Chapter 2

MATHEMATICAL BACKGROUND

This chapter presents the important mathematical definitions and concepts required in this monograph. They are presented in a logical order, with each definition building on earlier concepts. However, the broad goals of the analysis presented in this monograph should be reasonably clear with only a passing acquaintance of the mathematics in this chapter. For more background and context to this mathematical material, we recommend the following references [23, 33, 57–59, 74, 97].

1. Groups, Rings, and Fields

Groups, rings, and fields constitute the basic structures of abstract algebra. They are also the basic algebraic structures required for the definition and the algebraic analysis of the AES.

Groups

DEFINITION 2.1 Let G be a non-empty set with a binary operation $\circ: G \times G \to G$. We say that $(G, \circ)$ is a *group* if the following conditions hold.

- The operation $\circ$ is associative, that is $(g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$ for all $g_1, g_2, g_3 \in G$.
- There exists an element $e \in G$ such that $e \circ g = g \circ e = g$ for all $g \in G$. This element e is unique and is called the *identity element.*
- For every $g \in G$, there exists a unique element $g^{-1} \in G$ such that $g \circ g^{-1} = g^{-1} \circ g = e$. This element g^{-1} is called the *inverse* of g.

The *order* of a group $(G, \circ)$ is the cardinality of the set G and is often denoted by $|G|$. If the order of $(G, \circ)$ is finite, we say that G is a finite

group. Similarly, we say that an element $g \in G$ has finite order if there exists a positive integer m such that $g \circ \ldots \circ g = g^m = e$. In this case, the least such integer m is called the *order* of g and is denoted by $o(g)$, and so the inverse element $g^{-1} = g^{o(g)-1}$. For a finite group G, the order of any element divides the order of the group G.

DEFINITION 2.2 The group $(G, \circ)$ is said to be an *abelian* or *commutative* group if $g \circ g' = g' \circ g$ for all $g, g' \in G$.

The group operation $\circ$ is usually clear from the context. When this is the case, the symbol $\circ$ is omitted and the group $(G, \circ)$ denoted by G.

EXAMPLE 2.3 The set of integers $\mathbb{Z}$ under the operation of addition forms an abelian group. Similarly, if n is a positive integer, the set of integers $\mathbb{Z}_n = \{0, \ldots, n-1\}$ under the operation of addition modulo n forms an abelian group of order n. □

EXAMPLE 2.4 The set of integers $\mathbb{Z}_p^* = \{1, \ldots, p-1\}$ under the operation of multiplication modulo p forms an abelian group if p is prime. □

EXAMPLE 2.5 Suppose that G_1 and G_2 are groups, then $G = G_1 \times G_2$ is a group with operation defined as $(g_1, g_2) \circ (g_1', g_2') = (g_1 g_1', g_2 g_2')$. The group G is known as the *direct product* of G_1 and G_2. □

A non–empty subset $H \subset G$ is called a *subgroup* of G if H is itself a group under the same operation. For a finite group, *Lagrange's Theorem* states that the order of any subgroup divides the order of the group. A subgroup H of G is called a *normal subgroup* of G if $g^{-1}hg \in H$ for all $g \in G$ and $h \in H$. The notation $H < G$ and $H \lhd G$ is used to denote that H is a subgroup of G and that H is a normal subgroup of G respectively. A group that has no non–trivial normal subgroups is called a *simple group.*

If H is a subgroup of G, then the *right coset* of H in G defined by $g \in G$ is the set $Hg = \{hg | h \in H\}$. The set of right cosets, $\{Hg | g \in G\}$, forms a partition of the elements of G. We can also define *left cosets* of H in G in a similar manner. The set of right cosets of H in G and the set of left cosets of H in G have the same cardinality. This cardinality is known as the *index* of H in G and is denoted by $[G : H]$. If H is a normal subgroup of G, then the right coset and left coset defined by any $g \in G$ are identical, and $Hg = gH$ is simply called the *coset* of H in G defined by $g \in G$. In this case, the set of all cosets of H in G forms a group with binary operation $(Hg, Hg') \mapsto Hgg'$ for all $g, g' \in G$. This group is called the *quotient group* of G by H. This group has order $[G : H]$ and is denoted by G/H.

DEFINITION 2.6 Let S be a non-empty subset of G. Then the *group generated by* S is defined as the set of all finite products of form $g_1 \circ \ldots \circ g_k$, where either $g_i \in S$ or $g_i^{-1} \in S$.

The group generated by S is denoted by $\langle S \rangle$ and is the smallest subgroup of G which contains S. If $S = \{g\}$, then the group $\langle S \rangle = \langle g \rangle$ generated by a single element $g \in G$ is called the *cyclic group* generated by g. If g has finite order, then $\langle g \rangle = \{g, g^2, \ldots, g^{o(g)-1}, e\}$.

A permutation of a non-empty set $\mathcal{X}$ is a bijective mapping $\mathcal{X} \to \mathcal{X}$. The set of permutations of $\mathcal{X}$, under the operation of composition, forms a group known as the *symmetric group of* $\mathcal{X}$. We denote this group by $S_{\mathcal{X}}$. If $\mathcal{X}$ is finite with cardinality n, this group is also known as the symmetric group on n elements and is denoted by S_n. The order of the group S_n is $n!$. An element of the group S_n that permutes two elements of $\mathcal{X}$ and leaves the remaining elements fixed is called a *transposition*. An element $g \in S_n$ is said to be an *even* permutation if it can be expressed as a product of an even number of transpositions, otherwise g is said to be an *odd* permutation. The subset of S_n consisting of all even permutations is a normal subgroup of S_n, known as the *alternating group* on n elements and is denoted by A_n. For $n > 1$, the order of A_n is $\frac{1}{2}n!$. Furthermore, A_n is a simple group for $n \neq 4$.

DEFINITION 2.7 Let $\mathcal{X}$ be a non empty set and G a group. A *group action* of G on $\mathcal{X}$ is a mapping $G \times \mathcal{X} \to \mathcal{X}$, denoted by $(g, x) \mapsto g \cdot x$, such that the following two conditions hold.

- If e is the identity of G, then $e \cdot x = x$ for every $x \in \mathcal{X}$;
- $g \cdot (g' \cdot x) = (gg') \cdot x$ for all $g, g' \in G$ and for all $x \in \mathcal{X}$.

If there is a group action of a group G on a set $\mathcal{X}$, we say that the group G *acts on the set* $\mathcal{X}$. An example of a group action is the action of the symmetric group $S_{\mathcal{X}}$ on the set $\mathcal{X}$ defined by $(g, x) \mapsto g(x)$ for all permutations g of $S_{\mathcal{X}}$ and $x \in \mathcal{X}$.

If G is a group acting on the set $\mathcal{X}$, then the *orbit* of $x \in \mathcal{X}$ is defined to be $\{g \cdot x \mid g \in G\} \subset \mathcal{X}$. The orbits of $\mathcal{X}$ form a partition of $\mathcal{X}$. The *stabilizer* of an element $x \in \mathcal{X}$ is defined to be $G_x = \{g \in G | g \cdot x = x\}$ and is a subgroup of G. The number of elements in the orbit of $x \in \mathcal{X}$ is the index $[G : G_x]$. Furthermore, if $Fix(g)$ denotes the number of elements of $\mathcal{X}$ that are fixed by $g \in G$, then the number of orbits of G on $\mathcal{X}$ is

$$\frac{1}{|G|} \sum_{g \in G} Fix(g).$$

If the action of G on $\mathcal{X}$ has only one orbit, then for any pair of elements $x, x' \in \mathcal{X}$ there exists $g \in G$ such that $g \cdot x = x'$. In this case the action of G on $\mathcal{X}$ is said to be *transitive*. Furthermore, if for any pair of m-tuples $(x_1, \ldots, x_m), (x'_1, \ldots, x'_m) \in \mathcal{X}^m$ with distinct entries ($x_i \neq x_j$ and $x'_i \neq x'_j$) there exists $g \in G$ such that $g \cdot x_i = x'_i$, then the action is said to be *m-transitive*. The action is said to be *sharply m-transitive* if such an element $g \in G$ is unique.

If G acts on a set $\mathcal{X}$, then $\mathcal{Y} \subseteq \mathcal{X}$ is called a *block* of G if for every $g \in G$, we have either $g(\mathcal{Y}) = \mathcal{Y}$ or $g(\mathcal{Y}) \cap \mathcal{Y} = \emptyset$. The group G is said to be *primitive* if it has no non-trivial blocks, and *imprimitive* otherwise.

EXAMPLE 2.8 The symmetric group S_n acting on a set of n elements is a primitive and sharply n-transitive group. The alternating group A_n acting on a set of n elements is a primitive and sharply $(n-2)$-transitive $(n > 2)$ group. □

DEFINITION 2.9 Let $(G, \circ)$ and $(H, \cdot)$ be groups. A mapping $\phi: G \to H$ is a (group) *homomorphism* if, for all $g, g' \in G$,

$$\phi(g \circ g') = \phi(g) \cdot \phi(g').$$

An injective homomorphism is called a *monomorphism* and a surjective homomorphism is called an *epimorphism*. A bijective homomorphism $\phi: G \to H$ is called an *isomorphism*, and the groups G and H are said to be *isomorphic*, denoted by $G \cong H$. An isomorphism from G to itself is called an *automorphism* of G.

DEFINITION 2.10 If $\phi: G \to H$ is a homomorphism and e_H is the identity element of H, then the subset

$$\ker \phi = \{g \in G | \phi(g) = e_H\}$$

of G is called the *kernel* of the homomorphism ϕ.

We note that $\ker \phi$ is a normal subgroup of G and the *First Isomorphism Theorem* states that the quotient group $G/\ker \phi$ is isomorphic to the image of ϕ. Furthermore, any normal subgroup $H \lhd G$ is the kernel of the "natural" epimorphism $G \to G/H$ defined by $g \mapsto Hg$.

