# ALGEBRAS

KEITH CONRAD

## 1. Definitions and Examples

Let $k$ be a field. Both field extensions $K/k$ and matrix rings $\mathrm{M}_d(k)$ are examples of rings that are properly understood by taking into account the extra structure of a $k$-vector space. We want to discuss, in general, vector spaces over $k$ that have a ring structure compatible with the scalar multiplication by elements of $k$.

**Definition 1.1.** A *$k$-algebra* $A$ is a (possibly noncommutative) ring with identity that is also a $k$-vector space, such that for $\alpha \in k$ and $a, b \in A$,

$$\alpha(ab) = (\alpha a)b = a(\alpha b). \tag{1.1}$$

**Example 1.2.** A commutative ring containing $k$, such as a field extension, the polynomial ring $k[X, Y]$, or the formal power series ring $k[[X]]$ is a $k$-algebra.

**Example 1.3.** The ring $\mathrm{M}_d(k)$ under matrix addition and multiplication is a $k$-algebra. This is called a *matrix algebra* over $k$.

**Example 1.4.** The group ring $k[G]$ for a finite group $G$.

**Example 1.5.** The set $\mathrm{Hom}_k(V, V)$ of $k$-linear maps of a $k$-vector space $V$ to itself is a $k$-algebra under addition and composition of linear maps.

**Example 1.6.** The set $C([0, 1], \mathbf{R})$ of continuous functions $[0, 1] \to \mathbf{R}$ is an $\mathbf{R}$-algebra under the usual pointwise addition and multiplication of functions.

While a $k$-algebra may be noncommutative, (1.1) tells us that elements of $k$ are supposed to commute multiplicatively with all the elements of a $k$-algebra. The next example illustrates this point well.

**Example 1.7.** Since the center of the quaternions $\mathbf{H}$ is the real numbers, $\mathbf{H}$ is an $\mathbf{R}$-algebra. But $\mathbf{H}$ is not a $\mathbf{C}$-algebra using usual quaternionic multiplication with complex numbers since the only quaternions that commute with all quaternions are real numbers. Certainly $\mathbf{H}$ can be given the structure of a complex vector space, by letting scalar multiplication be left multiplication by the complex numbers. This is not the only type of complex vector space structure one can give to $\mathbf{H}$. Another one would be $z \cdot q \stackrel{\text{def}}{=} q\overline{z}$, where $q$ is a quaternion and $z$ is a complex number. But $\mathbf{H}$ is not a $\mathbf{C}$-algebra by either of these $\mathbf{C}$-vector space structures. Understand this example to keep clearly in mind the distinction between a $k$-vector space and a $k$-algebra.

You may have seen these examples before, treated just as rings. The existence of the field $k$ provides an extra basic structure, which is the main reason for isolating the concept of a $k$-algebra.

Our definition of a $k$-algebra easily extends to the notion of an $R$-algebra for a commutative ring $R$. That would be a ring $A$ with identity that is also an $R$-module such that

(1.1) is true for $\alpha \in R$ and $a$ and $b$ in $A$. For example, $\mathrm{M}_d(R)$ is an $R$-algebra and every ring is a $\mathbf{Z}$-algebra. We focus here on algebras over a field for simplicity (*e.g.*, bases are automatically available).

For a $k$-algebra $A$, $\alpha \mapsto \alpha \cdot 1_A$ for $\alpha \in k$ is a ring homomorphism of $k$ to the *center* of $A$, and it must be injective if $A \neq 0$ since ring homomorphisms from a field to a nonzero ring must be injective. Thus a more concrete way to think about a $k$-algebra is that it is a ring containing $k$ in the center. If we did not require our algebras to have a multiplicative identity then we could not make this concrete observation, but all algebras of interest to us will have an identity so we build that property into the definition. When defining things like $k$-algebra homomorphims, subalgebras, quotient algebras, and so forth, it is appropriate not to identify the scalar field $k$ with its image in a particular $k$-algebra.[1]

**Remark 1.8.** We included a multiplicative identity in our definition of a $k$-algebra, and that fits the needs of algebraists quite well. However, in analysis there are important examples of algebras "without identity," such as the real-valued continuous functions on $\mathbf{R}$ with compact support. If there could be a multiplicative identity, it should be the constant function 1, but that doesn't have compact support.

We say a $k$-algebra $A$ is *finite-dimensional* when it is finite-dimensional as a vector space over $k$, and we write $[A : k] = \dim_k(A)$. For example, $\mathbf{C}$ is a two-dimensional $\mathbf{R}$-algebra and an infinite-dimensional $\mathbf{Q}$-algebra, while $\mathrm{M}_d(k)$ is a $d^2$-dimensional $k$-algebra. The $\mathbf{R}$-algebra $C([0,1], \mathbf{R})$ is infinite-dimensional. We will often use the term "finite $k$-algebra" to mean "finite-dimensional $k$-algebra."

An algebra is called *commutative* when multiplication in the algebra is commutative.

A *subalgebra* of a $k$-algebra $A$ is a subset that is both a subring and a $k$-linear subspace, hence a $k$-algebra in its own right. For example, the $\mathbf{R}$-algebra $\mathbf{H}$ contains $\mathbf{C}$ as an $\mathbf{R}$-subalgebra. A *$k$-algebra homomorphism* $f\colon A \to B$ is a map between $k$-algebras that is both $k$-linear and a ring homomorphism. Unlike the $k$-linear maps from $A$ to itself, the subset of $k$-algebra homomorphisms of $A$ to itself is usually not even a vector space since the sum of two $k$-algebra homomorphisms need not be a $k$-algbera homomorphism (it need not be multiplicative). There is a natural notion of a *$k$-algebra isomorphism*. For example, if $V$ is a $d$-dimensional vector space over $k$, then choosing a $k$-basis of $V$ yields a $k$-algebra isomorphism

$$\mathrm{Hom}_k(V, V) \cong \mathrm{M}_d(k).$$

The point is that this is not just an isomorphism of rings; it's *$k$-linear* as well and we should pay attention to this extra structure in the isomorphism.

