# Introduction

### Semi-historical motivation

The algebraic theory of fields is intimately connected with the theory of numbers, and the study of this connection forms the basis of *algebraic number theory*. To just consider the simplest example possible, consider the number field $K = \mathbb{Q}(i)$. Its *ring of integers* is just the Gaussian integers $\mathcal{O}_K = \mathbb{Z}[i]$. For $\alpha \in K$, let $\overline{\alpha}$ denote its Galois conjugate (i.e., $\alpha \mapsto \overline{\alpha}$ is the nontrivial automorphism in $\mathrm{Gal}(K/\mathbb{Q})$, which is the same as complex conjugation for this choice of $K$.) We can define a *norm map* $N(\alpha) = N_{K/\mathbb{Q}}(\alpha) = \alpha\overline{\alpha}$ on $K$, which is a homomorphism from $K^\times$ into $\mathbb{Q}^\times$ when restricted to $K$, and takes values in $\mathbb{Z}$ for $\alpha \in \mathcal{O}_K$.

Note that for $\alpha = x + iy \in K$, $N(\alpha) = x^2 + y^2$, so the first connection of the algebra of $K$ with number theory is this: an element of $\mathbb{Z}$ is a sum of two integral squares if and only if it is a norm from $\mathbb{Z}[i]$. (Similarly, a rational number is a sum of two rational squares if and only if it is a norm from $\mathbb{Q}(i)$.) Since the norm is a multiplicative homomorphism, we see that if $m$ and $n$ are sums of two squares, so is $mn$. In fact, it is not hard to reduce the problem of determining which integers are sums of two squares to that of determining which primes are sums of two squares.

For a prime $p$, consider the ideal $p\mathcal{O}_K$ of $\mathcal{O}_K$. Since the ideals of $\mathcal{O}_K$ have unique factorization into prime ideals, one can easily see that either (i) $p\mathcal{O}_K$ is a prime ideal of $\mathcal{O}_K$, or (ii) $p\mathcal{O}_K = \mathfrak{p}_1\mathfrak{p}_2$ a product of two prime ideals. Case (i) means $p$ is not a sum of two squares, and case (ii) means $p$ is a sum of two squares. For this interpretation of case (ii), one can use the fact that $\mathbb{Z}[i]$ is a PID. This implies that $p$ is a sum of two squares if and only if $p$ is 2 or 1 mod 4. More generally, the arithmetic of quadratic fields is intimately related with the arithmetic of binary quadratic forms $q(x, y) = ax^2 + bxy + cy^2$. (See, e.g., my Number Theory II notes [Mar].)

Now one can try to generalize this by working with more general structures than quadratic fields. If one works with general number fields, one can still consider norm forms, but now these are higher degree: if $K/\mathbb{Q}$ is Galois of degree $n$, then the norm map can be viewed as a degree $n$ form in $n$ variables over $\mathbb{Q}$. On the other hand, what if one wants to consider quadratic forms in more variables?

Recall Hamilton's quaternions $\mathbb{H}$. This is a 4-dimensional vector space over $\mathbb{R}$ with basis $\{1, i, j, k\}$ which is made into a ring by defining multiplication on the basis elements subject to $i^2 = j^2 = k^2 = ijk = -1$ (of course 1 is the ring identity). Such a structure—a vector space over a field $F$ which is also a ring—is called an $F$-algebra. A perhaps more familiar example of an $F$-algebra is $M_n(F)$, the algebra of $n \times n$ matrices over $F$.

Hamilton's motiviation, I believe, was that just like the complex numbers $\mathbb{C}$ can be used

to study 2-dimensional space algebraically, one would like an algebra to study 3-dimensional space. There is no nice analogous 3-dimensional algebra, but $\mathbb{H}$ is an analogous 4-dimensional algebra. Namely, $\mathbb{H}$ is a noncommutative (e.g., $k = ij = -ji$) generalization of the complex numbers $\mathbb{C}$ (it is a division ring, or what is sometimes called a skew field—every nonzero element is invertible, unlike in matrix algebras). Even though $\mathbb{H}$ is 4-dimensional, one can use it to study 3-dimensional space, e.g., one can model $\mathbb{R}^3$ by the set of *pure quaternions* $\mathbb{H}^0 = \{yi + zj + wk : y, z, w, \in \mathbb{R}\}$.

