The characteristic polynomial and determinant are not ad hoc constructions

Skip Garibaldi*

30 April 2004

1 INTRODUCTION.

Most people are first introduced to the characteristic polynomial and determinant of a matrix in a linear algebra course as undergraduates. The determinant is usually defined as an alternating sum of products of entries of the matrix (as in Jacobi [16, sec. 4]) or as the unique map $M_n(F) \to F$ that is multilinear and alternating in the columns and has the value 1 at the identity matrix (as in Weierstrass [28] and the books by Hungerford [15], Lang [22], and Dummit and Foote [11]). As a student, I thought that these definitions were at best magical and at worst ad hoc. Where did the determinant come from? This paper gives definitions that I hope the reader will find more natural.

Admittedly, the determinant of a linear transformation on $\mathbb{R}^n$ is a natural enough object: its absolute value gives the factor by which the transformation enlarges volumes, and its sign says whether or not the map preserves orientation. These properties imply Weierstrass's axioms (see, for example, [13] or [23, sec. 5]).

Another good definition of the determinant—not so common at the undergraduate level—is in terms of the nth exterior power $\wedge^n F^n$, as in [5, chap. 3, sec. 8]. This also leads to the Weierstrass axioms.

But even these two "good" definitions have a taint of being special to matrices. (The first is even limited to matrices with real entries.) After all, analogues of the determinant are known for the quaternions, the octonions, finite-dimensional field extensions, ... It is not clear how to adapt the two good definitions to handle these algebras. As mathematicians, we should demand a definition that works simultaneously in all cases. We give such a definition of the characteristic polynomial in section 2; the constant term of this characteristic polynomial gives an analogue of the determinant. (One normally begins with a definition for the determinant and then defines the characteristic polynomial

*This article was published in the American Mathematical Monthly **111**, no. 9 (2004), 761–778. This version is very similar to the published version.

of a matrix a as $\det(xI - a)$. We work in the opposite direction here.) For n-by-n matrices, we derive Jacobi's alternating sum formula for the determinant. We also recover the known ad hoc formulas for the determinant for quaternions and finite-dimensional field extensions. Moreover, the product formula $\det(aa') = \det(a)\det(a')$ always holds.

The philosophy is the following. Consider the lines in $\mathbb{R}^2$ given by the equations

$$ax + by = c, \quad a'x + b'y = c'.$$

If the coefficients a, b, c, a', b', and c' are specific real numbers, the lines might be parallel or the same (degenerate case), but "typically" they intersect at exactly one point. If we treat the coefficients as independent indeterminates, we say that the lines are *generic*. Such lines intersect at the point

$$(x, y) = \left(\frac{b'c - c'b}{ab' - a'b}, \frac{ac' - a'c}{ab' - a'b} \right).$$

A typical n-by-n matrix has n distinct eigenvalues.[1] This is true in particular for a generic matrix γ, meaning one whose entries are indeterminates. For such a matrix, the traditional characteristic polynomial is just the minimal polynomial. To define the characteristic polynomial of a specific matrix a, we first find the minimal polynomial $\mathrm{minpoly}_\gamma(x)$ of our generic matrix. Plugging in specific values for the indeterminates in γ, we get a polynomial whose only indeterminate is x, and this is the characteristic polynomial of a. This method of defining the characteristic polynomial works for all finite-dimensional F-algebras, and the determinant is (up to a sign) the constant term of the characteristic polynomial.

The core of the idea—looking at the minimal polynomial of a generic element—goes back to the late 1800s (see, for example, [27, p. 241] and [26, p. 301]). All treatments that I have found, however, do not develop the properties of the general characteristic polynomial (as in [9, chap. 7]) or make use of known properties of the characteristic polynomial and determinant for matrices in studying the general characteristic polynomial (as in [1, sec. 10.3], [19], [18], and [17, sec. 5.18]). We use only elementary properties of matrices from the very nice paper [3] and the book [4].

Readers with an algebraic background may argue that one can obtain the characteristic polynomial of an n-by-n matrix over an arbitrary field F by applying the structure theory for finitely-generated torsion modules over a principal ideal domain (as is done in [14, sec. 6.7]). But if one is using that much algebra, the contents of this paper are not so far away and the results here are much stronger.

[1] If you randomly choose an n-by-n real matrix, the probability that you pick one with distinct complex eigenvalues is 100%. This is because the other real matrices are a set with Lebesgue measure zero. Topologically, amongst the n-by-n matrices, those with distinct eigenvalues form a dense open subset. Over $\mathbb{R}$ or $\mathbb{C}$ this is true in the usual topology; it is also true over any infinite field in the Zariski topology (see, for example, [25]).

2 THE CHARACTERISTIC POLYNOMIAL.

In this section, we define the characteristic polynomial of an element a in a finite-dimensional F-algebra A and give some of its basic properties. We begin with a definition.

Definition 2.1. Let F be a field. An *F-algebra* is a ring A with a multiplicative identity 1 ($\neq 0$) such that A is an F-vector space and

$$\alpha(ab) = (\alpha a)b = a(\alpha b) \qquad (\alpha \in F; a, b \in A). \tag{2.2}$$

(Alternately, A is a ring with identity such that there is a one-to-one homomorphism $F \to Z(A)$ that maps the identity in F to the identity in A.) All algebras that we consider will be *finite-dimensional* as vector spaces over F. Principal examples of F-algebras are the n-by-n matrices $M_n(F)$ and Hamilton's quaternions, which form an $\mathbb{R}$-algebra (see Example 4.1).

Note that the definition gives a copy of F inside the center of A, but there is no requirement that F be the entire center of A. For example, A may be taken to be a finite-dimensional field extension of F.

For clarity of exposition, we require that A be associative (for example, A cannot be the octonions). However, the definition of the characteristic polynomial that we present also works for the much broader class of strictly power-associative F-algebras, which includes the octonions and Jordan algebras (see section 7 for precise statements).

The main tool we need that may not be typically discussed in first-year graduate algebra is the tensor product $\otimes$. It is a canonical way to take a vector space V over the field F and produce a vector space over a larger field K. The new vector space is denoted by $K \otimes V$. Heuristically, one thinks of the elements of $K \otimes V$ as finite sums $\sum k_i v_i$ where the k_i are in K and the v_i are in V. The addition and scalar multiplication are the obvious ones. Formally, one writes such an element as $\sum k_i \otimes v_i$, although typically we will omit the symbol $\otimes$. The map $v \mapsto 1 \otimes v$ identifies V with a subset of $K \otimes V$. In particular, linearly independent elements of V are sent to linearly independent elements of $K \otimes V$. Moreover, if $v_1, v_2, \ldots, v_m$ span the F-vector space V, then their images span the K-vector space $K \otimes V$. Consequently, the dimension of V (over F) is the same as the dimension of $K \otimes V$ (over K).

If the vector space V has a multiplication—as is the case for the algebra A—then so does $K \otimes V$. For simple elements $k \otimes v$ and $k' \otimes v'$ of $K \otimes V$, put

$$(k \otimes v)(k' \otimes v') = kk' \otimes vv'$$

For general elements, expand $(\sum k_i \otimes v_i)(\sum k'_j \otimes v'_j)$ using the distributive law and apply the rule for simple elements. In this way, $K \otimes A$ is a K-algebra. For example, $K \otimes M_n(F)$ is isomorphic to $M_n(K)$.

(The previous two paragraphs are sufficient for understanding almost all of this paper, but the proofs of Lemmas 2.3 and 6.3 and Propositions 2.7 and 2.11 require a more general version. Let R be an integral domain containing F. We write $R \otimes A$ for the ring of finite sums $\sum r_i \otimes a_i$ with r_i in R and a_i in A, endowed with the product described in the preceding paragraph. It is naturally a subring of $K \otimes A$ for K the field of fractions of R. It is a free R-module whose rank is the dimension of A over F. The curious reader can find a detailed and more general introduction to the tensor product in [11, sec. 10.4].)