**Example 1.9.** The bilinear map $V \times V^\vee \to \mathrm{Hom}_k(V, V)$, which sends $(v, \varphi)$ to the function $w \mapsto \varphi(w)v$ induces a $k$-linear map $V \otimes_k V^\vee \to \mathrm{Hom}_k(V, V)$. Given a basis $e_1, \ldots, e_d$ of $V$, the simple tensors $e_i \otimes e_j^\vee$ correspond to the matrices with a 1 in the $(i, j)$ position and 0 elsewhere, so this $k$-linear map sends a basis to a basis and thus is a vector space isomorphism. Since $\mathrm{Hom}_k(V, V)$ is a $k$-algebra, we can transport the multiplication on it back to $V \otimes_k V^\vee$. Can you describe this ring structure on $V \otimes_k V^\vee$ directly?

Since multiplication by $k$ commutes with every element of $A$, for each $a \in A$ there is a $k$-algebra homomorphism

$$\mathrm{ev}_a\colon k[X] \to A$$

[1] If we replace the field $k$ with a commutative ring $R$ then the mapping $\alpha \mapsto \alpha \cdot 1_A$ from $R$ to an $R$-algebra with identity might not be injective. An example is thinking of $\mathbf{Z}/10\mathbf{Z}$ as a $\mathbf{Z}$-algebra.

given by $c_nX^n + \cdots + c_0 \mapsto c_na^n + \cdots + c_0 \cdot 1_A$, called *evaluation at* $a$. Of course, for $f(X) \in k[X]$, we often write $f(a)$ for $\mathrm{ev}_a(f)$. If $k$ were not in the center of $A$, then $\mathrm{ev}_a$ need not be a homomorphism! For example, take $A = \mathbf{H}$ and $k = \mathbf{C}$, which is an $\mathbf{R}$-subalgebra of $\mathbf{H}$ but is not in the center of $\mathbf{H}$. Note that $ev_j(X^2 + 1) = j^2 + 1 = 0$ and $\mathrm{ev}_j(X + i)\,\mathrm{ev}_j(X - i) = (j + i)(j - i) = 2ij \neq 0$.

Imitating the notion of algebraic field extensions, we say $a \in A$ is *algebraic* (over $k$) if $f(a) = 0$ for some nonzero $f \in k[X]$. We say $A$ is an *algebraic $k$-algebra* if every element of $A$ is algebraic over $k$. For example, a finite field extension $K/k$ is an algebraic $k$-algebra. More generally, a finite-dimensional $k$-algebra $A$ is algebraic, by the same proof as for finite field extensions: for $a \in A$ and $n = [A : k]$, the $n+1$ elements $1, a, \ldots, a^n$ satisfy a nontrivial $k$-linear relation, so $a$ is algebraic over $k$.

A *$k$-division algebra* is a division ring that is also a $k$-algebra. For example, $\mathbf{H}$ is an $\mathbf{R}$-division algebra, and a $\mathbf{Q}$-division algebra, although one usually views it only as an $\mathbf{R}$-division algebra. Our primary interest is in $k$-division algebras, such as fields containing $k$, so on some occasions we will prove results only for the case of division algebras when the proof for more general algebras is longer.

If $K/k$ is a field extension and $A$ is a $K$-algebra, then $A$ is also a $k$-algebra in the obvious way. The proof of the transitivity formula for dimensions of finite extensions of fields carries over to show that

$$[A : k] < \infty \Longleftrightarrow [A : k], [K : k] < \infty,$$

in which case $[A : k] = [A : K][K : k]$. For example, $\mathrm{M}_3(\mathbf{C})$ is a 9-dimensional $\mathbf{C}$-algebra, an 18-dimensional $\mathbf{R}$-algebra, and as a $\mathbf{Q}$-algebra it is infinite-dimensional.

We now introduce arguably the most important idea in this handout. Let $A$ be a $k$-algebra. To each $a \in A$ we associate the $k$-linear map $m_a \colon A \to A$ given by left multiplication by $a$:

$$x \mapsto ax.$$

The map $A \to \mathrm{Hom}_k(A, A)$ given by $a \mapsto m_a$ is a $k$-algebra homomorphism:

$$m_{a+b}(x) = (a + b)(x) = ax + bx = m_a(x) + m_b(x) = (m_a + m_b)(x),$$

$$(m_a \circ m_b)(x) = m_a(bx) = a(bx) = (ab)x = m_{ab}(x),$$

and

$$m_{\alpha a}(x) = (\alpha a)x = \alpha(ax) = \alpha(m_a(x)) = (\alpha m_a)(x),$$

so $m_a$ is additive in $a$, multiplicative in $a$ (note $m_1$ is the identity map on $A$), and $m_{\alpha a} = \alpha m_a$. Since $A$ has a multiplicative identity, we can recover $a$ from $m_a$ by evaluating at 1: $m_a(1) = a \cdot 1 = a$. Thus turning elements of $A$ into left multiplication maps on $A$ embeds $A$ in $\mathrm{Hom}_k(A, A)$ as a $k$-subalgebra.