EXAMPLE 2.11 Let H be the group $(\{-1, 1\}, \times)$, where $\times$ denotes the usual operation of integer multiplication. There exists a homomorphism from the symmetric group S_n onto H that maps every even permutation to 1 and every odd permutation to -1. The kernel of this homomorphism consists of all even permutations and so is the alternating group A_n. Thus the quotient group S_n/A_n is isomorphic to H. □

Isomorphic groups have identical algebraic structure and can be regarded as essentially the *same* algebraic object. Isomorphisms are often useful for solving problems that would otherwise be intractable. Thus obtaining alternative representations using isomorphisms is a common technique for the study and analysis of algebraic structures. We note however that constructing isomorphisms between two algebraic structures, and even constructing the inverse isomorphism of a known isomorphism, can often be a very difficult problem.

EXAMPLE 2.12 Let p be a prime number, and $\mathbb{Z}_{p-1}$ and $\mathbb{Z}_p^*$ denote the groups defined in Examples 2.3 and 2.4 respectively. The group $\mathbb{Z}_{p-1}$ is generated additively by the element $1 \in \mathbb{Z}_{p-1}$, and the group $\mathbb{Z}_p^*$ is generated multiplicatively by some $g \in \mathbb{Z}_p^*$. These groups are isomorphic, and an isomorphism between them can be defined by $m \mapsto g^m$, that is the exponentiation in $\mathbb{Z}_p^*$. The inverse isomorphism is known as the *discrete logarithm*, and the calculation of the discrete logarithm is generally believed to be a hard problem. The difficulty of computing this inverse isomorphism is the foundation of the security of many asymmetric cryptosystems, for example the *Digital Signature Standard* [93]. □

Rings

DEFINITION 2.13 Let R be a non-empty set with two associative binary operations $+, \cdot : R \times R \to R$. We say that $(R, +, \cdot)$ is a *ring* (*with unit*) if the following conditions hold.

- $(R, +)$ is an abelian group.
- The operation $\cdot$ is distributive over $+$, that is for all $r, r', r'' \in R$,

$$r \cdot (r' + r'') = r \cdot r' + r \cdot r'' \text{ and } (r' + r'') \cdot r = r' \cdot r + r'' \cdot r.$$

- There is an element $1 \in R$ such that $1 \cdot r = r \cdot 1 = r$ for all $r \in R$.

The identity element of the group $(R, +)$ is usually denoted by 0 and is called the *zero* of the ring $(R, +, \cdot)$. The element 1 is called the *identity element* of the ring $(R, +, \cdot)$.

DEFINITION 2.14 The ring $(R, +, \cdot)$ is a *commutative ring* if $r \cdot r' = r' \cdot r$ for all $r, r' \in R$, that is the operation $\cdot$ is commutative.

All rings considered in this monograph are commutative rings with unit. As with groups, we often assume that the operations $+$ and $\cdot$ are clear, and we denote the ring $(R, +, \cdot)$ simply by R. We also often denote $r \cdot r'$ simply by rr' for $r, r' \in R$.

A commutative ring R is called an *integral domain* if it contains no *zero-divisors*, that is $rr' \neq 0$ for all $r, r' \in R \setminus \{0\}$. A nonzero element r of a ring R is said to be *invertible* (or a *unit*) if there exists $r^{-1} \in R$ such that $r \cdot r^{-1} = r^{-1} \cdot r = 1$. The set of all invertible elements of R is denoted by R^* and forms a group under multiplication known as the *group of units* of R. If all nonzero elements of a ring R are invertible, then R is called a *division ring* and $R^* = R \setminus \{0\}$.

EXAMPLE 2.15 The set of integers $\mathbb{Z}$ under the operations of integer addition and multiplication forms a commutative ring. Similarly, the set of integers $\mathbb{Z}_n = \{0, \ldots, n-1\}$ under the operations of addition and multiplication modulo n forms a commutative ring. We note that $\mathbb{Z}_n$ is a division ring if and only if n is prime. □

DEFINITION 2.16 Let $(R, +, \cdot)$ be a ring and I a non empty subset of R. We say that I is an *ideal* of R, denoted by $I \lhd R$, if the following conditions hold.

- $(I, +)$ is a subgroup of $(R, +)$.
- For all $x \in I$ and $r \in R$, $x \cdot r \in I$ and $r \cdot x \in I$.

The *coset* of an ideal I in R defined by $r \in R$ is denoted by $I + r$ and defined to be the set $\{s + r | s \in I\}$. The cosets of an ideal $I \lhd R$ form a partition of the ring R. The set of all cosets of I forms a ring with addition and multiplication defined by $(I + r) + (I + r') = I + (r + r')$ and $(I + r)(I + r') = I + rr'$ respectively. This ring is denoted by R/I and is called the *quotient ring* or the *residue class ring modulo* I.

If S is a non-empty subset of R, then the *ideal generated by* S is denoted by $\langle S \rangle$ and consists of all finite sums of the form $\sum r_i s_i$, where $r_i \in R$ and $s_i \in S$. An ideal is said to be a *principal ideal* if it can be generated by one element $r \in R$. An integral domain in which every ideal is a principal ideal is called a *principal ideal domain*.

DEFINITION 2.17 If R and R' are rings, then $\phi : R \to R'$ is a (ring) *homomorphism* if the following conditions hold.

- $\phi(r + r') = \phi(r) + \phi(r')$ for all $r, r' \in R$.
- $\phi(r \cdot r) = \phi(r) \cdot \phi(r')$ for all $r, r' \in R$.

Different types of ring homomorphism are defined in a similar manner to group homomorphisms. The *kernel* $\ker \phi = \{r \in R | \phi(r) = 0\}$ of a ring homomorphism $\phi : R \to R'$ is an ideal of R. Furthermore, the quotient ring $R/\ker \phi$ is isomorphic to the image of R, and every ideal $I \lhd R$ is the kernel of the "natural" epimorphism $R \to R/I$ defined by $r \mapsto I + r$.

Fields

DEFINITION 2.18 A commutative division ring $\mathbb{F}$ is called a *field.*

Thus a field $\mathbb{F}$ is a ring $(\mathbb{F}, +, \cdot)$ such that both $(\mathbb{F}, +)$ and $(\mathbb{F} \setminus \{0\}, \cdot)$ are commutative groups.

EXAMPLE 2.19 The sets $\mathbb{Q}$ of rational numbers, $\mathbb{R}$ of real numbers, and $\mathbb{C}$ of complex numbers form fields under the usual operations of addition and multiplication. □

EXAMPLE 2.20 The set $\mathbb{Z}_n = \{0, \ldots n-1\}$ under addition and multiplication modulo an integer n is a field if and only if n is prime (Examples 2.3 and 2.4). □

If $\mathbb{F}$ is a field, we say that $\mathbb{F}$ has *positive characteristic* if there exists a positive integer m such that the m-fold sum $1 + \ldots + 1 = 0$. In this case, the least such integer m is called the *characteristic* of $\mathbb{F}$. If there is no such m, we say that $\mathbb{F}$ has *characteristic zero.* The infinite fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ all have characteristic zero, whilst the finite field $\mathbb{Z}_p$ has characteristic p. In fact, all finite fields have characteristic p for some prime p. We discuss further aspects of finite fields in Section 2.4.

2. Polynomial Rings

Polynomial rings are a special example of commutative ring that play an important role in the theory of finite fields. The algebraic analysis of the AES makes extensive use of polynomial rings.

Univariate polynomial rings

A *monomial* in the single variable or indeterminate x is the formal expression x^i for some $i \in \mathbb{N}$, that is some non-negative power of x. The *degree* of the monomial x^i is i.

DEFINITION 2.21 A univariate *polynomial* in the variable x over a field $\mathbb{F}$ is a finite linear combination over $\mathbb{F}$ of monomials in x, that is a formal expression of the form

$$c_d x^d + c_{d-1} x^{d-1} + \ldots + c_2 x^2 + c_1 x + c_0,$$

where d is a non-negative integer and $c_d, \ldots, c_0 \in \mathbb{F}$, with $c_d \neq 0$ if $d > 0$.

DEFINITION 2.22 The set of all univariate polynomials in the variable x over a field $\mathbb{F}$ forms a ring under the standard operations of polynomial

addition and multiplication. This ring is a principal ideal domain called the *univariate polynomial ring* over $\mathbb{F}$ and is denoted by $\mathbb{F}[x]$.

Let $f(x) \in \mathbb{F}[x]$ be a univariate polynomial. The *degree* of $f(x)$ is the maximum integer d such that $c_d \neq 0$, and is denoted by $\deg(f(x))$. If

$$f(x) = c_d x^d + \ldots + c_1 x + c_0,$$

then the summands $c_i x^i$ ($c_i \neq 0$) are called the *terms* of $f(x)$, and c_i is called the *coefficient* of the monomial x^i. Furthermore, we can define the *leading monomial*, *leading coefficient*, and *leading term* of $f(x)$ as x^d, c_d and $c_d x^d$ respectively. A polynomial $f(x)$ is a *monic* polynomial if its leading coefficient is 1.

The *evaluation* of the polynomial $f(x)$ at $a \in \mathbb{F}$ is defined as the element $\sum_{i=0}^{d} c_i a^i \in \mathbb{F}$ and is denoted by $f(a)$. We say that a is a *root* of $f(x)$ if $f(a) = 0$. A polynomial of degree d has at most d roots in $\mathbb{F}$.

THEOREM 2.23 *Univariate Division Algorithm.* Given $f(x)$ and $g(x) \in \mathbb{F}[x]$, then there exists $q(x), r(x) \in \mathbb{F}[x]$ with $\deg(r(x)) < \deg(g(x))$ such that $f(x) = q(x)g(x) + r(x)$. The univariate polynomial $r(x)$ is known as the *remainder* of the division of $f(x)$ by $g(x)$.

The well-known Euclidean algorithm to find the greatest common divisor of two polynomials is just the repeated application of Theorem 2.23.

