Chapter 2

Basics of associative algebras

Before defining quaternion algebras, we begin with some general preliminaries on algebras. This is largely to fix some definitions and to help us see how quaternion algebras fit into the general picture of algebraic structures. While some of the definitions may seem unmotivated at first, I hope this will be somewhat clarified by the examples. I urge you to think carefully about the definitions. A good definition is as important (and often harder to come up with) as a good theorem.

In the interest of time, many relevant details that I think of as "general algebraic facts" are left as exercises. This is a compromise forced on me by time constraints and my goals for the course, but I hope that I have treated things in a way that these exercises are both useful for learning the material and not overly demanding. The reader wanting more details can can consult another reference such as [Rei03], [Pie82], [BO13], [GS06] or [MR03]. There is also a new book [Bre14], which I haven't looked at closely but aims to present this material in a way requiring minimal prerequisites.

Algebras are be a generalization of field extensions. For instance, if K is a number field, then it can be regarded as a vector space over $\mathbb{Q}$ with a multiplication law for the vectors that is commutative. Algebras will be vector spaces over a field F with a multiplication law defined on the vectors, which we do not assume is commutative.

2.1 Algebras over fields

Let F be a field. We say A is an **(associative, unital) algebra** over F (or, for brevity, F-algebra) if A is a ring (containing $1 = 1_A$) which is an F-vector space, such that the F-action is compatible with multiplication in A in the sense that $(x \cdot a)b = x \cdot (ab) = a(x \cdot b)$ for all $a, b \in A$, $x \in F$.

Suppose that $A \neq 0$ (not the zero vector space over F). Then the multiplicative identity $1 = 1_A \in A$ must also be nonzero. Hence we may view F as a subring of A via $x \mapsto x \cdot 1_A$. Then the compatibility condition simply reads $xab = axb$ for all $a, b \in A$, $x \in F$, which is equivalent to the condition that $xa = ax$ for all $a \in A$, $x \in F$, i.e., F is contained in the **center** $Z(A) = \{z \in A : az = za \text{ for all } a \in A\}$ of A. In other words, a nonzero algebra is a ring containing F in its center which is also an F-vector space (where the scalar multiplication agrees and commutes with ring multiplication).

We often tacitly assume our algebras are nonzero.

Sometimes we will just say A is an algebra when we do not need to specify the field F.

Exercise 2.1.1. We say B is an **(F-)subalgebra** of A if it is a subring of A containing F.
(i) Check that an F-subalgebra is an F-algebra.
(ii) Suppose $A \neq 0$. Show $Z(A)$ is a subalgebra of A.

By the **dimension** $\dim A = \dim_F A$ of A, we mean the dimension of A as a vector space over F.

Unless otherwise stated, we assume all of our algebras are finite dimensional.

Example 2.1.1. Let E/F be an extension of fields of degree n. Then E is an F-algebra of dimension n. Note E has two extra properties that arbitrary algebras do not—it is commutative ($Z(E) = E$), and every nonzero element is invertible (i.e., has a multiplicative inverse).

Now let's look at examples both of algebras with one of these extra properties and with neither of these extra properties.

First, let's give the second bonus property a name. We say an F-algebra A is a **division algebra** if, as a ring, it is a division ring, i.e., if every nonzero element of A has a multiplicative (necessarily 2-sided) inverse. Note division rings are sometimes called skew fields, as the only condition lacking to be a field is commutativity of multiplication. In fact some authors use the term field to mean division ring, e.g. in [Wei95], or *corps* in French, e.g. [Vig80]. However this is not so common nowadays (at least in my circles).

Example 2.1.2. Let E/F and K/F be field extensions of degrees n and m. Let $A = E \oplus K$, the direct sum as both an F-vector space and as a ring, so addition and multiplication are component wise. Then A is an F-algebra (see below) of dimension $m+n$; it is commutative, but not a division algebra.

The above example is a special case of the direct sum for algebras, which we introduce below.

Exercise 2.1.2. Prove that Hamilton's quaternions $\mathbb{H}$, as given in the introduction (i.e., $\mathbb{H} = \mathbb{R}[i,j,k]/\langle i^2 = j^2 = k^2 = ijk = -1\rangle$), form a noncommutative division algebra over $\mathbb{R}$.

Example 2.1.3. The ring $M_n(F)$ of $n \times n$ matrices over F is an algebra (called a matrix algebra) of dimension n^2 over F. You can check the following facts (see exercises below if you don't know them already): If $n > 1$, then $M_n(F)$ is not commutative or a division

algebra for any field F. Moreover, the center $Z(M_F(F)) = F$.

Indeed, the examples of field extensions and matrix algebras are primary motivations for the definition of algebras—the notion of an algebra is a structure that encompasses both of these objects. More generally, one has algebras of functions, such as polynomial algebras, or C^*-algebras in analysis, but these are not finite-dimensional and will not enter into our study. However, certain infinite-dimensional operator algebras called Hecke algebras play an important role in modular forms, and we should encounter them later.

Exercise 2.1.3. Prove that $M_n(F)$ is not a division algebra for any $n > 1$.

Exercise 2.1.4. Prove that $M_n(F)$ is not commutative for any $n > 1$.

Exercise 2.1.5. Prove that $Z(M_n(F)) = F$. (Hint: check what it means to commute with the matrix E_{ij} with a 1 in the ij-position and 0's elsewhere.)

Subalgebras of matrix algebras are the prototypical example of (finite-dimensional) algebras. For instance, fields occur as subalgebras of matrix algebras.

Exercise 2.1.6. Consider the subalgebra A of $M_2(\mathbb{R})$ containing all elements of the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, where $a, b \in \mathbb{R}$. Show A is isomorphic to $\mathbb{C}$, as an $\mathbb{R}$-algebra.

Of course, isomorphic as algebras means what you think it means. Formally, we say a **homomorphism** of F-algebras $\phi : A \to B$ is an F-linear map which also is a ring homomorphism. (Recall, since we are working in the category of unital rings, this means we need $\phi(1_A) = 1_B$.) Further, it is an **isomorphism** if it is bijective, i.e., it is both a ring isomorphism and a vector space isomorphism.

Exercise 2.1.7. Which of the following are algebras over F? For those that are algebras, determine their dimension and center. For those that are not, state at least one property that fails. Below, assume $n > 1$, and that the ring operations are usual matrix addition and multiplication.

(i) The set of matrices of trace 0 in $M_n(F)$;
(ii) The set of matrices of determinant 1 in $M_n(F)$;
(iii) The set of diagonal matrices in $M_n(F)$;
(iv) The set of diagonal matrices in $M_n(F)$ whose lower right coordinate is 0;
(v) The set of diagonal matrices in $M_n(F)$ whose lower right coordinate is 1;
(vi) The set of upper triangular matrices in $M_n(F)$.

Recall from algebraic number theory, that one can represent a degree n field extension K/F in the space of $n \times n$ matrices over F by choosing an F-basis and letting K act on itself by left multiplication.

We can do the same for general algebras. Namely, fix a basis $e_1, \dots, e_n$ of A (as an F-vector space). An element α defines a linear operator $L_\alpha : A \to A$ via left multiplication $x \mapsto \alpha x$. Thus we can explicitly realize A as an algebra of $n \times n$ matrices over F with respect to our chosen basis. To write down this matrix representation of A, it suffices to write down matrices for $e_1, \dots, e_n$ and use linearity. This implies the following:

Proposition 2.1.1. *An n-dimensional F-algebra A can be realized (or represented) as a subalgebra of $M_n(F)$, i.e., there is an injective F-algebra homomorphism from A into $M_n(F)$.*

This says that algebras are relatively nice, well-behaved structures, and we can't get anything too weird.

Exercise 2.1.8. Let $F = \mathbb{R}$ and take $A = \mathbb{H}$. With respect to the usual basis $\{1, i, j, k\}$, write down the matrices for i, j, k acting by left multiplication. Using this, define an explicit embedding of $\mathbb{H}$ into $M_4(\mathbb{R})$.

Of course, this process does not necessarily give an "optimal" representation of A as a subalgebra of $M_n(F)$. For instance, if A is the $n \times n$ matrix algebra to start with, this process will realize A as a subalgebra of $n^2 \times n^2$ matrices.

Exercise 2.1.9. Let F be a field and $A = M_2(F)$. With respect to the basis $e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ and $e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, determine the matrices L_{e_i} for $i = 1, 2, \dots, 4$. Determine the image of A in $M_4(F)$ under the associated embedding.

Here is a better way to represent $\mathbb{H}$ in terms of matrices. Consider the ($\mathbb{R}$-)linear map from $\mathbb{H}$ to $M_2(\mathbb{C})$ given by

$$1 \mapsto \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \quad i \mapsto \begin{pmatrix} i & \\ & -i \end{pmatrix}, \quad j \mapsto \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}, \quad k \mapsto \begin{pmatrix} & i \\ i & \end{pmatrix}. \tag{2.1.1}$$

Exercise 2.1.10. Show that (2.1.1) defines an $\mathbb{R}$-algebra isomorphism

$$\mathbb{H} \simeq \left\{ \begin{pmatrix} \alpha & -\beta \\ \overline{\beta} & \overline{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C} \right\},$$

where bar denotes complex conjugation.

In any case, these matrix representations (or realizations) of algebras allow us to associate a useful set of invariants to objects in algebras. Let me briefly explain how we can do this with our non-optimal matrix embeddings—then we will come back an improve on this in Section 2.4.

For $\alpha \in A$, define the **non-reduced characteristic polynomial** (resp. **non-reduced minimal polynomial**[1]) of α to be the characteristic polynomial (resp. minimal polynomial) of L_α. The non-reduced minimal polynomial divides the non-reduced characteristic

[1]Really one can just define the minimal polynomial in the usual way—it's the minimal degree monic polynomial over F which annihilates α. I'm just doing this for symmetry's sake.

polynomial by the Cayley–Hamilton theorem, which states this for matrices. Similarly, we define the **non-reduced norm** (resp. **non-reduced trace**) of α to be the determinant (resp. trace) of L_α. We denote the non-reduced norm of α by $N^{\mathrm{nr}}_{A/F}(\alpha)$ and the non-reduced trace of α by $\mathrm{tr}^{\mathrm{nr}}_{A/F}(\alpha)$. From linear algebra, none of these invariants depend upon the choice of the basis $e_1, \dots, e_n$, which is why we call them invariants. (Put another way, they are invariant under conjugation: e.g., $N^{\mathrm{nr}}_{A/F}(\alpha) = N^{\mathrm{nr}}_{A/F}(\beta\alpha\beta^{-1})$ for any invertible $\beta \in A$.)

Exercise 2.1.11. Consider again $F = \mathbb{R}$ and $A = \mathbb{H}$. Write $\alpha = x + yi + zj + wk$, as in the introduction. Compute $N^{\mathrm{nr}}_{\mathbb{H}/\mathbb{R}}(\alpha)$ and $\mathrm{tr}^{\mathrm{nr}}_{\mathbb{H}/\mathbb{R}}(\alpha)$.

Exercise 2.1.12. Retering to Exercise 2.1.9, compute $N^{\mathrm{nr}}_{A/F}$ and $\mathrm{tr}^{\mathrm{nr}}_{A/F}$ of $\alpha \in A = M_2(F)$. How do they compare to the usual determinant and trace $\det\alpha$ and $\mathrm{tr}\,\alpha$ on $M_2(F)$?