*Assume from now on that* $n = [A : k] < \infty$. Choosing a $k$-basis of $A$ lets us express each $m_a$ as an $n$ by $n$ matrix $[m_a]$ acting on the left on column vectors of $k^n \cong A$ (as $k$-vector spaces). Thus $A$ acting on itself by left multiplication lets us view $A$ as matrices over $k$. (analogy: every finite group lives in some symmetric group by having the group act on itself by left multiplications). For $f(X) = \sum_i c_iX^i$ in $k[X]$ and $a$ and $x$ in $A$, $m_{f(a)}(x) = f(a)x = \sum_i c_i(a^ix) = \sum_i c_i(m_a)^i(x) = f(m_a)(x)$, so $m_{f(a)} = f(m_a)$ for all $a \in A$: a polynomial function of $m_a$ over $k$ is left multiplication on $A$ by that polynomial's value at $a$.

**Example 1.10.** If $\alpha \in k$, then with respect to any $k$-basis of $A$, $[m_\alpha]$ is the scalar diagonal matrix $\alpha \cdot I_n$, where $n = [A : k]$.

**Example 1.11.** $A = \mathbf{C}$, $k = \mathbf{R}$, basis $= \{1,i\}$. For $a = a_1 + a_2 i$, $[m_a]$ equals

$$\begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix}.$$

**Example 1.12.** $A = \mathbf{Q}(\sqrt{r})$ for $r$ a nonsquare rational, $k = \mathbf{Q}$, basis $= \{1,\sqrt{r}\}$. For $a = a_1 + a_2\sqrt{r}$, $[m_a]$ equals

$$\begin{pmatrix} a_1 & a_2 r \\ a_2 & a_1 \end{pmatrix}.$$

**Example 1.13.** $A = \mathbf{Q}(\theta)$ for $\theta$ a root of $X^3 - X - 1$, $k = \mathbf{Q}$, basis $= \{1,\ \theta,\ \theta^2\}$. For $a = a_1 + a_2\theta + a_3\theta^2$, $[m_a]$ equals

$$\begin{pmatrix} a_1 & a_3 & a_2 \\ a_2 & a_1 + a_3 & a_2 + a_3 \\ a_3 & a_2 & a_1 + a_3 \end{pmatrix}.$$

**Example 1.14.** $A = \mathbf{H}$, $k = \mathbf{R}$, basis $= \{1,i,j,k\}$ (of course, the basis element $k$ is not to be confused with our generic label for a field). For $a = a_1 + a_2 i + a_3 j + a_4 k$, $[m_a]$ equals

$$\begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{pmatrix}.$$

**Example 1.15.** Let $A = \mathrm{M}_d(k)$. Using the standard basis $E_{ij}$ ordered lexicographically, if $a = (a_{ij})$ then $[m_a]$ equals

$$(a_{ij}) \otimes I_d$$

in $\mathrm{M}_{d^2}(k) = \mathrm{M}_d(k) \otimes_k \mathrm{M}_d(k)$.

**Definition.** Let $A$ be an $n$-dimensional $k$-algebra, $a \in A$. The *characteristic polynomial* $\chi_a(X)$ of $a$ is the usual characteristic polynomial of the linear map $m_a$, namely

$$\chi_a(X) = \det(X \cdot I_n - m_a),$$

a monic polynomial in $k[X]$ of degree $n = [A : k]$.

By the Cayley-Hamilton theorem, $\chi_a(m_a) = 0$, so $m_{\chi_a(a)} = \chi_a(m_a) = 0 = m_0$. Since $m_{\chi_a(a)} = m_0$, we have proven the following theorem.

**Theorem 1.16.** *For a finite-dimensional $k$-algebra $A$, each $a$ in $A$ is a root of its characteristic polynomial $\chi_a(X) \in k[X]$.*

**Definition 1.17.** For a finite-dimensional $k$-algebra $A$, the *trace* of $a \in A$ is the trace of $m_a$. The *norm* of $a$ is the determinant of $m_a$.

That is, the trace of $a$ is $\mathrm{Tr}_{A/k}(m_a) \in k$, usually written as $\mathrm{Tr}_{A/k}(a)$, so $\mathrm{Tr}_{A/k}\colon A \to k$. (Sometimes you might see S or Sp used for trace since Spur is the German word for trace.) The norm of $a$ is $\det(m_a) \in k$, usually written as $\mathrm{N}_{A/k}(a)$, so $\mathrm{N}_{A/k}\colon A \to k$. Note that the constant term of $\chi_a(X)$ is $\chi_a(0) = (-1)^n \mathrm{N}_{A/k}(a)$ and the coefficient of $X^{n-1}$ is $-\mathrm{Tr}_{A/k}(a)$. So up to sign, the trace and norm of $a$ can be read from the characteristic polynomial, and in a field containing $k$ in which $\chi_a$ splits, the roots of $\chi_a$ have sum equal to $\mathrm{Tr}_{A/k}(a)$ and

product equal to $\mathrm{N}_{A/k}(a)$. In practice, there are more efficient ways of computing traces and norms. See Theorem 2.2 below.

Let's look at some examples.

By Example 1.10, $\mathrm{Tr}_{A/k}(\alpha) = n\alpha$ and $\mathrm{N}_{A/k}(\alpha) = \alpha^n$ for all $\alpha \in k$. In particular, $\mathrm{Tr}_{A/k}(1) = n = [A : k]$.