EXAMPLE 2.24 Suppose that

$$f(x) = x^6 + x^5 + x^3 + x^2 + x + 1 \text{ and } g(x) = x^4 + x^3 + 1$$

are polynomials in the univariate polynomial ring $\mathbb{Z}_2[x]$. We then have

$$x^6 + x^5 + x^3 + x^2 + x + 1 = x^2(x^4 + x^3 + 1) + (x^3 + x + 1),$$

so $f(x) = q(x)f(x) + r(x)$, where $q(x) = x^2$ and $r(x) = x^3 + x + 1$. □

A polynomial $f(x) \in \mathbb{F}[x]$ of positive degree is said to be *irreducible* in $\mathbb{F}[x]$ if there is no factorisation of the form $f(x) = p(x)q(x)$, where $p(x)$ and $q(x)$ are polynomials of positive degree in $\mathbb{F}[x]$. Every polynomial in $\mathbb{F}[x]$ can be written as the product of monic irreducible polynomials and some constant in $\mathbb{F}$, and this product is unique up to the order of the factors.

EXAMPLE 2.25 Let $f(x)$ be a polynomial in $\mathbb{F}[x]$ of degree d, and $\langle f(x) \rangle$ be the ideal generated by $f(x)$. The elements of the quotient ring $\frac{\mathbb{F}[x]}{\langle f(x) \rangle}$ can be written as polynomials

$$a_{d-1} x^{d-1} + \ldots + a_1 x + a_0$$

in $\mathbb{F}[x]$ of degree less than d. In this representation of the quotient ring, addition is simply polynomial addition. However, multiplication in the quotient ring is defined by applying Theorem 2.23. For two polynomials $g_1(x), g_2(x) \in \mathbb{F}[x]$, we know that there exists $q(x), r(x) \in \mathbb{F}[x]$ such that

$$g_1(x)g_2(x) = q(x)f(x) + r(x),$$

where $deg(r(x)) < deg(f(x)) = d$. In this representation of the quotient ring $\frac{\mathbb{F}[x]}{\langle f(x) \rangle}$, the product of $g_1(x)$ and $g_2(x)$ is $r(x)$. □

EXAMPLE 2.26 Let $f(x) = x^5 + x^4 + 1$ be a polynomial in the univariate polynomial ring $\mathbb{Z}_2[x]$. The product of the polynomials $(x^4 + x^3 + x^2 + 1)$ and $(x^4 + x^3 + x + 1)$ satisfies

$$\begin{aligned}(x^4 + x^3 + x^2 + 1)(x^4 + x^3 + x + 1) &= x^8 + x^4 + x^3 + x^2 + x + 1 \\ &= (x^3 + x^2 + x + 1)f(x) + 0.\end{aligned}$$

Thus in the quotient ring $R = \frac{\mathbb{Z}_2[x]}{\langle f(x) \rangle}$, the product of these two nonzero elements is 0, and R is not an integral domain. □

THEOREM 2.27 The quotient ring $\frac{\mathbb{F}[x]}{\langle f(x) \rangle}$ is a field if and only if $f(x)$ is irreducible in $\mathbb{F}[x]$.

The *Lagrange Interpolation Formula* is a well-known method for constructing a polynomial based on given values for evaluation of a function.

THEOREM 2.28 *Lagrange Interpolation Formula.* Given $n + 1$ pairs $(a_i, b_i) \in \mathbb{F} \times \mathbb{F}$, with $a_i \neq a_j$, there exists a unique polynomial $f(x) \in \mathbb{F}[x]$ of degree at most n with $f(a_i) = b_i$. This polynomial is given by

$$f(x) = \sum_{i=0}^{n} b_i \prod_{\substack{k=0 \\ k \neq i}}^{n} \left(\frac{x - a_k}{a_i - a_k} \right).$$

Multivariate polynomial rings

Let $\mathbb{N}^n = \{(\alpha_1, \ldots, \alpha_n) \mid \alpha_i \in \mathbb{N}\}$ denote the set of *multi-indices* of size n. A *monomial* in the variables $x_1, \ldots, x_n$ is a product of the form

$$x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n},$$

which we denote by X^α, $\alpha \in \mathbb{N}^n$. The *degree* of X^α is $d_\alpha = \sum_{i=1}^{n} \alpha_i$.

DEFINITION 2.29 A *polynomial* in n variables $x_1, \ldots, x_n$ over the field $\mathbb{F}$ is a finite linear combination over $\mathbb{F}$ of monomials in $x_1, \ldots, x_n$, that is a formal expression of the form

$$\sum_{\alpha \in N} c_\alpha X^\alpha,$$

where $c_\alpha \in \mathbb{F}$ and N is a finite subset of $\mathbb{N}^n$.

DEFINITION 2.30 The set of all polynomials in n variables over a field $\mathbb{F}$ forms a ring under the standard operations of polynomial addition and multiplication. This ring is called a *polynomial ring* over $\mathbb{F}$, and for variables $x_1, \ldots, x_n$ is denoted by $\mathbb{F}[x_1, \ldots, x_n]$.

Let $f = \sum c_\alpha X^\alpha \in \mathbb{F}[x_1, \ldots, x_n]$ be a multivariate polynomial over $\mathbb{F}$. The summands $c_\alpha X^\alpha$ ($c_\alpha \neq 0$) are called the *terms* of f, and c_α is said to be the *coefficient* of X^α. The *total degree* of f is the maximum of the degrees of all monomials of f. If all monomials of f have the same degree d, we say that f is *homogeneous* of degree d.

DEFINITION 2.31 Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial of total degree d. The polynomial f^h defined as

$$f^h = x_0^d \cdot f\left(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}\right)$$

is a homogeneous polynomial of degree d in $\mathbb{F}[x_0, x_1, \ldots, x_n]$, called the *homogenisation* of f.

DEFINITION 2.32 A total ordering $\prec$ on the set of monomials X^α (where $\alpha \in \mathbb{N}^n$) that is compatible with multiplication is called a *monomial ordering* in $\mathbb{F}[x_1, \ldots, x_n]$. An ordering is compatible with multiplication if $X^\alpha \prec X^\beta$ implies $X^\alpha X^\gamma \prec X^\beta X^\gamma$ for all multi–indices $\alpha, \beta, \gamma \in \mathbb{N}^n$.

We now define three common examples of monomial orderings.

DEFINITION 2.33 The *lex* (lexicographic) monomial ordering is defined by $X^\alpha \prec X^\beta$ if the left–most nonzero entry in the vector $\beta - \alpha \in \mathbb{Z}^n$ is positive.

DEFINITION 2.34 The *glex* (graded lexicographic) monomial ordering is defined by $X^\alpha \prec X^\beta$ if, firstly the degree of X^β is larger than the degree of X^α ($d_\beta > d_\alpha$), and secondly if $d_\beta = d_\alpha$ then the left–most nonzero entry in the vector $\beta - \alpha \in \mathbb{Z}^n$ is positive.

DEFINITION 2.35 The *grevlex* (graded reverse lexicographic) monomial ordering is defined by $X^\alpha \prec X^\beta$ if, firstly the degree of X^β is larger than the degree of X^α ($d_\beta > d_\alpha$), and secondly if $d_\beta = d_\alpha$ then the right–most nonzero entry in the vector $\beta - \alpha \in \mathbb{Z}^n$ is negative.

EXAMPLE 2.36 Some monomial orderings in $\mathbb{F}[x, y, z]$ are shown below.

$$\begin{array}{rl} \textit{lex} \text{ ordering:} & x^2y^3z^6 \prec x^2y^4z \text{ and } xy^3z \prec x^2yz^2 \\ \textit{glex} \text{ ordering:} & x^2y^4z \prec x^2y^3z^6 \text{ and } xy^3z \prec x^2yz^2 \\ \textit{grevlex} \text{ ordering:} & x^2y^4z \prec x^2y^3z^6 \text{ and } x^2yz^2 \prec xy^3z \end{array}$$

We can see that the pair of monomials x^2y^4z and $x^2y^3z^6$ and the pair of monomials x^2yz^2 and xy^3z are ordered differently under the various monomial orderings. □

Suppose the polynomial ring $\mathbb{F}[x_1, \ldots, x_n]$ has a monomial ordering $\prec$ and $f \in \mathbb{F}[x_1, \ldots, x_n]$ is a polynomial. The *leading monomial* of f is the maximal monomial of f with respect to the ordering $\prec$ and is denoted by $\mathrm{LM}(f)$. The *leading coefficient* of f is the coefficient of the leading monomial of f and is denoted by $\mathrm{LC}(f)$. The *leading term* of f is the term associated with the leading monomial and is denoted by $\mathrm{LT}(f)$, so $\mathrm{LT}(f) = LC(f)LM(f)$. The *multidegree* of f is the degree of the leading monomial of f and is denoted by $\mathrm{multideg}(f)$.

These concepts enable us to give a multivariate generalisation of the division algorithm for univariate polynomials (Theorem 2.23).

THEOREM 2.37 *Polynomial Division Algorithm.* Suppose that the polynomial ring $\mathbb{F}[x_1, \ldots, x_n]$ has a monomial ordering $\prec$ and that $(g_1, \ldots, g_s)$ is an ordered subset of $\mathbb{F}[x_1, \ldots, x_n]$. For any $f \in \mathbb{F}[x_1, \ldots, x_n]$, there exist $a_i, r \in \mathbb{F}[x_1, \ldots, x_n]$ such that

$$f = a_1g_1 + \ldots + a_sg_s + r,$$

where either $r = 0$, or $r \neq 0$ and no leading monomial of the polynomials g_i divides any of the monomials of r. Such a polynomial r is called a *remainder* of the division of f by the set $\{g_1, \ldots, g_s\}$. Furthermore, if $a_ig_i \neq 0$, then $\mathrm{multideg}(a_ig_i) \leq \mathrm{multideg}(f)$.

3. Linear Algebra

Linear algebra is at the heart of both the design and the analysis of the AES. Diffusion in the AES SP-network is achieved by a linear transformation. It is therefore not surprising to find linear algebra being used as a tool in the analysis of the cipher.

Vector spaces

DEFINITION 2.38 Let $(V, +)$ be an abelian group, $\mathbb{F}$ a field and $\cdot$ an operation $\mathbb{F} \times V \to V$. We say that V is a *vector space* over $\mathbb{F}$ if the following conditions hold.

