

0.2 Vector spaces

I'm going to begin this section at a rather basic level, giving the definitions of a field and of a vector space in much the same detail as you would have met them in a first linear algebra course. If you took the course at the sophomore level, you may have only used scalars from the field \mathbf{R} of real numbers. We will allow the scalars to come from any field, but it isn't too much of a jump. You should remember, of course, that a field can be finite (keep \mathbf{Z}_2 in mind). Most of the statements are taken from the appendices in the text, and are numbered accordingly.

DEFINITION A.1.1. A *field* is a set F together with two operations $+$ and \cdot for which the following conditions hold:

- (i) (*Closure*) for all $a, b \in F$, the sum $a + b$ and the product $a \cdot b$ again belong to F ;
- (ii) (*Associativity*) for all $a, b, c \in F$, $a + (b + c) = (a + b) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;
- (iii) (*Commutativity*) For all $a, b \in F$, $a + b = b + a$ and $a \cdot b = b \cdot a$;
- (iv) (*Distributive laws*) for all $a, b, c \in F$, $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$;
- (v) (*Existence of an additive identity*) there exists an element $0 \in F$ for which $a + 0 = a$ and $0 + a = a$, for all $a \in F$;
- (vi) (*Existence of a multiplicative identity*) there exists an element $1 \in F$, with $1 \neq 0$, for which $a \cdot 1 = a$ and $1 \cdot a = a$, for all $a \in F$;
- (vii) (*Existence of additive inverses*) for each $a \in F$, the equations $a + x = 0$ and $x + a = 0$ have a solution x in F , called the *additive inverse* of a , and denoted by $-a$;
- (viii) (*Existence of multiplicative inverses*) for each $a \in F$, with $a \neq 0$, the equations $a \cdot x = 1$ and $x \cdot a = 1$ have a solution x in F , called the *multiplicative inverse* of a , and denoted by a^{-1} .

In any field, we can identify the integer n with the element $n \cdot 1$. Although we can divide by any nonzero element of the field, we cannot necessarily divide by a nonzero "integer" since it is possible that $n \cdot 1 = 0$ even though $n \neq 0$. To deal with this problem we consider the *characteristic* of the field. (For details see Definition A.1.2 and Proposition A.1.3 in the text, together with Examples 1.2.8 and 1.3.5.)

Our goal is to study rings, and so we want to include the ring \mathbf{Z} of integers, as well as the ring

$$F[x] = \{a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n \mid a_i \in F, 0 \leq i \leq n\}$$

of all polynomials in one variable with coefficients in a field F . In \mathbf{Z} , only 1 and -1 have multiplicative inverses; in $F[x]$, only the nonzero constant polynomials have multiplicative inverses. In the definition of a ring we will keep all but condition (viii), which requires that each nonzero element must have a multiplicative inverse. Dropping this one condition gives the definition of a *commutative ring with identity*, stated in Definition 1.1.2 (a).

Even within linear algebra it is crucial to study matrices. In our study of rings, we want to include rings whose elements are matrices, and so we have to give up half of condition (iii) in the definition of a field, which is the requirement that multiplication must satisfy

the commutative law. Dropping both commutativity of multiplication and the existence of multiplicative inverses gives Definition 1.1.1, which defines an *associative ring with identity*.

In the definition of a field, if we drop only the commutativity of multiplication, we arrive at the notion of a *skew field*, also called a *division ring* (see Definition 1.1.6 (b)). Modules over skew fields turn out to have most of the properties of vector spaces over a field.

Here are some additional properties that follow immediately from the definition of a field. Actually, they hold in any ring, but we simply state them here for any elements a, b, c in the field F .

- (a) If $a + c = b + c$, then $a = b$. (b) If $a \cdot c = b \cdot c$ and $c \neq 0$, then $a = b$.
 (c) $a \cdot 0 = 0$ (d) $-(-a) = a$ (e) $(-a) \cdot (-b) = a \cdot b$

DEFINITION A.1.4. If F and E are fields such that F is a subset of E and the operations on F are those induced by E , then we say that F is a *subfield* of E and that E is an *extension field* of F .