Then the following elementary properties follow from the corresponding properties of determinant and trace:

Lemma 2.1.2. *The non-reduced norm map $N^{\mathrm{nr}}_{A/F} : A \to F$ is multiplicative*

$$N^{\mathrm{nr}}_{A/F}(\alpha\beta) = N^{\mathrm{nr}}_{A/F}(\alpha)N_{A/F}(\beta) = N^{\mathrm{nr}}_{A/F}(\beta\alpha)$$

and the trace map $\mathrm{tr}^{\mathrm{nr}}_{A/F} : A \to F$ is additive:

$$\mathrm{tr}^{\mathrm{nr}}_{A/F}(\alpha + \beta) = \mathrm{tr}^{\mathrm{nr}}_{A/F}(\alpha) + \mathrm{tr}^{\mathrm{nr}}_{A/F}(\beta)$$

for all $\alpha, \beta \in A$.

We will prefer to work with reduced norm and trace maps, which in the case of $\mathbb{H}$ are just given by the determinant and trace maps applied to the image of the embedding of $\mathbb{H}$ into $M_2(\mathbb{C})$ explained above. For $\mathbb{H}$, the reduced norm will be the quaternary quadratic form $N(x + yi + zj + wk) = x^2 + y^2 + z^2 + w^2$ described in the introduction.

Reduced norm and trace maps will be introduced in Section 2.4, which will generalize the embedding of $\mathbb{H}$ into $M_2(\mathbb{C})$ to more general algebras.

Direct sums and tensor products

We've seen some examples of algebras and know that any algebras can be constructed using matrices from Proposition 2.1.1. Now we describe some basic ways to construct new algebras from old ones. Since an F-algebra is a ring which is an F-module (vector space), we can try to extend the methods we know for constructing modules to the setting of algebras.

Proposition 2.1.3. *Let A, B be F-algebras. Then $A \oplus B$ and $A \otimes B = A \otimes_F B$ are also F-algebras of dimensions $\dim A + \dim B$ and $\dim A \cdot \dim B$, respectively.*

Proposition 2.1.4. *Let A be an F-algebra and K/F a field extension of possibly infinite degree. Then the* **extension of scalars** *$A \otimes_F K$ is a K-algebra of dimension $\dim_F A$.*

Here the direct sum of algebras, a priori just an F-module (vector space), is made into a ring with component-wise multiplication. Similarly, the tensor products, a priori just modules, can be made into rings—e.g., for $A \otimes B$, we define $(a \otimes b)(c \otimes d) = ac \otimes bd$ and extend this multiplication to all of $A \otimes B$ linearly. The dimension statements fall out of the dimension statements direct sum and tensor products of vector spaces, and the only thing to check is that these definitions of multiplication are valid and compatible with the vector space structure as required in the definition of an F-algebra. This is easy and I leave it to you.

Exercise 2.1.13. Prove the above two propositions.

Since we can embed A and B into matrix algebras, we can try to understand what direct sums and tensor products do at the level of matrices. Say $A \subset M_n(F)$ and $B \subset M_m(F)$ are subalgebras. Then it is easy to see that

$$A \oplus B \simeq \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} : a \in A, b \in B \right\} \subset M_{n+m}(F).$$

The tensor product can also be understood in terms of matrices. Say $a = (a_{ij}) \in M_n(F)$ and $b = b_{ij} \in M_m(F)$. Then the **Kronecker product** of a and b is the block matrix

$$a \odot b = \begin{pmatrix} a_{11}b & a_{12}b & \cdots & a_{1n}b \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}b & a_{n2}b & \cdots & a_{nn}b \end{pmatrix} \in M_{nm}(F).$$

(One defines the Kronecker product for non-square matrices similarly.) Usually, the Kronecker product is denoted with $\otimes$ instead of $\odot$, because it is the matrix realization of the tensor product. This is the content of the next exercise, and after doing this exercise, you can use $\otimes$ to denote Kronecker products.

Exercise 2.1.14. Show the map $a \odot b \mapsto a \otimes b$ gives an algebra isomorphism $M_{mn}(F) \simeq M_n(F) \otimes M_m(F)$.

In particular, tensoring two matrix algebras doesn't give us a new kind of algebra. Similarly, tensoring extension fields doesn't get us much new either.

Exercise 2.1.15. Let $F/\mathbb{Q}$ and $K/\mathbb{Q}$ be two quadratic extensions in $\mathbb{C}$ Show $F \otimes_{\mathbb{Q}} K$ is isomorphic to the compositum of FK if $F \neq K$ and $F \otimes_{\mathbb{Q}} K \simeq F \oplus F$ if $F = K$.

However, the extension of scalars will be very important for us. Just like with number fields F, we will pass to local completions F_v to apply local methods. For an algebra A, we will want to consider the "local completions" $A_v = A \otimes F_v$. This is important both for classifying algebras over number fields as well as understanding the arithmetic of algebras.

Tensor products of two non-commutative algebras will also be useful to consider as a tool to studying the general structure of algebras. In particular, the following will be useful.

Exercise 2.1.16. Let A and B be F-algebras. Show $Z(A \otimes_F B) \simeq Z(A) \otimes_F Z(B)$.

Algebras of small dimension

A basic problem, of course, is the classification of n-dimensional algebras up to isomorphism. We know by Proposition 2.1.1 that they can all be realized as subalgebras of $M_n(F)$. While we are only concerned with certain types of algebras in this class where the classification problem has a simple answer, to provide a little context we first consider general algebras of small dimension.

Clearly the only F-algebra of dimension 1 is F (up to isomorphism). (This is not true if we don't require unital, as one can also define an algebra with trivial product $ab = 0$.)

In dimension 2, we can say the following.

Proposition 2.1.5. *Let A be F-algebra of dimension 2. Then $A \simeq F \oplus F$, A is a quadratic field extension of F, or A contains a nonzero nilpotent element.*

(Recall a nilpotent element α of a ring is one such that $\alpha^n = 0$ for some $n \in \mathbb{N}$.)

Proof. First we note A must be commutative. Let $\{1, \alpha\}$ be a basis for A over F. From expanding terms, one sees that $(x+y\alpha)(x'+y'\alpha) = (x'+y'\alpha)(x+y\alpha)$ for any $x, y, x', y' \in F$. Hence A is commutative, as claimed. If every nonzero element of A is invertible, then A/F must be a quadratic field extension.

Suppose some $x + y\alpha$ is nonzero but not invertible. This implies $y \neq 0$, so by a change of basis if necessary, we can assume α itself is not invertible. From Proposition 2.1.1, we know that we can realize A as a subalgebra of $M_2(F)$. Write

$$\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with respect to some basis $\{e_1, e_2\}$. In fact, since the associated linear transformation L_α has a nontrivial kernel, we can choose e_2 such that $\alpha e_2 = 0$. This means $b = d = 0$ and $\alpha e_1 = ae_1 + ce_2$. If $a = 0$, then α is nilpotent.

So assume $a \neq 0$. Then we can replace e_1 with $e_1' = e_1 + \frac{c}{a}e_2$. Then $\alpha e_1' = \alpha e_1 = ae_1'$, so with respect the basis $\{e_1', e_2\}$, we see

$$\alpha = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence we get a realization of A as the subalgebra of diagonal matrices of $M_2(F)$, which is isomorphic to $F \oplus F$. □

More precisely, the proof shows us that when A contains nilpotent elements, A is isomorphic to the matrix algebra consisting of elements of the form $\begin{pmatrix} a & 0 \\ c & a \end{pmatrix}$.

The classification for 3-dimensional algebras is much more complicated, and I don't know where/if it is completely worked out. However, it at least has been in the case $F = \mathbb{C}$—see [FP09, Table 2] for a list of 22 (not necessarily unital) 3-dimensional complex associative

algebras. Of course, over a typical field F one gets infinitely many 3-dimensional (unital) algebras just via cubic field extensions or quadratic field extensions direct sum F.

We also remark that for a finite field F, one can show the number of n-dimensional F-algebras, up to isomorphism, is at most $(\#F)^{n^3}$ (see [Pie82, Prop 1.5]).

Simple and semisimple algebras

The above discussion suggests that you can have a wide variety of algebras even in quite small dimension. Not all of them are of equal interest however. Often it suffices to consider certain nice classes of algebras, such as *simple algebras*.

The definition of a simple algebra involves ideals. In case you haven't worked with ideals in noncommutative rings, let me recall the definition, which is similar to the commutative case. Let R be a ring, commutative or not. A subset I of R is a **two-sided ideal** (of R or in R) if (i) I is a group under addition, (ii) $RI \subset I$, and (iii) $IR \subset I$. If we only require that I satisfy (i) and (ii), we say I is a **left ideal** of R. Similarly, if we only require I satisfy (i) and (iii), we say I is a **right ideal** of R. A left or right ideal is called a **one-sided ideal**. Of course two-sided ideals are both left ideals and right ideals, and in the case R is commutative, all three of these notions are equivalent so we just say ideal. In the noncommutative case, if it is clear which of these types of ideal we are discussing (or if the discussion applies to all types of ideals), we will also just use the word "ideal" rather than repeating "two-sided ideal" or "left ideal" every time.

Every nonzero ring R has at least two two-sided ideals: $\{0\}$ and R. We call these the **trivial ideals**. An ideal other than R is called a **proper ideal**. A **maximal ideal** is a maximal proper ideal (with respect to inclusion among the collection of left, right or two-sided ideals). Similarly, a **minimal ideal** is a minimal nonzero ideal. We say the ring R is **simple** if it has no nontrivial ideals, i.e., no proper nonzero ideals.

We call a nonzero algebra A **simple** if it is simple as a ring.[2] We call a nonzero algebra A **semisimple** if it is a direct sum of (necessarily finitely many by our finite-dimensionality assumption) simple algebras. These definitions coincide with the usual notions of simple and semisimple for arbitrary (unital) rings, though semisimple is often defined as having trivial *Jacobson radical*

$$\operatorname{rad} A = \bigcap \mathfrak{m},$$

where $\mathfrak{m}$ ranges over all maximal left ideals of A.

> **Example 2.1.4.** Any division algebra A (and thus any field) is simple (as a ring or algebra).
>
> To see this, suppose $\mathcal{I}$ is a nonzero proper two-sided ideal of A. Let $\alpha \in \mathcal{I}$ and $\beta \in A - \mathcal{I}$ be nonzero elements. Then $\gamma = \beta\alpha^{-1} \in A$ by the division property and $\gamma\alpha = \beta \in \mathcal{I}$ by the definition of an ideal, a contradiction. Note this argument also applies to left ideals, and with a minor modification to right ideals, so any division algebra has no nonzero proper one-sided ideals either.

[2] This is different from the notion of being simple as a module (say over itself), as we will explain in the next section.

Example 2.1.5. Our earlier example $E \oplus K$ (Example 2.1.2) of the sum of two field extensions of F is semisimple, as E and K are simple by the previous example. Note that $E \oplus K$ is not itself simple—e.g., $E \oplus 0$ and $0 \oplus K$ are nontrivial two-sided ideals.

More generally, the direct sum of two nonzero algebras cannot be simple.

Exercise 2.1.17. Show that $M_n(F)$ is simple.

Recall our quasi-classification of 2-dimensional algebras in Proposition 2.1.5. One type is simple (quadratic field extensions), one is semisimple but not simple ($F \oplus F$), and the last type is neither:

Exercise 2.1.18. Let A be a 2-dimensional algebra over F with a nonzero nilpotent element α. Show that A is neither simple nor semisimple.

For many purposes, including ours, it suffices to consider semisimple algebras. Of course, understanding semisimple algebras just boils down to understanding simple algebras, and that is what we will focus on. However, semisimple algebras do play a role in the study of simple algebras—e.g., if A is a simple F algebra, one basic question is does the semisimple algebra $F^n = \oplus_{i=1}^n F$ embed in A?