- $a \cdot (v + v') = a \cdot v + a \cdot v'$ for all $v, v' \in V$ and $a \in \mathbb{F}$.
- $(a + a') \cdot v = a \cdot v + a' \cdot v$ for all $v \in V$ and $a, a' \in \mathbb{F}$.
- $(aa') \cdot v = a \cdot (a' \cdot v)$ for all $v \in V$ and $a, a' \in \mathbb{F}$.
- $1 \cdot v = v$ for all $v \in V$, where 1 is the identity element of $\mathbb{F}$.

In a vector space, an element of the set V is called a *vector* and an element of the field $\mathbb{F}$ is called a *scalar*. The operation $+$ is known as *vector addition* and the operation $\cdot$ as *scalar multiplication*. The identity element of the abelian group $(V, +)$ is called the *zero vector* and is usually denoted by 0. Furthermore, the symbol $\cdot$ is usually omitted if there is no danger of confusion.

EXAMPLE 2.39 The set $\mathbb{F}^n = \{(a_1, \ldots, a_n) \mid a_i \in \mathbb{F}\}$ forms a vector space over $\mathbb{F}$ with vector addition and scalar multiplication defined by

$$\begin{aligned} (a_1, \ldots, a_n) + (a'_1, \ldots, a'_n) &= (a_1 + a'_1, \ldots, a_n + a'_n), \text{ and} \\ a \cdot (a_1, \ldots, a_n) &= (aa_1, \ldots, aa_n). \end{aligned}$$ □

A subset U of a vector space V over a field $\mathbb{F}$ is a *subspace* of V if U is itself a vector space over $\mathbb{F}$. The notation $U < V$ is used to denote that U is a subspace of V. The intersection $U \cap U'$ of any two subspaces U and U' of V is a subspace of V. The sum of subspaces $U, U' < V$, defined by

$$U + U' = \{u + u' \mid u \in U, u' \in U'\},$$

is also a subspace of V. This definition extends in the obvious way to any finite sum of subspaces. If a vector space $V = U_1 + \ldots + U_m$ and the subspaces $U_1, \ldots, U_m$ have trivial pairwise intersections ($U_i \cap U_j = \{0\}$ for $i \neq j$), then V is said to be the *direct sum* of these subspaces and we write $V = U_1 \oplus \ldots \oplus U_m$. In this case, for any $v \in V$, there exist unique $u_i \in U_i$ such that $v = u_1 + \ldots + u_m$.

The set of all finite linear combinations of the vectors $v_1, \ldots, v_m \in V$,

$$\langle v_1, \ldots, v_m \rangle = \{a_1 v_1 + \ldots + a_m v_m \mid a_i \in \mathbb{F}, v_i \in V\},$$

is a subspace of V and is called the *subspace generated* by the set $\{v_1, \ldots, v_m\}$. A set of vectors $\{v_1, \ldots, v_m\}$ is said to *span* or to be a

spanning set of a subspace $U < V$ if for all $u \in U$ there exists $a_i \in \mathbb{F}$ such that $u = \sum_i a_i v_i$.

A *basis* for a vector space V is a minimal spanning set for V. Every vector in a vector space can be expressed as a unique linear combination over $\mathbb{F}$ of the elements of a basis. Any basis for a vector space V always has the same cardinality, which is called the *dimension* of the vector space V and is denoted by $\dim(V)$.

EXAMPLE 2.40 The multivariate polynomial ring $\mathbb{F}[x_1, \ldots, x_n]$ forms an infinite dimensional vector space over $\mathbb{F}$. The subset of $\mathbb{F}[x_1, \ldots, x_n]$ of all polynomials of degree at most 1 is a subspace of dimension $n+1$ with basis $\{1, x_1, \ldots, x_n\}$. □

A set of vectors $\{v_1, \ldots, v_m\}$ is said to be *linearly independent* if the expression $\sum_{i=1}^{m} a_i v_i = 0$ implies that $a_1 = \ldots = a_m = 0$. If $\{e_1, \ldots, e_n\}$ is a basis B for a vector space V of dimension n, then B is linearly independent and for any $v \in V$ there exist unique $a_i \in \mathbb{F}$ such that $v = a_1 e_1 + \ldots + a_n e_n$. Thus we can represent v with respect to the basis B by the n-tuple $(a_1, \ldots, a_n) \in \mathbb{F}^n$.

We can define cosets of subspaces in a similar manner to cosets of subgroups. In particular, the set of all cosets of U in V forms a vector space called a *quotient vector space* and is denoted by V/U.

Linear transformations

DEFINITION 2.41 A *linear transformation* or vector space *homomorphism* from a vector space V over a field $\mathbb{F}$ to a vector space U over $\mathbb{F}$ is a mapping $\psi: V \to U$ that satisfies the following two conditions.

- $\psi(v + v') = \psi(v) + \psi(v')$ for all $v, v' \in V$.
- $\psi(av) = a \cdot \psi(v) = a\psi(v)$ for all $v \in V$ and $a \in \mathbb{F}$.

A vector space *isomorphism* is a bijective linear transformation, and we use $V \cong U$ to denote that the vector spaces V and U are isomorphic.

EXAMPLE 2.42 Let V be a vector space over the field $\mathbb{F}$ of dimension n, and let $\{e_1, \ldots, e_n\}$ be a basis for V. Given $v \in V$, there exists a unique $(a_1, \ldots, a_n) \in \mathbb{F}^n$ such that $v = a_1 e_1 + \ldots + a_n e_n$. The mapping $V \to \mathbb{F}^n$ defined by $v \mapsto (a_1, \ldots, a_n)$ is a vector space isomorphism. Thus any two finite-dimensional vector spaces over the same field are isomorphic. □

EXAMPLE 2.43 Let V be a vector space of dimension n over the field $\mathbb{F}$, and $\alpha_1, \ldots, \alpha_n$ be elements of $\mathbb{F}$. Then every mapping $V \to \mathbb{F}$ of the form $a_1 e_1 + \ldots + a_n e_n \mapsto \alpha_1 a_1 + \ldots + \alpha_n a_n$ is a linear transformation

from V into $\mathbb{F}$, where $\mathbb{F}$ is considered as a one–dimensional vector space over $\mathbb{F}$. Furthermore, every linear transformation from $V \to \mathbb{F}$ is of this form. Such a transformation is known as a *linear functional* on V. □

Let $\psi: V \to V'$ be a linear transformation. Then the *kernel* of ψ is defined by $\ker\psi = \{v \in V | \psi(v) = 0\}$ and is a subspace of V. The *nullity* of ψ is the dimension of $\ker\psi$. The image of the linear transformation ψ is a subspace of V', and the *rank* of ψ is the dimension of $\psi(V)$. The *Rank–Nullity Theorem* states that

$$\dim(V) = \dim(\ker\psi) + \dim(\psi(V)).$$

The quotient vector space $V/\ker\psi$ is isomorphic to the image of V, so $\psi(V) \cong V/\ker\psi$. If $V' = V$, then the subspace $U < V$ is called a *ψ-invariant* subspace if $\psi(U) < U$. If ψ satisfies $\psi \circ \psi = \psi^2 = \psi$, then ψ is called a *projection* and $\psi(U)$ is a ψ-invariant subspace for any subspace $U < V$.

If $\psi: V \to V$ is a linear transformation, then $\sum_{i=0}^{d} a_i\psi^i$ is also a linear transformation on V. Furthermore, the set

$$I = \left\{ \sum_{i=0}^{d} a_i x^i \;\middle|\; \sum_{i=0}^{d} a_i\psi^i = 0 \right\}$$

is an ideal of the polynomial ring $\mathbb{F}[x]$. The *minimal polynomial* of the linear transformation ψ is defined as the unique monic polynomial $\min_\psi(x)$ that generates the principal ideal I. The minimal polynomial of ψ gives much information about both ψ and the ψ-invariant subspaces. For example, if $\min_\psi(x) = m_1(x)\ldots m_l(x)$ is the factorisation of the minimal polynomial of ψ into monic polynomials, then $m_i(\psi)$ has a natural interpretation as a linear transformation and the ψ-invariant subspaces are given by $\ker m_i(\psi)$.

DEFINITION 2.44 Suppose that V and V' are vector spaces over the field $\mathbb{F}$ and that $\psi: V \to V'$ is a linear transformation. If b is a vector in V', then the transformation $V \to V'$ defined by $v \mapsto \psi(v) + b$ is termed an *affine transformation.*

DEFINITION 2.45 Consider a mapping $\beta : V \times V \to V'$, where V and V' are vector spaces over the field $\mathbb{F}$. For $u \in V$, we can define the mappings $\beta'_u, \beta''_u: V \to V'$ by $v \mapsto \beta(u, v)$ and $v \mapsto \beta(v, u)$ respectively. The mapping β is called a *bilinear transformation* on V if β'_u and β''_u are linear transformations for all $u \in V$.

Matrices

DEFINITION 2.46 An $m \times n$ *matrix* over a field $\mathbb{F}$ is a rectangular array

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

with $a_{ij} \in \mathbb{F}$. The elements a_{ij} are called the *entries* of the matrix.

If A is an $m \times n$ matrix, the sequences $(a_{i1} \dots a_{in})$ are called the *rows* of A, and the sequences $(a_{1j} \dots a_{mj})$ are called the *columns* of A. Thus A has m rows and n columns. If $m = n$, then A is called a *square matrix* of order n. A *submatrix* of A is an $m' \times n'$ matrix ($m' \leq m$ and $n' \leq n$) obtained by taking a block of entries of M with m' rows and n' columns. The *transpose* of A is denoted by A^T and is the $n \times m$ matrix whose (i, j)-entry is given by a_{ji}.

EXAMPLE 2.47 Let $\mathcal{M}_{m\times n}(\mathbb{F})$ denote the set of all $m \times n$ matrices over $\mathbb{F}$. We can define the operation of addition of elements of $\mathcal{M}_{m\times n}(\mathbb{F})$ in the obvious way by adding the corresponding entries of the matrices. Similarly, we can define the scalar multiplication of a matrix $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ by an element $c \in \mathbb{F}$ to be the matrix obtained by simply multiplying every entry of A by c. Thus the set $\mathcal{M}_{m\times n}(\mathbb{F})$ forms a vector space over $\mathbb{F}$ of dimension mn. □

Let A be an $m \times n$ matrix and B be an $r \times s$ matrix over $\mathbb{F}$ defined as

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \dots & b_{1s} \\ \vdots & \ddots & \vdots \\ b_{r1} & \dots & b_{rs} \end{pmatrix}.$$

If $n = r$, we can define a multiplication of A by B. The *product* AB is the $m \times s$ matrix C whose entries are $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$.

