Foreword

Mathematics continually surprises and delights us with how useful its most abstract branches turn out to be in the real world. Indeed, physicist Eugene Wigner's memorable phrase[1] "The unreasonable effectiveness of mathematics" captures a critical aspect of this utility. Abstract mathematical ideas often prove to be useful in rather "unreasonable" situations: places where one, a priori, would not expect them at all! For instance, no one who was not actually following the theoretical explorations in multi-antenna wireless communication of the late 1990s would have predicted that division algebras would turn out to be vital in the deployment of multi-antenna communication. Yet, once performance criteria for space-time codes (as coding schemes for multi-antenna environments are called) were developed and phrased as a problem of design of matrices, it was completely natural that division algebras should arise as a solution of the design problem. The fundamental performance criterion ask for $n \times n$ matrices M_i such that the difference of any two of the M_i is of full rank. To anyone who has worked with division algebras, the solution simply leaps out: any division algebra of index n embeds into the $n \times n$ matrices over a suitable field, and the matrices arising from the embedding naturally satisfy this criterion.

But there is more. Not only did division algebras turn out to be the most natural context in which to solve the fundamental design problem above, they also proved to be the correct objects to satisfy various other performance criteria that were developed. For instance, a second, and critical, performance criterion called the coding gain criterion turned out to be naturally satisfied by considering division algebras over number fields and using natural $\mathbb{Z}$-orders within them that arise from rings of integers of maximal subfields. Other criteria (for instance "DMG optimality," "good shaping," "information-losslessness" to name just a few) all turned out to be satisfied by considering suitable orders inside suitable division algebras over number fields. Indeed, this exemplifies another phenomenon Wigner describes: after saying that "mathematical concepts turn up in entirely unexpected connections," he goes on to say that "they often permit an unexpectedly close and accurate description of the phenomena in these connections." The match between division algebras and the requirement of space-time codes is simply uncanny.

The subject of multi-antenna communication has several unsolved mathematical problems still, for instance, in the area of decoding for large numbers of antennas. Nevertheless, division algebras are already being deployed for practical two-antenna

[1] Eugene P. Wigner, The unreasonable effectiveness of mathematics in the natural sciences, Comm. Pure Appl. Math., **13** Feb. 1960, 1–14

systems, and codes based on them are now part of various standards of the Institute of Electrical and Electronics Engineers (IEEE). It would behoove a student of mathematics, therefore, to know something about the applicability of division algebras while studying their theory; in parallel, it is vital for a communications engineer working in coding for multiple-antenna wireless to know something about division algebras.

Berhuy and Oggier have written a *charming* text on division algebras and their application to multiple-antenna wireless communication. There is a wealth of examples here, particularly over number fields and local fields, with explicit calculations, that one does not see in other texts on the subject. By pairing almost every chapter with a discussion of issues from wireless communication, the authors have given a very concrete flavor to the subject of division algebras. The book can be studied profitably not just by a graduate student in mathematics, but also by a mathematically sophisticated coding theorist. I suspect therefore that this book will find wide acceptability in both the mathematics and the space-time coding community and will help cross-communication between the two. I applaud the authors' efforts behind this very enjoyable book.

B.A. Sethuraman

Northridge, California

Introduction

A central simple algebra over a field k is a finite dimensional k-algebra with center k which does not have any proper two-sided ideals. The most elementary example is the Hamilton quaternion algebra. More generally, a division ring with center k can be viewed as a central simple k-algebra, where the algebra structure is induced by the multiplication law. Central simple algebras and division rings have been extensively studied, and have appeared in many other areas of mathematics, such as ring theory, number theory, representation theory of finite groups, algebraic geometry or classification theory of quadratic forms. Surprisingly, they have recently been proven useful in coding theory.

The ambition of this book is to provide an introduction to the theory of central simple algebras accessible at a graduate level, starting from scratch and including fundamental concepts such as splitting fields, Brauer group, crossed product algebras, index and exponent, as well as algebras with involution. Even though most of our exposition is rather classical, we have tried to focus on explicit techniques and examples, most of them coming from coding theory. The codes presented in this book are there to illustrate the theory of central simple algebras, and do not give an exhaustive view of the work done on the theory of algebraic space-time coding.

The use of division algebras for space-time coding is usually attributed to the seminal work by B. A. Sethuraman et al. [**48**]. Number fields and cyclic algebras were discussed, which have been a favourite tool for space-time design (see for example [**12, 6, 40, 13, 32, 55**]). Other algebras have been explored, such as Clifford algebras [**27**], or crossed product algebras (e.g. [**57**]).

Alternative studies considered the use of maximal orders (e.g. [**56**]) or non-associative algebras (e.g. [**42**]).

Some surveys on coherent space-time coding [**36, 45**] and one survey on non-coherent space-time coding [**35**] are now available. These works are just representing a few of the different approaches studied so far in the area of space-time coding, which is still an active field of research. These are just pointers for the interested reader, and by no mean provide a complete list.

In Chapter I, we introduce the concept of a central simple k-algebra and give the first examples of such algebras, including quaternion algebras. We then explain how they can be embedded into matrix algebras, and how this result may be used in coding theory. In Chapter II, we have a closer look at the properties of quaternion algebras. We also prove that the only finite dimensional division $\mathbb{R}$-algebras are, up to isomorphism, $\mathbb{R}, \mathbb{C}$ or the Hamilton quaternion algebra $\mathbb{H}$. We then provide examples of quaternion based codes. The results presented in Chapter III are the

core of the theory. We first study the stability of central simple k-algebras under algebraic operations such as tensor product or base field extension. We then prove that any central simple k-algebra is isomorphic to a matrix algebra over a central division k-algebra, and establish that every k-automorphism of such an algebra is inner. We also focus on the structure of the centralizer of a simple subalgebra, which is a crucial tool in the study of maximal subfields and splitting fields of central simple algebras, which will be developed in Chapter IV. As an application of this theory, we define the reduced characteristic polynomial of an element of a central simple algebra, and introduce the concept of the reduced norm, which generalizes the determinant of a matrix. The latter can in turn be used to reinterpret code parameters. In Chapter V, we define the Brauer group $\mathrm{Br}(k)$ of a field k, which allows us to study globally all central simple k-algebras. We show that this group is an abelian torsion group, and use this result to define the exponent of a central simple k-algebra. We end this chapter by establishing the existence of a primary decomposition of a central simple k-algebra. In Chapter VI, we characterize central simple algebras which have a Galois maximal subfield. This leads to the notion of a crossed product algebra. We then present the standard results on these particular algebras. At the end of this chapter, crossed product algebras are used to construct families of codes. Chapter VII is devoted to cyclic algebras, that is, the case where the Galois maximal subfield is cyclic. At this occasion, an overview of the theory of central simple algebras over local and number fields is given without proofs. Explicit criteria to decide whether a given central simple algebra over a global field is division are established. Finally, these criteria are used to design codes based on cyclic division algebras. Chapter VIII focuses on central simple k-algebras of degree 4. We show that these algebras are crossed products over a biquadratic extension L/k, and a full description by generators and relations is given. We also provide a criterion to check if such an algebra is division in terms of the parameters defining the algebra when k is a number field, and applications to code constructions are given. In Chapter IX, the concept of a unitary involution on a central simple algebra is defined. The existence of unitary involutions is then investigated. We particularly focus on the case of crossed product algebras. We then explain how central simple algebras with a unitary involution may be used in coding theory via the construction of unitary codes, and we give various examples.