Here is one thing we can say about homomorphisms and simplicity.

Proposition 2.1.6. *Suppose $\phi : A \to B$ is an algebra homomorphism and A is simple. Then ϕ is injective. In particular,* $\dim_F A \leq \dim_F B$.

Proof. Consider $\ker \phi = \{a \in A : \phi(a) = 0\}$. This is a 2-sided ideal in A, and therefore must be either $\{0\}$ or A. However, the kernel can't be A because $\phi(1_A) = 1_B \neq 0$. The dimension statement follows because $\phi(A)$ is then a subalgebra of B of dimension $\dim_F A$. □

Exercise 2.1.19. Suppose $\phi : A \to B$ is an algebra homomorphism and B is simple. Must ϕ be surjective?

2.2 The Wedderburn structure theorem

Simple algebras have very simple (no pun intended) descriptions, though a complete classification up to isomorphism depends on the field F and this is considerably harder. In this section we will prove Wedderburn's famous theorem on the structure of simple algebras. To do this, we will apply some very simple (12% pun intended) module theory. As before A denotes an F-algebra.

We say M is a **left** (resp. **right**) **A-module** if it is a left (resp. right) module over A as a ring, i.e., M is an additive abelian group on which A acts from the left (resp. right) such that the A-action distributes and associates with addition in M and $1 \in A$ acts as the identity map. If it is understood or not important, we simply say A-module for left or

right A-module. Of course, if A is commutative, then we can think of left modules as right modules and vice versa just by writing the action of A on the other side of M. Note if A is a field, then an A-module is just an A-vector space.

If a general statement is true about left modules, then the analogue is also true for right modules, just writing things in the opposite order. One can formally show this by working with the **opposite algebra** A^{opp} where multiplication is defined in reverse order, i.e., A^{opp} is the same as A as a vector space, but the multiplication $a \cdot b$ is replaced by $b \times a := a \cdot b$. One easily checks this is also an algebra. Then left modules for A correspond to right modules for A^{opp}. Note A and A^{opp} are not isomorphic in general. One obvious case where they are isomorphic (in fact the same) is when A is commutative. Here is another case.

Exercise 2.2.1. If $A = M_n(F)$, show transpose defines an isomorphism of A^{opp} with A.

We will work more generally with matrix algebras over division rings. In case you haven't seem them, matrix algebras over division rings are defined just like matrix algebras over fields. As a vector space, $M_n(D)$ is just $D^{n \times n}$. Matrix multiplication is defined as usual, e.g., for $n = 2$,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix},$$

but now one needs to be careful about the order of terms in products like ae. In general, scalar multiplication by D is not commuative, and there are some difficulties when trying to define things like characteristic polynomials or determinants over D. However since we can realize D as a subalgebra of some $M_m(F)$, we can identify $M_n(D)$ with a subalgebra of $M_{mn}(F)$ to work with more familiar matrices over commutative fields.

The next exercise generalizes Exercises 2.1.5 and 2.2.1 to matrix algebras over division rings.

Exercise 2.2.2. For a division algebra D, show $Z(M_n(D)) \simeq Z(D)$ and $M_n(D)^{\mathrm{opp}} \simeq M_n(D^{\mathrm{opp}})$.

In light of the meta-equivalence of left and right modules, we will by default assume module without qualification means left module, but bear in mind analogous statements apply to right modules after switching A with A^{opp}.

We recall the following special case of Example 1.1.4.

Example 2.2.1. A itself is a left and right A-module. More generally, the left and right submodules of A are precisely the left and right ideals of A, so A-modules are can be viewed as a generalization of ideals.

Thus the meta-equivalence of left and right modules applies to ideals also.

Exercise 2.2.3. If A semisimple, show that any A-module M is a semisimple module, i.e., a direct sum of simple modules. (*Hint:* prove it for free A-modules and take quotients; cf. [Pie82, Sec 3.1] or [Rei03, Thm 7.1].)

Suppose M and N are (left) A-modules. Recall a map $\phi : M \to N$ is an (A-)module homomorphism if ϕ is a homomorphism of abelian groups such that $\phi(am) = a\phi(m)$ for $a \in A$. Denote by $\mathrm{Hom}_A(M, N)$ the set of all module homomorphisms from M to N and $\mathrm{End}_A(M) = \mathrm{Hom}_A(M, M)$. Under composition, $\mathrm{End}_A(M)$ is an F-algebra, which we call the **endomorphism algebra** (or **ring**) of M. Endomorphism algebras will be extremely important in understanding the structure of algebras.

Note that when $M = A$, we can also multiply on the right by A (i.e., A is both a left and right module over itself, or an A-A bimodule if you prefer). The right multiplication map $R_\alpha(m) = m\alpha \in \mathrm{End}_A(A)$ for any $\alpha \in A$—it is obviously an F-linear map from A to itself and $R_\alpha(a \cdot m) = am\alpha = a \cdot R_\alpha(m)$ for $a, m \in A$ by associativity. (On the other hand, if A is not commutative then left multiplication $L_\alpha(m) = \alpha m$ is a *right* A-module homomorphism from A to itself but not a left module homomorphism.) Thus right multiplication is a module homomorphism $R \in \mathrm{Hom}_A(A, \mathrm{End}_A(A))$ via $\alpha \mapsto R_\alpha$. It is natural to ask if this gives an embedding of A into $\mathrm{End}_A(A)$. In fact, it is an isomorphism:

Lemma 2.2.1. *As left A-modules, we have $\mathrm{End}_A(A) \simeq A$. However, as F-algebras we have $\mathrm{End}_A(A) \simeq A^{\mathrm{opp}}$.*

Proof. If $\alpha, \beta \in A$, we see $R_\alpha = R_\beta$ implies $\alpha = R_\alpha(1) = R_\beta(1) = \beta$, so the module homomorphism $R : A \to \mathrm{End}_A(A)$ is injective. To see it is surjective, let $\phi \in \mathrm{Hom}(A, \mathrm{End}_A(A))$ and use A-linearity: $\phi(x) = \phi(x \cdot 1) = x\phi(1)$ for any $x \in A$, so $\phi = R_\alpha$ where $\alpha = \phi(1)$. This prove the an isomorphism as A-modules.

For the statement about algebras, let $\times$ denote the multiplication in A^{opp}. Then, for $\alpha, \beta \in A$ (which is A^{opp} as a set) and $x \in A$,

$$R(\alpha \times \beta)x = R(\beta\alpha)x = x\beta\alpha = R(\alpha)R(\beta)x.$$

Hence $R : A^{\mathrm{opp}} \to \mathrm{End}_A(A)$ is an F-algebra morphism. By a similar argument to above, it is in fact an isomorphism. □

Recall that a simple A-module if it has no nonzero proper submodules. Our goal will be to study the structure of $A \simeq \mathrm{End}_A(A)$ (isomorphic as A-modules, but not as algebras in general) by decomposing it into simple A-modules.

When we consider A itself as a (left) A-module, there is a difference between being simple as an A-module and being simple as an F-algebra. Specifically, let A be a nonzero F-algebra, which we can also consider as a left A-module. The statement that A being simple as a ring/algebra means it has no nontrivial two-sided ideals, whereas A being simple an A-module means it has no nontrivial left ideals. Since any two-sided ideal is also a left ideal, A being simple as an A-module implies A is simple as an algebra. However, the converse is not true, as the following exercise shows.[3]

Exercise 2.2.4. Let $A = M_2(F)$ and $\mathcal{I} = \left\{ \begin{pmatrix} * & 0 \\ * & 0 \end{pmatrix} \right\}$ be the subset of A. Check that $\mathcal{I}$ is a simple left A-module. In particular, A is not simple as an A-module. However, we know $M_2(F)$ is simple as an algebra from Exercise 2.1.17.

Exercise 2.2.5. With notation as in Exercise 2.2.4, compute $\mathrm{End}_A(\mathcal{I})$.

On the other hand, we remark A being semisimple as an A-module is the same as being semisimple as an algebra (cf. Exercise 2.2.3).

It is easy to understand what the simple modules look like in terms of A. We will not actually use the next result, but I will just mention it for your edification.

Lemma 2.2.2. *Any simple left (resp. right) A-module M is isomorphic to $A/\mathfrak{m}$ (resp. $\mathfrak{m}\backslash A$) for a maximal left (resp. right) ideal $\mathfrak{m}$ of A.*

Proof. Say M is a simple left A-module. Fix a nonzero $m \in M$. Then the map $\phi(a) = am$ is a left A-module homomorphism from A (as a left A-module) to M. Since $1 \in A$, the image of ϕ contains m and thus is a nonzero left A-module. By simplicity, this means ϕ is surjective. Note that $\ker \phi$ is a left ideal, and ϕ defines an isomorphism of $A/\ker\phi$ with M. (The quotient $A/\ker\phi$, as an abelian group, is a left A-module and thus defines a quotient module.) Since $A/\ker\phi$ is has non nontrivial left ideals, $\ker\phi$ must be maximal. □

We are interested in the (say left) simple A-modules which are submodules of A, i.e., the left ideals of A which are simple as left A-modules. This means they do not properly contain any nonzero left ideals, i.e., they are minimal left ideals.

Here is a very simple (honestly, no pun intended) but very useful result.

Lemma 2.2.3 (Schur)**.** *Let M and N be (both left or right) A-modules and $\phi : M \to N$ be a nonzero homomorphism. If M is simple then ϕ is injective and if N is simple then ϕ is surjective.*

Compare this with Proposition 2.1.6 and Exercise 2.1.19 about homomorphisms of *simple algebras*, i.e., simple rings.

The above two lemmas are also true in the more general setting for modules over a ring R.

Corollary 2.2.4. *Let M be a (left or right) simple A-module. Then $\mathrm{End}_A(M)$ is a division algebra.*

Exercise 2.2.6. Prove Schur's lemma and deduce the above corollary.

[3] In algebra, there are a lot of similar notions, and one needs to keep track of their sometimes subtle differences with careful bookkeeping so as to minimize the number of false theorems one proves. *Exercise:* Find all the false theorems in these notes.

Exercise 2.2.7. Let M be a left or right A-module and $N = \bigoplus_{i=1}^n M$. Show that, $\mathrm{End}_A(N) \simeq M_n(\mathrm{End}_A(M)^{\mathrm{opp}})$ if M is a left A-module and $\mathrm{End}_A(N) \simeq M_n(\mathrm{End}_A(M))$ if M is a right A-module. In particular, if M is simple, conclude that $\mathrm{End}_A(N) \simeq M_n(D)$ for some division algebra D.

Exercise 2.2.8. Let A be a semisimple F-algebra. Show that the minimal left ideals of A are the same as the simple left A-modules which are submodules of A.

Lemma 2.2.5. *Let A be a simple F-algebra, and $\mathcal{I}$, $\mathcal{J}$ be minimal left ideals of A. Then $\mathcal{I} \simeq \mathcal{J}$ as A-modules and $\mathcal{J} \simeq \mathcal{I}\alpha$ for some $\alpha \in A$.*

Proof. Note $\mathcal{I}A$ and $\mathcal{J}A$ are nonzero two-sided ideal of A, thus $\mathcal{I}A = \mathcal{J}A = A$ by simplicity. Hence $\mathcal{I}\mathcal{J}A = \mathcal{I}A = A$ and $\mathcal{I}\mathcal{J} \neq 0$. Let $\alpha \in \mathcal{J}$ such that $\mathcal{I}\alpha \neq 0$. Then $\mathcal{I}\alpha \subset \mathcal{J}$ is a nonzero submodule of $\mathcal{J}$, thus $\mathcal{J} = \mathcal{I}\alpha$, and it is easy to see $\phi(x) = x\alpha$ is an A-module isomorphism from $\mathcal{I}$ to $\mathcal{J}$. □

Now we can prove our the main result of this section, which is a partial classification theorem for simple algebras—namely classification modulo the classification of division algebras over F.