DEFINITION 2.48 Suppose A is an $m \times n$ matrix over $\mathbb{F}$ and that $A_{i\cdot}$ denotes the i^{th} row of A. An *elementary row operation* on the matrix A is one of the following three types of operation.

- The replacement of $A_{i\cdot}$ by $cA_{i\cdot}$ where $c \in \mathbb{F}$ with $c \neq 0$.
- The replacement of $A_{i\cdot}$ by $A_{i\cdot} + cA_{j\cdot}$ where $c \in \mathbb{F}$ and $i \neq j$.
- The interchange of two rows of A.

An elementary row operation on the matrix A is equivalent to a mapping $A \mapsto PA$, where P is a $m \times m$ elementary row operation matrix.

Any $m \times n$ matrix that can be obtained from A by a series of elementary row operations is said to be *row-equivalent* to A. In particular, there is a special set of matrices called the *row-reduced echelon matrices*, and any matrix is row-equivalent to a unique row reduced echelon matrix.

The *rank* of an $m \times n$ matrix A is the number of linearly independent rows or columns (considered as vectors) of A. In particular, if A is a row-reduced echelon matrix, then the rank of A is the number of nonzero rows. We note that row-equivalent matrices have the same rank.

Let $\mathcal{M}_n(\mathbb{F})$ denote the set of all square matrices over $\mathbb{F}$ of order n. A matrix $A \in \mathcal{M}_n(\mathbb{F})$ with entries a_{ij} is a *symmetric matrix* if $A^T = A$, that is $a_{ij} = a_{ji}$, and A is a *diagonal matrix* if $a_{ij} = 0$ whenever $i \neq j$. The *identity matrix* is a diagonal matrix in which $a_{ii} = 1$ $(i = 1, \ldots, n)$ and is usually denoted by I. The identity matrix has the property that $AI = IA = A$ for any matrix $A \in \mathcal{M}_n(\mathbb{F})$. The square matrix A is an *invertible* or *non-singular* matrix if there exists an $n \times n$ matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$. A matrix is invertible if and only if it is row-equivalent to the identity matrix.

The *determinant* is a function $\det: \mathcal{M}_n(\mathbb{F}) \to \mathbb{F}$ on square matrices with special properties, and this function is widely used in the analysis of square matrices [59]. In particular, we have $\det(AB) = \det(A)\det(B)$, and a matrix A is invertible if and only if $\det(A) \neq 0$.

The set of $n \times n$ invertible matrices forms a group under the operation of matrix multiplication. This group is called the *general linear group* and is denoted by $\mathrm{GL}(n, \mathbb{F})$. The subset of all matrices that have determinant 1 forms a normal subgroup of $\mathrm{GL}(n, \mathbb{F})$. This subgroup is called the *special linear group* and is denoted by $\mathrm{SL}(n, \mathbb{F})$. Thus we have

$$\begin{aligned} \mathrm{GL}(n, \mathbb{F}) &= \{\ A \in \mathcal{M}_n(\mathbb{F}) \mid \det(A) \neq 0\ \}, \text{ and} \\ \mathrm{SL}(n, \mathbb{F}) &= \{\ A \in \mathcal{M}_n(\mathbb{F}) \mid \det(A) = 1\ \}. \end{aligned}$$

Matrices are often used to represent linear transformations between vector spaces and can be particularly useful for performing calculations with such mappings. For example, matrices provide an easy way of calculating the image of vectors under linear transformations or of calculating the composition of linear transformations. Furthermore, many properties of a linear transformation, such as its rank, minimal polynomial, invariant subspaces, can be easily obtained by analysing a matrix corresponding to that linear transformation.

Suppose that $\psi : V \to V'$ is a linear transformation between two vector spaces V and V' over a field $\mathbb{F}$ of dimensions n and m respectively. Suppose further that V has a basis $B = \{e_1, \ldots, e_n\}$ and V' has a basis $B' = \{e'_1, \ldots, e'_m\}$. Then there exist $a_{ij} \in \mathbb{F}$ such that $\psi(e_i) = \sum_{j=1}^{m} a_{ij} e'_j$ $(1 \leq i \leq n)$, and the matrix of the linear transformation ψ

with respect to the bases B and B' is defined as the $m \times n$ matrix A whose entries are a_{ij}. Any $v \in V$ is given by $v = \sum_{j=1}^{n} v_j e_j$ for some $v_i \in \mathbb{F}$. In this monograph, we represent vectors as *column vectors* or $n \times 1$ matrices. Thus the vector v is given by the column vector $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ with respect to the basis B, which we write as $(v_1, \ldots, v_n)^T$. The effect of the linear transformation ψ on the vector v is given by

$$\psi(v) = \sum_{j=1}^{n} v_j \psi(e_j) = \sum_{j=1}^{n} v_i \left(\sum_{i=1}^{m} a_{ij} e'_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} v_j \right) e'_i,$$

which is expressed in terms of matrices by the matrix multiplication

$$Av = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + \ldots + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + \ldots + a_{mn}v_n \end{pmatrix}.$$

The composition of linear transformations can also be easily computed using matrices. If $\psi : V \to V'$ and $\psi' : V' \to V''$ are linear transformations, and A and A' are the matrices associated with ψ and ψ' respectively, then $P = A'A$ is the matrix associated with the linear transformation $\psi' \circ \psi : V \to V''$. Thus the matrix of a composition of linear transformations is the product of the respective matrices.

We note that the matrix corresponding to a linear transformation is not unique as it depends on the basis chosen for the vector spaces. Suppose, as above, that we have a linear transformation $\psi : V \to V'$ between two vector spaces V and V' of dimension m and n respectively. If the linear transformation ψ is represented by an $m \times n$ matrix A with respect to one pair of bases and by another $m \times n$ matrix $\widetilde{A}$ with respect to another pair of bases, then there exist an invertible $n \times n$ matrix P and an invertible $m \times m$ matrix P' such that $\widetilde{A} = P'AP$. We say that the matrix $\widetilde{A}$ is obtained from A by a *change of basis.*

DEFINITION 2.49 Let A be an $n \times n$ matrix over the field $\mathbb{F}$. The *minimal polynomial* of the matrix A is the unique monic polynomial $\min_A(x) \in \mathbb{F}[x]$ of minimal degree such that $\min_A(A) = 0$. The *characteristic polynomial* of A is the polynomial $c_A(x) \in \mathbb{F}[x]$ defined by

$$c_A(x) = \det(xI - A).$$

We note that the minimal polynomial of a linear transformation is the same as the minimal polynomial of any of its associated matrices.

THEOREM 2.50 *Cayley–Hamilton Theorem.* The minimal polynomial of a matrix divides the characteristic polynomial.

EXAMPLE 2.51 Consider the matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

over the field $\mathbb{Z}_2$. It can be shown that the minimal polynomial of A is $\min_A(x) = x^3 + 1$ and the characteristic polynomial of A is

$$c_A(x) = x^4 + x^3 + x + 1.$$

We note that the minimal polynomial $\min_A(x)$ divides $c_A(x)$. Furthermore, $\min_A(x) = (x+1)(x^2+x+1)$ as a product of irreducible polynomials. Thus, if $\psi : V \to V'$ is a linear transformation associated with A, then the invariant subspaces of ψ are given by $\ker(\psi_1)$ and $\ker(\psi_2)$, where ψ_1 and ψ_2 are the linear transformations $V \to V'$ associated with the matrices $(A + I)$ and $(A^2 + A + I)$ respectively. □

Matrices are also widely used in coding theory, and most properties of linear codes can obtained by studying their generator and parity check matrices. Of special interest in the design and analysis of the AES are the matrices that arise from *maximal distance separable (MDS) codes* [76].

DEFINITION 2.52 An $m \times n$ matrix A is called an MDS matrix if and only if every square submatrix of A is invertible.

Linear systems and matrix complexity

Matrices can be used to represent systems of linear equations. Suppose we have such a system of m equations in n variables $x_1, \ldots, x_n$ given by

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ \vdots & \quad \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m, \end{aligned}$$

where a_{ij} and b_i are elements of a field $\mathbb{F}$. This equation system can be represented by the matrix equation

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

or equivalently $Ax = b$. The standard process of solving this equation system is to transform the matrix A to a row-reduced echelon matrix using elementary row operations. This corresponds to finding an invertible $m \times m$ matrix P such that PA is a row reduced echelon matrix. This allows us to obtain an equivalent matrix equation $PAx = Pb$, which gives us an immediate full solution for $x_1, \ldots, x_n$.

The simplest method of transforming the matrix A to a row reduced echelon matrix is known as *Gaussian reduction*. Performing Gaussian reduction on a square $n \times n$ matrix takes of the order of n^3 field operations. However, more sophisticated techniques for row-reducing a matrix can reduce this to less than cubic complexity.

DEFINITION 2.53 An $n \times n$ square matrix can be transformed to a row-reduced echelon matrix with complexity of the order of n^ω field operations. We call ω the *exponent of matrix reduction*. Thus $\omega = 3$ for Gaussian reduction. The smallest values of ω occur for row-reduction techniques for a *sparse matrix*, that is a matrix whose almost all entries are zero. The exponent of matrix reduction ω satisfies $2 < \omega \leq 3$.

Algebras

DEFINITION 2.54 Suppose $\mathcal{A}$ is a vector space over a field $\mathbb{F}$ with a multiplication operation $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$. If this multiplication operation is associative and a bilinear mapping on the vector space $\mathcal{A}$, then $\mathcal{A}$ is an (associative) $\mathbb{F}$-*algebra*, or more simply an *algebra*.