Now let $a_1, a_2, \ldots, a_m$ be an F-basis for A, let $R = F[t_1, t_2, \ldots, t_m]$ for (commuting) indeterminates $t_1, \ldots, t_m$, and let K be the quotient field of R. We call the element $\gamma = \sum_i t_i a_i$ of $K \otimes A$ a *generic element*. The powers $1, \gamma, \gamma^2, \ldots$ of γ live in the finite-dimensional K-vector space $K \otimes A$, so they are linearly dependent. That is, there is a nonzero polynomial $f(x)$ in $K[x]$ such that $f(\gamma) = 0$. Let $\mathrm{minpoly}_{\gamma/K}$ denote the nonzero monic polynomial in $K[x]$ of smallest degree such that $\mathrm{minpoly}_{\gamma/K}(\gamma) = 0$. It is called the *minimal polynomial* of γ over K.

Note that this polynomial is unique, for if $f(x)$ and $g(x)$ are monic polynomials of minimal degree such that $f(\gamma) = g(\gamma) = 0$, then $h(x) = f(x) - g(x)$ is a polynomial of smaller degree such that $h(\gamma) = 0$. This contradicts the minimality of f and g unless $h(x) = 0$.

Lemma 2.3. *The minimal polynomial* $\mathrm{minpoly}_{\gamma/K}$ *is in* $R[x]$*, not just* $K[x]$*.*

Proof. Consider the R-submodules A_j of $R \otimes A$ generated by $\{1, \gamma, \gamma^2, \ldots, \gamma^j\}$. They form an ascending chain $A_1 \subseteq A_2 \subseteq \cdots$. Since R is Noetherian (Hilbert's Basis Theorem) and $R \otimes A$ is a finitely-generated R-module, this chain must stabilize. That is, γ^{j+1} is in A_j for some j, so γ satisfies a monic polynomial f in $R[x]$.[2] Since $\mathrm{minpoly}_{\gamma/K}$ divides f in $K[x]$ and both are monic, $\mathrm{minpoly}_{\gamma/K}$ lies in $R[x]$ by Gauss's lemma. ☐

Definition 2.4. Express the element a of A in terms of the basis $a_1, a_2, \ldots, a_m$ as $a = \sum_i \alpha_i a_i$ with α_i in F. The substitution $t_i \mapsto \alpha_i$ defines a map $R[x] \to F[x]$. We call the image of $\mathrm{minpoly}_{\gamma/K}$ in $F[x]$ the *characteristic polynomial of* a and denote it by $\mathrm{chpoly}_{a,A/F}$ or simply chpoly_a.

Remark 2.5. It is immediate from the definition that

$$\deg(\mathrm{chpoly}_{a,A/F}) \leq \dim_F A$$

and that the degree of $\mathrm{chpoly}_{a,A/F}$ is the same for all a in A.

Example 2.6 (Upper-triangular matrices). Let A be the algebra of n-by-n upper-triangular matrices over F. Write E_{ij} for the matrix whose only nonzero entry is a 1 in the (i, j)-position. Fix a basis $a_1, a_2, \ldots, a_m$ for A over F consisting of E_{ij} such that $a_i = E_{ii}$ when $1 \leq i \leq n$. Let γ be the corresponding generic element defined earlier.

[2] A more direct argument would be: $R[\gamma]$ is an R-submodule of $R \otimes A$ and $R \otimes A$ is a finitely-generated R-module. Hence γ is integral over R [15, Theorem 8.5.3]. Unfortunately, the typical proof of this implication invokes determinants, so we use instead that R is Noetherian.

Let I_n denote the n-by-n identity matrix. For $i = 1, 2, \ldots, n$ the matrix $\gamma - t_i I_n$ has $n-1$ pivot columns—equivalently, $n-1$ leading 1s—in its row-reduced form, hence it has a nonzero kernel. That is, γ has an eigenvector in K^n with eigenvalue t_i. Since the t_i are distinct elements of K, these eigenvectors form a basis for K^n, so γ is similar in $M_n(K)$ to the diagonal matrix with diagonal entries $t_1, t_2, \ldots, t_n$. The minimal polynomial of γ is $\prod_{i=1}^n (x - t_i)$, since similar matrices have the same minimal polynomials.[3] By substitution, an upper-triangular matrix b has characteristic polynomial $\prod_{i=1}^n (x - b_{ii})$.

We now show that the characteristic polynomial is well defined in general.

Proposition 2.7. *The characteristic polynomial* $\mathrm{chpoly}_{a,A/F}$ *depends only on a, A, and F (and not on the choice of basis $a_1, a_2, \ldots, a_m$ for A).*

Proof. Suppose that we take another F-basis $b_1, b_2, \ldots, b_m$ of A with a corresponding generic element $\varepsilon = \sum_i t_i b_i$. We may write $b_i = \sum_j g_{ij} a_j$ for g an invertible matrix in $M_m(F)$. Let $f : R \to R$ be the F-algebra automorphism defined by

$$f(t_j) = \sum_i t_i g_{ij}.$$

Write a in A in terms of each basis as

$$a = \sum_j \alpha_j a_j = \sum_i \beta_i b_i. \tag{2.8}$$

We have a diagram

$$\begin{array}{ccc} R[x] & \xrightarrow{t_j \mapsto \alpha_j} & F[x] \\ {\scriptstyle f}\big\downarrow & & \big\| \\ R[x] & \xrightarrow{t_i \mapsto \beta_i} & F[x] \end{array}$$

with horizontal arrows the substitution maps. Equation (2.8) ensures that $\alpha_j = \sum_i \beta_i g_{ij}$ for all j, hence the diagram commutes.

If we begin with $\mathrm{minpoly}_{\gamma/K}$ in the upper left, substitution gives the member of $F[x]$ that is chpoly_a computed with respect to the basis $a_1, a_2, \ldots, a_m$. On the other hand, f extends in an obvious way to an automorphism of $R \otimes A$ such that

$$f(\gamma) = \sum_j f(t_j) a_j = \sum_j \Big(\sum_i t_i g_{ij}\Big) a_j = \sum_i t_i \Big(\sum_j g_{ij} a_j\Big) = \varepsilon.$$

Hence $f(\mathrm{minpoly}_{\gamma/K}) = \mathrm{minpoly}_{\varepsilon/K}$. The image of this in $F[x]$ is chpoly_a computed with respect to the basis $b_1, b_2, \ldots, b_m$. The commutativity of the diagram establishes the proposition. $\square$

[3]This argument may appear to be excessively long. It is included here to illustrate that we are not making use of determinants.

Undergraduates are asked to find the characteristic polynomial of a matrix whose entries are specific rational numbers. Of course their answer is the same whether they think of the matrix as living in $M_n(\mathbb{Q})$, $M_n(\mathbb{R})$, or $M_n(\mathbb{C})$. The characteristic polynomial from Definition 2.4 has the same property.

Lemma 2.9. *Let E be a field containing F, and fix a in A. The minimal polynomials and characteristic polynomials of a are the same over F and over E.*

Proof. Let d be the degree of the polynomial $\mathrm{minpoly}_{a/F}$. The elements $1_A, a, \dots, a^{d-1}$ are linearly independent over F, hence the elements $1 \otimes 1_A, 1 \otimes a, \dots, 1 \otimes a^{d-1}$ of $E \otimes A$ are linearly independent over E. Note that $1 \otimes a^i$ equals $(1 \otimes a)^i$ for all i, so the minimal polynomial $\mathrm{minpoly}_{(1\otimes a)/E}$ of a over E has degree at least d. Since this polynomial divides $\mathrm{minpoly}_{a/F}$, the two polynomials are the same.