By Example 1.11, $\chi_{a_1+a_2 i}(X) = X^2 - 2a_1X + a_1^2 + a_2^2$. Thus

$$\mathrm{Tr}_{\mathbf{C}/\mathbf{R}}(a_1 + a_2 i) = 2a_1, \quad \mathrm{N}_{\mathbf{C}/\mathbf{R}}(a_1 + a_2 i) = a_1^2 + a_2^2.$$

By Example 1.12, $\chi_{a_1+a_2\sqrt{r}}(X) = X^2 - 2a_1X + a_1^2 - ra_2^2$. Thus

$$\mathrm{Tr}_{\mathbf{Q}(\sqrt{r})/\mathbf{Q}}(a_1 + a_2\sqrt{r}) = 2a_1, \quad \mathrm{N}_{\mathbf{Q}(\sqrt{r})/\mathbf{Q}}(a_1 + a_2\sqrt{r}) = a_1^2 - ra_2^2.$$

By Example 1.13, $\mathrm{Tr}_{\mathbf{Q}(\theta)/\mathbf{Q}}(a_1 + a_2\theta + a_3\theta^2) = 3a_1 + 2a_3$ and

$$\mathrm{N}_{\mathbf{Q}(\theta)/\mathbf{Q}}(a_1 + a_2\theta + a_3\theta^2) = a_1^3 + a_2^3 - a_3^3 - a_1a_2^2 + a_1a_3^2 - a_2a_3^2 + 2a_1^2a_3 - 3a_1a_2a_3.$$

By Example 1.14, $\mathrm{Tr}_{\mathbf{H}/\mathbf{R}}(a_1 + a_2 i + a_3 j + a_4 k) = 4a_1$ and

$$\mathrm{N}_{\mathbf{H}/\mathbf{R}}(a_1 + a_2 i + a_3 j + a_4 k) = (a_1^2 + a_2^2 + a_3^2 + a_4^2)^2.$$

By Example 1.15, $\chi_{(a_{ij})}(X) = \det(X \cdot I_{d^2} - (a_{ij}) \otimes I_d) = \det((X \cdot I_d - (a_{ij})) \otimes I_d) = \det(X \cdot I_d - (a_{ij}))^d$, so $\mathrm{Tr}_{\mathrm{M}_d(k)/k}(a_{ij}) = d \sum a_{ii} = d \cdot \mathrm{tr}(a_{ij})$ and $\mathrm{N}_{\mathrm{M}_d(k)/k}(a_{ij}) = \det(a_{ij})^d$. This gives the relation between the characteristic polynomial of a matrix in $\mathrm{M}_d(k)$ as usually defined in linear algebra and as defined when viewing $\mathrm{M}_d(k)$ as a $k$-algebra: one is the $d$-th power of the other.

For $a, b \in A$ and $\alpha \in k$, recall we saw that $m_{a+b} = m_a + m_b$, $m_{ab} = m_a m_b$, and $m_{\alpha a} = \alpha m_a$. Thus, taking traces and determinants on these equations, we find that $\mathrm{Tr}_{A/k} \colon A \to k$ is $k$-*linear* (hence identically zero or surjective, since $k$ is a one-dimensional $k$-vector space) and $\mathrm{N}_{A/k} \colon A \to k$ is multiplicative. Obviously, $\mathrm{N}_{A/k} \colon A^\times \to k^\times$. More generally, we have

**Theorem 1.18.** *$a \in A^\times$ if and only if $\mathrm{N}_{A/k}(a) \neq 0$.*

*Proof.* Let $\chi_a(X) = X^n + c_{n-1}X^{n-1} + \cdots + c_0$, where $c_0 = \pm \mathrm{N}_{A/k}(a) \in k$. Since $\chi_a(a) = 0$, we have

$$(a^{n-1} + c_{n-1}a^{n-2} + \cdots + c_1)a = -c_0,$$

so if $c_0 \neq 0$ then we can divide by it and show $a$ is invertible. □

For $a \in A$, $k[a]$ is a commutative $k$-subalgebra of $A$. If $a \in A^\times$, then $k[a]$ is *not* neccesarily a field. For example, consider $A = k[Y]/(Y^2 - 1)$ with $a$ being the class of $Y$. Then $k[a] = A$ is not a field, but $a^2 = 1$ so $a$ is a unit in $A$. When $A$ is a division ring, we'll see below that $k[a]$ is a field.

Let $M_a = M_a(X)$ be the nonzero monic polynomial of least degree in $k[X]$ with $a$ as a root. (Don't confuse this with the notation $\mathrm{M}_d(k)$ for a matrix algebra over $k$.) Recall our standing assumption that $[A : k]$ is finite, so $M_a$ does exist since $A/k$ is algebraic. We call $M_a$ the *minimal polynomial* of $a$. As usual, if $f \in k[X]$ then $f(a) = 0$ if and only if $M_a \mid f$ in $k[X]$. In particular, $M_a \mid \chi_a$. However, the usual proof that minimal polynomials are irreducible if the $k$-algebra is a field does *not* work in general, since typically in $A$ one has $xy = 0$ without $x$ or $y$ equaling 0. If $A$ is a division ring then the proof for fields does work, so $M_a$ is always irreducible when $A$ is a division ring (left as exercise for the reader). As a consequence, let's show that if $A$ is a division ring, then $k[a]$ is a field for each $a \in A$. It is certainly a commutative ring, so we only have to show that every nonzero element of $k[a]$

has a multiplicative inverse. A typical element has the form $f(a)$ for $f \in k[X]$. If it is not zero, then $f(X)$ is not divisible by $M_a(X)$ in $k[X]$. Since $A$ is a division ring, $M_a(X)$ is irreducible in $k[X]$, so there are $g_1, g_2 \in k[X]$ such that $fg_1 + M_a g_2 = 1$ in $k[X]$. Taking $X = a$ we find that $f(a)g_1(a) = 1$ in $k[a]$. Thus, we may write $k(a)$ for $k[a]$ if $A$ is a division ring.