We would like to sincerely thank N. Markyn, S. Pumpluen, A. Quéguiner-Mathieu, B.A. Sethuraman, J.-P. Tignol, T. Unger and R. Vehkalahti for their careful reading of substantial parts of this book. Their pertinent comments enabled us to dramatically improve the quality of the exposition.

CHAPTER I

Central simple algebras

This chapter contains the necessary definitions and background on central simple algebras. After some preliminaries on k-algebras and tensor products, we introduce central simple algebras and give some examples. We then show how to identify central simple algebras with matrix subalgebras. As a first illustration, we explain how central simple algebras may be used in coding theory, and examples of code constructions are presented.

I.1. Preliminaries on k-algebras

In the sequel, k will denote an arbitrary field.

Definition I.1.1. A **k-algebra** is a pair (A, μ), where A is a k-vector space and $\mu : A \times A \longrightarrow A$ is a k-bilinear map, called the **product law** of A. We write aa' for $\mu(a, a')$, and call it the **product** of the elements a and a'.

A k-algebra A is called **associative** (resp. **commutative**, resp. **unital**) if the product law is associative (resp. commutative, resp. has a unit element 1_A).

Examples I.1.2.

(1) The ring of polynomials $k[X]$ is a commutative, associative and unital k-algebra.
(2) If L/k is a field extension, then L is a commutative, associative and unital k-algebra.

□

Definition I.1.3. A **k-algebra morphism** is a k-linear map $f : A \longrightarrow B$ satisfying

$$f(aa') = f(a)f(a') \text{ for all } a, a' \in A.$$

If A and B are unital, we require in addition that $f(1_A) = 1_B$. A **k-algebra isomorphism** is a k-algebra morphism which is bijective. In this case, the inverse map f^{-1} is also a k-algebra morphism.

Definition I.1.4. A **subalgebra** of a k-algebra A is a linear subspace B of A which is closed under the product. If A is unital, we require in addition that $1_A \in B$. It is unital, (resp. associative, resp. commutative) whenever A is.

Examples I.1.5.

(1) The intersection of an arbitrary family of subalgebras of a k-algebra A is again a subalgebra of A.

(2) The image of any k-algebra morphism $f : A \longrightarrow B$ is a subalgebra of B.

□

DEFINITION I.1.6. The **center** of a k-algebra A is by definition the set

$$Z(A) = \{z \in A \mid az = za \text{ for all } a \in A\}.$$

It is a commutative subalgebra of A whenever A is associative.

EXAMPLE I.1.7. The matrix algebra $\mathrm{M}_n(k)$, consisting of $n \times n$ matrices with entries from k, is a unital k-algebra with center k (we identify k with the set of scalar matrices). □

REMARK I.1.8. If A is an associative unital k-algebra, then addition and product naturally endow A with a ring structure. In particular, every subalgebra of A is also a subring, and every k-algebra morphism is also a ring morphism. Moreover if $1_A \neq 0_A$ (i.e. A is not zero), k identifies with a subalgebra of $Z(A)$ (hence a subalgebra of A).

Indeed the k-bilinearity of the product law and the properties of 1_A imply that we have

$$(\lambda{\cdot}1_A)a = 1_A(\lambda{\cdot}a) = (\lambda{\cdot}a)1_A = a(\lambda{\cdot}1_A)$$

for all $a \in A$ and $\lambda \in k$, so $k{\cdot}1_A \subset Z(A)$. One may verify that $k{\cdot}1_A$ is a k-subalgebra of $Z(A)$. Hence the map

$$\begin{aligned} k &\longrightarrow Z(A) \\ \lambda &\longmapsto \lambda{\cdot}1_A \end{aligned}$$

is a non-trivial k-algebra morphism, which is injective since k is a field. □

In this book, all k-algebras will implicitly be assumed to be unital, associative, and finite-dimensional over k. Moreover, we will systematically identify k and $k{\cdot}1_A$.

DEFINITION I.1.9. A **division** k-algebra is a k-algebra which is also a division ring (that is, every non-zero element is invertible).

At this stage, it may be worth making a few remarks on subalgebras of finite dimensional division algebras generated by a single element.

Let D be a finite dimensional division k-algebra, and let $d \in D$. We denote by $k[d]$ the smallest subalgebra of D containing d, and by $k(d)$ the smallest division subalgebra of D containing d. Clearly, we have

$$k[d] = \{P(d) \mid P \in k[X]\}.$$

Since D is finite dimensional over k, so is $k[d]$. Therefore, the successive powers of d cannot be linearly independent, and the evaluation morphism

$$ev_d : \begin{aligned} k[X] &\longrightarrow D \\ P &\longmapsto P(d) \end{aligned}$$

cannot be injective. Hence, its kernel is generated by a unique monic polynomial $\mu_{d,k} \in k[X]$, and we have an isomorphism of k-algebras

$$k[X]/(\mu_{d,k}) \cong_k k[d].$$

Since D has no zero divisors, $k[d]$ is an integral domain and $(\mu_{d,k})$ is a prime ideal, hence maximal. Thus $k[d]$ is a field, $k[d] = k(d)$ and we have

$$[k(d) : k] = \deg(\mu_{d,k}).$$

Moreover, $\mu_{d,k}$ is irreducible since it generates a maximal ideal of $k[X]$, and $\mu_{d,k}(d) = 0$.

We will use these facts without further reference from now on.

Definition I.1.10. Let D be a division k-algebra, and let $d \in D$. The polynomial $\mu_{d,k}$ is called **the minimal polynomial** of $d \in D$ over k.

We now recall the main properties of the tensor product of k-algebras.

If A and B are k-algebras, their **tensor product** $A \otimes_k B$ may be viewed as the k-vector space spanned by the symbols $a \otimes b, a \in A, b \in B$ subject to the relations:

$$(a + a') \otimes b = a \otimes b + a' \otimes b$$
$$a \otimes (b + b') = a \otimes b + a \otimes b'$$
$$(\lambda a) \otimes b = a \otimes (\lambda b) = \lambda(a \otimes b)$$

for all $a, a' \in A, b, b' \in B, \lambda \in k$. The symbols $a \otimes b$ are called **elementary tensors**.