The F-basis $a_1, a_2, \dots, a_m$ of A gives an E-basis $1 \otimes a_1, 1 \otimes a_2, \dots, 1 \otimes a_m$ of $E \otimes A$, and the generic element constructed from this E-basis is the image of γ in $E(t_1, t_2, \dots, t_m) \otimes A$. Since the minimal polynomials of γ over K and over $E(t_1, \dots, t_m)$ are the same by the preceding paragraph, we get

$$\mathrm{chpoly}_{a,A/F} = \mathrm{chpoly}_{(1\otimes a),(E\otimes A)/E}$$

by substitution. ∎

In general, we write

$$\mathrm{chpoly}_a(x) = x^n - c_1(a)x^{n-1} + \cdots \\ \cdots + (-1)^{n-1}c_{n-1}(a)x + (-1)^n c_n(a). \quad (2.10)$$

The elements $c_1(a)$ and $c_n(a)$ play the roles of the trace and determinant of a.

Proposition 2.11. *Let A be a finite-dimensional F-algebra. Then the following statements are true:*

(1) (Cayley-Hamilton) $\mathrm{chpoly}_a(a) = 0$ *for each a in A.*

(2) *If $\varphi \colon A \to A$ is a ring automorphism or anti-automorphism that restricts to an automorphism of F, then $\varphi(c_i(a)) = c_i(\varphi(a))$ for all a in A.*

(3) *The c_i satisfy $c_i(\alpha a) = \alpha^i c_i(a)$ for all α in F and a in A.*

(4) *The mapping $c_1 \colon A \to F$ is F-linear.*

(5) *If B is a subalgebra of A and b is in B, then* $\mathrm{chpoly}_{b,B/F}$ *divides* $\mathrm{chpoly}_{b,A/F}$ *in $F[x]$.*

We will observe in Theorem 3.4 that our notion of characteristic polynomial on $M_n(F)$ is the same as the usual one. Then Proposition 2.11 contains many results that one typically proves in a linear algebra course. For example, (2) tells us that that similar matrices have the same characteristic polynomial and that $\det(a^t) = \det(a)$ for a in $M_n(F)$, where a^t denotes the transpose of a.

Proof. To establish (1), write $a = \sum_i \alpha_i a_i$. Then $\operatorname{chpoly}_a(a)$ is obtained by making the substitution $t_i \mapsto \alpha_i$ in $\operatorname{minpoly}_{\gamma/K}(\gamma)$. (This substitution defines a homomorphism $R \otimes A \to A$.) Since $\operatorname{minpoly}_{\gamma/K}(\gamma) = 0$ in $R \otimes A$, we have $\operatorname{chpoly}_a(a) = 0$ in A.

We next prove (2). The map φ extends naturally to an automorphism $g \mapsto {}^{\varphi}g$ of $K[x]$ by applying φ to the coefficients of g. Similarly, φ extends to an automorphism or anti-automorphism of $R \otimes A$ such that ${}^{\varphi}g(\varphi(u)) = \varphi(g(u))$ for every g in $R[x]$ and u in $R \otimes A$. We have: $g(\gamma) = 0$ *if and only if* ${}^{\varphi}g(\varphi(\gamma)) = 0$. Hence

$$ {}^{\varphi}\operatorname{minpoly}_{\gamma/K} = \operatorname{minpoly}_{\varphi(\gamma)/K} . $$

The diagram

$$\begin{array}{ccc} R[x] & \xrightarrow{t_i \mapsto \alpha_i} & F[x] \\ {\scriptstyle\varphi}\big\downarrow & & \big\downarrow{\scriptstyle\varphi} \\ R[x] & \xrightarrow{t_i \mapsto \varphi(\alpha_i)} & F[x] \end{array}$$

commutes. Beginning with $\operatorname{minpoly}_{\gamma/K}$ in the upper left and going clockwise, we obtain chpoly_a in the upper right $F[x]$, and then ${}^{\varphi}\operatorname{chpoly}_a$ in the lower right $F[x]$. Going counterclockwise, we obtain ${}^{\varphi}\operatorname{minpoly}_{\gamma/K} = \operatorname{minpoly}_{\varphi(\gamma)/K}$ in the lower left. The image in the lower right $F[x]$ is the characteristic polynomial of $\varphi(a)$ (computed with respect to the basis $\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_m)$ of A, but that is irrelevant by Proposition 2.7). The commutativity of the diagram gives the desired equality $\operatorname{chpoly}_{\varphi(a)} = {}^{\varphi}\operatorname{chpoly}_a$.

Turning to (3) and (4), we suppose first that α is not 0. If we write the minimal polynomial of γ/K as $\sum_{i=0}^{n} c_i x^i$ for c_i in R, then the minimal polynomial of $\alpha\gamma/K$ is $\sum_{i=0}^{n} c_i \alpha^{n-i} x^i$. Thus

$$ \alpha^n \operatorname{minpoly}_{\gamma/K}(x) = \operatorname{minpoly}_{\alpha\gamma/K}(\alpha x) \qquad (\text{in } R[x]) $$

and

$$ \alpha^n \operatorname{chpoly}_a(x) = \operatorname{chpoly}_{\alpha a}(\alpha x) \qquad (\text{in } F[x]). $$

This gives

$$ \alpha^n c_i(a) = \alpha^{n-i} c_i(\alpha a) \qquad (0 \le i \le n; \alpha \ne 0). \tag{2.12} $$

In particular, (2.12) holds when α is an indeterminate. Since $c_i \colon A \to F$ is given by a polynomial in the coordinates of a with respect to some basis $a_1, a_2, \ldots, a_m$, this polynomial is homogeneous of degree i. This gives (3) and (4).

For (5), fix a basis $b_1, b_2, \ldots, b_r$ of B and extend it to a basis $a_1, a_2, \ldots, a_m$ of A with $a_i = b_i$ when $1 \le i \le r$. By analogy, set $S = F[t_1, t_2, \ldots, t_r]$, let L denote the quotient field of S, and let ε be the generic element $\sum_{i=1}^{r} t_i b_i$ in $S \otimes A$.

Write $b = \sum_i \beta_i b_i$ with β_i in F. Define a map $\phi \colon R \to S$ by sending t_j to 0 when $r < j \le m$. The image of $\operatorname{minpoly}_{\gamma/K}$ under the composition

$$ R[x] \xrightarrow{\phi} S[x] \xrightarrow{t_i \mapsto \beta_i} F[x] $$

is $\mathrm{chpoly}_{b,A/F}$. Similarly, the image of $\mathrm{minpoly}_{\varepsilon/L}$ is $\mathrm{chpoly}_{b,B/F}$.

The homomorphism ϕ extends naturally to a map $R \otimes A \to S \otimes A$ such that $\phi(\gamma) = \varepsilon$. We have

$$0 = \phi(0) = \phi(\mathrm{minpoly}_{\gamma}(\gamma)),$$

which equals the polynomial $\phi(\mathrm{minpoly}_{\gamma})$ in $S[x]$ evaluated at $\phi(\gamma) = \varepsilon$. Consequently, $\mathrm{minpoly}_{\varepsilon/L}$ divides $\phi(\mathrm{minpoly}_{\gamma/K})$ in $L[x]$. Since S is a unique factorization domain, $\mathrm{minpoly}_{\varepsilon/L}$ divides $\phi(\mathrm{minpoly}_{\gamma/K})$ in $S[x]$. Consequently, the image $\mathrm{chpoly}_{b,B/F}$ of $\mathrm{minpoly}_{\varepsilon/L}$ in $F[x]$ divides the image $\mathrm{chpoly}_{b,A/F}$ of $\phi(\mathrm{minpoly}_{\gamma/K})$. □

3 MATRICES.

In this section, we observe that the characteristic polynomial as defined in the previous section agrees with the usual linear algebra notion of characteristic polynomial in the case where $A = M_n(F)$.

Everyone knows the next lemma, but maybe not the clean proof:

Lemma 3.1. *If T is a linear transformation on an F-vector space of positive dimension n, then T satisfies a nonzero polynomial of degree at most n.*

Proof. We sketch the nice proof from [7] that proceeds by induction on n. The case $n = 1$ is clear. We assume that $n > 1$. Let v be a nonzero vector in the vector space V. The $n + 1$ vectors

$$v, T(v), T^2(v), \ldots, T^n(v)$$

must be linearly dependent, so there is a nonzero polynomial $g(x)$ in $F[x]$ of degree $\leq n$ such that $g(T)v = 0$.