The next definition we need to review is that of a *vector space over the field F* . Our ultimate goal is to study the analog over a ring R , which we will call a *left R -module*. If you make a careful comparison between Definitions 2.1.1 and A.1.5, you will see that we can use exactly the same axioms in both cases. The only change is that the scalars are allowed to come from a ring instead of a field. Of course, this does make things much more complicated, because we lose the ability to divide through by a scalar.

If you don't want to wait for examples of modules, you can skip ahead in the text to examples 2.1.1–2.1.3 and 2.1.5. Example 2.1.1 take note of the fact that the ring R is a module over itself; Example 2.1.2 just shows that a vector space is an example of a module; Example 2.1.3 shows how to think of an abelian group A as a module over the ring \mathbf{Z} by defining a scalar multiplication $n \cdot a$ in terms of the addition in A .

You may also find it interesting to read Example 2.1.5, which shows that the vector space over F defined by using all column vectors with n components is actually a module over the ring $M_n(F)$ of all $n \times n$ matrices over F . To see why there is a "scalar multiplication", you can think of the column vector as a matrix, and then you can multiply it on the left by an $n \times n$ matrix.

DEFINITION A.1.5. A *vector space* over the field F is a set V on which two operations are defined, called addition and scalar multiplication, and denoted by $+$ and \cdot respectively. The operations must satisfy the following conditions:

- (i) (*Closure*) for all $a \in F$ and all $u, v \in V$, the sum $u + v$ and the scalar product $a \cdot v$ are uniquely defined and belong to V ;
 (ii) (*Associativity*) for all $a, b \in F$ and all $u, v, w \in V$, $u + (v + w) = (u + v) + w$ and $a \cdot (b \cdot v) = (a \cdot b) \cdot v$;
 (iii) (*Commutativity of addition*) for all $u, v \in V$, $u + v = v + u$;
 (iv) (*Distributive laws*) for all $a, b \in F$ and all $u, v \in V$, $a \cdot (u + v) = (a \cdot u) + (a \cdot v)$ and $(a + b) \cdot v = (a \cdot v) + (b \cdot v)$;
 (v) (*Existence of an additive identity*) there exists an element 0 in V for which $v + 0 = v$ and $0 + v = v$ for all $v \in V$;

(vi) (*Existence of additive inverses*) for each $v \in V$, the equations $v + x = 0$ and $x + v = 0$ have a solution $x \in V$, denoted by $-v$;

(vii) (*Unitary law*) for all $v \in V$, $1 \cdot v = v$.

Using 0 for the additive identity of V , and $-v$ for the additive inverse of $v \in V$, we have the following results: $0 + v = v$, $a \cdot 0 = 0$, and $(-a)v = a(-v) = -(av)$ for all $a \in F$ and $v \in V$. The proofs involve the distributive laws, which give the only connection between addition and scalar multiplication.

One of the most basic examples of a vector space (which turns out to be very important in Galois theory) comes up in the situation in which E is an extension field of the field F . Then E is certainly an abelian group, and because F is a subfield of E it is possible to multiply elements of E by elements of F . This shows that E is actually a vector space over F . The next example is probably the one that is the most familiar to you.

EXAMPLE. For any field F , the set F^n of n -tuples is a vector space over F .

The n -tuples can be written either as a row vector or as a column vector. Addition of n -tuples is defined component by component:

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n).$$

Similarly, scalar multiplication is defined componentwise:

$$a \cdot (x_1, x_2, \dots, x_n) = (ax_1, ax_2, \dots, ax_n).$$

It is not hard to check that all of the necessary axioms are satisfied. In each case it comes down to the fact that the field axioms hold in each component.