Since you are already familiar with the minimal polynomial for elements in a field extension but perhaps not as familiar with the characteristic polynomial for such elements, there may be some confusion between them. One way to understand the difference is by looking at the degree. The characteristic polynomial of each element $a \in A$ has the same degree, $[A : k]$, while the minimal polynomial has varying degree, equal to $[k[a] : k]$. If $A = k[a]$, then $\deg(M_a) = [A : k] = \deg(\chi_a)$ and both $M_a$ and $\chi_a$ are monic with $M_a \mid \chi_a$, so $M_a = \chi_a$. In general, as we will see in the next section, for division algebras (in particular, for field extensions) the degree is essentially the only difference between the minimal and characteristic polynomials, since one is a power of the other. The fact that characteristic polynomials have a uniform degree for all elements accounts in part for why we will use them sometimes instead of minimal polynomials when associating an element of an algebra with a polynomial having that element as a root.

## 2. Basic Results

First we say something about division algebras over algebraically closed fields and over finite fields. What are the finite division algebras over $\mathbf{C}$? If $D$ is a finite division algebra over $\mathbf{C}$, choose $x \in D$. Since $x$ is algebraic over $\mathbf{C}$, it is a root of some polynomial $f(X) \in \mathbf{C}[X]$. Let $f(X) = (X - z_1) \cdot \cdots \cdot (X - z_n)$, so (since $\mathbf{C}$ is in the center of $D$) we have $0 = f(x) = (x - z_1) \cdot \cdots \cdot (x - z_n)$. Since $D$ is a division ring, one of the factors $x - z_i$ is zero, so $x = z_i \in \mathbf{C}$. Thus, the only finite division algebra over $\mathbf{C}$ is $\mathbf{C}$. (Why can't we take $D = \mathbf{H}$? It contains $\mathbf{C}$, of course, but it is not a $\mathbf{C}$-algebra, since $\mathbf{C}$ is not in the center, which was the crucial property one needs in order to conclude that $f(x) = \prod_i (x - z_i)$.) In fact, we've actually proven that the only algebraic $\mathbf{C}$-division algebra (possibly infinite-dimensional over $\mathbf{C}$, a priori!) is $\mathbf{C}$. The exact same argument shows that the only algebraic division algebra over an algebraically closed field is the field itself. Since the finite-dimensional division algebras over a finite field are precisely the finite division rings, we mention a famous theorem due to Wedderburn: every finite division ring is commutative (see [2, §7.2] for a proof). Thus, there are no noncommutative finite-dimensional division algebras over a finite field.

Now we prove some theorems. Let's show that in a $k$-division algebra, the characteristic polynomial is a power of the minimal polynomial. In particular, the next result applies to finite field extensions of $k$.

**Theorem 2.1.** *If $D$ is a finite-dimensional $k$-division algebra, then for all $a \in D$,*

$$\chi_a = M_a^{\deg(\chi_a)/\deg(M_a)} = M_a^{[D:k]/[k[a]:k]}.$$

*Thus,* $\mathrm{N}_{D/k}(a) = (-1)^{[D:k]} M_a(0)^{[D:k]/[k[a]:k]}$.

*Proof.* If $a \in k$, then $\chi_a(X) = (X - a)^n$ and $M_a(X) = X - a$, so the result is clear. Thus assume $a \notin k$, so

$$M_a(X) = X^m + c_{m-1}X^{m-1} + \cdots + c_0,$$

with $m > 1$. By the definition of $M_a$, $\{1, a, \ldots, a^{m-1}\}$ is a $k$-basis of $k[a]$. Since $k$ lies in the center of $D$ and $D$ is a division ring and algebraic over $k$, $k[a]$ is a *field*, so $D$ can be

viewed as a, say, left vector space over $k[a]$. Since $k \subset k[a] \subset D$, $s = \dim_{k[a]}(D) < \infty$. Let $v_1, \ldots, v_s$ be a $k[a]$-basis of $D$. Then

$$D = \bigoplus_{j=1}^{s} k[a]v_j = \bigoplus_{j=1}^{s} \bigoplus_{i=0}^{m-1} ka^i v_j,$$

so to compute $\chi_a(X)$ we use as an ordered $k$-basis of $D$ the set

$$\{v_1, av_1, \ldots, a^{m-1}v_1; \ldots; v_s, av_s, \ldots, a^{m-1}v_s\}.$$

Note $[D : k] = s[k[a] : k]$. Since

$$a \cdot a^{m-1}v_j = -c_0 v_j - c_1 a v_j - \cdots - c_{m-1}a^{m-1}v_j,$$

we see that with respect to the above basis, $[m_a]$ is a block diagonal matrix with $s$ $m$ by $m$ diagonal blocks $B$ equal to

$$\begin{pmatrix} 0 & 0 & 0 & & -c_0 \\ 1 & 0 & 0 & \ldots & -c_1 \\ 0 & 1 & 0 & & -c_2 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & -c_{m-1} \end{pmatrix}.$$

Thus

$$\begin{aligned} \chi_a(X) &= \det(X \cdot I_m - B)^s \\ &= (X^m + c_{m-1}X^{m-1} + \cdots + c_0)^s \text{ by induction on } m \\ &= M_a(X)^s. \end{aligned}$$

Now set $X = 0$ and use the equation $\mathrm{N}_{D/k}(a) = (-1)^{[D:k]}\chi_a(0)$ to get the formula for $\mathrm{N}_{D/k}(a)$. □

All our concepts apply to a finite extension of fields. We now recover a more concrete definition of trace and norm when $A$ is a field extension of $k$ generated by one element. In particular, the next result applies whenever $k$ has characteristic 0 or is a finite field or is algebraically closed.

