[x]$-module, M_T. The original R-module structure is encoded in the way the constant polynomials act $a_0 \longrightarrow a_0 I$, so they are multiplication by a_0.

(Question 8.19) Let $p, q \in R[x]$ and $T \in \mathrm{End}_R(M)$. Show that $p(T) \circ q(T) = (p \cdot q)(T)$ where $p \cdot q$ is the product polynomial in $R[x]$. This is the heart of showing that the above defines an $R[x]$-module structure on M.

(Question 8.20) Suppose T is the linear operator $\dfrac{d}{dx} : C^\infty \longrightarrow C^\infty$. Show that

(1) T^n is the linear operator $f \longrightarrow \dfrac{d^n f}{dx^n}$

(2) If $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{R}[x]$, then $p(T)$ is the differential operator

$$f \longrightarrow a_n \frac{d^n f}{dx^n} + a_{n-1}\frac{d^{n-1} f}{dx^{n-1}} + \cdots + a_1 \frac{df}{dx} + a_0 f$$

(Question 8.21) Suppose $\mathbf{A} \in M_{2\times 2}(\mathbb{R})$ and $T \in \mathrm{Hom}(\mathbb{R}^2, \mathbb{R}^2)$ is $T(\mathbf{v}) = \mathbf{A}\mathbf{v}$. Show that

(1) $T^3(\mathbf{v}) = \mathbf{A}^3\mathbf{v}$

(2) If $p(x) = x^2 + 1$, then $p(T)$ is multiplication by $\mathbf{A}^2 + \mathbf{I}_2$

(3) If $\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ then $p(T)$ is the zero homomorphism.

One of our principal goals in this course is to understand the module V_T for a given $T \in \mathrm{End}_{\mathbb{F}}(V)$ when V is a vector space over a field $\mathbb{F}$. Note, however, that $\mathbb{F}[x]$ is a ring, so to do so we will need to use our understanding of modules and not just vector spaces.