The construction in the example, using n -tuples, can also be given for modules. In this case it is called a *direct sum* of modules. (See Section 2.2 of the text for the definition and a discussion of direct sums of modules.)

DEFINITION. Given a vector space V over a field F , a subset W of V is called a *subspace* if W is a vector space over F under the operations already defined on V .

PROPOSITION. A subset W of a vector space V is a subspace of V iff (i) W is nonempty; (ii) if $u, v \in W$, then $u + v \in W$; and (iii) if $v \in W$ and $a \in F$, then $av \in W$.

There is a corresponding notion of a submodule of a module, with the obvious definition. In the case of a ring R , which we can view as a module over itself, any submodule is called a *left ideal*. The ring R is also a *right* module over itself, because we can multiply by the scalars on the right as well as the left. In this case we get the notion of a *right ideal*, and then a subset of R that is both a left ideal and a right ideal is called a *two-sided ideal*, or simply an *ideal*. These definitions are given on page 68 of the text.

After defining the notions of vector spaces and subspaces, the next step is to identify the functions that can be used to relate one vector space to another. These functions should

respect the algebraic structure of the vector spaces, so it is reasonable to require that they preserve addition and scalar multiplication.

DEFINITION A.1.12. Let V and W be vector spaces over the field F . A *linear transformation* from V to W is a function $f : V \rightarrow W$ such that $f(au + bv) = af(u) + bf(v)$ for all vectors $u, v \in V$ and all scalars $a, b \in F$.

If a linear transformation is one-to-one and onto, it is called a *vector space isomorphism*, or simply an *isomorphism*. You can think of a vector space isomorphism $f : V \rightarrow W$ as just giving a way to rename the elements of V so that they look just like the elements of W , since any isomorphism preserves all of the algebraic structure. One of the powerful ideas of abstract algebra is to think of isomorphic objects as simply being one and the same.

Corresponding to the definition of a linear transformations between vector spaces, we have the concept of an *R -homomorphism* between left R -modules. You can see, in Definition 2.1.6, that to state the general definition we only have to change the field F to a ring R . Of course, this rather innocuous change causes a great many problems. Or maybe I should just say that it makes life much more interesting.

With this additional definition, you can read the next example in the text, Example 2.1.6, which turns out to have very important implications for linear algebra. This example starts with a vector space V over a field F , and a single linear transformation $T : V \rightarrow V$. To help understand what T is doing, we can think of the action of T as a multiplication, so that we define $T \cdot v = T(v)$, for every $v \in V$. But from our point of view, a multiplication like this should be defined for an entire ring, not just for a single element T . We can form powers T^n , using composition of functions, and linear combinations $a_0I + a_1T + \cdots + a_nT^n$, where $a_i \in F$ and I is the identity transformation on V . These can be made into a ring, and used for the “scalars”, but it is easier to let the ring of scalars be the polynomial ring $F[x]$; we just have to substitute T in place of x . These remarks lead to the multiplication in Example 2.1.6: given a polynomial $a_0 + a_1x + \cdots + a_nx^n \in F[x]$ and a vector $v \in V$, we define

$$(a_0 + a_1x + \cdots + a_nx^n) \cdot v = [a_0I + a_1T + \cdots + a_nT^n](v) = a_0v + a_1T(v) + \cdots + a_nT^n(v).$$

This multiplication, together with a structure theorem for certain modules, can be used to find standard forms for the matrix associated with T (see Section 2.7).

We next turn to the definition of a basis. The fact every vector space has a basis is certainly the single most important property of vector spaces.

DEFINITION A.1.6 (a). Let $S = \{v_1, \dots, v_n\}$ be a set of vectors in the vector space V over the field F . Any vector of the form $v = \sum_{i=1}^n a_i v_i$, for $a_i \in F$, is called a *linear combination* of the vectors in S . The set S is said to *span* V if each element of V can be expressed as a linear combination of the vectors in S .