Theorem 2.2.6 (Wedderburn)**.** *Let A be a simple F-algebra. Then $A \simeq M_n(D)$ where D is a division algebra over F. Furthermore, this n and D are uniquely determined (D up to isomorphism) by A.*

Proof. Since A is semisimple as a left A-module (Exercise 2.2.3), it decomposes as a direct sum $A \simeq \bigoplus_{i=1}^n M_i$ of simple left A-modules. Identifying A with the direct sum $\bigoplus M_i$, we can view each M_i as a minimal left ideal $\mathcal{I}_i$ of A. (In the case where $A = M_n(F)$, we can take minimal left ideals to be the subspaces $\mathcal{I}_i$ consisting of matrices which are zero off the i-th column.) By Lemma 2.2.5, all $\mathcal{I}_i$ are isomorphic (as A-modules) to a single minimal left ideal $\mathcal{I}$, i.e., $A \simeq \bigoplus_{i=1}^n \mathcal{I}$.

Then $\mathrm{End}_A(A) = \mathrm{End}_A(\oplus\mathcal{I}) = M_n(\mathrm{End}_A(\mathcal{I})^{\mathrm{opp}})$ by Exercise 2.2.7. However $\mathrm{End}_A(\mathcal{I})$ is a division algebra D by the Corollary 2.2.4, thus

$$A \simeq \mathrm{End}_A(A)^{\mathrm{opp}} \simeq M_n(\mathrm{End}_A(\mathcal{I})^{\mathrm{opp}})^{\mathrm{opp}} \simeq M_n(D^{\mathrm{opp}})^{\mathrm{opp}} \simeq M_n(D),$$

as F-algebras using Lemma 2.2.1 for the first isomorphism and Exercise 2.2.2 for the last. (We could have avoided the business with opposites if we started with $\mathcal{I}$ being a minimal *right* ideal and $D = \mathrm{End}_A(\mathcal{I})$ and worked with right A-modules; cf. Exercise 2.2.7.)

Now we want to prove uniqueness. Recall by Lemma 2.2.5, all minimal left ideals of A are isomorphic. Since the isomorphism class of $\mathcal{I}$ determines n and isomorphism class of D in the above procedure to write $A \simeq M_n(D)$, we see this procedure results in a unique n and D (up to isomorphism) given A. However, to show that $M_{n_1}(D_1) \simeq M_{n_2}(D_2)$ implies $n_1 = n_2$ and $D_1 \simeq D_2$ for division algebras D_1, D_2 requires a little more. We need to know the above procedure for $M_n(D)$ gives back n and D, rather than some n' and D'.

So fix a division algebra D and $n \in \mathbb{N}$, and set $A = M_n(D)$. Let $e \in A$ be the matrix which is 1 in the upper left entry and 0 elsewhere. It is easy to check that $\mathcal{I} = Ae$ must be a minimal left ideal of A. Note $\mathcal{I}$ is just the set of matrices in A which are 0 off the first column. Then $A = \mathcal{I}e_1 \oplus \cdots \oplus \mathcal{I}e_n$ where e_i is a matrix with a 1 in the first entry of the i-th column and 0 elsewhere (so $e_1 = e$ and $\mathcal{I}e_i$ is the set of matrices which are 0 off the i-th column). Thus we recover the n we started with in the above procedure.

Now it suffices to show that the above procedure recovers D. One can show that, as F-algebras, right multiplication defines an isomorphism of $\mathrm{End}_A(\mathcal{I})$ with D^{opp}. This is similar in spirit to Lemma 2.2.1, and the details are Exercise 2.2.9. Thus $A \simeq M_n((D^{\mathrm{opp}})^{\mathrm{opp}}) \simeq M_n(D)$ as above. □

Outside of these notes, the term "Wedderburn's theorem" sometimes refers to another theorem, like a generalization of the above result known as the Artin–Wedderburn theorem, or Wedderburn's theorem on finite division rings (that they are all fields). To specify, the above theorem (or a generalization) is sometimes called Wedderburn's structure theorem.

The endomorphism algebra $\mathrm{End}_A(\mathcal{I})$ that came up in the proof of Wedderburn's theorem is easy to understand in terms of matrices. The following exercise completes the proof of the uniqueness part of Wedderburn's theorem, and is a generalization of Exercise 2.2.5.

Exercise 2.2.9. Let D be a division algebra over F, $A = M_n(D)$, and $\mathcal{I}$ be the minimal left ideal consisting of matrices which are zero in every entry off the first column. For $\delta \in D$, show that right multiplication by $\mathrm{diag}(\delta, \dots, \delta)$ defines an endomorphism $R_\delta : \mathcal{I} \to \mathcal{I}$. Show moreover, that $\delta \mapsto R_\delta$ defines an isomorphism (as F-algebras) of D^{opp} with $\mathrm{End}_A(\mathcal{I})$.

The following exercise gives a kind of duality of endomorphism algebras, and implicitly arose in the proof of Wedderburn's theorem.

Exercise 2.2.10. Let A be a simple F-algebra and $\mathcal{I}$ a minimal left ideal. Put $D' = \mathrm{End}_A(\mathcal{I})$. Viewing $\mathcal{I}$ as a left D'-module, show $\mathrm{End}_{D'}(\mathcal{I}) \simeq A$ as F-algebras. That is, we have $\mathrm{End}_{\mathrm{End}_A(\mathcal{I})}(\mathcal{I}) \simeq A$.

The above exercises will be also important in the proof of the Skolem–Noether theorem in the next section.

Of course, we haven't really completed a classification of simple F-algebras (even modulo the classification of division algebras), because we have not actually shown that $M_n(D)$ is a simple F-algebra! We have only done this for $n = 1$ from Example 2.1.4 and when $D = F$ in Exercise 2.1.17. But it is true.

Theorem 2.2.7. *Let D be a division algebra over F. Then $M_n(D)$ is a simple F-algebra for any $n \in \mathbb{N}$.*

This is basically an exercise in linear algebra over division rings, generalizing Exercise 2.1.17, so I will leave it to you:

Exercise 2.2.11. Prove the above theorem.

2.3 Central simple algebras and Skolem–Noether

We can do one more reduction to understanding simple algebras, by understanding the centers of algebras. We say an F-algebra A is **central** if $Z(A) = F$. Then Wedderburn's structure theorem implies

Corollary 2.3.1. *Let A be a simple F-algebra. Then there exists a field extension K/F of finite degree such that A is a central simple K-algebra.*

Proof. By Wedderburn's theorem, $A \simeq M_n(D)$ for some D. Recall $Z(M_n(D)) = Z(D)$ by Exercise 2.2.2. Now $Z(D)$ must be commutative, and therefore it is a field K which must contain F (and is a finite extension by finite-dimensionality). Note the action of scalar multiplication by K makes A into a K-linear space, and thus a K-algebra. Since A is simple as an F-algebra it is simple as a K-algebra (simplicity as a ring is independent of the underlying field as an algebra) and it is central over K. □

This corollary says it suffices to understand central simple algebras, and thus by Wedderburn's theorem, to understand central division algebras.

The corollary says that we can extend the algebra structure of A to an algebra over a field contained in the center. The following exercise shows we can extend the vector space structure to fields not contained in the center.

Exercise 2.3.1. Let A be a CSA over F and K a subalgebra which is a field. Show that A is a K-vector space but not a K-algebra.

A central simple algebra is in some sense an even more basic object than a simple algebra, and is abbreviated **CSA**. Many elegant results (e.g., extension of scalars and the Skolem–Noether theorem) are true for CSAs, but not true for arbitrary simple algebras.

Proposition 2.3.2. *Let A be a CSA over F. Then:*

(i) If B is a simple F-algebra then $A \otimes_F B$ is also simple F-algebra. Hence if B is a CSA, so is $A \otimes_F B$.

(ii) If K/F a field extension, then the extension of scalars $A \otimes_F K$ is a CSA over K.

Here K/F need not be finite degree. On the other hand, even if K/F is finite degree and A is a simple algebra over F, $A \otimes_F K$ need not be a simple algebra over K. For instance, if K/F is quadratic and $A \simeq K$, then $A \otimes_F K \simeq K \otimes_F K \simeq K \oplus K$ (cf. Exercise 2.1.15), which is semisimple but not simple (as an F-algebra or as a K-algebra).

Proof. Note that it essentially suffices to show the first part of (i) by Exercise 2.1.16. (Technically to deal with infinite degree extensions K/F in (ii), one should show (i) and Exercise 2.1.16 without our default assumption that B is not be finite dimensional. This is still possible if B is *artinian*, in particular, if $B = K$ is a field—see [Rei03].) For this, roughly, one can take a nonzero 2-sided ideal $\mathcal{I}$ in $A \otimes B$ and show it must contain an element of

the form $1 \otimes b$, and in fact $1 \otimes 1$. However, I'm not convinced the proof is particularly, enlightening, and will not give it. See, e.g., [Rei03, Sec 7b], [BO13, Sec III.1] or [MR03, Prop 2.8.4] for details. □

Exercise 2.3.2. Let A be a CSA over F. Show $(\alpha \otimes \beta)x \mapsto \alpha x \beta$ defines an algebra isomorphism of $A \otimes A^{\mathrm{opp}}$ with $\mathrm{End}_F(A)$.

Recall $\mathrm{End}_F(A)$ denote the endomorphisms of A as an *F-module*. This is different than the space of F-algebra homomorphisms from A to A. When A is simple, the Skolem–Noether theorem below tells us about F-algebra homomorphisms from A to A, which are just the automorphisms together with the zero map. However the F-algebra "endomorphisms" of A do not form a ring—if you add two of them, then their sum would send 1_A to $2 \cdot 1_A$.

The main result we want to prove in this section is about embedding simple algebras into central simple algebras. For instance if $A = M_n(D)$ is central over F and K/F is a field extension, we might want to know in what ways does K embed in A. For instance, what are the embeddings of $\mathbb{C}$ into Hamilton's quaternions $\mathbb{H}$? The following theorem tells us that either K does not embed in A or it embeds in an essentially unique way (unique up to conjugation).

Theorem 2.3.3 (Skolem–Noether)**.** *Let A and B be simple F-algebras, and assume that A is central. If $\phi, \psi : B \to A$ are algebra homomorphisms, then there exists $\alpha \in A^\times$ such that $\psi(\beta) = \alpha\phi(\beta)\alpha^{-1}$ for all $\beta \in B$.*

As another consequence (taking $B = A$), this says that any algebra automorphism of a CSA A is inner, i.e., given by conjugation of an element of $A^\times$.

Proof. Recall from Proposition 2.1.6 that ϕ, ψ must be injective, hence $\phi(B)$ and $\psi(B)$ are subalgebras of A isomorphic to B.

Let M be a simple A-module and $D = \mathrm{End}_A(M)$, which is a division algebra from Corollary 2.2.4. For instance, we can take $M = \mathcal{I}$, where $\mathcal{I}$ is a minimal left ideal. Explicitly, if $A = M_n(D')$ then we can take $\mathcal{I}$ to be the ideal of matrices in A which are zero in every entry not in the first column. Then $D' \simeq D^{\mathrm{opp}}$ as F-algebras by Exercise 2.2.9.