Analogous to the case of quadratic fields (or more specifically $\mathbb{C}/\mathbb{R}$), one defines an involution of $\mathbb{H}$ by $\overline{x + yi + zj + wk} = x - yi - zj - wk$, and subsequently a norm

$$N(\alpha) = \alpha\overline{\alpha} = x^2 + y^2 + z^2 + w^2, \quad \alpha = x + yi + zj + wk.$$

The norm is a multiplicative map into $\mathbb{R}_{\geq 0}$, i.e., $N(\alpha\beta) = N(\alpha)N(\beta)$. Since $\mathbb{R}$ is commutative, multiplicativity implies $N(\alpha\beta) = N(\beta\alpha)$ even though $\alpha\beta$ may not equal $\beta\alpha$ (this is analogous to the determinant equality $\det(AB) = \det(BA)$ for square matrices $A, B$). The norm is natural from the geometric point of view: if $\alpha_m = y_m i + z_m j + w_m k$ is a pure quaternion for $m = 1, 2$, then

$$N(\alpha_1 - \alpha_2) = (y_1 - y_2)^2 + (z_1 - z_2)^2 + (w_1 - w_2)^2$$

is simply the square of the Euclidean distance between $\alpha_1$ and $\alpha_2$ (viewed as elements of $\mathbb{R}^3$).

Returning to the notion of quadratic forms, we see that the norm map on $\mathbb{H}$ gives the quaternary (4-variable) quadratic form $x^2 + y^2 + z^2 + w^2$ over $\mathbb{R}$. The multiplicativity of the norm map means that the product of two sums of 4 squares is again a sum of 4 squares. Restricting the norm map to $\mathbb{H}^0 \simeq \mathbb{R}^3$ gives a ternary quadratic form $y^2 + z^2 + w^2$. However it is not true that the product of two sums of 3 squares is again a sum of 3 squares—the issue is roughly that the product of two pure quaternions is not in general a pure quaternion.

In number theory, we are interested in questions like, what *integers* are a sum of 4 squares (of integers)? For this we want to work with notions of algebraic and integral elements, like $K = \mathbb{Q}(i)$ and $\mathcal{O}_K = \mathbb{Z}[i]$. So instead of working with the classical Hamilton quaternions $\mathbb{H}$, it makes sense to work with a rational analogue

$$B = \mathbb{H}_{\mathbb{Q}} = \{x + yi + zj + wk : x, y, z, w \in \mathbb{Q}\}.$$

This is a 4-dimensional $\mathbb{Q}$-algebra. The obvious choice for the analogue of integers is

$$\mathbb{Z}[i, j, k] = \{x + yi + zj + wk : x, y, z, w \in \mathbb{Z}\}.$$

These are called *Lipschitz integers*. However, it turns out that the larger set of *Hurwitz integers*

$$\mathcal{O}_B = \mathbb{Z}[i, j, \frac{1 + i + j + k}{2}]$$

is nicer to work with, and it is more akin to the ring of integers of a number field. These analogues of rings of integers are called *orders*. Just like with Gaussian integers, these particular orders possess a kind of of unique factorization, and one can use this to prove Lagrange's theorem that every positive integers is the sum of 4 squares. One can also use quaternions to reprove Gauss's theorem about which numbers are sums of 3 squares.