Informally, we can regard an algebra as a vector space that is also a ring. The dimension of the algebra $\mathcal{A}$ is the dimension of $\mathcal{A}$ as a vector space. The subset $\mathcal{A}' \subset \mathcal{A}$ is a *subalgebra* of $\mathcal{A}$ if $\mathcal{A}'$ is an algebra in its own right, and $\mathcal{A}'$ is an *ideal subalgebra* if it is also an ideal of the ring $\mathcal{A}$. We can also classify mappings between two algebras in the usual way, so an *algebra homomorphism* is a mapping that is both a ring homomorphism and a vector space homomorphism.

EXAMPLE 2.55 The ring of polynomials $\mathbb{F}[x_1, \ldots, x_n]$ is a vector space over $\mathbb{F}$ (Example 2.40). Thus $\mathbb{F}[x_1, \ldots, x_n]$ forms an $\mathbb{F}$-algebra, known as a *polynomial algebra*. □

EXAMPLE 2.56 The set $\mathcal{M}_n(\mathbb{F})$ of $n \times n$ matrices over $\mathbb{F}$ forms a vector space over $\mathbb{F}$ of dimension n^2 (Example 2.47). Matrix multiplication is an associative bilinear mapping on $\mathcal{M}_n(\mathbb{F})$. Thus $\mathcal{M}_n(\mathbb{F})$ forms an $\mathbb{F}$-algebra of dimension n^2. The set $\mathcal{D}_n(\mathbb{F})$ of $n \times n$ diagonal matrices over $\mathbb{F}$ forms a subalgebra of $\mathcal{M}_n(\mathbb{F})$ of dimension n. Such algebras are known as *matrix algebras*. □

4. Finite Fields

The design of the AES is based around finite fields. All the operations used by the AES are described by algebraic operations on a finite field of even characteristic. In this section, we discuss the properties of finite fields relevant to the specification and algebraic analysis of the AES.

Finite fields and subfields

The set $\mathbb{Z}_p = \{0, \ldots, p-1\}$ with addition and multiplication operations defined modulo p forms a finite field if and only if p is prime (Example 2.20). This field is called the *Galois field* of order p and is denoted by $\mathrm{GF}(p)$. The Galois field $\mathrm{GF}(p)$ plays a fundamental role in the theory of finite fields.

DEFINITION 2.57 Suppose that $\mathbb{F}$ and $\mathbb{K}$ are two fields. If $\mathbb{F} \subset \mathbb{K}$, then $\mathbb{F}$ is said to be a *subfield* of $\mathbb{K}$, or equivalently $\mathbb{K}$ is said to be an *extension field* of $\mathbb{F}$.

THEOREM 2.58 A finite field of characteristic p (prime) has a unique minimal subfield isomorphic to $\mathrm{GF}(p)$.

If $\mathbb{K}$ is a extension field of the field $\mathbb{F}$, then $\mathbb{K}$ is also a vector space over $\mathbb{F}$. The dimension of this vector space is the *degree* of the extension. If $\mathbb{F}$ has order q and $\mathbb{K}$ is an extension field of $\mathbb{F}$ of degree d, then $\mathbb{K}$ has order q^d. As every finite field has prime characteristic p, it follows from Theorem 2.58 that every finite field has order p^n for some prime p and some positive integer n.

THEOREM 2.59 For every prime number p and every positive integer n, there exists a finite field of order p^n. Furthermore, any two finite fields of order p^n are isomorphic.

Thus finite fields of order p^n are unique up to isomorphism. This field is called the *Galois field* of order p^n and denoted by $\mathrm{GF}(p^n)$. A subfield of $\mathrm{GF}(p^n)$ has order p^d, where d is a divisor of n. Furthermore, there is exactly one subfield of order p^d for every divisor d of n. For example, the finite field $\mathrm{GF}(2^8)$ has $\mathrm{GF}(2^4)$, $\mathrm{GF}(2^2)$, and $\mathrm{GF}(2)$ as proper subfields.

THEOREM 2.60 The multiplicative group $\mathrm{GF}(q)^*$ is a cyclic group of order $q-1$.

A generator of the multiplicative group $\mathrm{GF}(q)^*$ is called a *primitive element* of the field $\mathrm{GF}(q)$. The number of primitive elements in $\mathrm{GF}(q)$ is $\varphi(q-1)$, where $\varphi(m)$ is *Euler's totient function*, which gives the number of positive integers less than or equal to m and coprime to m.

Explicit construction of finite fields

Theorem 2.27 provides a method of constructing a finite field as a quotient ring. Suppose $\mathbb{F}$ is a finite field of order $q = p^n$ and $f(x) \in \mathbb{F}[x]$ is an irreducible polynomial of degree d. The quotient ring $\mathbb{K} = \frac{\mathbb{F}[x]}{\langle f(x) \rangle}$ is a field of order $q^d = p^{nd}$, which is an extension field of degree d of $\mathbb{F}$. In the manner given in Example 2.25, its elements can be represented as

$$a_{d-1}x^{d-1} + \ldots + a_2x^2 + a_1x + a_0,$$

where $a_i \in \mathbb{F}$. Addition and multiplication are then as described in Example 2.25. Theorem 2.59 states that any finite field of order p^{nd} is isomorphic to $\mathbb{K}$.

We can also construct $\mathrm{GF}(p^{nd})$ directly as an *extension field* of $\mathbb{F}$. Let θ denote a root of the irreducible polynomial $f(x)$ of degree d. The set $\mathbb{F}(\theta)$ of all quotients (with nonzero denominator) of polynomials in θ with coefficients in $\mathbb{F}$ is the smallest field containing both θ and $\mathbb{F}$. Furthermore, $\mathbb{F}(\theta)$ is the extension field obtained by *adjoining* θ to $\mathbb{F}$. This extension field $\mathbb{F}(\theta)$ has p^{nd} elements and so is isomorphic to $\mathrm{GF}(p^{nd})$. The elements of $\mathbb{F}(\theta)$ are given by

$$a_{d-1}\theta^{d-1} + \ldots + a_2\theta^2 + a_1\theta + a_0,$$

where $a_i \in \mathbb{F}$. If the element θ is a generator of the multiplicative group of $\mathbb{F}(\theta)$, then the polynomial $f(x)$ is called a *primitive polynomial.*

EXAMPLE 2.61 The polynomial $m(x) = x^8 + x^4 + x^3 + x + 1 \in \mathrm{GF}(2)[x]$ is irreducible. If θ is a root of $m(x)$, then

$$\mathrm{GF}(2)(\theta) \cong \frac{\mathrm{GF}(2)[x]}{\langle m(x) \rangle} \cong \mathrm{GF}(2^8).$$

The elements of the quotient ring $\frac{\mathrm{GF}(2)[x]}{\langle m(x) \rangle}$ are given, for $a_i \in \mathbb{F}$, by

$$a_7x^7 + \ldots + a_2x^2 + a_1x + a_0;$$

whereas the elements of extension field $\mathbb{F}(\theta)$ are given, for $b_i \in \mathbb{F}$, by

$$b_7\theta^7 + \ldots + b_2\theta^2 + b_1\theta + b_0.$$

We note that $m(x)$ is not primitive, since the order of $\theta \in \mathbb{F}(\theta)$ is 51. □

Irreducible polynomials over a field $\mathbb{F}$ of order q are the basic tools for the construction of all finite extensions of $\mathbb{F}$. If $\mathbb{K}$ is an extension of $\mathbb{F}$ of order q^n, then Theorem 2.60 shows that $a^{q^n-1} - 1 = 0$ for all

nonzero $a \in \mathbb{K}$. Thus the polynomial $x^{q^n} - x$ has all q^n elements of $\mathbb{K}$ as a root. The field $\mathbb{K} \cong \mathrm{GF}(q^n)$ is known as the *splitting field* of the polynomial $x^{q^n} - x$. This polynomial can be used to obtain all irreducible polynomials over $\mathbb{F}$ with the required degree.

THEOREM 2.62 Let $\mathbb{F}$ be a finite field of order q. Then the polynomial $x^{q^n} - x \in \mathbb{F}[x]$ is the product of all monic irreducible polynomials in $\mathbb{F}[x]$ whose degree divides n.

The number of irreducible polynomials in $\mathbb{F}[x]$ of degree n is given by

$$\frac{1}{n} \sum_{d|n} \mu(d) q^{\frac{n}{d}} ,$$

where μ is the *Möbius function*, defined by $\mu(1) = 1$, $\mu(n) = (-1)^k$ if n is the product of k distinct primes, and 0 otherwise. The number of primitive polynomials of degree n is $\frac{1}{n}\varphi(q^n - 1)$, where φ is Euler's totient function.

EXAMPLE 2.63 There are $\frac{1}{8}\left(\mu(1)2^8 + \mu(2)2^4 + \mu(4)2^2 + \mu(8)2^1\right) = 60$ irreducible polynomials of degree 8 in $\mathrm{GF}(2)[x]$, of which $\frac{1}{8}\varphi(2^8 - 1) = 16$ are primitive polynomials. □

DEFINITION 2.64 A field $\mathbb{F}$ is said to be *algebraically closed* if every polynomial in $\mathbb{F}[x]$ has a root in $\mathbb{F}$. The *algebraic closure* of a field $\mathbb{F}$ is the smallest extension field $\mathbb{K}$ of $\mathbb{F}$ such that $\mathbb{K}$ is algebraically closed.

Representations of a finite field

Let $\mathbb{F}$ be a field and $\mathbb{K} = \mathbb{F}(\theta)$ be an extension field of $\mathbb{F}$ of degree d. The most common way to describe the elements of $\mathbb{K}$ is to regard all elements as vectors in the vector space $\mathbb{K}$ of dimension d over the $\mathbb{F}$. Every element in $\mathbb{K}$ can be written uniquely as

$$a_{d-1}\theta^{d-1} + \ldots + a_2\theta^2 + a_1\theta + a_0,$$

where $a_i \in \mathbb{F}$. Thus the set $\{\theta^{d-1}, \ldots, \theta^2, \theta, 1\}$ forms a basis of $\mathbb{K}$ as a d-dimensional vector space over $\mathbb{F}$. This basis is called a *polynomial basis* for the field $\mathbb{K}$.