Then we can make M a $D \otimes B = D \otimes_F B$-module in two ways:

$$(\delta \otimes \beta)m = \delta(\phi(\beta)m)$$

and

$$(\delta \otimes \beta)m = \delta(\psi(\beta)m),$$

where we extend this action to $D \otimes B$ linearly. Explicitly, if $M = \mathcal{I}$ and we identify $D = D'$ as sets, then the first action, say, is just

$$(\delta \otimes \beta)x = \phi(\beta) \cdot x \cdot \begin{pmatrix} \delta & & \\ & \ddots & \\ & & \delta \end{pmatrix},$$

where each $\cdot$ is just matrix multiplication in $M_n(D')$.

Since A is central, D is central. Therefore, by Proposition 2.3.2, we know $D \otimes B$ is a simple algebra. Hence any module over $D \otimes B$ is a direct sum of simple modules (Exercise 2.2.3). This implies any two finitely-generated modules of the same dimension (as F-vector spaces) of the $D \otimes B$ are isomorphic (cf. Lemma 2.2.5 and Exercise 2.2.7). Thus there is an isomorphism $\sigma : M \to M$ such that

$$\delta(\phi(\beta) \cdot \sigma(m)) = \sigma(\delta(\psi(\beta) \cdot m)), \quad \delta \in D,\ \beta \in B,\ m \in M.$$

Taking $\beta = 1$ shows $\sigma \in \mathrm{End}_D(M)$. But from Exercise 2.2.10, we know $\mathrm{End}_D(M) \simeq A$—explicitly, with $M = \mathcal{I}$, σ is just left matrix multiplication by some $\alpha \in A$. Thus

$$\delta(\phi(\beta) \cdot \alpha \cdot m) = \alpha \cdot \delta(\psi(\beta) \cdot m), \quad \delta \in D,\ \beta \in B,\ m \in M.$$

In particular, for $\delta = 1$ we get

$$\phi(\beta) \cdot \alpha \cdot m = \alpha \cdot \psi(\beta) \cdot m, \quad \beta \in B,\ m \in M$$

Since A acts faithfully on M (recall $A \simeq \mathrm{End}_D(M)$, or just think in terms of matrices), this means $\phi(\beta) \cdot \alpha = \alpha \cdot \psi(\beta)$ for all $\beta \in B$, as desired. Note also α is invertible because it represents an isomorphism $\sigma : M \to M$. □

Here's an example of an application. Consider the question: determine the elements $\alpha \in \mathbb{H}$ such that $\alpha^2 = -1$. We know $\pm i, \pm j, \pm k$ are possibilities. What about others? Suppose α is an element such that $\alpha^2 = -1$. Consider the 2-dimensional algebra $B = \mathbb{R}[\alpha] \subset \mathbb{H}$. Then $B \simeq \mathbb{C}$, so this question is tantamount to determining the embeddings of $\mathbb{C}$ into $\mathbb{H}$. Specifically, the set of such α are precisely the elements which are conjugate to i in $\mathbb{H}$. Realizing $\mathbb{H}$ in $M_2(\mathbb{C})$ as in Exercise 2.1.10, we see the set of such α will be precisely the matrices of the form

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \begin{pmatrix} i & \\ & -i \end{pmatrix} \begin{pmatrix} \overline{a} & -b \\ \overline{b} & a \end{pmatrix} = \begin{pmatrix} (|a|^2 - |b|^2)i & -2abi \\ 2\overline{ab}i & -(|a|^2 - |b|^2)i \end{pmatrix}, \tag{2.3.1}$$

where

$$a, b \in \mathbb{C}, \quad |a|^2 + |b|^2 = 1.$$

Here we just conjugated i by norm 1 (i.e., determinant 1) elements of $\mathbb{H}$ because any element of $\mathbb{H}^\times$ can be made norm 1 by multiplying by an element of the center $\mathbb{R}^\times$ (and the center does nothing for conjugation).

Mild warning: for a division algebra D over a field F we do not in general have $D^\times = F^\times D^1$, where D^1 denotes the norm 1 elements. It happens to be true for $\mathbb{H}/\mathbb{R}$ because the image $N_{\mathbb{H}/\mathbb{R}}(\mathbb{H}^\times) = \mathbb{R}_{>0} = N_{\mathbb{H}/\mathbb{R}}(\mathbb{R}^\times)$, i.e., in effect because every positive real has a real square root.

2.4 Splitting fields for simple algebras

Let A be a simple algebra over F. In this section, we will be concerned with two related questions:

(1) What fields K embed in A?

(2) How can we find an "optimal matrix representation" of A, i.e., an embedding $A \hookrightarrow M_n(K)$ where K/F is a field extension and n is as small as possible?

Answers to these questions will provide important insight into the structure of A. The first question is analogous to understanding what abelian subgroups are contained in a given group. The second question (somewhat analogous to realizing a finite group inside S_n with n as small as possible) will be important for many things, such as defining a reduced norm map in a way that is compatible with the determinant for matrix algebras.

As a first example, if $F = \mathbb{R}$ and $A = \mathbb{H}$, then $\mathbb{C}$ embeds in $\mathbb{H}$ (in many ways, but all conjugate by Skolem–Noether). On the other hand $\mathbb{H}$ embeds in $M_2(\mathbb{C})$ (Exercise 2.1.10), and since $\mathbb{H}$ is not a field itself, this is clearly the smallest dimensional matrix algebra over a field into which $\mathbb{H}$ embeds.

For the first question, we may as well just ask what are the largest fields which embed in A. Here, we need to be a little careful about what we mean by embed. Let us say a subalgebra $K \subset A$ is a **subfield** of A if it is also a field. Let us say a subset $K \subset A$ is a **quasi-subfield**[4] if it is a field under the ring operations of A, i.e., if K is a subfield of A in the category of rngs (not a typo—a rng is a "ring without identity," i.e., a (potentially) non-unital ring). Note proper subfields of F (regarding F as a field, not an F-algebra) are examples of quasi-subfield of A which are not "subfields," but we will also see below that there are quasi-subfields containing F which are not subfields of A (i.e., not subalgebras) because the identity elements do not coincide.

We say a subfield K of A is a **maximal subfield** if it is maximal among subfields with respect to inclusion. We often abuse terminology and say an extension K/F is a subfield or quasi-subfield or maximal subfield of A if it is isomorphic to one. Note $\mathbb{C}$ is the unique-up-to-isomorphism maximal subfield of $\mathbb{H}$ (though as we saw in the last section, it can be realized in $\mathbb{H}$ in infinitely many ways, which are all conjugate). Maximal subfields exist (though are not unique in general) because K must contain F, and chains of subfields of A must terminate by finite dimensionality. In fact, since A is simple, $Z(A)$ is a field by Corollary 2.3.1, so maximal subfields must contain $Z(A)$. Thus, for this question, we can replace F by $Z(A)$ and assume A is a CSA.

In fact, the crucial case is where A is a central division algebra, because it is easy to say what fields are contained in matrix algebras.

Proposition 2.4.1. *The quasi-subfields of $M_n(F)$ containing F are (up to isomorphism) precisely the field extensions of F of degree at most n. The field extensions K/F of degree dividing n are in fact subfields.*

Proof. Let K/F be a field extension of degree $m \leq n$. Then K embeds as a subalgebra of $M_m(F)$ by Proposition 2.1.1, and is thus a subfield of $M_m(F)$. If $m|n$, we can realize K as a subfield of $M_n(F)$ by composition with the algebra embedding $M_m(F) \hookrightarrow M_n(F)$ given

[4] I do not know of standard terminology for this distinction. Patent pending.

by the block diagonal embedding

$$\alpha \mapsto \begin{pmatrix} \alpha & & & \\ & \alpha & & \\ & & \ddots & \\ & & & \alpha \end{pmatrix}.$$

Otherwise, we can use the map of rngs (again, not a typo)

$$\alpha \mapsto \begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix}$$

(which is not a homomorphism of *unital* rings) to realize K as a quasi-subfield of $M_n(F)$.

To see that there are no other quasi-subfields, suppose K is a quasi-subfield of $M_n(F)$. We can write $K = F(\alpha)$ by the primitive element theorem. Then the non-reduced minimal polynomial of α as degree at most n, thus K must be of degree at most n. □

We will see below that for a quasi-subfield K of $M_n(F)$ is a subfield if and only if $K \supset F$ and $[K : F]|n$.

When $n = 2$, this says the maximal subfields of $M_n(F)$ are precisely the quadratic extensions of F, unless F contains all its square roots. (If F is algebraically closed, e.g., $F = \mathbb{C}$, F is always the unique subfield of $M_n(F)$. Similarly if $F = \mathbb{R}$ and $n > 1$, then $\mathbb{C}$ is always the unique maximal subfield of $M_n(\mathbb{R})$.) The following exercise gives an explicit realization of these subfields.

Exercise 2.4.1. Let $K = F(\sqrt{\delta})$ be a quadratic field extension of F. Show

$$a + b\sqrt{\delta} \mapsto \begin{pmatrix} a & \delta b \\ b & a \end{pmatrix}$$

defines an isomorphism of K with a subfield of $M_2(F)$. On the other hand, show that if δ is a square in F, then

$$\left\{ \begin{pmatrix} a & \delta b \\ b & a \end{pmatrix} : a, b \in F \right\}$$

is a subalgebra of $M_2(F)$ isomorphic to $F \oplus F$.

On the other hand, there is essentially no distinction between quasi-subfields and subfields of division algebras.

Exercise 2.4.2. Let D be a division algebra over F and K a quasi-subfield of D which contains F. Show K is a subfield of D.

Now let's consider the second question, about finding an optimal matrix representation of A. Again the crucial case is where A is a (central) division algebra. To see why, recall by Wedderburn's theorem we know $A \simeq M_m(D)$ for a unique m and a division algebra D/F (unique up to isomorphism). So if we have an embedding $D \hookrightarrow M_n(K)$, then we get an

embedding of $A \hookrightarrow M_{mn}(K)$. While it's not obvious at this stage, using tools we develop one can show that if this was an optimal matrix representation for D, the corresponding matrix representation for A is also optimal. The way we will obtain matrix representations is through *splitting fields*.

Let A be a CSA over F. We say a field $K \supset F$ **splits** A, or is a **splitting field** for A, if $A \otimes_F K \simeq M_n(K)$, as K-algebras. (We do not require $K \subset A$, or even K/F to be finite degree.) If F itself is a splitting field for A, i.e., $A \simeq M_n(F)$ for some n, then we simply say A is **split**.

Exercise 2.4.3. Consider the Hamilton quaternion algebra $\mathbb{H}$ over $\mathbb{R}$, with subfield $\mathbb{C}$. Show $\mathbb{H}$ splits over $\mathbb{C}$, i.e., $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \simeq M_2(\mathbb{C})$.

The point is that if K splits A, then we can identify A as a subalgebra of $A \otimes_F K \simeq M_n(K)$ (as F-algebras), and thus get a matrix realization.

Exercise 2.4.4. Let A, B be simple F-algebras. Show the map $\alpha \mapsto \alpha \otimes 1_B$ defines an algebra embedding of A into $A \otimes_F B$.

The connection between the two questions we posed will be that for a central division algebra D, maximal subfields K are the same as splitting fields. In particular, maximal subfields of division algebras must all have the same dimension, in contrast to the case of maximal subfields of matrix algebras. We will also deduce that any division algebra has square dimension.