A famous theorem of Frobenius says that the only finite-dimensional division algebras over $\mathbb{R}$ are $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$. (If one allows non-associative algebras, one also gets the octonions $\mathbb{O}$, which are pretty cool too, and are related to quadratic forms in 8 variables.) However, over $\mathbb{Q}$, there are a lot more. In particular, there are infinitely many quaternion division algebras (dimension 4 division algebras) $B$ over any number field $F$. Consequently, we can use other quaternion algebras $B$ to study other ternary and quaternary quadratic forms over $\mathbb{Q}$ and more general number fields. For instance, there is a quaternion algebra

$$B' = \mathbb{Q} \oplus \mathbb{Q}i' \oplus \mathbb{Q}j' \oplus \mathbb{Q}k'$$

over $\mathbb{Q}$ with

$$(i')^2 = -2, \quad (j')^2 = -3, \quad (k')^2 = -6.$$

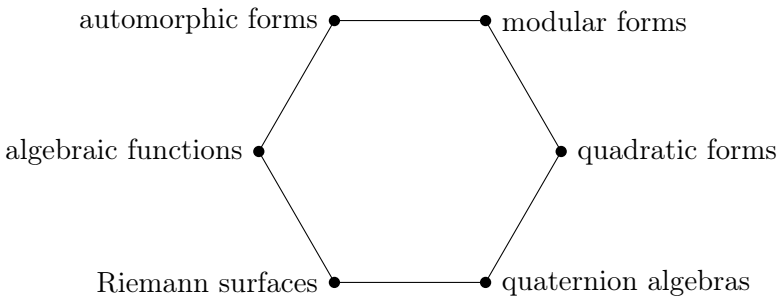
Studying its arithmetic can be used to show that the quaternary quadratic form $x^2 + 2y^2 + 3w^2 + 6z^2$ also represents all positive integers.[1]

## Goals

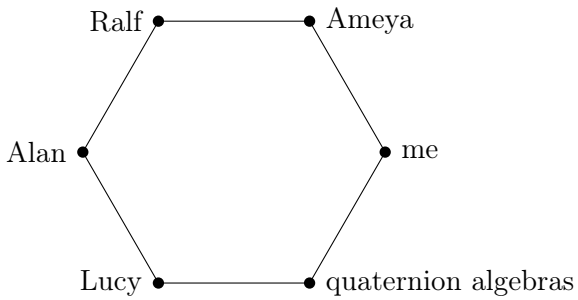
Nowadays, quaternion algebras have applications to many different subjects and problems besides just quadratic forms in 3 and 4 variables. In number theory, they are also important in the theory of modular forms via theta series and the Jacquet–Langlands correspondence, which is my principal interest in this subject. (Modular forms and theta series are also closely related to, and historically arise from the study of, quadratic forms, but of perhaps greater interest now is relations with objects such as elliptic curves.) Not unrelated to this, quaternion algebras are used to construct an important class of examples of algebraic groups. For instance, the norm 1 subgroup of $\mathbb{H}$ modulo $\pm 1$ is isomorphic to the Lie group SO(3) of rotations of Euclidean 3-space. (The norm 1 subgroup of $\mathbb{H}$ itself is isomorphic to the special unitary group SU(2).) Quaternion algebras lead to the important notion of inner forms of algebraic groups. Thus understanding quaternion algebras is useful for the theory of algebraic groups as well as automorphic representations. Quaternion algebras are also directly related to important objects in algebraic geometry, such as supersingular elliptic curves and Shimura curves. There are also applications to hyperbolic geometry and error correcting codes. Moreover, many of these applications have generalizations upon considering more general division algebras and even non-associative algebras such as the octonions.

The great number theorist Eichler famously illustrated connections between important topics with a hexagon:

[1] Quaternion algebras are not necessarily the best way to answer the specific question: does a quaternary quadratic form represent all integers? This is, at least in part, because one does not get arbitrary quaternary quadratic forms out of quaternion algebras. Rather, a better case for the utility of quaternion algebras for quadratic forms is to be made either for the case of ternary quadratic forms or for studying *composition* of quaternary quadratic forms. I chose to discuss the relation with quaternary quadratic forms rather than ternary ones in this introduction because it's a bit simpler and makes a closer analogy with the relationship between binary quadratic forms and quadratic fields. In any case, applications to quadratic forms are not a major goal of this course.