Set $U = \ker g(T)$. The linear transformations T and $g(T)$ commute, so $T(U) \subseteq U$ and T induces a linear transformation $T_{V/U}$ on V/U. By induction, $T|_U$ satisfies a polynomial $m_U(x)$ with $\deg(m_U) \leq \dim U$, and $T_{V/U}$ satisfies a polynomial $m_{V/U}(x)$ with $\deg(m_{V/U}) \leq \dim V/U = \dim V - \dim U$. Then $m_{V/U}(T)V \subseteq U$ and T satisfies the polynomial $m_U(x)\, m_{V/U}(x)$. Moreover,

$$\deg(m_U \cdot m_{V/U}) = \deg(m_U) + \deg(m_{V/U}) \leq \dim U + \dim(V/U) = \dim V. \quad \square$$

Corollary 3.2. *The characteristic polynomial (in the sense of this paper) of a matrix in $M_n(F)$ has degree at most n.*

Proof. Lemma 3.1 applies in particular to an element γ of $M_n(K)$ considered as a linear transformation of K^n. Substitution gives the corollary. □

Fix an algebraic closure $\overline{F}$ of F. For a in $M_n(F)$, we call λ in $\overline{F}$ an *eigenvalue* of a if the kernel of $(\lambda I_n - a)$ is nonzero. Let U_λ denote the corresponding generalized eigenspace; i.e., U_λ is the set of vectors v in $\overline{F}^n$ that belong to the kernel of $(\lambda I_n - a)^r$ for some natural number r. The *multiplicity* $m(\lambda)$ of an eigenvalue λ is $\dim_{\overline{F}} U_\lambda$.

Example 3.3 (Upper-triangular matrices). The multiplicity of an eigenvalue λ of an upper-triangular matrix b is the number of times λ appears as a diagonal entry of b. This can be proved directly from the definitions (see, for example, [4, Theorem 8.10]).

Theorem 3.4. *For a in $M_n(F)$ the characteristic polynomial of a (as in Definition 2.4) factors in $\overline{F}[x]$ as*

$$(x-\lambda_1)^{m(\lambda_1)}(x-\lambda_2)^{m(\lambda_2)}\cdots(x-\lambda_k)^{m(\lambda_k)}, \tag{3.5}$$

where $\lambda_1, \lambda_2, \dots, \lambda_k$ are the distinct eigenvalues of a.

Proof. Let b be an upper-triangular matrix, and let B denote the subalgebra of $M_n(F)$ consisting of all upper-triangular matrices. The characteristic polynomial $\mathrm{chpoly}_{b,B/F}$ of b as an element of B was computed in Example 2.6: it is given by (3.5) as observed in Example 3.3. According to Proposition 2.11(5), $\mathrm{chpoly}_{b,B/F}$ divides the characteristic polynomial $\mathrm{chpoly}_{b,M_n(F)/F}$ of b as an element of $M_n(F)$. Since both polynomials are monic and have degree n (Example 2.6, Corollary 3.2), the theorem holds for upper-triangular matrices.

Since the characteristic polynomial of the given matrix a is unchanged under scalar extension, we may assume that F is algebraically closed, i.e., $F = \overline{F}$. Here we need to invoke one somewhat sophisticated result from linear algebra: since F is algebraically closed, a is similar to an upper-triangular matrix b [3, Theorem 6.2]. But the theorem holds for b by the preceding paragraph. Since similarity changes neither the characteristic polynomial (Proposition 2.11(2)) nor the eigenvalues, the theorem holds for a. $\square$

In [3], Axler develops many of the typical properties of matrices (e.g., the existence of eigenvalues and the decomposition with respect to generalized eigenspaces) over an algebraically closed field without use of the determinant. For example, in section 5 of that paper he defines the characteristic polynomial to be exactly the product displayed in the theorem. Logically, one could replace his section 5 with this paper.

Corollary 3.6. *For a in $M_n(F)$ the minimal polynomial* $\mathrm{minpoly}_a$ *and the characteristic polynomial* chpoly_a *have the same irreducible factors in $F[x]$.*

Proof. Irreducible polynomials in $F[x]$ are determined (up to scalar factors) by their roots in an algebraic closure $\overline{F}$. Thus we may assume that F is algebraically closed.

As in the proof of Theorem 3.4, a is similar to an upper-triangular matrix. Since conjugation changes neither the characteristic nor the minimal polynomial, we may assume that a is upper-triangular.

By Theorem 3.4, every irreducible factor of the characteristic polynomial is of the form $(x-\lambda_i)$ where λ_i is a diagonal entry in a, say $\lambda_i = a_{ii}$. The (i,i)-entry of $\mathrm{minpoly}_a(a) = 0$ is 0, but it is also $\mathrm{minpoly}_a(\lambda_i)$. Therefore, $(x-\lambda_i)$ divides $\mathrm{minpoly}_a$.

Since $\mathrm{minpoly}_a$ divides chpoly_a by the Cayley-Hamilton theorem (Proposition 2.11(1)), the corollary is proved. $\square$

Proposition 3.7. *For a and a' in $M_n(F)$, the following statements are true:*

(1) *a is invertible if and only if $c_n(a) \neq 0$.*

(2) (Jacobi formula) $c_n(a) = \sum_{\sigma \in S_n} (\operatorname{sign} \sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$.

(3) $c_n(aa') = c_n(a)c_n(a')$.

(4) $c_1(a) = a_{11} + a_{22} + \cdots + a_{nn}$.

Proof. For (1), we note that the matrix a is invertible if and only if the kernel of a is zero [4, Proposition 3.17], which happens if and only if 0 is not an eigenvalue of A. By Theorem 3.4, this is true if and only if $c_n(a) \neq 0$.

We now follow [3, sec. 9]. Write $d(a)$ for the right-hand side of (2). A straightforward rearrangement of terms as in [6, p. 179] or [4, Theorem 10.31] shows that $d(aa') = d(a)d(a')$. Therefore, (2) implies (3).

We now prove (2). Suppose first that a is upper-triangular. Then both sides of (2) are just the product of the diagonal entries of a, hence (2) holds in this case.

Now consider the general case. Since $c_n(a)$ and $d(a)$ are unchanged if we enlarge our base field, we may assume that F is algebraically closed, in which event a is similar to an upper-triangular matrix b, i.e., $b = gag^{-1}$ for some g in $M_n(F)$. Then

$$c_n(b) = d(b) = d(gag^{-1}) = d(g)d(ag^{-1}) = d(ag^{-1})d(g) = d(ag^{-1}g) = d(a).$$

Since $c_n(a) = c_n(b)$ by Proposition 2.11(2), we have proved (2).

As to (4), by Proposition 2.11(4) there exist β_{ij} in F such that

$$c_1(a) = \sum_i \sum_j \beta_{ij} a_{ij}$$

for every matrix a. In the case where a is upper-triangular, the proof of Theorem 3.4 shows that the characteristic polynomial of a is as in Example 2.6. In particular, (4) holds. This gives

$$\beta_{ij} = \begin{cases} 0 & \text{if } i < j, \\ 1 & \text{if } i = j. \end{cases}$$

A symmetric argument with lower-triangular matrices shows that $\beta_{ij} = 0$ when $i > j$. $\square$

Definition 3.8. For a an element of a finite-dimensional F-algebra A, we define the *trace of* a to be

$$\operatorname{tr}_{A/F}(a) = c_1(a)$$

and the *determinant of* a to be

$$\det_{A/F}(a) = c_n(a).$$

If there is no danger of ambiguity, we write simply $\det_A$ or $\det$ instead of $\det_{A/F}$ and similarly for the trace. The trace $\mathrm{tr}_{M_n(F)}$ and determinant $\det_{M_n(F)}$ are the usual trace and determinant from linear algebra by Proposition 3.7—or by Theorem 3.4, depending on your point of view.