Remark 2.4.2 (*Terminology and connection with algebraic groups*)**.** The term split agrees with the use of the term for algebraic groups. Namely, we can consider $G = A^\times$ as an *algebraic group* over F. Roughly, an algebraic group is a matrix group defined by polynomials over F such as the **general linear group** $\mathrm{GL}_n(F) = M_n(F)^\times$ or the **special linear group** $\mathrm{SL}_n(F) = \{g \in M_n(F) : \det g = 1\}$. Inside $G = A^\times$, we can look at a maximal **(algebraic) torus** T, which is just by definition a maximal abelian algebraic subgroup of G. We say a torus T is *split* if $T \simeq (F^\times)^m$ for some m, and G is **split** if it contains a split maximal torus. Then A being split (a matrix algebra) is the same as $G = A^\times$ being split as an algebraic group. In particular, if $A = M_n(F)$, the diagonal subgroup $(F^\times)^n$ is a maximal torus, which is split). Other maximal tori in $A^\times = \mathrm{GL}_n(F)$ will be the multiplicative groups of direct sums of field extensions, where the degrees sum to n—e.g., $K^\times$ is also a maximal torus in $\mathrm{GL}_n(F)$ for any degree n field extension K/F by Proposition 2.4.1.

For example, if we take $A = \mathbb{H}$ and $F = \mathbb{R}$, so $G = \mathbb{H}^\times$, then that any maximal torus in G turns out to be isomorphic to $\mathbb{C}^\times$. Neither G nor A is split over $\mathbb{R}$, and we see a torus $T = \mathbb{C}^\times \simeq S^1 \times \mathbb{R}^\times$ is topologically a (bi)infinite cyclinder. However, when we tensor up to $\mathbb{C}$ we see $A \otimes_{\mathbb{R}} \mathbb{C} \simeq M_2(\mathbb{C})$ so $G_{\mathbb{C}} := (A \otimes_{\mathbb{R}} \mathbb{C})^\times \simeq \mathrm{GL}_2(\mathbb{C})$. Thus we can take $(\mathbb{C})^2$ to be a maximal torus $T_{\mathbb{C}}$. If we take the real points of $T_{\mathbb{C}}$ we just get $T_{\mathbb{C}}(\mathbb{R}) = \mathbb{R}^\times \times \mathbb{R}^\times$, i.e., we have "split" the circle S^1 and turned it into the straight line minus a point $\mathbb{R}^\times$. Hence there is a geometric meaning of the term split.

If you haven't see this stuff before, you might wonder why an algebraic torus is called a torus. In our example above, the torus wasn't topologically a torus but an open cyclinder. However, you do get topological tori for some groups. If we instead work with

$G' = H^\times/\mathbb{R}^\times = G/Z(G)$, then a maximal torus T' is isomorphic to $\mathbb{C}^\times/\mathbb{R}^\times \simeq S^1$. This is a circle, which is the 1-dimensional version of a torus (the n-dimensional topological torus being $(S^1)^n$, with $n = 2$ giving the usual torus). One (or two, depending on how you count) of the main families of real Lie groups are the **special orthogonal groups** $\mathrm{SO}(n) = \left\{A \in \mathrm{GL}(n,\mathbb{R}) : {}^tAA = I, \det A = 1\right\}$. (These groups are also algebraic groups over $\mathbb{R}$, being defined by polynomial equations.) In fact, $G' \simeq \mathrm{SO}(3)$ and $T' \simeq \mathrm{SO}(2)$. The next higher dimensional analogue of G' in this framework is $\mathrm{SO}(4)$, which has as a maximal torus $\mathrm{SO}(2) \times \mathrm{SO}(2) \simeq S^1 \times S^1$, the usual torus.

To obtain the connection between maximal subfields and splitting fields, it will be useful to look at centralizers. Suppose A, B are F-algebras and B is a subalgebra of A. The **centralizer** of B in A is

$$C_A(B) = \{\alpha \in A : \alpha\beta = \beta\alpha \text{ for all } \beta \in B\}.$$

It is easy to check that $C_A(B)$ is also a subalgebra of F which contains $Z(A)$. Further $B \subset C_A(B)$ if and only if B is commutative. Note $C_A(Z(A)) = A$ and $C_A(A) = Z(A)$.

Exercise 2.4.5. Let A be an F-algebra and B a subalgebra. View A as a left $(A \otimes B^{\mathrm{opp}})$-module via $(\alpha \otimes \beta)x \mapsto \alpha x \beta$ (cf. Exercise 2.3.2). Show right multiplication defines an F-algebra isomorphism of $\mathrm{End}_{A\otimes B^{\mathrm{opp}}}(A) \simeq C_A(B)^{\mathrm{opp}}$.

This is essentially [Pie82, Lem 12.7]. When A is central and $B = F$, this just says $\mathrm{End}_A(A) \simeq A^{\mathrm{opp}}$, which was Lemma 2.2.1.

Theorem 2.4.3 (Double centralizer theorem (DCT))**.** *Let A be a CSA over F and B be a simple subalgebra. Then:*

(1) $C_A(B)$ *is a simple subalgebra of* A *and* $Z(C_A(B)) = Z(B)$*;*

(2) $\dim A = \dim B \cdot \dim C_A(B)$*;*

(3) $C_A(C_A(B)) = B$*; and*

(4) if B *is central, then* $A \simeq B \otimes C_A(B)$ *as* F*-algebras.*

Proof. By Proposition 2.3.2, we know $C := A \otimes B^{\mathrm{opp}}$ is simple. The equality $Z(C_A(B)) = Z(B)$ is straightforward, which proves (1).

Let $\mathcal{I}$ be a minimal left ideal of C. Then, as the proof of the Wedderburn structure theorem, we know $C \simeq \mathrm{End}_C(\bigoplus \mathcal{I})^{\mathrm{opp}} \simeq M_n(D^{\mathrm{opp}})$, where $D = \mathrm{End}_C(\mathcal{I})$ is a division algebra. Thus

$$\dim C = \dim A \cdot \dim B = n^2 \dim D = n \dim \mathcal{I}$$

(as F-vector spaces). Now A embeds in C by Exercise 2.4.4 so we can write $A \simeq \bigoplus_{i=1}^m \mathcal{I}$ and $\dim A = m \dim \mathcal{I}$. Combining with the equation for $\dim C$, we see

$$\dim B = \frac{n}{m}.$$

On the other hand, $C_A(B)^{\mathrm{opp}} \simeq \mathrm{End}_C(A) \simeq M_m(D^{\mathrm{opp}})$, using Exercise 2.4.5 for the first isomorphism. So $\dim C_A(B) = m^2 \dim D = \frac{m^2 \dim \mathcal{I}}{n}$, which gives (2).

It is straightforward from the definition that $B \subset C_A(C_A(B))$. Applying (2) to $B' = C_A(B)$ shows $\dim C_A(C_A(B)) = \dim B$ and we get (3).

Now suppose B is a CSA. Since B and $C_A(B)$ commute, the map $\beta \otimes \gamma \mapsto \beta\gamma$ defines an algebra homomorphism $\phi : B \otimes C_A(B) \to A$: it's clearly F-linear—it respects ring multiplication as

$$\phi(\beta \otimes \gamma)\phi(\beta' \otimes \gamma') = \beta\gamma\beta'\gamma' = \beta\beta'\gamma\gamma' = \phi((\beta \otimes \gamma)(\beta' \otimes \gamma')).$$

Then ϕ injective as A is simple. It must be surjective by dimension from (2). □

Corollary 2.4.4. *With notation as in the theorem,* $\dim B | \dim A$.

We won't actually need (4) for our results on splitting fields, but it is a nice structural result further describing how B is dual to its centralizer in A. Here's a generalization.

> **Exercise 2.4.6.** Let A be a CSA over F and B a subalgebra with center K. Show $C_A(K) \simeq B \otimes_K C_A(B)$ as K-algebras.

Now we can prove our main result for this section.

Theorem 2.4.5. *Let D be a central division F-algebra. A subfield $K \subset D$ is maximal if and only if K splits D, and $\dim_F D = [K : F]^2$. In particular, D has a splitting field and $\dim_F D$ is square.*

Note this does not mean any splitting field for D is a maximal subfield of D. If K is maximal subfield then any extension $L \supset K$ will also be a splitting field (recall splitting fields need not be subfields) as $D \otimes_F L \simeq (D \otimes_F K) \otimes_K L$, say as L-algebras. However we will just be interested in splitting fields which are subfields.

Proof. Let B be a simple subalgebra of D. For $\delta \otimes \beta \in D \otimes B$, we consider the action on D^{opp} given by $(\delta \otimes \beta) \cdot x = \beta x \delta$, $x \in D^{\mathrm{opp}}$. Note, for $\gamma \in C_D(B)$,

$$(\delta \otimes \beta) \cdot x = \beta\gamma x\delta = \gamma\beta x\delta = \gamma(\delta \otimes \beta) \cdot x.$$

That is, $\delta \otimes \beta$ yields a $C_D(B)$-linear operator on D, and in fact defines an algebra homomorphism $\phi : D \otimes B \to \mathrm{End}_{C_D(B)}(D^{\mathrm{opp}})$ (viewing D^{opp} as a left $C_D(B)$-module via left multiplication). Then ϕ is injective because $D \otimes B$ is simple. Note $C_D(B)$ must also be a division algebra, so as left $C_D(B)$-modules, $D^{\mathrm{opp}} \simeq \bigoplus_{i=1}^r C_D(B)$, where $r = \frac{\dim D}{\dim C_D(B)} = \dim B$, using (2) of the double centralizer theorem for the last equality. Thus $\mathrm{End}_{C_D(B)}(D^{\mathrm{opp}}) \simeq M_r(C_D(B))$, and hence has dimension $\dim C_D(B) r^2 = \dim D \cdot \dim B$ over F (again by the DCT). Looking at dimensions shows ϕ is surjective, and we get an isomorphism

$$D \otimes B \simeq \mathrm{End}_{C_D(B)}(D^{\mathrm{opp}}) \simeq M_r(C_D(B)). \tag{2.4.1}$$

Let $K \subset D$ be a subfield and put $n = [K : F]$. Note K is maximal if and only if $C_D(K) = K$ (otherwise adjoining an element of $C_D(K)$ would give another subfield containing K).

First suppose K is maximal. Then (2.4.1) reads

$$D \otimes K \simeq M_n(C_D(K)) \simeq M_n(K),$$

by (3) of the DCT. Hence K splits D and $\dim D = n^2$.

Conversely, suppose K is a splitting field. Then (2.4.1) gives $M_r(C_D(K)) \simeq D \otimes K \simeq M_m(K)$ for some m. Considering dimensions shows $\dim D = m^2$. Since, $C_D(K)$ is a division algebra the isomorphism $M_r(C_D(K)) \simeq M_m(K)$ forces $r = m$ and $C_D(K) = K$ by Wedderburn's theorem. Hence K is maximal. □

Corollary 2.4.6. *All maximal subfields of a division algebra D with center F have the same degree.*

Corollary 2.4.7. *Let A be a CSA over F. Then $\dim A = n^2$ for some $n \in \mathbb{N}$, and there exists a degree d field extension K/F for some $d|n$ such that $A \otimes_F K \simeq M_n(K)$. We call $n = \deg A$ the* **degree** *of A over F.*

So the dimension of a CSA is always a square. This is not true for non-central simple algebras A/F. Instead, we have that $\dim_F A$ is a square times $\dim_F Z(A)$.

Proof. By Wedderburn's theorem, $A \simeq M_r(D)$ for some D. Since $\dim D = d^2$ for some d by the theorem, $\dim A = (rd)^2$. Since D splits over a field K/F of degree D, so does $M_r(D)$ as $M_r(D) \otimes K \simeq M_r(D \otimes K)$ (see exercise below). □

Exercise 2.4.7. Let D be a division algebra over F and K a field extension of F. Show $M_r(D \otimes K) \simeq M_r(D) \otimes K$ as K-algebras.

We can use the above ideas to determine the subfields of $M_n(F)$.