**Theorem 2.2.** *Let $k$ be a field and $K = k(\theta)$ a finite separable extension of degree $n$, so there are $n$ distinct embeddings $\sigma_1, \ldots, \sigma_n$ of $K$ into an algebraic closure $\overline{k}$ of $k$. For each $a \in K$,*

$$\chi_a(X) = \prod_i (X - \sigma_i(a)),$$

*so $\mathrm{Tr}_{K/k}(a) = \sum_i \sigma_i(a)$ and $\mathrm{N}_{K/k}(a) = \prod_i \sigma_i(a)$.*

*Proof.* Let

$$f(X) = (X - \sigma_1(a)) \cdot \cdots \cdot (X - \sigma_n(a)).$$

When $K = k(a)$, the result is clear since $\chi_a = M_a$. In general, $\chi_a = M_a^{[K:k(a)]}$ and every embedding $k(a) \to \overline{k}$ has $[K : k(a)]$ distinct extensions to $K \to \overline{k}$. Thus, letting $\tau$ run over

the distinct embeddings $k(a) \to \overline{k}$,

$$\begin{aligned} f(X) &= (\prod_{\tau} X - \tau(a))^{[K:k(a)]} \\ &= M_a^{[K:k(a)]} \\ &= \chi_a(X). \end{aligned}$$

□

For example, if $k = \mathbf{Q}$ and $K = \mathbf{Q}(\sqrt{r})$ for $r$ a nonsquare rational, then we can let $\sigma_1(\sqrt{r}) = \sqrt{r}$ and $\sigma_2(\sqrt{r}) = -\sqrt{r}$, so by the theorem,

$$\chi_{a_1+a_2\sqrt{r}}(X) = (X - (a_1 + a_2\sqrt{r}))(X - (a_1 - a_2\sqrt{r})) = X^2 - 2a_1X - (a_1^2 - ra_2^2),$$

which coincides with what we found before by doing a matrix calculation.

**Corollary 2.3.** *For $a \in \mathbf{F}_{q^n}$,*

$$\mathrm{Tr}_{\mathbf{F}_{q^n}/\mathbf{F}_q}(a) = a + a^q + \cdots + a^{q^{n-1}} \text{ and } \mathrm{N}_{\mathbf{F}_{q^n}/\mathbf{F}_q}(a) = a \cdot a^q \cdot \cdots \cdot a^{q^{n-1}}.$$

Since the trace map from $\mathbf{F}_{q^n}$ to $\mathbf{F}_q$ is a polynomial function of degree $q^{n-1}$ and $\mathbf{F}_{q^n}$ has size $q^n$, the trace map is not identically zero, so it is *onto* since it is $\mathbf{F}_q$-linear. Letting $\gamma$ be a generator of $\mathbf{F}_{q^n}^\times$, we see $\mathrm{N}_{\mathbf{F}_{q^n}/\mathbf{F}_q}(\gamma) = \gamma^{(q^n-1)/(q-1)}$ has (multiplicative) order $q-1$, so the norm map from $\mathbf{F}_{q^n}^\times$ to $\mathbf{F}_q^\times$ is *onto*.

We now consider a finite-dimensional $L$-algebra $A$ with $K$ a subfield of $L$ such that $[L:K] < \infty$. We have finite-dimensional algebras $A/L$, $A/K$, and $L/K$. The next theorem is called the transitivity of the trace.

**Theorem 2.4.** $\mathrm{Tr}_{A/K} = \mathrm{Tr}_{L/K} \circ \mathrm{Tr}_{A/L}$. *In particular, if $a \in L$ then $\mathrm{Tr}_{A/k}(a) = [A : L]\mathrm{Tr}_{L/K}(a)$. If $A$ is a field, or even a division ring, then $\mathrm{Tr}_{A/k}(a) = [A : K[a]]\mathrm{Tr}_{K[a]/K}(a)$.*

*Proof.* Let $(e_1, \ldots, e_m)$ be an ordered $L$-basis of $A$ and $(f_1, \ldots, f_n)$ be an ordered $K$-basis of $L$. Thus as an ordered $K$-basis of $A$ we can use

$$(e_1f_1, \ldots, e_1f_n; \ldots; e_mf_1, \ldots, e_mf_n).$$

For $a \in A$, let

$$ae_j = \sum_{i=1}^m c_{ij}e_i, \quad c_{ij}f_s = \sum_{r=1}^n b_{ijrs}f_r,$$

for $c_{ij} \in L$ and $b_{ijrs} \in K$. Thus $a(e_jf_s) = \sum_i \sum_r b_{ijrs}e_if_r$. So

$$[m_a]_{A/L} = (c_{ij}), \quad [m_{c_{ij}}]_{L/K} = (b_{ijrs}), \quad [m_a]_{A/K} = ([m_{c_{ij}}]_{L/K}).$$

Thus

$$\begin{aligned} \mathrm{Tr}_{L/K}(\mathrm{Tr}_{A/L}(a)) &= \mathrm{Tr}_{L/K}(\sum_i c_{ii}) \\ &= \sum_i \mathrm{Tr}_{L/K}(c_{ii}) \\ &= \sum_i \sum_r b_{iirr} \\ &= \mathrm{Tr}_{A/K}(a). \end{aligned}$$

□

We use Theorem 2.2 to establish a transitivity formula for norms in two cases.

