Introduction to Coding Theory **Técnico–ULisboa, 2026/2027**

Basics of finite fields and linear codes

Lecturer: João Ribeiro

For a more detailed discussion of these concepts, see these excellent lectures notes of Forney and Sections 2.1, 2.2, 2.3, and 5.1 of the Essential Coding Theory book.

Definition 1 (Group) *A pair $G = (S, *)$ with S a set and $* : G \times G \to G$ an operation is a* group *if it satisfies the following properties:*

- ***(Closure)*** *For every $x, y \in S$ we have $x * y \in S$.*
- ***(Associativity)*** *The operation $*$ is associative. That is, $x * (y * z) = (x * y) * z$ for all $x, y, z \in S$.*
- ***(Existence of identity)*** *There is an* identity *$e \in S$ such that $x * e = e * x = x$ for all $x \in S$.*
- ***(Existence of inverses)*** *For every $x \in S$ there is $y \in S$ such that $x * y = y * x = e$. We call y the* inverse *of x and write y as x^{-1}.*

The group is said to be abelian *if $*$ is commutative, i.e., $x * y = y * x$ for all $x, y \in S$.*

For example, the set of integers $\mathbb{Z}$ with the usual addition is a group, but the set of integers with the usual multiplication is not a group. Also, for any integer $n \geq 1$, the set $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ with the operation being addition modulo n is a group. Formally, we should always denote a group as a pair $G = (S, *)$. However, the operation we are considering is often clear from context, and in those cases we may simply refer to the set S as the group.

Another example of a group is the set of real numbers $\mathbb{R}$ together with the usual addition. We may wonder whether we can get a group by combining $\mathbb{R}$ with the usual multiplication. This is not a group, since 0 has no inverse as $0 \cdot x = 0$ for all $x \in \mathbb{R}$. Nevertheless, the set $\mathbb{R} \setminus \{0\}$ together with the usual multiplication is a group – for every $x \in \mathbb{R} \setminus \{0\}$ we know that there exists an inverse $1/x \in \mathbb{R} \setminus \{0\}$ such that $x \cdot 1/x = 1$. Furthermore, the addition and multiplication are commutative and interact nicely, in the sense that

$$a \cdot (b + c) = a \cdot b + a \cdot c,$$

which is called the distributive property. additionmarizing:

- $\mathbb{R}$ with the usual addition is an abelian group, and the identity is 0;
- $\mathbb{R} \setminus \{0\}$ with the usual multiplication is an abelian group, and the identity is 1;
- The two operations satisfy the distributive property.

Intuitively, a *field* is any set S together with two operations $+$ and $\cdot$ that behave like the reals.

Definition 2 (Field) *A triple $(S, +, \cdot)$ with S a set and $+ : S \times S \to S$ and $\cdot : S \times S \to S$ operations is said to be a* field *if it satisfies the following properties:*

- *$(S, +)$ is an abelian group. Denote its identity element by 0;*
- *$(S \setminus \{0\}, \cdot)$ is an abelian group. Denote its identity element by 1;*
- *The two operations satisfy the distributive property.*

As we saw above, $\mathbb{R}$ together with the usual addition and multiplication is a field. The set of complex numbers $\mathbb{C}$ with the usual addition and multiplication of complex numbers is also a field. And the set of rational numbers $\mathbb{Q}$ together with the usual addition and multiplication over $\mathbb{R}$ is also a field. But are there more fields? In particular, are there *finite* fields (those having a finite set of elements)?

Theorem 1 *For any prime p we have that $\mathbb{Z}_p$ together with addition and multiplication modulo p is a field.*

Proof: We prove only that every $\alpha \in \mathbb{Z}_p \setminus \{0\}$ has a multiplicative inverse modulo p.

First, we show that

$$\alpha \cdot \beta = 0 \pmod{p} \tag{1}$$

implies that $\alpha = 0 \pmod{p}$ or $\beta = 0 \pmod{p}$. To see this, note that Equation (1) is equivalent to p dividing $\alpha \cdot \beta$. But since p is prime, this means that p divides at least one of α and β. This implies that for any distinct $\beta, \gamma \in \mathbb{Z}_p$ we have

$$\alpha \cdot \beta \neq \alpha \cdot \gamma \pmod{p}.$$

In turn, this means that for every $\alpha \neq 0 \pmod{p}$ the map $\beta \mapsto \alpha \cdot \beta$ is a bijection, and so there exists β such that $\alpha \cdot \beta = 1 \pmod{p}$. ∎

Theorem 1 shows that for every prime p there exists a field of size p. We use $\mathbb{F}_p$ to denote this field, and we call the size of $\mathbb{F}_p$ its *order*, and we say that a field with finite order is a *finite field*. Are there fields of other orders? It turns out that finite fields only exist for prime power orders.