The product on $A \otimes_k B$ is the unique product law satisfying

$$(a \otimes b)(a' \otimes b') = aa' \otimes bb' \text{ for all } a, a' \in A, b, b' \in B.$$

If $(e_i)_{i \in I}$ and $(e'_j)_{j \in J}$ are k-bases of A and B as k-vector spaces, then $(e_i \otimes e'_j)_{(i,j) \in I \times J}$ is a k-basis of $A \otimes_k B$. In particular $A \otimes_k B$ is finite-dimensional as a k-vector space if and only A and B are, and in this case we have

$$\dim_k(A \otimes_k B) = \dim_k(A) \dim_k(B).$$

Moreover, if $\varphi : A \longrightarrow C$ and $\psi : B \longrightarrow C$ are two morphisms of unital k-algebras satisfying

$$\varphi(a)\psi(b) = \psi(b)\varphi(a) \text{ for all } a \in A, b \in B,$$

there exists a unique morphism $h : A \otimes_k B \longrightarrow C$ of unital k-algebras satisfying

$$h(a \otimes 1_B) = \varphi(a) \text{ and } h(1_A \otimes b) = \psi(b) \text{ for all } a \in A, b \in B.$$

In particular, if $f : A \longrightarrow A'$ and $g : B \longrightarrow B'$ are two morphisms of unital k-algebras, there exists a unique k-algebra morphism

$$f \otimes g : A \otimes_k B \longrightarrow A' \otimes_k B'$$

satisfying

$$(f \otimes g)(a \otimes b) = f(a) \otimes g(b) \text{ for all } a \in A, b \in B.$$

If f and g are isomorphisms, so is $f \otimes g$.

Finally, if A and B are unital, the k-algebra morphisms

$$\begin{array}{c} A \longrightarrow A \otimes_k B \\ a \longmapsto a \otimes 1_B \end{array} \quad \text{and} \quad \begin{array}{c} B \longrightarrow A \otimes_k B \\ b \longmapsto 1_A \otimes b \end{array}$$

are injective.

Now let L/k be an arbitrary field extension. If A is a k-algebra and B is an L-algebra, then $A \otimes_k B$ has a natural structure of L-algebra, where the structure of L-vector space is defined on elementary tensors by

$$\lambda\cdot(a \otimes b) = a \otimes \lambda b \text{ for all } \lambda \in L, a \in A, b \in B.$$

In particular, $A \otimes_k L$ has a natural structure of an L-algebra. Moreover, $A \otimes_k L$ is finite dimensional over L if and only if A is finite dimensional over k. In this case, we have

$$\dim_L(A \otimes_k L) = \dim_k(A).$$

If A and B are unital, we have a natural isomorphism of L-algebras

$$(A \otimes_k L) \otimes_L B \cong_L A \otimes_k B.$$

Similarly, $B \otimes_k A$ and $L \otimes_k A$ have a natural structure of L-algebras, and we have an isomorphism of L-algebras

$$B \otimes_L (L \otimes_k A) \cong_L B \otimes_k A.$$

If now A and B are two unital k-algebras, we have a natural L-algebra isomorphism

$$(A \otimes_k B) \otimes_k L \cong_L (A \otimes_k L) \otimes_L (B \otimes_k L).$$

Finally, if $k \subset K \subset L$ is a tower of field extensions, we have

$$(A \otimes_k K) \otimes_K L \cong_L A \otimes_k L.$$

The justification of the tensor product properties described above is quite lengthy, so we leave the details for now. For the sake of completeness, the reader may find full constructions and proofs in Appendix A.

We end this section with an elementary lemma.

LEMMA I.1.11. *Let A be a k-algebra, let $n \geq 1$ be an integer and let L/k be a field extension. Then the following properties hold:*

(1) *we have a natural k-algebra isomorphism $\mathrm{M}_n(k) \otimes_k A \cong_k \mathrm{M}_n(A)$;*
(2) *we have a natural L-algebra isomorphism $\mathrm{M}_n(k) \otimes_k L \cong_L \mathrm{M}_n(L)$.*

Proof.

(1) The k-algebra morphisms

$$\begin{array}{c} \mathrm{M}_n(k) \longrightarrow \mathrm{M}_n(A) \\ M \longmapsto M \end{array} \quad \text{and} \quad \begin{array}{c} A \longrightarrow \mathrm{M}_n(A) \\ a \longmapsto aI_n \end{array}$$

have commuting images, and therefore there is a unique k-algebra morphism $\varphi : \mathrm{M}_n(k) \otimes_k A \longrightarrow \mathrm{M}_n(A)$ satisfying

$$\varphi(M \otimes a) = aM, \text{ for all } M \in \mathrm{M}_n(k), a \in A.$$

Since $\mathrm{M}_n(k) \otimes_k A$ and $\mathrm{M}_n(A)$ have the same dimension over k, it suffices to prove that φ is surjective. Let E_{ij} be the matrix with coefficient 1 at row i and column j and coefficients 0 elsewhere. For any matrix $M' = (m'_{ij}) \in \mathrm{M}_n(A)$, we have

$$\varphi\big(\sum_{i,j} E_{ij} \otimes m'_{ij}\big) = M',$$

which proves the surjectivity of φ.

(2) By (1), we have an isomorphism of k-algebras $\mathrm{M}_n(k) \otimes_k L \cong_k \mathrm{M}_n(L)$. One may check that this isomorphism is also L-linear. □

REMARK I.1.12. In particular, we have a natural isomorphism

$$\mathrm{M}_m(k) \otimes_k \mathrm{M}_n(k) \cong_k \mathrm{M}_{mn}(k)$$

which maps $M \otimes N$ onto the Kronecker product of M and N. □

I.2. Central simple algebras: the basics

We now define the main object of this book.

DEFINITION I.2.1. Let k be a field. A k-algebra A is **simple** if it has no non-trivial two-sided ideals.

The next lemma gives an elementary but very useful property of simple algebras.

LEMMA I.2.2. *Let k be a field, and let $\phi : A \longrightarrow B$ be a k-algebra morphism. If A is simple, then ϕ is injective. If moreover A and B are finite dimensional over k and $\dim_k(A) = \dim_k(B)$, then ϕ is an isomorphism.*

Proof. Assume that A is simple. Since $\ker(\phi)$ is a two-sided ideal of A, we have $\ker(\phi) = (0)$ or A. The latter case cannot happen since $\phi(1) = 1$. Hence ϕ is injective; the last part is clear. □

We now give examples of simple algebras.

EXAMPLES I.2.3.

(1) Any division ring D is a simple $Z(D)$-algebra.

(2) Let k be an arbitrary field. Then $\mathrm{M}_n(k)$ is a simple k-algebra.

Indeed, let J be a non-zero ideal of $\mathrm{M}_n(k)$, and let $M = (m_{ij})_{i,j}$ be a non-zero element of J. Fix two integers r, s such that $m_{rs} \neq 0$. For all $i = 1, \ldots, n$, we have

$$m_{rs}^{-1} E_{ir} M E_{si} = E_{ii},$$

and therefore,

$$I_n = \sum_i E_{ii} = \sum_i m_{rs}^{-1} E_{ir} M E_{si} \in J$$

since J is a two-sided ideal. Hence J contains a unit, so $J = \mathrm{M}_n(k)$.

(3) Similar arguments show that if D is a division k-algebra, then $\mathrm{M}_r(D)$ is a simple k-algebra for all $r \geq 1$.

□

We now give our first concrete example of a simple k-algebra. Let k be a field of characteristic different from 2.