DEFINITION A.1.6 (b). Let $S = \{v_1, \dots, v_n\}$ be a set of vectors in the vector space V over the field F . The vectors in S are said to be *linearly dependent* if one of the vectors can

be expressed as a linear combination of the others. If not, then S is said to be a *linearly independent* set.

DEFINITION A.1.9 (a). A subset of the vector space V is called a *basis* for V if it spans V and is linearly independent.

PROPOSITION A.1.10. Let S be a nonempty subset of the vector space V . Then S is a basis for V if and only if each vector in V can be written uniquely as a linear combination of vectors in S .

THEOREM. Every vector space has a basis.

This important theorem is proved in the text in Theorem A.2.5, as an illustration of the use of Zorn's lemma. As a consequence of more general results for modules, it is proved again in Corollary 2.3.4 of the text, which contains the stronger result that any nonzero vector space over a *skew* field has a basis.

The notion of a basis can be generalized to R -modules, as in Definition 2.2.1, and a module is called a *free module* if it has a basis. One of the important differences between modules and vector spaces is that a module need not be free. It is not hard to find examples of modules that are not free, since most modules tend to be a long way from having a basis.

Here is the most elementary example of a module that is not free. Let R be the ring $\mathbf{Z}_4 = \{0, 1, 2, 3\}$. The subset $M = \{0, 2\}$ is closed under addition and multiplication by any element of R , so it is a submodule of R , when R is thought of as a module over itself. No basis can contain 0, so the only possibility for a basis for M is the set $\{2\}$. But if M had a basis consisting of a single element, then M would be forced to have the same number of elements as R , and this clearly isn't the case. You can see that M is a perfectly good module, but it is too small to be a free module over this particular ring. Actually, if all nonzero left R -modules are free it forces R to be a skew field (see the exercises in Section 2.3 of the notes).

Although it is rare for a module to have a basis, it always has a spanning set. Even though in most cases we can't find an *independent* spanning set, it is still important to know when we can find a *finite* spanning set. In Definition 2.1.5 (b) a module is said to be *finitely generated* if it has a finite spanning set. If it can be spanned by a single element, it is called *cyclic* (see Definition 2.1.5 (a)). One of the famous theorems that is proved in the text is the Hilbert basis theorem (Theorem 2.4.10), which in its more elementary form states that if F is a field, then over the polynomial ring $F[x_1, x_2, \dots, x_n]$, every submodule of a finitely generated module is again finitely generated.

The next theorem makes an even stronger statement than just asserting the existence of a basis in any vector space. It shows that any linearly independent set of vectors can be "extended" to a basis for the vector space. As an important consequence, we get Theorem A.2.6, which shows that every subspace of a vector space has a "complement."

THEOREM A.2.5. Let F be a field, and let V be any vector space over F . Then every linearly independent subset of V is contained in a basis for V .

THEOREM A.2.6. Let V be a vector space over the field F , and let W be any subspace of V . Then there exists a subspace Y of V such that each element $v \in V$ can be written uniquely in the form $v = w + y$ for some $w \in W$ and $y \in Y$.

Once again there are problems when we move from vector spaces to modules, because we can lose the existence of complements. For example, the submodule (or ideal) $M = \{0, 2\}$ in the ring \mathbf{Z}_4 does not have a complement. To see this, just note that there is no other proper nonzero submodule, since 1 and 3 are invertible in the ring, and therefore can't belong to a proper submodule.

Thus Theorem A.2.6 raises an important question for modules. Can we find conditions on the ring that guarantee that every proper submodule of every nonzero module has a complement? Modules that have this crucial property are called *completely reducible* in Definition 2.2.12. The answer to the question is contained in the Artin-Wedderburn theorem (see Theorem 3.3.2, Corollary 3.3.4, and Theorem 2.3.6), which shows that every nonzero left R -module is completely reducible if and only if the ring R can be written as a finite direct sum of matrix rings over skew fields.