**Theorem 2.5.** *(i) Let $L/K$, $K/k$ be finite extension fields, where $k$ has characteristic 0 or is finite or is algebraically closed. Then $\mathrm{N}_{L/k} = \mathrm{N}_{K/k} \circ \mathrm{N}_{L/K}$.*

*(ii) If $D$ is a finite-dimensional $K$-division algebra and $k$ is a subfield of $K$ such that $[K : k]$ is finite, then for $a \in K$, $\mathrm{N}_{D/k}(a) = \mathrm{N}_{K/k}(\mathrm{N}_{D/K}(a))$.*

*Proof.* (i) Fix an algebraic closure $\overline{k}$ of $k$. Fix an embedding $k \hookrightarrow \overline{k}$. Let $\tau_1, \ldots, \tau_m$ be the distinct embeddings of $K$ into $\overline{k}$ extending $\psi$. Each $\tau_i$ has $n = [L : K]$ distinct extensions $\sigma_{i1}, \ldots, \sigma_{in}$ to embeddings $L \to \overline{k}$. Thus

$$\begin{aligned} \mathrm{N}_{K/k}(\mathrm{N}_{L/K}(a)) &= \prod_{i=1}^{m} \tau_i(\mathrm{N}_{L/K}(a)) \\ &= \prod_{i=1}^{m} \tau_i(\tau_i^{-1} \prod_{j=1}^{m} \sigma_{ij}(a)) \\ &= \prod_{i,j} \sigma_{ij}(a) \\ &= \mathrm{N}_{L/k}(a). \end{aligned}$$

(ii) Let $M_a(X)$ be the minimal polynomial of $a$ in $k[X]$. By Theorem 2.1,

$$\begin{aligned} \mathrm{N}_{K/k}(\mathrm{N}_{D/K}(a)) &= \mathrm{N}_{K/k}(a^{[D:k]}) \\ &= (\mathrm{N}_{K/k}(a))^{[D:k]} \\ &= ((-1)^{[K:k]} M_a(0)^{[K:k]/[k[a]:k]})^{[D:K]} \\ &= (-1)^{[D:k]} M_a(0)^{[D:k]/[k[a]:k]} \\ &= \mathrm{N}_{D/k}(a). \end{aligned}$$

□

**Remark 2.6.** It is true more generally that for $K/k$ a finite extension of fields and $A$ a finite-dimensional $K$-algebra that $\mathrm{N}_{A/k} = \mathrm{N}_{K/k} \circ \mathrm{N}_{A/K}$; see [1, Appendix B]. We will only need here the transitivity of the norm map in the cases where we have proven it above.

We now describe how the characteristic polynomial varies over a division algebra.

**Theorem 2.7.** *Let $D$ be a finite $k$-division algebra, $a \in D$. In $\overline{k}[X]$, let*

$$\chi_a(X) = (X - r_1) \cdot \cdots \cdot (X - r_n).$$

*For $g(X) \in k[X]$,*

$$\chi_{g(a)}(X) = (X - g(r_1)) \cdot \cdots \cdot (X - g(r_n)).$$

*In particular, $\chi_{a+1}(X) = \chi_a(X-1)$ and $\chi_{a^m}(X) = (X-r_1^m)\cdot\cdots\cdot(X-r_n^m)$, so $\mathrm{N}_{D/k}(a+1) = (-1)^{[D:k]}\chi_a(-1)$ and $\mathrm{Tr}_{D/k}(a^m) = \sum_i r_i^m$.*

*Proof.* Let $f(X) = (X-g(r_1))\cdot\cdots\cdot(X-g(r_n))$. The coefficients are symmetric polynomials in $r_1, \ldots, r_n$, so by the symmetric function theorem $f(X) \in k[X]$. Let $M_{g(a)}(X) \in k[X]$ be the minimal polynomial of $g(a)$ over $k$, so $M_{g(a)}$ is irreducible in $k[X]$ (since $D$ is a division ring). The fields $k(a)$ and $k(r_i)$ are isomorphic $k$-algebras, so $M_{g(a)}$ is the minimal polynomial for $g(r_i)$ over $k$, since $M_{g(a)}$ is irreducible monic in $k[X]$ and has $g(r_i)$ as a root.

Let $\pi(X)$ be an irreducible monic factor of $f$ in $k[X]$. Then $\pi(g(r_i)) = 0$ for some $i$, so $M_{g(a)} \mid \pi$ in $k[X]$. Thus $\pi = M_{g(a)}$, since both are monic and irreducible. Since $f$ is monic, we thus see that $f$ is a power of $M_{g(a)}$. By Theorem 2.1, $\chi_{g(a)} \in k[X]$ is a power of $M_{g(a)}$ with degree $[D : k] = n = \deg(f)$, so $f = \chi_{g(a)}$. □

The following theorem could have been part of the previous one, but we have chosen to isolate it because its statement is worth remembering on its own. It gives a basic connection between the norm and trace.