Let $a, b \in k^\times$, and consider the k-linear subspace $(a, b)_k$ of $\mathrm{M}_4(k)$ generated by the matrices

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, i = \begin{pmatrix} 0 & a & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & a \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$j = \begin{pmatrix} 0 & 0 & b & 0 \\ 0 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, ij = \begin{pmatrix} 0 & 0 & 0 & -ab \\ 0 & 0 & b & 0 \\ 0 & -a & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Straightforward computations show that these matrices are linearly independent over k, and that we have

$$i^2 = a, j^2 = b, (ij)^2 = -ab \text{ and } ji = -ij.$$

It easily follows that $(a,b)_k$ is a k-subalgebra of $\mathrm{M}_4(k)$ of dimension 4 over k.

Definition I.2.4. Let k be a field of characteristic different from 2. The k-algebra $(a,b)_k$ is called a **quaternion k-algebra**.

Proposition I.2.5. *Let k be a field of characteristic different from 2. For every $a, b \in k^\times$, the k-algebra $(a,b)_k$ is a simple k-algebra, with center isomorphic to k.*

Proof. Let us first determine the center of $(a,b)_k$.

Let $q_1 = x + yi + zj + tij \in (a,b)_k$ and assume that $q_1 \in Z((a,b)_k)$. Then we have

$$iq_1 = i(x + yi + zj + tij) = xi + ay + zij + taj$$

and

$$q_1 i = (x + yi + zj + tij)i = xi + ay - zij - taj.$$

Since by assumption $iq_1 = q_1 i$, we have therefore $z = t = 0$ and thus $q_1 = x + yi$. Since we have $jq_1 = q_1 j$, we get $xj - yij = xj + yij$ in a similar way, so $y = 0$ and $q_1 = x \in k$. Hence $Z((a,b)_k) = k$.

Let us prove now that $(a,b)_k$ is simple. For, let I be a non-zero two-sided ideal of $(a,b)_k$, and let $q_1 = x + yi + zj + tij \in I, q_1 \neq 0$. We then have

$$\frac{1}{2}(iq_1 - q_1 i) = zij + taj \in I \text{ and } \frac{1}{2}(iq_1 + q_1 i) = xi + ay \in I.$$

Since by assumption x, y, z or t is non-zero, it follows that $zij + taj$ or $xi + ay$ is non-zero. Assume for example that $q_2 = zij + taj$ is not zero, that is $z \neq 0$ or $t \neq 0$. We have

$$\frac{1}{2}(jq_2 - q_2 j) = -bzi \in I \text{ and } \frac{1}{2}(jq_2 + q_2 j) = tab \in I.$$

If $t \neq 0$, then $tab \in k^\times$ is a unit of $(a,b)_k$; if $z \neq 0$, then $-bzi \in k^\times$ is a unit of $(a,b)_k$ (with inverse $-(abz)^{-1}i$). In both cases, I contains a unit, so $I = (a,b)_k$. The case $xi + ay \neq 0$ may be dealt with in a similar way and is left to the reader. □

Remark I.2.6. Later on, we will see a criterion to decide whether or not $(a,b)_k$ is a division algebra. For the moment, let us just point out that it can actually be a division algebra for some well-chosen values of a and b. For example, if $k = \mathbb{R}$ and $a = b = -1$, we obtain the Hamilton quaternion algebra $\mathbb{H}$, which is known to be a division ring. We will recover this fact in the next chapter. □

Definition I.2.7. A k-algebra A is called **central** if $Z(A) = k$. A **central simple** k-algebra is a k-algebra which is central and simple.

Examples I.2.8.

(1) The k-algebra $\mathrm{M}_n(k)$ is central simple.

(2) If D is a division ring, then its center $Z(D)$ is a field and D is a central $Z(D)$-algebra.

(3) If D is a central division k-algebra, then $\mathrm{M}_r(D)$ is a central simple k-algebra for all $r \geq 1$.

Indeed, the fact that $\mathrm{M}_r(D)$ is a simple k-algebra was already pointed out in Example I.2.3 (3). Now if $M \in Z(\mathrm{M}_r(D))$, the equality

$$E_{ij}M = ME_{ij} \text{ for all } i, j$$

shows that M is diagonal and that $m_{11} = \cdots = m_{rr}$. Let us denote this common value by d. The fact that $M \in Z(\mathrm{M}_r(D))$ then easily implies that $d \in Z(D) = k$. Hence $M \in Z(D)$, so $Z(\mathrm{M}_r(D)) = k$ (where k is identified with the set of scalar matrices).

(4) If L/k is a field extension, then L is a simple k-algebra which is not central.

(5) By Proposition I.2.5, any quaternion k-algebra is a central simple k-algebra.

□

Definition I.2.9. We say that a central simple k-algebra is **split** if it is isomorphic to a matrix algebra.

Matrix algebras are in some sense the simplest examples of central simple algebras. Even if not all simple algebras are split, they can be naturally viewed as subalgebras of matrix algebras, as we proceed to show now. This property is particularly interesting for explicit computations. Let us give a definition first.

Definition I.2.10. Let A be a k-algebra. A **subfield** of A is a commutative subalgebra L of A which is also a field. In particular, A is a right L-vector space. Moreover, L contains k since it is a k-algebra. However, notice that A may not be an L-algebra (unless $L = k$), since L does not necessarily commute with all the elements of A.

We may now state the next result.

Lemma I.2.11. *Let A be a k-algebra, and let L be a subfield of A. For all $a \in A$, the map*

$$\ell_a : \begin{array}{c} A \longrightarrow A \\ z \longmapsto az \end{array}$$

is an endomorphism of the right L-vector space A, and the map

$$\phi : \begin{array}{c} A \longrightarrow \mathrm{End}_L(A) \\ a \longmapsto \ell_a \end{array}$$

is a k-algebra morphism. In particular, if A is simple, ϕ is injective.

Proof. Recall that the structure of L-vector space on $\mathrm{End}_L(A)$ is defined by

$$\begin{array}{c} \mathrm{End}_L(A) \times L \longrightarrow \mathrm{End}_L(A) \\ (u, \lambda) \longmapsto u\lambda, \end{array}$$

where

$$(u\lambda)(z) = u(z)\lambda \text{ for all } z \in A.$$

Let us check that ℓ_a is an endomorphism of the right L-vector space A and that the map

$$\phi\colon \begin{array}{c} A \longrightarrow \mathrm{End}_L(A) \\ a \longmapsto \ell_a \end{array}$$

is a k-algebra morphism. We have

$$\ell_a(z+z') = a(z+z') = az + az' = \ell_a(z) + \ell_a(z'),$$

for all $z, z' \in A$. Moreover, for all $\lambda \in L$, we have

$$\ell_a(z\lambda) = a(z\lambda) = (az)\lambda = \ell_a(z)\lambda = (\ell_a\lambda)(z).$$

Hence ℓ_a is an endomorphism of the right L-vector space A.