Our next step is to look at the concept of the dimension of a vector space. Once again, it turns out that life is much more complicated for modules over a ring than it is for vector spaces over a field. But modules do have various weaker properties that can still be used to prove some of the results that hold for vector spaces.

THEOREM A.1.7. Let V be a vector space, let $S = \{u_1, u_2, \dots, u_m\}$ be a set that spans V , and let $T = \{v_1, v_2, \dots, v_n\}$ be a linearly independent set. Then $n \leq m$, and V can be spanned by a set of m vectors that contains T .

COROLLARY A.1.8. Any two finite subsets that both span V and are linearly independent must have the same number of elements.

DEFINITION A.1.9 (b). If V has a finite basis, then it is said to be *finite dimensional*, and the number of vectors in the basis is called the *dimension* of V , denoted by $\dim(V)$.

THEOREM. Any n -dimensional vector space over the field F is isomorphic to F^n .

EXAMPLE. The field \mathbf{C} of complex numbers is a two dimensional vector space over the field \mathbf{R} .

EXAMPLE. If F is a field, then the set of polynomials $F[x]$ with coefficients in F is an infinite dimensional vector space over F .

The above theorem stating that any n -dimensional vector space is isomorphic to F^n has a direct generalization to modules. It is proved in Proposition 2.2.5 (a) that any R -module with a basis consisting of n elements is isomorphic to the module of n -tuples of elements from R , denoted by R^n , where addition and scalar multiplication are defined component by component.

It is Corollary A.1.8 that justifies the definition of the dimension of a vector space over a field, since it shows that the number of elements in a basis is an invariant of the vector space. When we look at free modules over a ring, this important result may fail to hold. The ring R is said to have *invariant basis number* if the following condition holds: if M is a left R module with a basis consisting of n elements, then every other basis for M also has n elements. There are exercises in the notes that look at this question—see exercise sets in Section 2.5 and Section 2.6. In particular, the second of these exercises gives an example of a ring that does not have invariant basis number.

Now I want to discuss some of the good properties of vector spaces that come from the ability to give a definition of dimension. First of all, if a vector space V has finite dimension, then so does each of its nonzero subspaces. The next proposition follows from Theorem A.2.5.

PROPOSITION. If V is a vector space with $\dim(V) = n$, then $\dim(W) < n$ for any proper subspace W .

As a consequence of the proposition, if $V_0 \supset V_1 \supset V_2 \supset \cdots \supset V_{k-1} \supset V_k$ is any chain of *distinct* subspaces of V , and $\dim(V) = n < \infty$, then we must have $k \leq n$. For example, if $\dim(V) = 1$, then the only subspaces are V and $\{0\}$, and $V \supset \{0\}$ is the longest chain of distinct subspaces.

The one-dimensional subspaces play an important role for vector spaces, since any one-dimensional vector space is just a copy of the field. To translate to modules, we focus on the fact that a one-dimensional vector space has no proper nonzero submodules. The general definition is given in Definition 2.1.9: a nonzero module M is called *simple* if its only submodules are M and $\{0\}$. To see how simple modules can be described explicitly in terms of the ring, refer to Proposition 2.1.11 (a) and Proposition 2.1.8 (a). The simple modules play an important role, especially in representation theory.

Over a field, every vector space is isomorphic to a direct sum of one-dimensional subspaces. Over a ring, it need not be true that every module is isomorphic to a direct sum of simple submodules. (Just look at \mathbf{Z}_4 as a module over itself.) According to Definition 2.3.1 (b), a module is called *semisimple* if it is isomorphic to a direct sum of simple submodules. As something of a surprise, it turns out that semisimple modules are the same as completely reducible modules, as shown by Corollary 2.2.3.

This section has previewed a long list of definitions, much too long to remember the first time you read it. I hope that you will come back to this section as you meet these definitions in the text, so that you can keep the new theory in perspective. It will also help to keep in mind what happens for abelian groups, and that is the next topic.