Which is to say that quaternion algebras play an important role in the topics labelling the other 5 vertices. Personally, I think algebraic groups should be one of the vertices, maybe replacing algebraic functions. At OU, quaternion algebras are related to the interests of Lucy Lifschitz, Ameya Pitale, Alan Roche, Ralf Schmidt and me. If you like superficial coincidences, you could make another, more personal hexagon for OU:




In this course, we'll study algebras as they relate to number theory, specifically with a view towards applications to modular forms and automorphic representations. The main focus will be on quaternion algebras, but we will develop at least some of the theory in the more general context of (central simple) algebras, with an eye towards understanding how things may generalize. (Whereas quaternion algebras arise in the study of automorphic forms on GL(2) and GSp(4), considering automorphic forms on GL($n$) leads to more general central simple algebras.) Along the way, we'll also learn more about related aspects of algebra and number theory, such as algebraic number theory and algebraic groups, local-global phenomena, adelic analysis, and some aspects of the theories of quadratic forms, modular forms and automorphic representations.

Specifically, our first goal for the course will be to understand the structure of (finite-dimensional) associative algebras and specifically the classification of quaternion algebras over number fields. The classification is understood using local-global principals, which means we will need to also study quaternion algebras over $p$-adic fields. Here there are some relations with class field theory, though we won't focus on this.

Our second goal for the course is to delve into the arithmetic of quaternion algebras, meaning the integral structures (i.e., orders like $\mathcal{O}_B$ above), ideal theory and factorizations. Here again we will initially work in the context of more general algebras before specializing to get precise results for quaternion algebras. This will yield some applications to quadratic forms, as well as being crucial for the next part.

The third goal for the course is to explain some connections with modular forms and automorphic representations. We won't present a complete theory of modular forms on quaternion algebras, but study the, in some sense, simplest type of quaternionic modular forms, which will correspond to weight 2 holomorphic (elliptic or Hilbert) modular forms (via the so-called Jacquet–Langlands correspondence). Our approach (particularly the treatment of Hecke operators) may also be useful for readers interested in modular or automorphic forms. We will explain the Jacquet–Langlands correspondence in our simple setting and describe some applications. The Jacquet–Langlands correspondence and applications will be treated in a seminar-type expository style, and we will not attempt to prove it.

At the end of the course, time permitting, we will discuss some arithmetic of octonion algebras, which are related to quadratic forms in an analogous way to quaternion algebras, but in higher dimensions.

### Prerequisites

I will expect students to be familiar with the material from a standard graduate algebra class (primarily: groups, rings, fields, and Galois theory—exposure to modules is helpful, but I will review definitions and basic results) and know some basic algebraic number theory (number fields, rings of integers, ideals and class groups, but not class field theory). Some exposure to $p$-adic numbers would be helpful, but I include a brief review of $p$-adic fields. Some basic point-set topology will also be used.

It is not necessary to be familiar with modular forms or automorphic representations in advance, though any familiarity beforehand will be helpful for the introduction to the Jacquet–Langlands correspondence. There is one subsection where I will explain the passage from the usual representation-theoretic statement of the Jacquet–Langlands to the classical version in which I assume familiarity with automorphic representations, but the reader not familiar with these things can just skip to the statement in classical language.

### To the reader

There are exercises woven throughout the text. It is intended that you do the most of exercises before proceeding further in the notes. Many of them are very simple, and just meant as a quick example that I think is a waste of electrons for me to write, but not a waste of graphite for you to work out. Of course, there are more challenging (as well as more tangential) exercises interspersed throughout, but I will usually not differentiate which problems I think are harder or easier, or more important or less important (and you may have a different opinion than I).

There are likely many misprints and some (I hope just minor) errors throughout the text. I would appreciate it if you inform me of the mistakes you find so that I can correct them.

As a caution to the reader, we mark off certain assumptions and warnings in an enviroment like this:

This is just a warning. In the event of an actual emergency, instructions to save the foundations of mathematics will be displayed here.

## References

There are many treatments of quaternions and quaternion algebras. Here are some references that I used and some other standard references, which you may find helpful throughout the course or in further studies.

For the basic facts about the arithmetic of quaternion algebras, Vigneras' book [Vig80] (in French) is the canonical reference. Vigneras refers to Reiner's book [Rei03] for proofs at some points, which treats more general algebras, which we will also study to some extent. Much of the arithmetic structure of quaternion algebras is also discussed in 3 chapters of the book by Maclachlan and Reid [MR03] (in English), which applies the theory to hyperbolic 3-manifolds.