Clearly, we have $\ell_1 = \mathrm{Id}_A$. Moreover, for every $a, a', z \in A$ and $\xi \in k$, we have

$$\ell_{a+a'}(z) = (a+a')z = az + a'z = \ell_a(z) + \ell_{a'}(z),$$

$$\ell_{aa'}(z) = aa'z = a(a'z) = \ell_a(a'z) = (\ell_a \circ \ell_{a'})(z),$$

$$\ell_{a\xi}(z) = (a\xi)z = a(\xi z) = a(z\xi) = (az)\xi = (\ell_a\xi)(z).$$

We then get

$$\ell_{a+a'} = \ell_a + \ell_{a'}, \ell_{aa'} = \ell_a \circ \ell_{a'}, \ell_{a\xi} = \ell_a\xi,$$

and the result follows. The last part comes from Lemma I.2.2. □

Remark I.2.12. Let A be a **simple** k-algebra. Let us choose a basis of the right L-vector space A, and set

$$m = \dim_L(A) = \frac{\dim_k(A)}{[L:k]}.$$

Then composing the injective k-algebra morphism ϕ defined in the previous lemma with the isomorphism $\mathrm{End}_L(A) \cong_k \mathrm{M}_m(L)$ gives rise to an **injective** k-algebra morphism

$$\varphi_{A,L}\colon \begin{array}{c} A \hookrightarrow \mathrm{M}_m(L) \\ a \longmapsto M_a, \end{array}$$

where M_a is the matrix of left multiplication by a in the chosen L-basis of A.

For example, if $L = k$, we obtain an injection $\varphi_{A,k} : A \hookrightarrow \mathrm{M}_d(k)$, where $d = \dim_k(A)$. □

The reader will find plenty of examples of computations of such an injection $\varphi_{A,L}$ in the next section and the other chapters, as we will use it to provide all code constructions presented in this book.

I.3. Introducing space-time coding

Coding theory deals with the problem of transmitting data reliably over a communication channel which is noisy. The coding problem addressed depends on the characteristics of the channel. In classical coding theory, the channel involves a transmitter and a receiver, with between the two of them a discrete channel given by

$$y = x + v,$$

where x, y, v are n-dimensional vectors over a finite field F (typically of characteristic 2). The vector x is the transmitted signal, the vector v is the noise, and y is the noisy received vector. Transmission takes place during n time slots. A linear code over F is a subspace of F^n, where n is called the length of the code. Encoding consists of mapping a string of data of length k into a redundant coded version of length n. We call k the rank or dimension of the code. Vectors in F^n are called codewords. The minimum distance of a code is the smallest Hamming distance between two distinct codewords, where the Hamming distance counts the number of entries in which the two codewords differ. Since the minimum distance is a performance parameter of the code, a fundamental problem in the design of such codes is to find codes with large rank and minimum distance with respect to the block length.

The above channel model was introduced for wired discrete channels. There exists an analogous model for continuous channels, given by

$$y = x + v,$$

where x, y, v are now n-dimensional vectors over the complex field $\mathbb{C}$, and the noise vector v has random independent and identically distributed (i.i.d.) Gaussian entries with zero mean and unit variance. This is called a Gaussian channel, and it has a corresponding coding theory of its own. A further generalization appeared with the introduction of wireless communication. Now transmitter and receiver are both equipped with one antenna (see Fig. 1). Let $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ be the signal to be transmitted. At time t, $t = 1, \ldots, n$ the transmit antenna sends x_t, which will reach the receive antenna via different paths, due to the nature of the wireless environment. This is taken into account in the channel model, given by

$$y_t = x_t h_t + v_t, \ t = 1, \ldots, n,$$

where the coefficients h_t and v_t model respectively fading (coming from the signal propagation through multipaths) and noise. The wireless channel from the transmitter to the receiver during n time slots can thus be modeled as follows:

$$y = xH + v,$$

where $y \in \mathbb{C}^n$ is the received vector, and H is a diagonal $n \times n$ matrix called the **fading matrix** or **channel matrix**. All noise and fading coefficients are assumed to be i.i.d. complex Gaussian random variables with zero mean and unit variance.

In order to transmit more and more data in wireless environments, systems with multiple antennas at both transmitter and receiver have been introduced. They are commonly called Multiple Input Multiple Output (MIMO) systems or channels. Let us first consider a channel with two transmit and two receive antennas. At time t, the first and second antennas respectively send x_{1t} and x_{2t}. Both signals will be

received by the two receive antennas, and will follow different paths to reach each of them. The signals y_{1t}, y_{2t} sensed by each receive antenna are

$$\begin{aligned} y_{1t} &= h_{11}x_{1t} + h_{12}x_{2t} + v_{1t} \\ y_{2t} &= h_{21}x_{1t} + h_{22}x_{2t} + v_{2t} \end{aligned}$$

where h_{ji} denotes the fading from the ith transmit antenna to the jth receive antenna, and v_{jt} denotes the noise at the jth receive antenna at time t. Note that in the above equations, the fading coefficients h_{ji} should depend on t. However, it is reasonable to assume that the environment does not change so fast, and that there is a period of time T during which the channel (that is h_{ji}) remains constant. This period T is called a **coherence interval**, and the length of T depends on the channel considered.

For example, let us assume that the channel stays approximately constant over a period of length $T = 2$, and the transmission starts at time $t = 1$. The first antenna transmits at time $t = 1$ and $t + 1 = 2$ the signals x_{11} and x_{12} respectively. Similarly, the second antenna transmits at time t and $t + 1$ the signals x_{21} and x_{22} respectively. The first antenna receives consecutively a signal which is the sum of the two transmitted signals with fading and some noise, that is

$$\begin{aligned} y_{11} &= h_{11}x_{11} + h_{12}x_{21} + v_{11} \\ y_{12} &= h_{11}x_{12} + h_{12}x_{22} + v_{12}. \end{aligned}$$

Similarly, the second antenna gets

$$\begin{aligned} y_{21} &= h_{21}x_{11} + h_{22}x_{21} + v_{21} \\ y_{22} &= h_{21}x_{12} + h_{22}x_{22} + v_{22}. \end{aligned}$$

This can be written in a matrix equation as

$$\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} + \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}.$$

This model can be generalized to the case where we have M transmit antennas and N receive antennas. At time t, the M antennas each send one signal. Those M signals can be collected and written as a vector $x_t = (x_{1t}, \ldots, x_{Mt})^t$. Each x_{it} will be received by all the N antennas. Thus x_{it} follows N different paths, each corresponding to a given fading denoted by h_{ji}, $j = 1, \ldots, N$ to reach its N destinations. Now, each receive antenna will sense a signal, which is the sum of noisy and faded copies of the signals transmitted by all antennas. Let us now consider T instances of the transmission, where we recall that T is the coherence time interval, during which the channel is assumed to be constant. The model for transmission with multiple antennas over a coherence time T can be summarized as follows:

$$\mathbf{Y}_{N\times T} = \mathbf{H}_{N\times M}\mathbf{X}_{M\times T} + \mathbf{V}_{N\times T},$$

where all matrices have coefficients in $\mathbb{C}$, and their dimensions are written as subscript. The t^{th} column of the matrix $\mathbf{X}$ contains the vector x_t, sent at time t. The matrices $\mathbf{H}$ and $\mathbf{V}$ are random matrices whose coefficients are complex i.i.d. Gaussian random variables.