Exercise 2.4.8. Let K/F be a field extension. Show K is (isomorphic to) a subfield of $M_n(F)$ if and only $[K : F]|n$. (*Suggestion:* use the DCT.)

The following exercise essentially says that for CSAs—say $A = M_n(D)$ where D is division of degree d—the optimal matrix representations of A are into $M_{nd}(K)$ where $[K : F] = d$ (cf. (2.5.1)).

Exercise 2.4.9. Let $A = M_n(D)$ where D is a central division algebra over F of degree d. Show that any splitting field K of A which is a subfield of A must satisfy $[K : F] \geq d$.

Remark 2.4.8. An important result is not just that splitting fields K/F exist, but that we can take K to be a *separable* extension. This requires some additional work to prove. However, since we are concerned with local and global fields F of characteristic 0, we get separability automatically in our cases of interest.

If A is a CSA over F with $A \simeq M_n(D)$ for D a division algebra, we call $\deg D$ the **(Schur) index** of A, denoted by $\operatorname{ind} A$. The above corollary says that any CSA A has a splitting field K which is a subfield with $[K : F] = \operatorname{ind} A$. We can take K to be a subfield of A. This does not mean any subfield K of A of degree $\operatorname{ind} A$ over F splits A. For instance if $A = M_2(D)$ where D is a degree 2 division algebra, so $\operatorname{ind} A = 2$, there are subfields K of A of degree 2 over F which are not contained in D, and these need not split A.

One can show that any the degree of any splitting field (subfield or not) of A must be a multiple of the index of A, i.e., we cannot find splitting fields of smaller degree than what we can get by using subfields. (See, e.g., [GS06, Sec 4.5].)

2.5 Reduced norm and trace

With the above results about splitting fields in mind, now we will define reduced norms and traces.

Let A be a CSA over F of degree n, and let K be a splitting field for A. To be more explicit, by Wedderburn's theorem, we can assume $A = M_m(D)$ where D is a (central) division algebra over F degree $d = \frac{n}{m}$, and can take K to be a maximal subfield of D. Then we have an algebra embedding

$$\iota : A \hookrightarrow A \otimes K \simeq M_n(K), \tag{2.5.1}$$

using Exercise 2.4.4 for the hookarrow. To be clear, ι denotes an embedding $A \hookrightarrow M_n(K)$.

For $\alpha \in A$, define the **reduced characteristic polynomial** (resp. **(reduced) minimal polynomial**, resp. **(reduced) norm**, resp. **(reduced) trace**) to be the characteristic polynomial (resp. minimal polynomial, resp. determinant, resp. trace) of $\iota(\alpha) \in M_n(K)$. Let us temporarily denote these by p_α^ι, m_α^ι, $N^\iota(\alpha)$ and $\operatorname{tr}^\iota(\alpha)$. A priori, these polynomials are just polynomials defined over K and depend on ι. In fact they are defined over F and do not depend on ι.

Note that if $A = M_n(F)$, then $K = F$ so the reduced norm, trace, etc. agree with the usual notions for matrices (over fields). The reason we make the above definitions for general CSAs is that there are issues which arise when trying to generalize determinants and characteristic polynomials to matrix algebras over skewfields. There are theories of "noncommutative determinants" to address this, but we will not pursue that approach.

Lemma 2.5.1. *The polynomials p_α^ι and m_α^ι are polynomials of degree at most nd defined over F. Furthermore, the quantities p_α^ι, m_α^ι, N^ι and tr^ι do not depend on the choice ι or K.*

Proof. Consider two embeddings $\iota, \iota' : A \hookrightarrow M_n(K)$. We can extend these embeddings to isomorphisms $\iota_K, \iota'_K : A \otimes K \xrightarrow{\sim} M_n(K)$ via $\iota_K(\alpha \otimes x) = \iota(\alpha)x$ for $\alpha \in A$, $x \in K$, and similarly for ι'. Now by Skolem–Noether applied to K-algebras, ι_K and ι'_K must be conjugate by some $g \in \mathrm{GL}_n(K) = M_n(K)^\times$. Thus $\iota'(\alpha) = g\iota(\alpha)g^{-1}$ for some $g \in \mathrm{GL}_n(K)$ so $p_\alpha^\iota = p_\alpha^{\iota'}$, and similarly for minimal polynomials. This shows our given quantities don't depend on ι given K.

To show the coefficients of the characteristic and minimal polynomials lie in F, fix K, ι and let us assume K/F is Galois so that the set of fixed points of $\mathrm{Gal}(K/F)$ acting on K is

just F. (If K/F is not Galois, we can just replace K by the Galois closure L and extend ι to a map $A \hookrightarrow M_n(L)$ without changing the characteristic polynomial.) Let $\sigma \in \mathrm{Gal}(K/F)$. Set $\iota' = \sigma \circ \iota$. If $p_\alpha^\iota(x) = \sum c_i x^i$, then $p_\alpha^{\iota'}(x) = \sum c_i^\sigma x^i$. Thus $p_\alpha^\iota = p_\alpha^{\iota'}$ implies the coefficients of $p_\alpha^\iota(x)$ are Galois invariant, and thus lie in F. The same is true for minimal polynomials.

Finally, suppose we have two embeddings $\iota : A \hookrightarrow M_n(K)$ and $\iota' : A \hookrightarrow M_n(K')$, where K and K' are maximal subfields of D. Let $L = KK'$ be the compositum, and we can extend ι, ι' to isomorphism $\iota_L, \iota'_L : A \otimes L \xrightarrow{\sim} M_n(L)$. Then p_α^ι and $p_\alpha^{\iota'}$ must agree with the characteristic polynomials of $\iota_L(\alpha)$ and $\iota'_L(\alpha)$, which must be the same by our previous argument. Hence the reduced characteristic polynomial is also independent of K. This implies reduced norms and traces do not depend on K. Similarly for minimal polynomials. □

Consequently we will denote the reduced characteristic and minimal polynomials and reduced norm and trace simply by p_α, m_α, $N_{A/F}(\alpha) = N(\alpha)$ and $\mathrm{tr}_{A/F}(\alpha) = \mathrm{tr}(\alpha)$. I may also drop parentheses for the reduced norm and traces, and simply call them the norm and trace as we will work with these rather than the non-reduced ones. Note the (reduced) minimal polynomial must be the minimum degree monic polynomial over F which annihilates α, and therefore agrees with the "non-reduced" minimal polynomial.

Exercise 2.5.1. Let $F = \mathbb{R}$, $A = \mathbb{H}$ and write $\alpha = x + yi + zj + wk \in \mathbb{H}$. Use the embedding $\mathbb{H} \hookrightarrow M_2(\mathbb{C})$ from (2.1.1) to compute $N(\alpha) = x^2 + y^2 + z^2 + w^2$ and $\mathrm{tr}\,\alpha = 2x$.

Lemma 2.5.2. *Let A be a CSA of degree n. For $\alpha \in A$, p_α is a degree n polynomial over F and the reduced norm and reduced trace give maps $N : A \to F$, $\mathrm{tr} : A \to F$. We have the following properties of the reduced norm and trace maps:*

(1) for $\alpha, \beta \in A$, $N(\alpha\beta) = N(\alpha)N(\beta)$ and $\mathrm{tr}(\alpha + \beta) = \mathrm{tr}\,\alpha + \mathrm{tr}\,\beta$;

(2) for $\alpha \in A$, $N(\alpha) \neq 0$ if and only if $\alpha \in A^\times$; and

(3) for $x \in F$, $N_{A/F}(x) = x^n$ and $\mathrm{tr}_{A/F}(x) = nx$.

Proof. We already showed the assertion about p_α, which implies that the reduced norm and trace are F-valued, as they are, up to signs, coefficients of p_α. The first property follows from multiplicativity of determinant and additivity of trace for matrices. The third property follows because $\iota(x) = \mathrm{diag}(x, \dots, x)$.

For (2), if $\alpha \in A^\times$, then $1 = N(1) = N(\alpha)N(\alpha^{-1})$ implies $N(\alpha) \neq 0$. If $N(\alpha) \neq 0$, then the characteristic polynomial $p_\alpha(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_0$ has nonzero constant term. By the Cayley–Hamilton theorem, $p_\alpha(\alpha) = 0$ (and thus the minimal polynomial divides the characteristic polynomial), i.e.,

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

Dividing by $-c_0$ gives an inverse to α. □

This lemma implies that the reduced characteristic polynomial, reduced norm and reduced trace are different from the non-reduced versions whenever $D \neq F$, because they will be different on elements of F (e.g., for $x \in F$, $N^{\mathrm{nr}}_{A/F}(x) = x^n$).

A consequence of (1) is that the reduced norm and trace define group homomorphisms

$$N : A^\times \to F^\times, \quad \mathrm{tr} : A \to F,$$

and thus are analogous to norms and traces for number field extensions. A consequence of (2) is that A is a division algebra (i.e., $n = 1$) if and only if the norm is nonzero on all nonzero elements.

As mentioned in the introduction, the reduced norm will provide a link between quaternion algebras and quadratic forms, generalizing the case of $\mathbb{H}$ in the exercise above.

Exercise 2.5.2. Let D/F be a central division algebra of degree n, and K be a subfield of D. Show that for any $x \in K$,

$$\mathrm{tr}_{D/F} x = \frac{n}{[K:F]} \mathrm{tr}_{K/F} x, \quad N_{D/F}(x) = N_{K/F}(x)^{\frac{n}{[K:F]}}.$$

(The quantities on the left denote reduced trace and norm on D, and on the right are trace and norm of extensions of number fields.) In particular, if K is a maximal subfield of D which contains x, then the reduced trace and norm agree with the trace and norm for extensions of algebraic number fields.

2.6 Simple algebras over $\mathbb{R}$ and $\mathbb{C}$

We call $\mathbb{R}$-algebras *real* algebras and $\mathbb{C}$-algebras *complex* algebras. Historically, these were the algebras of the most interest (and still are of great interest). As an easy application of the theory we have developed thus far, we can now classify the real and complex simple (and hence also semisimple) algebras. On the other hand, the classification over number fields still requires a lot of work, and we will not prove the full classification in this course, but will discuss it in the next section.

In this section, we will classify the real and complex simple algebras. By Wedderburn's theorem, it suffices to classify the real and complex division algebras. Besides the result itself being appealing, it arises in the classification of central simple algebras over number fields because this classification uses local methods (cf. Section 2.7). So the real and complex classification will describe the possibilities at archimedean places.

The complex case is simpler, so we begin with that.

Proposition 2.6.1. *The only complex division algebra is* $\mathbb{C}$.

Proof. Let D be a complex division algebra. Fix $\alpha \in D$ and let p be its minimal polynomial. Then $p(\alpha) = 0$ by definition of minimal polynomial. But since p factors into linear factors over $\mathbb{C}$, we have $\alpha - z = 0$ for some $z \in \mathbb{C}$ (assuming $\alpha \neq 0$). Hence $\alpha \in \mathbb{C}$. □

Corollary 2.6.2. *Any simple* $\mathbb{C}$*-algebra is isomorphic to a complex matrix algebra* $M_n(\mathbb{C})$.

The classification of simple algebras over $\mathbb{R}$ follows from the following famous result.

Theorem 2.6.3 (Frobenius)**.** *Let D be a real division algebra, i.e., a division algebra over $\mathbb{R}$. Then D is isomorphic to $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$.*

Proof. We use the fact that the only field extensions $K/\mathbb{R}$ of finite degree are $\mathbb{R}$ and $\mathbb{C}$. If D is not central over $\mathbb{R}$, then by Corollary 2.3.1 D must be central over $\mathbb{C}$, hence by the previous result $D = \mathbb{C}$.