4 QUATERNIONS.

Example 4.1. Hamilton's quaternions—usually denoted by $\mathbb{H}$—are defined to be the ring constructed by taking the complex numbers $\mathbb{C}$ and adjoining an element j such that $j^2 = -1$, j commutes with real numbers, and $ij = -ji$. It has $\mathbb{R}$-basis 1, i, j, and k, where $k = ij$. A lot of interesting information about the quaternions can be found in [12, chap. 7]. (Note that $\mathbb{H}$ is an $\mathbb{R}$-algebra but not a $\mathbb{C}$-algebra, since $\mathbb{C}$ is not in the center of $\mathbb{H}$.)

Set ϕ to be the $\mathbb{R}$-linear map $\mathbb{H} \to M_2(\mathbb{C})$ defined by

$$1 \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i \mapsto \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad j \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad k \mapsto \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

This extends to an isomorphism $\phi\colon \mathbb{C} \otimes \mathbb{H} \to M_2(\mathbb{C})$ (as $\mathbb{C}$-algebras).

Since the characteristic polynomial is unchanged when we enlarge our base field (Lemma 2.9), the characteristic polynomial of a quaternion q has degree 2, just like a matrix in $M_2(\mathbb{C})$. That is,

$$\mathrm{chpoly}_q(x) = x^2 - \mathrm{tr}_{\mathbb{H}}(q)\, x + \det_{\mathbb{H}}(q).$$

We now determine $\mathrm{tr}_{\mathbb{H}}$ and $\det_{\mathbb{H}}$. Every quaternion can be written as $q = r + si + uj + vk = z + wj$ for some real numbers r, s, u, and v and complex numbers $z = r + si$ and $w = u + vi$. We have

$$\phi(q) = \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}.$$

Because the characteristic polynomial is unchanged when we enlarge our base field, we find that

$$\mathrm{tr}_{\mathbb{H}}(q) = \mathrm{tr}_{M_2(\mathbb{C})}(\phi(q)) = z + \overline{z} = 2r$$

and

$$\det_{\mathbb{H}}(q) = \det_{M_2(\mathbb{C})}(\phi(q)) = z\overline{z} + w\overline{w} = r^2 + s^2 + u^2 + v^2.$$

Example 4.2 (Matrices over the quaternions). Write $M_2(\mathbb{H})$ for the set of 2-by-2 matrices with entries in $\mathbb{H}$. The obvious addition and multiplication make it into a 16-dimensional $\mathbb{R}$-algebra.

In [8], Cayley defined a determinant $\mathrm{Cdet}\colon M_2(\mathbb{H}) \to \mathbb{H}$ by

$$\mathrm{Cdet} \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} = q_{11}q_{22} - q_{21}q_{12}.$$

He noted that his determinant had some unsavory properties. For example,

$$\operatorname{Cdet}\begin{pmatrix} q & q' \\ q & q' \end{pmatrix} = 0 \qquad (q, q' \in \mathbb{H}),$$

whereas

$$\operatorname{Cdet}\begin{pmatrix} i & i \\ j & j \end{pmatrix} = ij - ji = 2ij \neq 0.$$

We can constrast this with the determinant that we have just defined. Just as for $\mathbb{H}$, there is an isomorphism $\phi_2 \colon \mathbb{C} \otimes M_2(\mathbb{H}) \to M_4(\mathbb{C})$ such that

$$\phi_2 \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} = \begin{pmatrix} \phi(q_{11}) & \phi(q_{12}) \\ \phi(q_{21}) & \phi(q_{22}) \end{pmatrix}.$$

(Recall that $\phi(q)$ is a 2-by-2 complex matrix for every q in $\mathbb{H}$.) If m in $M_2(\mathbb{H})$ has a repeated row or column, then so does $\phi_2(m)$, hence

$$\det_{M_2(\mathbb{H})}(m) = \det_{M_4(\mathbb{C})}(\phi_2(m)) = 0.$$

More generally, our trace and determinant have the nice properties of the usual trace and determinant for matrices as given in Proposition 2.11.

For a more comprehensive discussion of various types of determinants for $M_2(\mathbb{H})$, see [2]. Aslaksen refers to our $\det_{M_2(\mathbb{H})}$ as the *Study determinant*.

Example 4.3 (Central simple algebras). A typical topic for a first-year graduate algebra course is Wedderburn's description of simple Artinian rings: they are isomorphic to $M_r(D)$ for D a skew field. We write F for the center of D (which is necessarily a field), and suppose that D is finite-dimensional over F. Such an algebra $M_r(D)$ is called *central simple*. We have just seen two examples of these, with $F = \mathbb{R}$, $D = \mathbb{H}$, and $r = 1, 2$.

The trace $\operatorname{tr}_{M_r(D)}$ and determinant $\det_{M_r(D)}$ are called the *reduced trace* and *reduced norm*, respectively. They are usually constructed by "Galois descent" as in [10, p. 145] or [24], using facts special to central simple algebras. Here we get them as a consequence of the general theory of the characteristic polynomial from section 2.

5 MORE PROPERTIES OF THE CHARACTERISTIC POLYNOMIAL.

Here we discuss the example of finite-dimensional field extensions (Example 5.2). We also establish some additional nice properties of the characteristic polynomial (Corollary 5.3).

Write $\operatorname{End}_F(A)$ for the set of F-linear maps $A \to A$. It is an F-algebra, with function composition as multiplication. It is isomorphic to $M_m(F)$.

For a in A, write L_a for the element of $\operatorname{End}_F(A)$ defined by

$$L_a(b) = ab \qquad (b \in A).$$

The map $a \mapsto L_a$ defines an F-algebra homomorphism called the *left regular representation* of A. This homomorphism is injective: if $L_a = 0$, then $L_a(a') = 0$ for all a' in A, whence $0 = L_a(1_A) = a \cdot 1_A = a$.

Proposition 5.1. *Let a be an element in a finite-dimensional F-algebra A. The minimal polynomial* $\mathrm{minpoly}_{a/F}$ *divides the characteristic polynomial* $\mathrm{chpoly}_{a,A/F}$ *of a, which in turn divides the characteristic polynomial* $\mathrm{chpoly}_{L_a,\mathrm{End}_F(A)/F}$*, all in $F[x]$. All three polynomials have the same irreducible factors in $F[x]$.*

Proof. Since $\mathrm{chpoly}_a(a) = 0$ by the Cayley-Hamilton theorem (Proposition 2.11(1)), the minimal polynomial of a divides the characteristic polynomial chpoly_a. Because the left regular representation is injective, we see that chpoly_a divides chpoly_{L_a} by 2.11(5).

We are reduced to showing that chpoly_{L_a} and $\mathrm{minpoly}_a$ have the same irreducible factors. Now the left regular representation is injective, so a and L_a have the same minimal polynomials. That is, we need show only that chpoly_{L_a} and $\mathrm{minpoly}_{L_a}$ have the same irreducible factors. Since $\mathrm{End}_F(A)$ is isomorphic to $M_m(F)$ for $m = \dim_F A$, an appeal to Corollary 3.6 completes the proof. □

The proposition gives us the power to handle another example.

Example 5.2 (Finite-dimensional field extensions). Let A be an extension field of F of finite dimension m. Every element a of A corresponds to an element L_a in $\mathrm{End}_F(A) \cong M_m(F)$, and the characteristic polynomial of a divides the characteristic polynomial of L_a by Proposition 5.1. The trace and norm of a are usually defined to be the trace and determinant of L_a.