John Voight's book *The arithmetic of quaternion algebras* is still in progress, but will be hopefully completed soon and currently an incomplete version may be downloaded on the author's website. When finished, it should provide a rather comprehensive reference for the arithmetic of quaternion algebras.

For the general structure of (central) simple algebras, with some specific information about quaternion algebras, there are several other books such as Pierce [Pie82], Gille–Szamuely [GS06] and Berhuy–Oggier [BO13] (the latter explains applications to error-correcting codes). However, only Reiner's book seriously considers the arithmetic of these algebras. Weil's book [Wei95] also treats some aspects of central simple algebras (including some aspects of arithmetic), where they are used to establish class field theory.

Conway and Smith's cute little book [CS03] discusses some aspects of arithmetic and geometry of quaternions and octonions, including a study of factorization theory which is lacking from most treatments. However, it is light on the algebraic theory (being more focused on geometry) and does not consider general quaternion algebras, which are important for connections with modular forms.

Much of the arithmetic theory, with a view towards applications to modular forms, is also summarized in various papers, e.g., Pizer's papers [Piz80], [Piz76] and [Piz77] and Gross's paper [Gro87] for working over $\mathbb{Q}$, or Dembele–Voight's paper [DV13] for working over totally real fields.

## Remarks on notation

Here are some comments that probably go without saying.

I am not French, so the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ do not include 0.

For us, ring means *unital* ring, i.e., rings have 1. For rings we allow $1 = 0$, but for fields we do not. What a double standard.

For a nonzero ring $R$ (so $1 \neq 0$), $R^\times$ denotes the unit (or multiplicative) group of invertible elements under multiplication.

For sets $A$ and $B$, $A \subset B$ means $A$ is contained in $B$, not that $A$ is properly contained in $B$. Of course $a < b$ does not mean $a \leq b$. Another double standard.

For an extension of number fields $K/F$, we denote the norm map by $N_{K/F}$, which we often abbreviate at $N$ if the fields $K$ and $F$ are clear. We will follow a similar convention for $p$-adic fields.

Positive (for real numbers) will mean $\geq 0$ and strictly positive means $> 0$, and similarly for negatives. By this logic non-negative should mean not $\leq 0$, i.e., strictly positive, but of course it just means positive, and I may occasionally say non-negative to emphasize I don't mean strictly positive. This creates the further conundrum that not non-negative is not (not (not negative)) (which is not to say not non-negative equals not not negative). Crap. This is not going well. Maybe we should just stop the course now. Or not not stop, which I can no longer be sure is the same as stop.

The following conventions I will explain again in the text, but this is a heads up so you don't get a heart attack.[2]

By a $p$-adic field, I mean a finite extension of some $\mathbb{Q}_p$ (rather than just some $\mathbb{Q}_p$), i.e., a nonarchimedean local field of characteristic 0.

If $F$ is a number field or a $p$-adic field, I may denote the ring of integers of $F$ by $\mathcal{O}_F$ or $\mathfrak{o}_F$. The latter usage will be to help you avoid confusion when we start using $\mathcal{O}$'s to denote orders over $\mathfrak{o}_F$ inside algebras over $F$.

The localization of a ring $\mathcal{O}$ at an ideal $\mathfrak{p}$ will be denote by $\mathcal{O}_{(\mathfrak{p})}$, not $\mathcal{O}_{\mathfrak{p}}$.

An ideal *in* an order $\mathcal{O}$ means an an integral ideal, where as an ideal *of* an order $\mathcal{O}$ means a fractional ideal. The term "ideal" by itself will by default mean fractional ideal.

For $R$ a ring, $R^{\times n}$ denotes the $n$-th powers in $R^\times$.

Unless otherwise stated, all of our algebras over a ring or field $R$ will be associative (except for the chapter on octonion algebras), unital and finite dimensional.

[2]Medical disclaimer: This is not a panacea to prevent all heart attacks. I'm not that kind of doctor. You still can't eat KFC 6 times a day.