Coding for the above MIMO channel consists of designing the codewords $\mathbf{X}$ as a function of the data (or **information symbols**), which typically adds redundancy,

similarly to the classical case. The set of codewords is called a **codebook** or simply a code. We will typically consider linear codes, that is, the encoding map from the information symbols to a codeword $\mathbf{X}$ will be linear. Since the data is encoded during **time** (we consider a time interval of T slots) and **space** (since we have M antennas), codes for multiple antennas systems are often called **space-time codes**.

We are now left with discussing how to design good space-time codes. In what follows, we denote by $\mathbf{X}^*$ the Hermitian conjugate of $\mathbf{X}$. Coding should be done so as to help the receiver to recover the transmitted signal $\mathbf{X}$ from the received signal $\mathbf{Y}$, despite the fading and noise. When there is no fading, a transmitted signal $\mathbf{X}$ will only be affected by noise. Geometrically, $\mathbf{X}$ can be seen as a point in an MT-dimensional space, and the received signal $\mathbf{Y}$ lies within a ball centered in $\mathbf{X}$ of radius given by the variance of the noise. In this case, the decoder which knows all the possible codewords can compute $\|\mathbf{X}-\mathbf{Y}\|^2$ for all possible $\mathbf{X}$ in the codebook, where the norm is the Frobenius norm: $\|A\|^2 = \mathrm{Tr}(AA^*)$, and Tr denotes the trace. It then decides that its estimate $\hat{\mathbf{X}}$ of $\mathbf{Y}$ is given by the matrix which minimizes $\|\mathbf{X}-\mathbf{Y}\|^2$. If the codewords are designed such that there is only one codeword in a ball of radius the variance of the noise, then the decoder will with high probability get the right estimate. The situation is different in case of fading.

Let us for now assume that the receiver has the knowledge of the channel $\mathbf{H}$. This is called the **coherent** case. The **non-coherent** case considers the scenario when the receiver does not know the channel, and will be discussed later on. A decoding rule is obtained as follows. Let $\mathcal{C}$ denote the codebook. The receiver knows $\mathbf{Y} = \mathbf{HX} + \mathbf{V}$, the codebook, and an estimate of $\mathbf{H}$. It thus computes the "faded" codebook $\{\mathbf{HX} \mid \mathbf{X} \in \mathcal{C}\}$ by multiplying every codeword by $\mathbf{H}$. It then chooses as decoded codeword the one which minimizes the distance between $\mathbf{HX}$ and $\mathbf{Y}$. We thus have that the decoded codeword $\hat{\mathbf{X}}$ is given by

$$\hat{\mathbf{X}} = \min_{\mathbf{X}\in\mathcal{C}} \|\mathbf{HX}-\mathbf{Y}\|^2.$$

An error will occur if the decoded codeword $\hat{\mathbf{X}}$ is different from the transmitted codeword $\mathbf{X}$. A way of formalizing the reliability of a channel is thus to compute its **pairwise probability of error**, namely, the probability of sending $\mathbf{X}$ and decoding erroneously $\hat{\mathbf{X}} \neq \mathbf{X}$. We write such probability $\mathbb{P}(\mathbf{X} \to \hat{\mathbf{X}})$. In [**52**], the following upper bound on this probability of error has been computed:

$$\mathbb{P}(\mathbf{X} \to \hat{\mathbf{X}}) \leq \left((\prod_{i=1}^{r} \lambda_i)^{1/r} c(\rho) \right)^{-rN}$$

where N is the number of receive antennas, r is the rank of the matrix $A(\mathbf{X},\hat{\mathbf{X}}) := (\mathbf{X}-\hat{\mathbf{X}})(\mathbf{X}-\hat{\mathbf{X}})^*$, and λ_i, $i = 1,\ldots,r$ are the non-zero eigenvalues of $A(\mathbf{X},\hat{\mathbf{X}})$. Furthermore, $c(\rho)$ is a constant that depends on channel parameters, and importantly on the **signal-to-noise ratio (SNR)** of the channel at the receiver, denoted by ρ, and defined by

$$\rho = \frac{\mathbb{E}[||\mathbf{HX}||^2]}{\mathbb{E}[||\mathbf{V}||^2]}.$$

We indeed expect the probability of error to depend on how strong the signal is compared to the noise occurring over the channel.

We call the negative exponent of $c(\rho)$ given here by rN in the above expression the **diversity order** of the pairwise error probability. The higher the diversity order is, the smaller the upper bound will be. Since $0 \leq r \leq M$, the best diversity order is MN, which is obtained when the matrix $A(\mathbf{X}, \hat{\mathbf{X}})$ is full rank.

Design criteria to build the codebook $\mathcal{C}$ are derived from the above upper bound. Codes that have parameters that minimize the bound will give the best performance. The design criteria are summarized as follows:

(1) **The rank criterion**: in order to achieve the maximum diversity MN, the matrix $A(\mathbf{X}, \hat{\mathbf{X}}) = (\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^*$ has to be full rank for any pair of codewords $\mathbf{X} \neq \hat{\mathbf{X}}$. Codes that achieve the maximal diversity are called **fully diverse**.

(2) **The determinant criterion**: once a given diversity is obtained, the minimum of the determinant of $A(\mathbf{X}, \hat{\mathbf{X}})$ taken over all pairs of distinct codewords must be maximized. If the code is not fully diverse, then the determinant is understood as the product of the non-zero eigenvalues $\lambda_1, \ldots, \lambda_r$. The product $\big(\prod_{i=1}^{r} \lambda_i\big)^{1/r}$ is called the **coding gain**. Note that the obvious solution which consists in scaling a whole codebook by multiplying it by a constant does not work: it only increases the signal-to-noise ratio ρ.

Since the coefficients of the codewords are complex, we may see them as included in a subfield L of $\mathbb{C}$. Furthermore, since rectangular codewords can be obtained from square ones by removing the appropriate number of rows or columns, we can suppose that $T = M$. Moreover, in order to obtain a decoding process with optimal performances, we need $M = N$ (see [**36**] for more details). From now on, we will therefore assume that $T = M = N = n$.

Looking at the above design criteria and remarks, we can summarize the coding problem as follows: find a family $\mathcal{C}$ of matrices in $\mathrm{M}_n(L)$ such that the matrix $(\mathbf{X}' - \mathbf{X}'')(\mathbf{X}' - \mathbf{X}'')^*$ is full rank, for all $\mathbf{X}' \neq \mathbf{X}'' \in \mathcal{C}$, or equivalently, such that

$$\det(\mathbf{X}' - \mathbf{X}'') \neq 0, \text{ for all } \mathbf{X}' \neq \mathbf{X}'' \in \mathcal{C}.$$

In this case, the previous estimation of the probability error may be rewritten as

$$\mathbb{P}(\mathbf{X}' \to \mathbf{X}'') \leq \frac{c(\rho)^{-n^2}}{|\det(\mathbf{X}' - \mathbf{X}'')|^{2n}}.$$

Let us note here that the cardinality of $\mathcal{C}$ plays a role, since higher coding gain can be obtained with a smaller cardinality. The cardinality of the code is often normalized and expressed in terms of its **rate**.