If A is separable over F, then by the Theorem of the Primitive Element $A = F[\theta]$ for some θ in A. The minimal polynomial of θ has degree m. Since it divides the characteristic polynomial of L_θ by Proposition 5.1 and the latter polynomial has degree m, we find that

$$\mathrm{chpoly}_{\theta,A/F} = \mathrm{chpoly}_{L_\theta}.$$

The characteristic polynomials chpoly_a have the same degree for all a in A. Accordingly,

$$\mathrm{chpoly}_a = \mathrm{chpoly}_{L_a} \qquad (a \in A).$$

In particular, for finite separable field extensions, our trace and determinant agree with the usual trace and norm.

If A is not separable over F, there can be some disagreement. For example, let $F = \mathbb{F}_2(u,v)$ be the field of rational functions in indeterminates u and v, where $\mathbb{F}_2$ is the field with two elements. The field $A = F(\sqrt{u}, \sqrt{v})$ is a purely inseparable extension of dimension 4 over F, with basis 1, $\sqrt{u}$, $\sqrt{v}$, and $\sqrt{uv}$. The generic element

$$\gamma = t_1 \cdot 1 + t_2\sqrt{u} + t_3\sqrt{v} + t_4\sqrt{uv}$$

has minimal polynomial

$$\gamma^2 - (t_1^2 + t_2^2 u + t_3^2 v + t_4^2 uv).$$

Here the characteristic polynomial of each element a in A divides but is not equal to the characteristic polynomial of L_a.

Proposition 5.1 also allows us to prove that many nice properties of the characteristic polynomial of a matrix hold as well for characteristic polynomials of elements of A.

Recall (Remark 2.5) that the characteristic polynomial chpoly_a has the same degree for every a in A. We say that A *has degree* n if chpoly_a has degree n for each a in A.

Corollary 5.3. *Let A be a finite-dimensional F-algebra of degree n. For a in A, the following assertions hold:*

(1) $c_i(1_A) = \binom{n}{i}$. *In particular,* $\mathrm{tr}_A(1_A) = n$ *and* $\det_A(1_A) = 1$.

(2) *a is invertible if and only if* $\det_A(a) \neq 0$.

(3) *a is nilpotent if and only if* $\mathrm{chpoly}_a = x^n$.

Proof. For (1) we observe that the minimal polynomial of 1_A is $x - 1$. Proposition 5.1 thus gives

$$\mathrm{chpoly}_{1_A} = (x-1)^n.$$

To treat the ($\Leftarrow$) direction of (2), we first define the *adjoint* of a, denoted by $\mathrm{adj}\, a$, as follows:

$$\mathrm{adj}\, a = (-1)^{n+1}[a^{n-1} - c_1(a)a^{n-2} + \cdots + (-1)^{n-1}c_{n-1}(a)].$$

Then

$$a \cdot \mathrm{adj}\, a = (-1)^{n+1}[\mathrm{chpoly}_a(a) - (-1)^n c_n(a)] = \det_A(a)1_A.$$

This shows that a is invertible when $\det_A(a)$ is not zero. For the converse, assume that a is invertible. Then L_a is invertible with inverse $L_{a^{-1}}$. By Proposition 3.7(1), the constant term $\det_{\mathrm{End}_F(A)}(L_a)$ of chpoly_{L_a} is not zero. Because $\mathrm{chpoly}_{a,A/F}$ and $\mathrm{chpoly}_{L_a,\mathrm{End}_F(A)/F}$ have the same irreducible factors in $F[x]$, the constant term $\det_A(a)$ of $\mathrm{chpoly}_{a,A/F}$ is not zero.

Finally, we consider (3). The element a is nilpotent if and only if it satisfies the polynomial x^r for some natural number r, which is the case if and only if a has minimal polynomial x^p for some natural number p. Since the minimal polynomial and characteristic polynomial have the same irreducible factors (Proposition 5.1), this holds if and only if the characteristic polynomial of a is x^n. □

Remark 5.4. One might be tempted to accept the traditional definition of characteristic polynomial for matrices, and then define the characteristic polynomial of a in A as $\mathrm{chpoly}_{L_a,\mathrm{End}_F(A)/F}$. But this definition is unsatisfactory for the quaternions, because their traditional characteristic polynomial—derived in Example 4.1—has degree two, but chpoly_{L_a} would have degree four. And there is another problem: there is no strong mathematical reason to prefer the left regular representation over the right regular representation (defined in the obvious manner as $a \mapsto R_a$), and the characteristic polynomials of L_a and R_a may differ. Adrian Wadsworth points out that the upper-triangular matrices from Example 2.6 provide an example of this difficulty. In particular, the generic element γ has

$$\mathrm{chpoly}_{L_a}(\gamma) = (x - t_1)^n (x - t_2)^{n-1} \cdots (x - t_{n-1})^2 (x - t_n)$$

and

$$\mathrm{chpoly}_{R_a}(\gamma) = (x - t_1)(x - t_2)^2 \cdots (x - t_{n-1})^{n-1} (x - t_n)^n.$$

6 THE PRODUCT FORMULA FOR DETERMINANTS.

In this section, we prove that the usual product formula for determinants of matrices holds for an arbitrary finite-dimensional F-algebra A:

Theorem 6.1. *For every a and a' in A, it is true that*

$$\det_A(aa') = \det_A(a) \det_A(a').$$

We postpone the proof until the end of the section. A more cryptic—but also more powerful—proof can be found in [19, chap. 6, sec. 5].

Remark 6.2. Recall that $\det_A$ is both a function $A \to F$ and an element of $R = F[t_1, \ldots, t_m]$. To evaluate the function $\det_A$ at a in A, we write a in terms of the basis $a_1, a_2, \ldots, a_m$ as $a = \sum_i \alpha_i a_i$ and make the substitution $t_i \mapsto \alpha_i$ in the polynomial $\det_A$. In this manner, we may view each polynomial in R as a function $A \to F$.

Lemma 6.3. *Suppose that F is infinite. Let f in R be such that $f(1_A) = 1$ and $f(aa') = f(a)\, f(a')$ for all a and a' in A. If a member g of R divides f and has $g(1_A) = 1$, then $g(aa') = g(a)\, g(a')$ for all a and a' in A.*

Proof. Since $f(1_A) = g(1_A) = 1$, the polynomials f and g are nonzero elements of R. Suppose that f or g belongs to the group $F^\times$ of units in R. Then g is a unit, hence $g = 1$ and the lemma holds.

We may assume that f and g are nonzero nonunits. Since R is a unique factorization domain, we can write $f = f_1 f_2 \cdots f_r$, where each factor f_k is irreducible in R. After multiplying the f_k by elements of $F^\times$ if necessary, we may assume that $f_k(1_A) = 1$ for every k. The fact that g divides f and $g(1_A) = 1$ means that g is a product $g = \prod_{\ell \in L} f_\ell$ for some subset L of $\{1, \ldots, n\}$. To prove the lemma, it suffices to prove that $f_k(aa') = f_k(a) f_k(a')$ for $k = 1, \ldots, n$ and for all a, a' in A.

Set $R' = F[u_1, \ldots, u_m, v_1, \ldots, v_m]$ for independent indeterminates $u_1, \ldots, u_m$, $v_1, \ldots, v_m$. Define elements μ and ν of $R' \otimes A$ by

$$\mu = \sum_i u_i a_i, \quad \nu = \sum_j v_j a_j.$$

As in Remark 6.2, every element of R defines a map $R' \otimes A \to R'$. For example, to find $f(\mu)$ we take f and replace t_i with u_i. Similarly, we get $f(\nu)$ via the substitution $t_j \mapsto v_j$.

We claim that $f(\mu\nu) = f(\mu)f(\nu)$. Let $\delta = f(\mu\nu) - f(\mu)f(\nu)$ in R'. Every element of R' defines a function $F^{2m} \to F$ by plugging in for the u_i and v_j. Making the substitutions $u_i \mapsto \alpha_i$ and $v_j \mapsto \beta_j$ in δ, where α_i and β_j belong to F, we obtain

$$f(ab) - f(a)f(b)$$

for $a = \sum \alpha_i a_i$ and $b = \sum \beta_j a_j$, which is 0 by hypothesis. That is, δ gives the map $F^{2m} \to F$ that is identically 0. Since F is infinite, $\delta = 0$. (This can be proved by induction on the number of variables appearing in the polynomial δ, using that fact that a nonzero polynomial in one variable has only finitely many roots; see [21, Proposition 1.3a].) This establishes the claim.