Theorem 2 *If there exists a field $\mathbb{F}_q$ of order q, then $q = p^r$ for some prime p and integer $r \geq 1$. Furthermore, for every prime p and integer $r \geq 1$ there exists a field of order p^r.*

1 Polynomials

Fix a field $\mathbb{F}$. A nonzero polynomial in a variable x over $\mathbb{F}$ is an expression of the form

$$f(x) = f_0 + f_1 x + f_2 x^2 + \cdots + f_d x^d$$

for some integer $d \geq 1$, where $f_0, \ldots, f_d \in \mathbb{F}$ and $f_d \neq 0$. We say that f has degree d, and write its degree as $\deg f$. additions and multiplications of polynomials are the usual ones. Namely,

$$f(x) + g(x) = \sum i = 0^{\max(\deg f, \deg g)} (f_i + g_i) x^i$$

and

$$f(x) \cdot g(x) = \sum_{i=0}^{\deg f + \deg g} \left(\sum_{j=0}^{i} f_j \cdot g_{i-j} \right) x^i,$$

where we take $g_i = 0$ whenever $i < 0$ or $i > \deg g$. In particular, $\deg(f+g) \leq \max(\deg f, \deg g)$ and $\deg f \cdot g = \deg f \cdot \deg g$.

We denote the set of all such polynomials by $\mathbb{F}[x]$.

The evaluation of f at $\beta \in \mathbb{F}$ is given by

$$f(\beta) = \sum_{i=0}^{d} \alpha_i \cdot \beta^i,$$

where all operations are performed in $\mathbb{F}$, and so $f(\beta) \in \mathbb{F}$ too.

1.1 Basic properties of polynomials

Just like there is a division theorem for integers, stating that for any two positive a and b we can write

$$a = q \cdot b + r$$

for unique q and remainder $r < b$, there is also a division theorem for polynomials. In particular, $a = r \pmod{b}$.

Theorem 3 *For any two polynomials $f(x), g(x) \in \mathbb{F}[x]$ with $g(x) \neq 0$ there exist unique polynomials $q(x), r(x) \in \mathbb{F}[x]$ such that $\deg r < \deg g$ and*

$$f(x) = q(x)g(x) + r(x).$$

We say that g divides f if the remainder $r(x) = 0$. More generally, we write $f(x) = r(x) \pmod{g(x)}$.

For example, we can use Theorem 3 to prove that if $f(\beta) = 0$, then the polynomial $x - \beta$ divides f. We call β a *root* of f. Indeed, by Theorem 3 we can write

$$f(x) = q(x)(x - \beta) + r(x)$$

with $\deg r = 0$. This means r is constant, and, since $f(\beta) = q(\beta) \cdot 0 + r(\beta)$, we conclude that $r(x) = f(\beta) = 0$.

We can also use Theorem 3 to prove another basic but extremely useful fact about polynomials over fields.

Theorem 4 *Every non-zero polynomial $f(x) \in \mathbb{F}[x]$ of degree d has at most d roots.*

Proof: We prove this by induction.

When $d = 0$ this is immediate.

Now, let $d > 0$ and suppose that the statement is true for polynomials of degree $d-1$. Fix an arbitrary polynomial $f \in \mathbb{F}[x]$ of degree d. If f does not have roots then we are done, so assume that f has a root α. As we saw above, this means that we can write

$$f(x) = q(x)(x - \alpha)$$

for some $q(x)$. Note that $q(x)$ has degree $d-1$. Furthermore, if $\beta \neq \alpha$ is also a root of f, then

$$q(\beta)(\beta - \alpha) = f(\beta) = 0,$$

and so $q(\beta) = 0$ since $\beta - \alpha \neq 0$ and $\mathbb{F}$ is a field. In other words, all roots of f distinct from α are also roots of q. Since q has at most $d-1$ roots by the induction hypothesis, we conclude that f has at most d roots. ■

2 Finite fields from irreducible polynomials

In most of the course it will be fine for you to just think about prime order fields, which as we saw above are just integers modulo a prime. However, it is sometimes good to have a basic idea of how to construct finite fields of all prime *power* orders.