Definition I.3.1. The **rate** R of the code $\mathcal{C}$ is defined by

$$R = \frac{1}{M} \log_2 |\mathcal{C}|.$$

The difficulty in building fully-diverse matrices clearly comes from the nonlinearity of the determinant. Not much can be said about the determinant of the difference of two matrices. In order to overcome this obstacle, one natural solution is to look

for a linear codebook $\mathcal{C}$, namely one that satisfies

$$\mathbf{X}', \mathbf{X}'' \in \mathcal{C} \Rightarrow \mathbf{X}' \pm \mathbf{X}'' \in \mathcal{C}.$$

This indeed simplifies the design criterion to

$$\det(\mathbf{X}) \neq 0, \ \mathbf{0} \neq \mathbf{X} \in \mathcal{C}.$$

We thus restrict our attention to additive subgroups D of $\mathrm{M}_n(L)$. The condition that $\det(\mathbf{X}) \neq 0$ now reads that we need all the non-zero matrices in D to be invertible.

Using the work presented in Section I.2, we now illustrate how fully diverse space-time codes can be obtained by taking D to be a simple algebra. Let A be a simple k-algebra, and let L be a subfield of A. We will restrict ourselves to k-algebras A such that $k \subset \mathbb{C}$ since $L \subset \mathbb{C}$. By Remark I.2.12, we have an injective k-algebra morphism

$$\varphi_{A,L} : \begin{array}{c} A \hookrightarrow \mathrm{M}_m(L) \\ a \longmapsto M_a, \end{array}$$

where $m = \dim_L(A)$ (and can thus chosen to be n, the number of transmit antennas).

Definition I.3.2. We will call a **code** or an **algebra based code**, any set $\mathcal{C} \subset \mathrm{M}_n(\mathbb{C})$ of matrices satisfying

$$\mathcal{C} \subset \mathcal{C}_{A,L} = \{M_a \mid a \in A\},$$

where A is a simple k-algebra and L is a subfield of A.

We could define a slightly more general notion of algebra based code by replacing $\varphi_{A,L}$ by any injective k-algebra morphism $\varphi : A \longrightarrow \mathrm{M}_m(L)$. However, in the following (as well as in the existing literature), all the examples considered will make use of the map $\varphi_{A,L}$.

Remark I.3.3. It is clear that if A is a k-division algebra (thus simple), then any code $\mathcal{C}_{A,L} = \{\mathbf{X} = M_a \mid a \in A\}$ is fully diverse, and therefore so is any algebra based code $\mathcal{C}$.

Indeed, if $\mathbf{X}', \mathbf{X}'' \in \mathcal{C}_{A,L}, \mathbf{X}' \neq \mathbf{X}''$, then there exist $a', a'' \in A, a' \neq a''$ such that

$$\mathbf{X}' = M_{a'} \text{ and } \mathbf{X}'' = M_{a''}.$$

Now the map

$$\varphi_{A,L} : \begin{array}{c} A \hookrightarrow \mathrm{M}_m(L) \\ a \longmapsto M_a, \end{array}$$

is a k-algebra morphism, hence in particular a group morphism. We then have

$$\mathbf{X}' - \mathbf{X}'' = M_{a'} - M_{a''} = M_{a'-a''}.$$

Now since A is a k-division algebra and $a' - a'' \neq 0$, $a' - a''$ is a unit of A. Since $\varphi_{A,L}$ is a ring morphism, it maps units to units, and thus $\mathbf{X}' - \mathbf{X}'' = \varphi_{A,L}(a' - a'')$ is an invertible matrix, implying that $\mathcal{C}_{A,L}$ (and therefore $\mathcal{C}$) is fully diverse. □

Example I.3.4. Let us start with the easy case where we consider a field extension F/k of degree m, where $F = k(\theta)$ and θ has minimal polynomial

$$\mu_{\theta,k}(X) = \mu_0 + \mu_1 X + \cdots + \mu_{m-1} X^{m-1} + X^m.$$

Now take $A = F$, and $L = k$. We thus have

$$\varphi_{F,k}\colon \begin{array}{c} F \hookrightarrow \mathrm{M}_m(k) \\ a \longmapsto M_a. \end{array}$$

Then a k-basis for F is given by $\{1, \theta, \ldots, \theta^{m-1}\}$. Let $a \in F$, that is, $a = a_0 + a_1\theta + \cdots + a_{m-1}\theta^{m-1}$. One can compute that

$$M_a = \begin{pmatrix} a_0 & -\mu_0 a_{m-1} & \cdots \\ a_1 & a_0 - \mu_1 a_{m-1} & \cdots \\ a_2 & a_1 - \mu_2 a_{m-1} & \cdots \\ \vdots & \vdots & \cdots \\ a_{m-1} & a_{m-2} - \mu_{m-1} a_{m-1} & \cdots \end{pmatrix}.$$

By Remark I.3.3, the code

$$\mathcal{C}_{F,k} = \{M_a \mid a \in F\}$$

is fully diverse, since F is a field. Since F/k is finite, and thus F is an algebraic extension of k, this can also be seen as follows. We have that

$$\det(\mathbf{X}) = \det(\varphi_{F,k}(a)) = N_{F/k}(a),$$

where $N_{F/k}(a)$ denotes the norm map. Since F is a field, we have

$$N_{F/k}(a) = 0 \iff a = 0.$$

In the particular case where $\mu_{\theta,k}(X) = X^m - \lambda$, for some $\lambda \in k^\times$, we can write explicitly

$$M_a = \begin{pmatrix} a_0 & \lambda a_{m-1} & \lambda a_{m-2} & \ldots & \lambda a_1 \\ a_1 & a_0 & \lambda a_{m-1} & \ldots & \lambda a_2 \\ a_2 & a_1 & a_0 & \ldots & \lambda a_3 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ a_{m-2} & a_{m-3} & a_{m-4} & \ldots & \lambda a_{m-1} \\ a_{m-1} & a_{m-2} & a_{m-3} & \ldots & a_0 \end{pmatrix}.$$

We then obtain the following explicit description of the code $\mathcal{C}_{F,k}$:

$$\mathcal{C}_{F,k} = \left\{ \begin{pmatrix} a_0 & \lambda a_{m-1} & \lambda a_{m-2} & \ldots & \lambda a_1 \\ a_1 & a_0 & \lambda a_{m-1} & \ldots & \lambda a_2 \\ a_2 & a_1 & a_0 & \ldots & \lambda a_3 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ a_{m-2} & a_{m-3} & a_{m-4} & \ldots & \lambda a_{m-1} \\ a_{m-1} & a_{m-2} & a_{m-3} & \ldots & a_0 \end{pmatrix}, \ a_0, \ldots, a_{m-1} \in k \right\}.$$

Note that the elements $a_0, a_1, \ldots, a_{m-1}$ are the information symbols to be sent over the channel. This code is valid if there are m transmit antennas at the transmitter end. Transmission takes place over m periods of time. At time $t = 1$, each of the m transmit antennas sends one information symbol a_i, that is, the first column is sent. During the $m - 1$ other time slots, the $m - 1$ other columns are sent. The first column contains the data. The other columns contain the redundancy that protects the data. Such a code has been proposed in [**47**]. □

Note that in the above code, we are sending m^2 coefficients for communicating only m information symbols. We can now define another notion of **rate**, similar this time to the one found in classical coding theory.