Thus

$$\prod_{k=1}^{r} f_k(\mu\nu) = f(\mu\nu) = f(\mu)f(\nu) = \prod_{k=1}^{r} f_k(\mu)f_k(\nu)$$

in R'. Recall that we obtain $f_k(\mu)$ and $f_k(\nu)$ by substituting one set of indeterminates for another in f_k. Hence, since f_k is prime in R, the polynomials $f_k(\mu)$ and $f_k(\nu)$ are prime in R' for all k. Therefore the prime factorization of $f_k(\mu\nu)$ in the unique factorization domain R' is a product of $f_p(\mu)$ and $f_q(\nu)$ for some ps and qs. Substituting for the v_j so that ν is sent to 1_A has the effect that

$$f_k(\mu\nu) \mapsto f_k(\mu), \quad f_p(\mu) \mapsto f_p(\mu), \qquad f_q(\nu) \mapsto f_q(1_A) = 1$$

for all p and q. Hence the only irreducible factor of $f_k(\mu\nu)$ amongst the $f_p(\mu)$ is $f_k(\mu)$. Similarly, substituting for the u_i so that $\mu \mapsto 1_A$, we have $f_k(\mu\nu) \mapsto f_k(\nu)$. It follows that

$$f_k(\mu\nu) = f_k(\mu)f_k(\nu) \qquad (k = 1, \dots, n).$$

By doing the substitution $\mu \mapsto a$ and $\nu \mapsto a'$, we obtain $f_k(aa') = f_k(a)f_k(a')$ for all a and a' in A. □

Proof of Theorem 6.1. Since the determinant of an element of A is unchanged when we enlarge the base field (Lemma 2.9), we may assume that F is infinite. Let $\gamma = \sum t_i a_i$ be a generic element of A as in section 2, and consider the element L_γ in $\mathrm{End}_K(K \otimes A)$.

The minimal polynomial $\mathrm{minpoly}_\gamma$ divides $\mathrm{chpoly}_{L_\gamma, \mathrm{End}_K(K\otimes A)/K}$ in $K[x]$ by Proposition 5.1. By Lemma 2.3, both polynomials actually lie in $R[x]$ for $R = F[t_1, \dots, t_m]$, so $\mathrm{minpoly}_\gamma$ divides $\mathrm{chpoly}_{L_\gamma}$ in $R[x]$ by Gauss's lemma. Substituting 0 for x defines a surjection $R[x] \to R$ that sends $\mathrm{minpoly}_\gamma$ to $(-1)^n \det_A$ for $n = \deg(\mathrm{minpoly}_\gamma)$ and $\mathrm{chpoly}_{L_\gamma}$ to the function $a \mapsto (-1)^m \det_{\mathrm{End}_F(A)}(L_a)$. Since substitution is a ring homomorphism, $\det_A$ divides $a \mapsto \det_{\mathrm{End}_F(A)}(L_a)$ in R.

We have

$$\det_{\mathrm{End}_F(A)}(L_{aa'}) = \det_{\mathrm{End}_F(A)}(L_a L_{a'}) = \det_{\mathrm{End}_F(A)}(L_a)\ \det_{\mathrm{End}_F(A)}(L_{a'}),$$

where the last equality comes from Proposition 3.7(3) since $\mathrm{End}_F(A)$ is isomorphic to $M_m(F)$. Note that

$$\det_{\mathrm{End}_F(A)}(L_{1_A}) = \det_{M_m(F)}(1_{M_m(F)}) = 1$$

and $\det_A(1_A) = 1$ by Proposition 5.3(1). Because F is infinite, Lemma 6.3 gives

$$\det_A(aa') = \det_A(a)\ \det_A(a')$$

as desired. □

To summarize: We defined the characteristic polynomial of an element in a finite-dimensional F-algebra in section 2. We defined the determinant $\det_A$ to be the constant term of this polynomial (Definition 3.8). In the case $A = M_n(F)$, we found that $\det_{M_n(F)}$ is given by the Jacobi formula (Proposition 3.7(2)), hence the product formula holds for $\det_{M_n(F)}$ (Proposition 3.7(3)). Finally, we used Proposition 5.1 to prove the product formula for $\det_A$ (Theorem 6.1).

7 MISCELLANEOUS REMARKS.

This section is a survey of related results. It is necessarily briefer and more technical than the rest of the paper.

The "usual" definition of the characteristic polynomial.

For a in A, we claim that the formula

$$\mathrm{chpoly}_a = \det(x \cdot 1_A - a) \tag{7.1}$$

holds in $F[x]$. But what does the expression "$\det(x \cdot 1_A - a)$" mean? For matrices, the determinant is given by the Jacobi formula involving the entries of a matrix, a formula that makes sense whether the entries are elements of F or polynomials in $F[x]$. The same reasoning holds for our more general notion of determinant.

One way to prove (7.1) is as follows. First prove it for the algebra A of upper-triangular matrices from Example 2.6. Then prove (7.1) for $A = M_n(F)$ by reducing to the upper-triangular case as in the proof of Theorem 3.4. This implies (7.1) for general A by the arguments in [19, p. 225].

The characteristic polynomial of a product.

If a and b are n-by-n matrices, it is well known that the products ab and ba have the same characteristic polynomials. We can prove this statement directly for general A. Let c_i be any of the coefficients of the characteristic polynomial as in (2.10). Then $c_i \colon A \to F$, and we want to show that $c_i(ab) = c_i(ba)$ for all i and for all a and b in A. Fix i.

First assume that b is invertible. Then

$$c_i(ab) = c_i(b(ab)b^{-1}) = c_i(ba)$$

by Proposition 2.11(2). Now the set U of pairs (a, b) in $A \times A$ such that b is invertible is a nonempty open subset of $A \times A$ in the Zariski topology by Corollary 5.3. The map $A \times A \to F$ defined by

$$(a, b) \mapsto c_i(ab) - c_i(ba)$$

is given by a polynomial in the coordinates of a and b and is zero on U, hence is identically zero on $A \times A$. This proves that $c_i(ab) = c_i(ba)$ for all a and b in A. (This argument also works for $A = M_n(F)$ with the traditional definition of the characteristic polynomial.)

Octonions, Jordan algebras, etc.

The results of sections 2 and 5 hold for many nonassociative algebras A. Specifically, suppose that A is an F-algebra in the sense of Definition 2.1, except that the multiplication need not be associative. (We still require that elements of F associate with every pair of elements of A, as in (2.2).) We say that A is *power-associative* if the value of a product $aa \cdots a$ does not depend on where one puts the parentheses. We say that A is *strictly power-associative* if $K \otimes A$ is power-associative for every field extension K of F. For example, the octonions are strictly power-associative. So is every Jordan algebra over a field of characteristic different from 2. When F is infinite, the algebra A is strictly power-associative if and only if it is power-associative.

The statements of section 2, Proposition 5.1, and parts (1) and (3) of Corollary 5.3 hold verbatim if we merely require that A be strictly power-associative and finite-dimensional. The proofs go through with essentially no change. In section 5, the left regular representation $A \to \mathrm{End}_F(A)$ is just an injection of vector spaces. But the restriction to the subring generated by F and a does preserve multiplication for every a in A, and this is enough to prove Proposition 5.1 and parts (1) and (3) of Corollary 5.3. The statement of Proposition 5.3(2) has to be modified slightly (see [19, Corollary 2, p. 227] for a correct version).

The product formula (Theorem 6.1) does not hold in this level of generality. For example, it fails when A is a 27-dimensional exceptional Jordan algebra.[4]

The Pfaffian.