So suppose $Z(D) = \mathbb{R}$ and let K be a maximal subfield of D. If $K = \mathbb{R}$, then D is split so $D = \mathbb{R}$. If $K = \mathbb{C}$, by Theorem 2.4.5, we see D is a degree 2 division algebra containing $\mathbb{C}$. Now apply the exercise below. □

There are more elementary proofs of Frobenius's's theoremseseses,[5] e.g., R.S. Palais's's' note in the *Monthly* (Apr 1968), but I wanted to show the utility of splitting fieldseses.

Exercise 2.6.1. Show any 4-dimensional real division algebra D is isomorphic to $\mathbb{H}$. (*Suggestion:* One approach is to first observe you can write $D = \mathbb{C} \oplus \mathbb{C}j$ for some $j \in D$ such that $j^2 = -1$, and then determine what right multiplication by i does to $\mathbb{C}j$.)

You can do the above exercise without much theory. This exercise is also a consequence of general structure theory of quaternion algebras we will develop later.

Corollary 2.6.4. *Any simple $\mathbb{R}$-algebra is isomorphic to $M_n(\mathbb{R})$, $M_n(\mathbb{C})$ or $M_n(\mathbb{H})$ for some n.*

Proof. This follows immediately from combining Frobenius's theorem and Wedderburn's theorem. □

2.7 The local-global principle and CSAs over number fields

The local-global classification of CSAs over number fields was one of the big theorems in algebra in the early 20th century, and closely tied to the development of class field theory.[6] We will not have time to prove the full classification of CSAs over number fields, but in this section I will briefly summarize the main results. In the subsequent chapters, we will go through the classification in detail for degree 2 CSAs, i.e., quaternion algebras, modulo the proof of Hasse–Minkowski. (There is only 1 degree 1 CSA/F up to isomorphism of course.)

Let F be a number field and A be a CSA of degree n over F, i.e., of dimension n^2. For any place v of F, we consider the local algebra

$$A_F = A \otimes_F F_v,$$

which is a CSA over F_v of degree n by Proposition 2.3.2. It is clear that $A \simeq A'$ implies $A_v \simeq A'_v$ for all v. It is not at all obvious that the converse is true.

Theorem 2.7.1 (Albert–Brauer–Hasse–Noether, local-global principle)**.** *Let A, A' be CSAs over F. Then $A \simeq A'$ if and only if $A_v \simeq A'_v$ for all places v of F.*

[5] Just like Wedderburn, Frobenius has a bunch of famous theorems. If you say called Frobenius's theorem on real division algebras, it should be clear you mean this one (and you should mean this one).

[6] See Peter Roquette's article, *The Brauer–Hasse–Noether theorem in historical perspective*, for a nice exposition of the historical development of this classification.

Pierce [Pie82, Sec 18.4] calls this "The most profound result in the theory of central simple algebras." This also sometimes just called the Brauer–Hasse–Noether theorem, though Albert (an American mathematician) played a role in its proof. Had the four mathematicians been on the same continent or lived in a time with more advanced travel and communication options, the correspondence described in Roquette's article leads me to believe the original proof would have been a 4-way collaboration.

One method of proof is to reduce to the case of *cyclic algebras* and use Hasse's norm theorem. Cyclic algebras of degree n are CSAs which can be constructed in a certain concrete way in terms of matrices over a degree n cyclic extension K/F (i.e., K/F is Galois with cyclic Galois group). Specifically, let K/F be a cyclic extension of degree n, σ a generator of the Galois group, and $b \in K^\times$. Then the **cyclic algebra** $(K/F, \sigma, b)$ is the degree n CSA over F with generated by an element y and the extension K subject to the relations

$$y^n = b, \quad \alpha y = y\alpha^\sigma, \qquad \text{for all } \alpha \in K.$$

Explicitly we can realize this inside $M_n(K)$ by taking

$$y = \begin{pmatrix} 0 & 0 & \cdots & 0 & b \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \in \mathrm{GL}_n(F) \subset \mathrm{GL}_n(K),$$

and embedding K in $M_n(K)$ via $\alpha \mapsto \mathrm{diag}(\alpha, \alpha^\sigma, \dots, \alpha^{(n-1)\sigma})$. One can show the index of $(K/F, \sigma, b)$ is just the order of b in $F^\times/F^{(n)}$.

Exercise 2.7.1. Let $n = 2$. Show the cyclic algebra $(K/F, \sigma, b)$ is a division algebra if b is not a norm from K and split if b is a norm from K.

Albert–Brauer–Hasse–Noether (or some subset) eventually proved that any CSA over a number field or p-adic field (or $\mathbb{R}$ or $\mathbb{C}$) is a cyclic algebra, which one deduces as a consequence of the above local–global principle (and Roquette says that the authors really considered this as their main result).

Hasse's norm theorem is the following local–global principle for norms:

Theorem 2.7.2 (Hasse's norm theorem)**.** *Let K/F be a cyclic extension of number fields. Then $x \in F$ is a norm from K if and only if it is a norm in F_v from K_v for all v.*

The proof of Hasse's norm theorem is long, but the reduction of Albert–Brauer–Hasse–Noether to this theorem is relatively short (see, e.g., [Pie82, Sec 18.4]). One can also prove the local global principle by using zeta functions—e.g., Weil uses the zeta function approach in [Wei95, Thm XI.2] to prove the above ABHN local-global principle in the case $A' = M_n(F)$.

Thus, to classify CSAs over number fields, it suffices to (i) classify CSAs over local fields, and (ii) determine when local CSAs can be patched together to make a global CSA.

While to answer (i) it suffices to classify division algebras over local fields by Wedderburn's theorem, the answer is nicer to explain in terms of CSAs. We have already explained the classification of CSAs over archimedean fields in Section 2.6: over $\mathbb{R}$, one just gets $M_n(\mathbb{R})$ and $M_{n/2}(\mathbb{H})$; over $\mathbb{C}$, just $M_n(\mathbb{C})$.

Theorem 2.7.3. *Let v be a nonarchimedean valuation. Then the CSAs of degree n over F_v are, up to isomorphism, in bijection with $\mathbb{Z}/n\mathbb{Z}$. Under this correspondence, the index of a CSA corresponding to $a \in \mathbb{Z}/n\mathbb{Z}$ is the order of a in $\mathbb{Z}/n\mathbb{Z}$. Thus 0 corresponds to $M_n(F_v)$ and the elements of $(\mathbb{Z}/n\mathbb{Z})^\times$ correspond to division algebras.*

When $n = 2$, this says that the only quaternion algebras over a p-adic field are (up to isomorphism) the unique division algebra of dimension 4 and $M_2(F_v)$.

This classification is generally proven using the Brauer group of F_v. The **Brauer group** of a field k is the collection $Br(k)$ of CSAs (up to isomorphism) over k modulo the equivalence $M_n(D) \sim M_m(D)$, i.e., two CSAs are Brauer equivalent if they have the same index. The group law is tensor product.

Exercise 2.7.2. Let A be a CSA over k. Show $A \otimes A^{\mathrm{opp}} \sim k$ (Brauer equivalence), and show that tensor product makes $Br(k)$ into an abelian group.

Exercise 2.7.3. Show $Br(\mathbb{R}) \simeq \mathbb{Z}/2\mathbb{Z}$ and $Br(\mathbb{C}) \simeq \{1\}$.

The above local classification can be deduced from the following result, which is nowadays typically proved by cohomological methods.

Theorem 2.7.4. *For nonarchimedean v, $Br(F_v) \simeq \mathbb{Q}/\mathbb{Z}$.*

It is a theorem that, over p-adic fields (or number fields), the exponent of a CSA in the Brauer group (also called the *period* of the CSA) is the same as the index. (Over general fields the *period-index theorem* says the exponent or period divides the index, but Brauer constructed examples to show they need not be equal.)

Via this isomorphism with the Brauer group, each CSA A_v of degree n over F_v corresponds to a rational number of the form $\frac{a}{n}$ where $0 \le a < n$. Then Theorem 2.7.3 is essentially just the "degree n" part of Theorem 2.7.4.

The rational number $\frac{a}{n}$ is called the **(Hasse) invariant** of A_v and denoted $\operatorname{inv} A_v$. This invariant will play an important role in the global classification. For v archimedean, $\operatorname{inv} A_v = 0$ if $A_v \simeq M_n(\mathbb{R})$ or $A_v \simeq M_n(\mathbb{C})$, and $\operatorname{inv} A_v = \frac{1}{2}$ if $A_v \simeq M_{n/2}(\mathbb{H})$. Note for any v, $\operatorname{inv} A_v = 0$ if and only if split and more generally the order of $\operatorname{inv} A_v$ in $\mathbb{Q}/\mathbb{Z}$ equals $\deg A_v$.

See, e.g., [Pie82, Chap 17] for a detailed exposition of these facts.

Now let us describe the complete classification of CSAs over F.

We say A is **unramified** or **split** at v, or A_v is unramified if $A_v \sim F$ (Brauer equivalence), i.e., if $A_v \simeq M_n(F)$ is split, i.e., if $\operatorname{inv} A_v = 0$. Otherwise A is **ramified** at v. Let

ram A denote the set of places of F at which A is ramified.[7]

Theorem 2.7.5. *Let F be a number field. Then*

(1) Any CSA A/F is unramified at almost all places.

(2) Every CSA A/F of degree n satisfies

$$\sum_v \operatorname{inv} A_v = \sum_{v \in \operatorname{ram} A} \operatorname{inv} A_v \in \mathbb{Z}.$$

(3) Given any finite set S of places of F and $a_v \in \left\{\frac{0}{n}, \frac{1}{n}, \dots, \frac{n-1}{n}\right\}$ a local Hasse invariant for each $v \in S$ such that $\sum_{v \in S} a_v \in \mathbb{Z}$, there exist a CSA A/F of degree n which is unramified at each $v \notin S$ such that $\operatorname{inv} A_v = a_v$ for each $v \in S$.

In the third part of the theorem, it is understood that at a real place each Hasse invariant must be 0 or $\frac{1}{2}$ and at each complex place each Hasse invariant must be 0 (or one can just assume S does not contain complex places).

By the Albert–Brauer–Hasse–Noether theorem, this gives a complete classification of CSAs over number fields as the conditions in (3) determine A up to isomorphism (the local Hasse invariants determine A_v up to isomorphism). In particular, if A is a CSA which is not split (i.e., $A \not\simeq M_n(F)$), then it must be ramified at at least 2 places.

When $n = \deg A = 2$, each $\operatorname{inv} A_v$ is either 0 or $\frac{1}{2}$, with the latter happening precisely when A_v is ramified, i.e., a degree 2 division algebra. The condition (2) that the invariants must sum to an integer is simply that A is ramified at a (finite) even number of places. Part (3) of the theorem says that given any set S consisting of an even number of non-complex places, there is a quaternion algebra A which is ramified precisely at $v \in S$, i.e., A_v is division if and only if $v \in S$.

Exercise 2.7.4. Fix three distinct places v_1, v_2, v_3 of F (archimedean or not).
(i) Count the number of CSAs of degree n with $\operatorname{ram} A = \{v_1, v_2\}$.
(ii) Count the number of CSAs of degree n with $\operatorname{ram} A = \{v_1, v_2, v_3\}$.

[7] Like with number fields, being ramified is something that can happen at only finitely many places as stated in the theorem. However, unlike CSAs, for extension of number fields K/F being unramified and split at v are not the same—we have (infinitely many) inert places too. So you may not want to think of ramification for CSAs as exactly corresponding to that for number fields now, though we will explain an analogy between these two notions of ramification when we examine division algebras over local fields more closely (at least in the quaternion case).