DEFINITION I.3.5. We call **rate** (or sometimes **throughput**) the ratio of information symbols per coefficients sent.

The rate of the above code is thus $m/m^2 = 1/m$.

EXAMPLE I.3.6. Let $\mathbb{H} = (-1,-1)_{\mathbb{R}}$ be the Hamilton quaternion algebra. Notice that $\mathbb{H}$ contains a subfield isomorphic to $\mathbb{C}$, namely $\mathbb{R}(i)$, where i denotes one of the generators of $\mathbb{H}$. In particular, one may consider the element $\zeta_8 = \dfrac{1+i}{\sqrt{2}}$, which is a primitive 8-th root of unity. Let

$$A = \mathbb{Q}(\zeta_8) \oplus j\mathbb{Q}(\zeta_8).$$

Then one may check easily that A is a division $\mathbb{Q}(i)$-algebra of dimension 4. We thus have

$$\varphi_{A,\mathbb{Q}(i)} : \begin{array}{c} A \hookrightarrow \mathrm{M}_4(\mathbb{Q}(i)) \\ a \longmapsto M_a. \end{array}$$

A $\mathbb{Q}(i)$-basis for A is given by $(1, \zeta_8, j, j\zeta_8)$.

Let $a = a_1 + \zeta_8 a_2 + j a_3 + j\zeta_8 a_4 \in A$. Using the fact that we have

$$z\zeta_8 = \zeta_8 z \text{ and } zj = j\overline{z} \text{ for all } z \in \mathbb{R}(i),$$

one can compute that

$$M_a = \begin{pmatrix} a_1 & ia_2 & -\bar{a_3} & -\bar{a_4} \\ a_2 & a_1 & i\bar{a_4} & -\bar{a_3} \\ a_3 & ia_4 & \bar{a_1} & \bar{a_2} \\ a_4 & a_3 & -i\bar{a_2} & \bar{a_1} \end{pmatrix}.$$

Similarly as in the previous example, matrices M_a, $a \in A$, can be used to define a code $\mathcal{C}$ as follows:

$$\mathcal{C} = \left\{ \begin{pmatrix} a_1 & ia_2 & -\bar{a_3} & -\bar{a_4} \\ a_2 & a_1 & i\bar{a_4} & -\bar{a_3} \\ a_3 & ia_4 & \bar{a_1} & \bar{a_2} \\ a_4 & a_3 & -i\bar{a_2} & \bar{a_1} \end{pmatrix},\ a_1, a_2, a_3, a_4 \in \mathbb{Q}(i) \right\}.$$

This construction (together with further improvements) has been proposed in [**20**]. It is a codebook designed for 4 transmit antennas and has rate $4/16 = 1/4$. It is fully diverse by Remark I.3.3, since $A = \mathbb{Q}(\zeta_8) \oplus j\mathbb{Q}(\zeta_8)$ is a division algebra, as already noticed above. □

As it can be seen from these two examples, code constructions require an explicit presentation of the algebra considered. We will thus see further code examples once more simple algebras have been studied.

Let us consider now the general case:

EXAMPLE I.3.7. Let $\mathcal{C}$ be an algebra based code

$$\mathcal{C} \subset \mathcal{C}_{A,L} = \{\mathbf{X} = \varphi_{A,L}(a),\ a \in \mathcal{A}\} \subset \mathrm{M}_n(\mathbb{C}),$$

where A is a simple k-algebra, and L is a subfield of A.

Let us compute the rate of $\mathcal{C}$, that we will denote by $r(\mathcal{C})$.

The information symbols that we would like to transmit are elements of k, which may be used to define elements of A, after the choice of a k-basis of A. Each element of a may then carry $\dim_k(A)$ information symbols. However, an $m \times m$ matrix may contain m^2 information symbols, so we have

$$r(\mathcal{C}) = \frac{\dim_k(A)}{m^2},$$

where $m = \dim_L(A)$. Since $\dim_k(A) = \dim_L(A)[L : K] = m[L : K]$, this may be rewritten as

$$r(\mathcal{C}) = \frac{[L : k]^2}{\dim_k(A)},$$

so we should choose L/k such that $[L : k]$ is as large as possible. If A is a central division k-algebra, there exists a subfield L of A containing k such that $[L : k] = \deg(A)$ (as it will be shown in Chapter IV).

In particular, we obtain a code with a rate equal to 1 in this case, which is the best possible value. □

Exercises

1. Show that a k-subalgebra of a simple k-algebra is not necessarily simple.

2. Let k be a field. Let $n \geq 2$ be an integer. Assume that $\mathrm{char}(k)$ is prime to n, and that $\mu_n \subset k$. Let $\zeta_n \in k^\times$ be a primitive n^{th}-root of 1, and let $a, b \in k^\times$. Let $e, f \in \mathrm{M}_n(\overline{k})$ be the matrices defined by

$$e = \begin{pmatrix} 0 & & & a \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}, f = \begin{pmatrix} \beta & & & \\ & \zeta_n^{-1}\beta & & \\ & & \ddots & \\ & & & \zeta_n^{-(n-1)}\beta \end{pmatrix},$$

   where $\beta \in \overline{k}$ satisfies $\beta^n = b$.

   (a) Show that we have $e^n = a, f^n = b$ and $ef = \zeta_n fe$.

   (b) Let $\{a, b\}_{n,\zeta_n}$ be the k-subalgebra of $\mathrm{M}_n(\overline{k})$ generated by e and f. Show that $\{a, b\}_n$ is a central simple k-algebra of dimension n^2.

   (c) Assume that $L = k(f) \cong_k k(\sqrt[n]{b})$ has degree n. Compute the left multiplication matrix of an element of $\{a, b\}_{n,\zeta_n}$ with respect to the L-basis $(1, e, \ldots, e^{n-1})$.

3. Let p be a prime number, and let k be a field of characteristic p. Let $a \in k^\times$ and $b \in k$. Let $e, f \in \mathrm{M}_p(\overline{k})$ be the matrices defined by

$$e = \begin{pmatrix} 0 & & & a \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}, f = \begin{pmatrix} \beta & & & \\ & \beta - 1 & & \\ & & \ddots & \\ & & & \beta - p + 1 \end{pmatrix},$$

where $\beta \in \overline{k}$ satisfies $\beta^p - \beta = b$.

(a) Show that we have $e^p = a$, $f^p - f = b$ and $ef = fe + e$.

(b) Let $(a, b]_p$ be the k-subalgebra of $\mathrm{M}_n(\overline{k})$ generated by e and f. Show that $(a, b]_p$ is a central simple k-algebra of dimension p^2.

(c) Assume that $L = k(f) \cong_k k(\wp^{-1}(b))$ has degree p, where

$$\wp : \begin{array}{c} \overline{k} \longrightarrow \overline{k} \\ x \longmapsto x^p - x \end{array}$$

is the Weierstrass function. Compute the left multiplication matrix of an element of $(a, b]_p$ with respect to the L-basis $(1, e, \ldots, e^{n-1})$.