Suppose that the characteristic of F is not 2. Write $\mathrm{Skew}_n(F)$ for the vector space of *skew-symmetric* n-by-n matrices, i.e., matrices a such that $a^t = -a$. If n is odd, then $\det(-a) = -\det(a)$ by Proposition 2.11(3). Since $\det(a^t) = \det(a)$ by Proposition 2.11(2), all n-by-n skew-symmetric matrices have determinant 0.

[4] These algebras correspond to Lie algebras of type F_4 in the Killing-Cartan classification. Their determinants are 27-dimensional cubic forms associated with Lie algebras of type E_6.

Suppose next that n is even. There is a polynomial map $\mathrm{Pf}\colon \mathrm{Skew}_n(F) \to F$ such that

$$(\mathrm{Pf}(a))^2 = \det_{M_n(F)}(a) \tag{7.2}$$

called the *Pfaffian* (see, for example [22, chap. 15, sec. 9]). It seems a bit mysterious! (Note that equation (7.2) determines the Pfaffian only up to sign. Classically one chooses an invertible skew-symmetric matrix S with $\det(S) = 1$ and fixes the sign of Pf so that $\mathrm{Pf}(S) = 1$.)

In fact, the Pfaffian exists as a consequence of the characteristic polynomial as defined in section 2. Fix a matrix S as in the preceding paragraph and define a multiplication $\cdot$ on $\mathrm{Skew}_n(F)$ by

$$a \cdot b = \frac{1}{2}(aS^{-1}b + bS^{-1}a).$$

This makes $\mathrm{Skew}_n(F)$ into a Jordan F-algebra with identity element S. These algebras arise naturally in the classification of central simple Jordan algebras.[5] They are associated with nondegenerate skew-symmetric bilinear forms on F^n.

As described in the previous subsection, the theory developed in section 2 delivers a characteristic polynomial and a determinant for this Jordan algebra. One finds that the determinant *is* the Pfaffian:

$$\det_{\mathrm{Skew}_n(F)}(a) = \mathrm{Pf}(a)$$

(see [19, pp. 230–232]).

Matrices over rings.

We have given a uniform construction of an analogue of the characteristic polynomial (hence the determinant) for every algebra over a field F. But the characteristic polynomial and determinant are typically defined for $M_n(R)$, where R is merely a commutative ring with 1. Our definition also gives a characteristic polynomial in that case.

First, consider $M_n(\mathbb{Z})$ as a $\mathbb{Z}$-algebra. Since $\mathbb{Z}$ is a Noetherian unique factorization domain and $M_n(\mathbb{Z})$ is a free $\mathbb{Z}$-module of finite rank, the results in section 2 hold with $A = M_n(\mathbb{Z})$ and $F = \mathbb{Z}$. (In Proposition 2.11(5), we require that B be a direct summand of A.) The argument in Lemma 2.9 gives that the characteristic polynomial of a matrix a in $M_n(\mathbb{Z})$ is the same as the characteristic polynomial of a considered as a matrix in $M_n(\mathbb{Q})$, which is the usual characteristic polynomial by Theorem 3.4.

A coefficient c_i of the characteristic polynomial is given by an element of $\mathbb{Z}[t_1, t_2, \ldots, t_m]$, and it is evaluated on a matrix a as described in Remark 6.2. For example, undergraduates are taught how to evaluate $c_n = \det$, which is given by the Jacobi formula with respect to the standard basis of $M_n(\mathbb{Z})$. In this manner, one can evaluate c_i on a matrix in $R \otimes M_n(\mathbb{Z}) = M_n(R)$ for any

[5] See [19, chap. 5, sec. 7]. They correspond to Lie algebras of type C in the Killing-Cartan classification.

commutative ring R with 1. This gives a characteristic polynomial—the usual one—for every a in $M_n(R)$.

ACKNOWLEDGMENTS. I thank Adrian Wadsworth, Eric Brussel, and the referees for their comments, which greatly improved the structure and exposition of this paper.

References

[1] A. A. Albert, *Modern Higher Algebra*, University of Chicago Press, Chicago, 1937.

[2] H. Aslaksen, Quaternionic determinants, *Math. Intelligencer* **18** (1996) 57–65.

[3] S. Axler, Down with determinants!, this MONTHLY **102** (1995) 139–154.

[4] ———, *Linear Algebra Done Right*, 2nd ed., Springer-Verlag, New York, 1997.

[5] N. Bourbaki, *Algebra I: Chapters 1–3*, Springer-Verlag, Berlin, 1998.

[6] J. G. Broida and S. G. Williamson, *A Comprehensive Introduction to Linear Algebra*, Addison-Wesley, Redwood City, CA, 1989.

[7] M. D. Burrow, The minimal polynomial of a linear transformation, this MONTHLY **80** (1973) 1129–1131.

[8] A. Cayley, On certain results relating to quaternions, *Philosophical Magazine* **26** (1845) 141–145; also in *Mathematical Papers*, vol. 1, Johnson Reprint Corp., New York, 1963, pp. 123–126.

[9] L. E. Dickson, *Algebras and Their Arithmetics*, University of Chicago Press, Chicago, 1923.

[10] P. K. Draxl, *Skew Fields*, London Math. Soc. Lecture Note Series, vol. 81, Cambridge University Press, Cambridge, 1983.

[11] D. Dummit and R. Foote, *Abstract algebra*, 2nd ed., Prentice-Hall, Upper Saddle River, NJ, 1999.

[12] H.-D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, J. Neukirch, A. Prestel, and R. Remmert, *Numbers*, Springer-Verlag, New York, 1991.

[13] J. Hannah, A geometric approach to determinants, this MONTHLY **103** (1996) 401–409.

[14] I. N. Herstein, *Topics in Algebra*, 2nd ed., Xerox College Publishing, Lexington, MA, 1975.

[15] T. Hungerford, *Algebra*, 2nd ed., Springer-Verlag, New York, 1980.

[16] C. G. J. Jacobi, De formatione et proprietatibus determinantium, *J. Reine Angew. Math.* **22** (1841) 285–318; also in *Ges. Werke*, vol. 3, Chelsea, New York, 1969, pp. 355–392.

[17] N. Jacobson, *The Theory of Rings*, American Mathematical Society, New York, 1943.

[18] ———, Generic norm of an algebra, *Osaka Math. J.* **15** (1963) 23–50; also in *Coll. Math. Papers*, vol. 2, Birkhäuser, Boston, pp. 489–515.

[19] ———, *Structure and Representations of Jordan Algebras*, American Mathematical Society, Providence, 1968.

[20] E. Knobloch, From Gauß to Weierstraß: Determinant theory and its historical evaluation, in *The Intersection of History and Mathematics*, Science Networks, vol. 15, Birkhäuser, Boston, 1994, pp. 51–66.

[21] E. Kunz, *Introduction to Commutative Algebra and Algebraic Geometry*, Birkhäuser, Boston, 1985.

[22] S. Lang, *Algebra*, revised 3rd ed., Springer-Verlag, New York, 2002.

[23] P. D. Lax, *Linear Algebra*, John Wiley & Sons, New York, 1997.

[24] R. S. Pierce, *Associative Algebras*, Springer-Verlag, New York, 1982.

[25] J. A. Rosoff, A topological proof of the Cayley-Hamilton theorem, *Missouri J. Math. Sci.* **7** (1995), 63–67.

[26] G. Scheffers, Zurückführung complexer Zahlensysteme auf typische Formen, *Math. Annalen* **39** (1891) 293–390.

[27] E. Study, Über Systeme von complexen Zahlen, *Gott. Nachr.* **9** (1889) 237–268.

[28] K. Weierstrass, *Zur Determinantentheorie*, note based on lectures given in 1886/87; posthumously in *Mathematische Werke*, vol. 3, J. Knoblauch, ed., Mayer and Müller, Berlin, 1903, pp. 271–286.

Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322
http://www.mathcs.emory.edu/~skip
skip@member.ams.org