EXAMPLE 2.65 Suppose θ is a root of $x^8 + x^4 + x^3 + x + 1 \in \mathrm{GF}(2)[x]$, and let $\mathbb{K}$ be the field $\mathrm{GF}(2)(\theta)$ (Example 2.61). Any multiplication mapping $\mathbb{K} \to \mathbb{K}$ is a linear transformation of $\mathbb{K}$ as a vector space over $\mathrm{GF}(2)$. The squaring mapping in $\mathbb{K}$ is also a linear transformation. If we let T_θ and S denote the matrices that correspond to multiplication

by θ and squaring with respect to the polynomial basis $\{\theta^7, \ldots, \theta^2, \theta, 1\}$, then we have

$$T_\theta = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad S = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

□

There are other bases which are used for the field $\mathbb{K}$ when considered as a vector space over $\mathbb{F}$, such as the *normal basis* $\{\beta, \beta^q, \beta^{q^2}, \ldots, \beta^{q^{d-1}}\}$ for suitable $\beta \in \mathbb{K}$. This representation is particularly useful when performing the exponentiation of elements in $\mathbb{K}$ and may offer implementation advantages in some situations.

There are also methods of describing an element of the finite field $\mathbb{K}$ of order q^n which depend on logarithmic functions of $\mathbb{K}$ rather than the vector space aspect of $\mathbb{K}$. Suppose β is a primitive element of $\mathbb{K}$ and $a = \beta^i$ $(0 \leq i < q^n - 1)$. The *discrete logarithm* is a function $\log_\beta : \mathbb{K}^* \to \mathbb{Z}_{q^n-1}$ defined by $\log_\beta a = \log_\beta \beta^i = i$. We can thus represent the nonzero elements $a \in \mathbb{K}$ by $\log_\beta a \in \mathbb{Z}_{q^n-1}$. If we adopt the convention that the discrete logarithm of 0 is denoted by ∞, then we can represent an element of $\mathbb{K}$ by an element of $\overline{\mathbb{Z}}_{q^n-1} = \mathbb{Z}_{q^n-1} \cup \{\infty\}$.

The *Zech* or *Jacobi logarithm* offers another logarithmic method for describing a finite field element. The Zech logarithm is based on the function $Z: \overline{\mathbb{Z}}_{q^n-1} \to \overline{\mathbb{Z}}_{q^n-1}$ given by

$$Z(n) = \log_\beta(\beta^n + 1),$$

so $\beta^{Z(n)} = \beta^n + 1$ with the convention that $\beta^\infty = 0$. The definition can be extended to all integers by working modulo $q^n - 1$. The Zech logarithm of β^n can now be defined to be $Z(n)$. We have the following identities concerning this function Z:

$$\begin{aligned} Z(Z(n)) &= n, \\ Z(2n) &= 2Z(n), \\ Z(-n) &= Z(n) - n. \end{aligned}$$

This function is of interest since it can be used to calculate the sum of two powers of β, since

$$\beta^m + \beta^n = \beta^n(\beta^{m-n} + 1) = \beta^n \beta^{Z(m-n)} = \beta^{n+Z(m-n)}.$$

Functions in a finite field

DEFINITION 2.66 Let $\mathbb{F}$ be a finite field of order q and $\mathbb{K}$ be an extension field of $\mathbb{F}$ of degree d. The elements $a, a^q, a^{q^2}, \ldots, a^{q^{d-1}}$ are the *conjugates* of $a \in \mathbb{K}$ with respect to $\mathbb{F}$.

THEOREM 2.67 Suppose $\mathbb{K}$ is an extension of a field $\mathbb{F}$ of degree d. Any element $a \in \mathbb{K}$ is a root of an irreducible polynomial $f(x) \in \mathbb{F}[x]$ of degree n dividing d. The roots of $f(x)$ are the conjugates of a.

We now consider some functions of interest on finite fields.

DEFINITION 2.68 Let $\mathbb{F}$ be a finite field of order q and $\mathbb{K}$ be an extension field of $\mathbb{F}$ of degree d. The *trace* function on $\mathbb{K}$ with respect to $\mathbb{F}$ is the function $\mathrm{Tr}: \mathbb{K} \to \mathbb{F}$ defined by

$$\mathrm{Tr}(a) = a + a^q + a^{q^2} + \ldots + a^{q^{d-1}}.$$

Thus the trace of an element $a \in \mathbb{K}$ is the sum of all conjugates of a. The trace function is a linear functional on $\mathbb{K}$, considered as a vector space over $\mathbb{F}$ (Example 2.43). In fact, any linear functional on $\mathbb{K}$ is of the form $a \mapsto \mathrm{Tr}(\beta a)$ for some $\beta \in \mathbb{K}$.

DEFINITION 2.69 Let $\mathbb{F}$ be a finite field of order q and $\mathbb{K}$ be an extension field of $\mathbb{F}$ of degree d. The *norm* function on $\mathbb{K}$ with respect to $\mathbb{F}$ is the function $\mathrm{N}: \mathbb{K} \to \mathbb{F}$ defined by

$$\mathrm{N}(a) = a\ a^q\ a^{q^2}\ \ldots\ a^{q^{d-1}} = a^{\frac{q^d-1}{q-1}}.$$

Thus the norm of an element $a \in \mathbb{K}$ is the product of all conjugates of a. The norm function is a group homomorphism $\mathbb{K}^* \to \mathbb{F}^*$ between the multiplicative groups of the fields $\mathbb{K}$ and $\mathbb{F}$.

DEFINITION 2.70 A *linearised polynomial* $f(x) \in \mathbb{K}[x]$ is a polynomial given by

$$f(x) = a_0 x + a_1 x^q + a_2 x^{q^2} + \ldots + a_{d-1} x^{q^{d-1}},$$

where $a_i \in \mathbb{K}$. Thus a linearised polynomial $f(x)$ is a polynomial whose evaluation $f(a)$ for any $a \in \mathbb{K}$ gives a linear combination of the d conjugates of a.

Linearised polynomials are linear transformations on $\mathbb{K}$, when considered as a vector space over $\mathbb{F}$. Conversely, any linear transformation of $\mathbb{K}$ over $\mathbb{F}$ can be expressed as a linearised polynomial.

EXAMPLE 2.71 Any linear transformation of $\mathrm{GF}(2^8)$ as a vector space over $\mathrm{GF}(2)$ can be represented by a (linearised) polynomial of the form $f(x) = a_0x^{2^0} + a_1x^{2^1} + a_2x^{2^2} + \ldots + a_7x^{2^7}$, where $a_i \in \mathrm{GF}(2^8)$. □

We now consider the field $\mathrm{GF}(p^d)$ as an extension field of $\mathrm{GF}(p)$, where p is prime. The mapping $\tau\colon \mathrm{GF}(p^d) \to \mathrm{GF}(p^d)$ defined by $a \mapsto a^p$ maps a to one of its conjugates with respect to $\mathrm{GF}(p)$. This mapping satisfies

$$\tau(a + a') = \tau(a) + \tau(a') \quad \text{and} \quad \tau(aa') = \tau(a)\tau(a').$$

Thus τ is a field automorphism of $\mathrm{GF}(p^d)$, known as the *Frobenius* automorphism. The set of all automorphisms of $\mathrm{GF}(p^d)$ under the operation of composition is the cyclic group of order d generated by τ. We note that τ fixes all elements of the subfield $\mathrm{GF}(p)$ of $\mathrm{GF}(p^d)$. Thus the automorphisms of $\mathrm{GF}(p^d)$ are also linear transformations over $\mathrm{GF}(p)$.

5. Varieties and Gröbner Bases

A large part of this monograph is concerned with expressing an AES encryption as a system of polynomial equations and considering methods of solution for such equations. In this section, we give a brief overview of the basic concepts used to analyse such equation systems.

Varieties

An *affine subset* of a vector space V is a coset or translate $U + u$ of some subspace $U < V$. The *affine space* based on V is the geometrical space given by considering certain geometrical properties of the affine subsets of V [58]. Thus we can usually identify the n-dimensional affine space over a field $\mathbb{F}$ with $\mathbb{F}^n$. The projective space $PG(n, \mathbb{F})$ is the geometrical space given by considering the one-dimensional subspaces of the $(n + 1)$-dimensional vector space $\mathbb{F}^{n+1}$. Thus we can represent an element of $PG(n, \mathbb{F})$ by a nonzero vector $(a_0, a_1, \ldots, a_n) \in \mathbb{F}^{n+1}$, where all nonzero scalar multiples of $(a_0, a_1, \ldots, a_n)$ represent the same element of $PG(n, \mathbb{F})$.

DEFINITION 2.72 Let $\mathbb{F}$ be a field and $\mathbb{F}^n$ denote the n-dimensional affine space over $\mathbb{F}$, and suppose that $f_1, \ldots, f_m$ are polynomials in $\mathbb{F}[x_1, \ldots, x_n]$. The *affine variety* defined by $f_1, \ldots, f_m$ is the subset of $\mathbb{F}^n$ given by

$$\{\ (a_1, \ldots, a_n) \in \mathbb{F}^n \mid f_i(a_1, \ldots, a_n) = 0 \quad \text{for } i = 1, \ldots, m\ \}.$$

This variety is denoted by $\mathcal{V}(f_1, \ldots, f_m)$.