**Theorem 2.8.** *Let $D$ be an $n$-dimensional $k$-division algebra, for example a field extension of degree $n$. For $a \in D$, let $\chi_a(X) = X^n + c_{n-1}X^{n-1} + \cdots + c_0 \in k[X]$. For all $y \in k$,*

$$\begin{aligned} \mathrm{N}_{D/k}(1+ay) &= 1 - c_{n-1}y + \cdots + (-1)^{n-1}c_1y^{n-1} + (-1)^n c_0 y^n \\ &= 1 + \mathrm{Tr}_{D/k}(a)y + \cdots + \mathrm{N}_{D/k}(a)y^n, \end{aligned}$$

*so for fixed $a \in D$, $y \mapsto \mathrm{N}_{D/k}(1+ay)$ is a polynomial function in $y \in k$.*

*Proof.* We may assume $y \neq 0$. In a splitting field of $\chi_a$, let $\chi_a(X) = \prod_i (X - r_i)$. By Theorem 2.7, $\chi_{1+ay}(X) = \prod_i (X - (1 + yr_i))$. Since $y \in k$, $y$ commutes with all elements of $D$. So

$$\begin{aligned} \mathrm{N}_{D/k}(1+ay) &= (-1)^n \chi_{1+ay}(0) \\ &= (-1)^n \prod_i (-1 - yr_i) \\ &= (-y)^n \prod_i (-1/y - r_i) \\ &= (-y)^n \chi_a(-1/y) \\ &= (-y)^n ((-1/y)^n + c_{n-1}(-1/y)^{n-1} + \cdots + c_1(-1/y) + c_0) \\ &= 1 - c_{n-1}y + \cdots + (-1)^{n-1}c_1y^{n-1} + (-1)^n c_0 y^n. \end{aligned}$$

□

Let $A$ be a finite-dimensional $k$-algebra, say $n = [A : k]$. By choosing a $k$-basis of $A$, we get an isomorphism $A \cong k^n$ of $k$-vector spaces, allowing us to view the norm map $\mathrm{N}_{A/k} \colon A \to k$ as a map $k^n \to k$. The next theorem shows that this is a polynomial map. Note that the proof uses none of the theorems we have proven; it only uses the definition of the norm map.

**Theorem 2.9.** *For a $k$-basis $e_1, \ldots, e_n$ of $A$, there is a $P \in k[X_1, \ldots, X_n]$ such that*

$$\mathrm{N}_{A/k}(x_1e_1 + \cdots + x_ne_n) = P(x_1, \ldots, x_n).$$

*Proof.* Let $e_ie_j = \sum_{r=1}^n c_{ijr}e_r$, $c_{ijr} \in k$. For $a = x_1e_1 + \cdots + x_ne_n$ with $x_i \in k$,

$$\begin{aligned} ae_j &= \sum_r x_r e_r e_j \\ &= \sum_i (\sum_r x_r c_{rji}) e_i, \end{aligned}$$

so with respect to the basis $(e_1, \ldots, e_n)$, $[m_a] = (\sum_r x_r c_{rji})$. Let

$$P(X_1, \ldots, X_n) = \det(\sum_r c_{rji} X_r) \in k[X_1, \ldots, X_n],$$

so $\mathrm{N}_{A/k}(a) = P(x_1, \ldots, x_n)$. □

This concludes the results on algebras that we will need, but we make some additional remarks for the interested reader. Since $a \mapsto m_a$ is an injective $k$-algebra homomorphism from $A$ to $\mathrm{Hom}_k(A, A)$, we see that the minimal polynomial $M_a$ of $a \in A$ equals the minimal polynomial of the linear map $m_a$. By definition, the characteristic polynomial $\chi_a$ of $a$ equals the characteristic polynomial of the linear map $m_a$. Thus, properties of minimal or charactersitic polynomials of elements of a finite-dimensional $k$-algebra ought to follow from properties of minimal or characteristic polynomials of matrices.

As an example of this principle, we show that Theorem 2.7 applies to all finite $k$-algebras, not just division algebras. The characteristic polynomial of the linear map $m_a$ is $\prod_{i=1}^n (X - r_i)$. By [3, Theorem 3.10, Chapter XIV], the characteristic polynomial of the linear map $g(m_a)$ is $\prod_{i=1}^n (X - g(r_i))$. Since $g(m_a) = m_{g(a)}$, we are done. This implies that Theorem 2.8 is also true for finite $k$-algebras, since it is a formal consequence of Theorem 2.7.

Finally, note that in a finite $k$-algebra that is not a division algebra, $\chi_a$ is *not* necessarily a power of $M_a$. For the relation between them when $A$ is a matrix algebra (a fairly general class of examples), see [3, §§2-3, Chapter XIV], where $q_a$ and $P_a$ are written in place of $M_a$ and $\chi_a$. See especially Theorem 2.1 and Theorem 3.5. The possible reducibility of $M_a$ in $k[X]$ when $A$ is not a division algebra can be viewed as the basic obstruction to $\chi_a$ being a power of $M_a$.

**Exercise.** Let $A$ be a finite $k$-algebra, $a \in A$. Write

$$\chi_a(X) = X^n + c_{n-1}(a)X^{n-1} + \cdots + c_1(a)X + c_0(a),$$

so we view the coefficients of $\chi_a$ as functions of $a$. For example, $c_{n-1}(a) = -\mathrm{Tr}_{A/k}(a)$ and $c_0(a) = (-1)^n \mathrm{N}_{A/k}(a)$.

1. Show $c_i(ab) = c_i(ba)$ for $0 \leq i \leq n-1$ and $a, b \in A$.
2. Show this is also true for $M_a(X)$ in place of $\chi_a(X)$. How does this generalize to a family $f_a(X)$ of polynomials such that $f_a(X) \mid \chi_a(X)$ for all $a$?

## References

[1] J.W.S. Cassels, *Local Fields*, Cambridge University Press, New York, 1986.
[2] I.N. Herstein, *Topics In Algebra*, 2nd ed., Wiley, New York, 1975.
[3] S. Lang, *Algebra*, 3rd ed., Addison-Wesley, New York, 1993.