Thus the affine variety of Definition 2.72 describes the set of solutions in $\mathbb{F}$ of the polynomial equation system

$$f_1(x_1, \ldots, x_n) = 0, \ \ldots \ , f_m(x_1, \ldots, x_n) = 0.$$

EXAMPLE 2.73 Consider the polynomial ring $\mathbb{R}[x, y]$ in two variables, and let $f(x, y) = x^2 + y^2 - 1$ and $g(x, y) = x - 1$ be two polynomials in $\mathbb{R}[x, y]$. The affine variety $\mathcal{V}(f)$ consists of the points in the circle of radius 1 in $\mathbb{R}^2$ and is the solution set of the equation $x^2 + y^2 = 1$. The affine variety $\mathcal{V}(f, g) = \{(1, 0)\} \in \mathbb{R}^2$ is the set of solutions to $f(x, y) = g(x, y) = 0$. □

DEFINITION 2.74 Let $\mathrm{PG}(n, \mathbb{F})$ denote the projective space of dimension n. Suppose that $f_1, \ldots, f_m$ are homogeneous polynomials in the polynomial ring $\mathbb{F}[x_0, x_1, \ldots, x_n]$. The *projective variety* defined by the polynomials $f_1, \ldots, f_m$ is the subset of $\mathrm{PG}(n, \mathbb{F})$ given by

$$\{ \ (a_0, a_1, \ldots, a_n) \in \mathrm{PG}(n, \mathbb{F}) \mid f_i(a_0, a_1, \ldots, a_n) = 0 \text{ for } i = 1, \ldots, m \ \}.$$

The projective space $\mathrm{PG}(n, \mathbb{F})$ can be partitioned into two subsets U and H, where

$$\begin{aligned} U &= \{ \ (a_0, a_1, \ldots, a_n) \in \mathrm{PG}(n, \mathbb{F}) \mid a_0 \neq 0 \ \}, \text{ and} \\ H &= \{ \ (0, a_1, \ldots, a_n) \in \mathrm{PG}(n, \mathbb{F}) \ \}. \end{aligned}$$

The subset U can be identified with the affine space $\mathbb{F}^n$ by using the mapping

$$(a_0, a_1, \ldots, a_n) \mapsto \left(\frac{a_1}{a_0}, \ldots, \frac{a_n}{a_0} \right).$$

Furthermore, the subset H can be identified with the projective space $\mathrm{PG}(n-1, \mathbb{F})$ by using the mapping $(0, a_1, \ldots, a_n) \mapsto (a_1, \ldots, a_n)$. Thus the projective space $\mathrm{PG}(n, \mathbb{F})$ can be partitioned into an affine space U and a projective space H of smaller dimension. The projective part H is known as the *hyperplane at infinity* of $\mathrm{PG}(n, \mathbb{F})$.

Given a projective variety $\mathcal{W} \in \mathrm{PG}(n, \mathbb{F})$, the set $\mathcal{V} = \mathcal{W} \cap U$ can be considered as an affine variety of $\mathbb{F}^n$ and is called the *affine portion* of $\mathcal{W}$. Thus every projective variety $\mathcal{W}$ can be seen as consisting of an affine variety $\mathcal{V}$ together with its points at infinity $\mathcal{W} \cap H$. Theorem 2.75 summarises the relationship between an affine and a projective variety.

THEOREM 2.75 Let $\mathcal{V} \subset \mathbb{F}^n$ be the affine variety defined by the polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$. If f_i^h denotes the homogenisation of the polynomial f_i, then the variety $\mathcal{W}$ defined by the polynomials $f_1^h, \ldots, f_m^h \in \mathbb{F}[x_0, x_1, \ldots, x_n]$ is a projective variety of $\mathrm{PG}(n, \mathbb{F})$, of which the affine portion is $\mathcal{W} \cap U = \mathcal{V}$.

The above definitions of affine and projective varieties are given in terms of a finite set of polynomials. However, Theorem 2.76 shows that varieties are in fact defined by polynomial ideals.

THEOREM 2.76 Let I be an ideal of $\mathbb{F}[x_1, \ldots, x_n]$. If $\mathcal{V}(I)$ denotes the set

$$\{ (a_1, \ldots, a_n) \in \mathbb{F}^n \mid f(a_1, \ldots, a_n) = 0 \text{ for } f \in I \},$$

then $\mathcal{V}(I)$ is an affine variety. Furthermore, if $I = \langle f_1, \ldots, f_m \rangle$, then $\mathcal{V}(I) = \mathcal{V}(f_1, \ldots, f_m)$.

Similarly, a projective variety can be defined by a *homogeneous ideal* of $\mathbb{F}[x_0, x_1, \ldots, x_n]$, that is an ideal which is generated by homogeneous polynomials.

Gröbner bases

Theorem 2.76 means that the problem of finding the solutions of a polynomial equation system over a field $\mathbb{F}$ is often studied in the context of commutative algebra and polynomial ideals. The solution set of a particular system

$$f_1(x_1, \ldots, x_n) = 0, \ \ldots \ , f_m(x_1, \ldots, x_n) = 0$$

can be found by computing the variety $\mathcal{V}(I)$, where $I = \langle f_1, \ldots, f_m \rangle$. In particular, any generating set of I can be used to compute $\mathcal{V}(I)$. The *Hilbert Basis Theorem* states that any ideal $I \lhd \mathbb{F}[x_1, \ldots, x_n]$ is finitely generated. A Gröbner basis of the polynomial ideal I is a particular type of generating set of I and can be particularly useful in obtaining various properties of I, including the variety $\mathcal{V}(I)$.

DEFINITION 2.77 Suppose that $\mathbb{F}[x_1, \ldots, x_n]$ is a polynomial ring over a field $\mathbb{F}$ with a monomial ordering and that $I \lhd \mathbb{F}[x_1, \ldots, x_n]$ is a non-trivial ideal. We let $\mathrm{LT}(I)$ denote the set of all leading terms of elements of I and $\langle \mathrm{LT}(I) \rangle$ denote the ideal generated by the monomials in $\mathrm{LT}(I)$. A finite set $G = \{g_1, \ldots, g_s\} \subset I$ is said to be a *Gröbner basis* of I if

$$\langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_s) \rangle = \langle \mathrm{LT}(I) \rangle.$$

Thus G is a Gröbner basis of I if and only if the leading term of any polynomial in I is divisible by at least one of the leading terms $\{\mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_s)\}$.

Every non trivial ideal $I \lhd \mathbb{F}[x_1, \ldots, x_n]$ has a Gröbner basis, which is a generating set or basis for the ideal I. If G is a Gröbner basis of I

and $f \in I$, then the set $G \cup \{f\}$ satisfies Definition 2.77 and is also a Gröbner basis of I. Thus an ideal does not have a unique Gröbner basis.

DEFINITION 2.78 A *reduced Gröbner basis* for I is a Gröbner basis G such that the leading coefficient of every polynomial in G is 1 and none of the monomials of any $f \in G$ is divisible by the leading term of any other polynomial in G. Thus in a reduced Gröbner basis G, no monomial of $f \in G$ belongs to the ideal $\langle\ LT(G \setminus \{f\})\ \rangle$.

Every non–trivial ideal I of $\mathbb{F}[x_1, \ldots, x_n]$ has a unique reduced Gröbner basis (with respect to a specific monomial ordering). We can obtain the reduced Gröbner basis for I from a Gröbner basis G for I by dividing or reducing each $f \in G$ by the set $G \setminus \{f\}$.

EXAMPLE 2.79 We consider the ring of real polynomials in three variables $\mathbb{R}[x, y, z]$ with the *grevlex* ordering. The set

$$\{z^6 - x^2y,\ yz^4 + x,\ xy^2 + z^2\}$$

is a (reduced) Gröbner basis for the ideal of $\mathbb{R}[x, y, z]$ generated by these three polynomials. By contrast, consider the set

$$G = \{xy^2 + zx,\ y^2z + z^2 - y\}$$

and the ideal I generated by these two polynomials. We have

$$xy = z(xy^2 + xz) - x(y^2z + z^2 - y),$$

so $xy \in I$. However, xy is not divisible by the leading term of either polynomial in G (xy^2 or y^2z). Thus G is not a Gröbner basis for the ideal I. □

Theorem 2.80 gives a sufficient condition in terms of the *greatest common divisor* of pairs of leading monomials for identifying whether a set is a Gröbner basis of a polynomial ideal.

THEOREM 2.80 Suppose $G \subset \mathbb{F}[x_1, \ldots, x_n]$ is a set of polynomials such that $\gcd(\mathrm{LM}(f), \mathrm{LM}(g)) = 1$ for all distinct $f, g \in G$. Then G is a Gröbner basis for the ideal $\langle G \rangle$.

Thus, if the leading monomials of all polynomials in a set G are pairwise coprime, then G is a Gröbner basis for the ideal generated by the polynomials of G. However, Example 2.79 shows that the condition of Theorem 2.80 is not necessary for a set G to be a Gröbner basis of $\langle G \rangle$.

Gröbner bases are an extremely powerful concept, with many applications in commutative algebra, algebraic geometry, and computational

algebra. For example, Gröbner bases can be used to solve the ideal membership problem, that is to decide whether a polynomial f is in an ideal $I \lhd \mathbb{F}[x_1, \ldots, x_n]$. A polynomial f is in I if and only if f reduces to zero with respect to a Gröbner basis of I, that is the division of f by a Gröbner basis of I has remainder zero (Theorem 2.37).

The main relevance of Gröbner bases to cryptology is the problem of solving polynomial equation systems. If we have such a system

$$f_1(x_1, \ldots, x_n) = 0, \ldots, f_m(x_1, \ldots, x_n) = 0,$$

then we can find its solution set by computing the Gröbner basis for the ideal $I = \langle f_1, \ldots, f_m \rangle$ and computing the associated variety $\mathcal{V}(I)$. The Gröbner basis of I provides implicit solutions to the equation system over the algebraic closure of the field $\mathbb{F}$. A particularly useful monomial ordering for finding solutions to this polynomial equation system in $\mathbb{F}$ is the *lex* ordering, which is an example of an *elimination ordering*.

It is worth noting that equation systems arising in cryptography often display many properties. Cryptographic equation systems are often defined over a small finite field $\mathrm{GF}(q)$ and the solutions of cryptographic interest lie in this field. In this case, we could add the field relations $x_i^q - x_i$ to the original equation system. In this way the solutions of the extended equation system are restricted to the base field $\mathrm{GF}(q)$. Furthermore, cryptographic equation systems often have a unique solution $(a_1, \ldots, a_n) \in \mathrm{GF}(q)^n$. In this case, the reduced Gröbner basis of the ideal corresponding to the extended equation system would be $\{x_1 - a_1, \ldots, x_n - a_n\}$.

We discuss some methods and algorithms for computing a Gröbner basis of an ideal $I \lhd \mathbb{F}[x_1, \ldots, x_n]$ in Section 6.1.

http://www.springer.com/978-0-387-24363-4

Algebraic Aspects of the Advanced Encryption Standard

Cid, C.; Murphy, S.; Robshaw, M.

2006, VIII, 148 p., Hardcover
ISBN: 978-0-387-24363-4