Definition 3 *A polynomial $f \in \mathbb{F}[x]$ is* irreducible *if $f(x) = g(x)h(x)$ implies that either $\deg g = 0$ or $\deg h = 0$.*

For example, degree-1 polynomials are irreducible. Also, $1 + x + x^2$ is irreducible over $\mathbb{F}_2$, but $1 + x^2 = (1 + x)^2$ is not.

Irreducible polynomials play the same role for $\mathbb{F}[x]$ as prime numbers do for the integers $\mathbb{Z}$. In particular, all polynomials can be written into a product of irreducible polynomials, and taking $\mathbb{F}[x]$ modulo an irreducible polynomial yields a field (the proof follows along the lines of the argument that $\mathbb{Z}_p$ is a field for p prime, replacing "prime" by "irreducible polynomial").

Theorem 5 *Let p be a prime and $E(x) \in \mathbb{F}_p[x]$ an irreducible polynomial of degree s. Then, the set*

$$\mathbb{F}_{E(x)} = \{f(x) \in \mathbb{F}_p[x] : \deg f < s\}$$

together with addition and multiplication modulo $E(x)$ is a field.

Note that the number of polynomials in $\mathbb{F}_p[x]$ of degree at most $s-1$ is exactly p^s. Therefore, there exists a field of size p^s whenever there exists an irreducible polynomial over $\mathbb{F}_p$ of degree s. Since

for every prime p and integer $s \geq 1$ there exists an irreducible polynomial over $\mathbb{F}_p$ of degree s (we do not show this), there exist finite fields of all prime power orders.

Furthermore, it turns out that all finite fields of the same order are *isomorphic*. This justifies talking about *the* finite field $\mathbb{F}_q$.

3 q-ary linear codes

Just like we defined binary linear codes, we can also define linear codes over any finite field $\mathbb{F}_q$. Consider the set $\mathbb{F}_q^n$ of length-n vectors with entries from $\mathbb{F}_q$. We can see this set as a vector space by considering the following operations:

- **(Vector addition)** If $u, v \in \mathbb{F}_q^n$, then $u + v = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$ with $+$ addition in $\mathbb{F}_q$.
- **(Scalar multiplication)** If $u \in \mathbb{F}_q^n$ and $\alpha \in \mathbb{F}_q$, then $\alpha u = (\alpha u_1, \alpha u_2, \ldots, \alpha u_n)$ with $\cdot$ multiplication in $\mathbb{F}_q$.

We can then generalize the notion of linear independence. Given a set of vectors $v_1, \ldots, v_k \in \mathbb{F}_q^n$, define

$$\operatorname{span}(v_1, \ldots, v_k) = \left\{ \sum_{i=1}^{k} \alpha_i v_i : \alpha_1, \ldots, \alpha_k \in \mathbb{F}_q \right\}.$$

In other words, $\operatorname{span}(v_1, \ldots, v_k)$ is the set of all linear combinations of $v_1, \ldots, v_k$.

Definition 4 (Linear independence) *A collection of vectors $v_1, \ldots, v_k$ is said to be* linearly independent *for every $i \in \{1, \ldots, k\}$ we have $v_i \notin \operatorname{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$. If this does not hold, we say that $v_1, \ldots, v_k$ are* linearly dependent*.*

In words, $v_1, \ldots, v_k$ are linearly independent if we cannot write any v_i as a linear combination of the remaining vectors. If this holds, then all linear combinations of these vectors are distinct, and so

$$|\operatorname{span}(v_1, \ldots, v_k)| = q^k.$$

3.1 Linear codes

A q-ary code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is said to be linear if it is a linear subspace of $\mathbb{F}_q^n$. That is:

- If $c, c' \in \mathcal{C}$, then $c + c' \in \mathcal{C}$, where $+$ is coordinate-wise addition in $\mathbb{F}_q$.
- If $c \in \mathcal{C}$ and $\alpha \in \mathbb{F}_q$, then $\alpha \cdot c \in \mathcal{C}$, where $\cdot$ is coordinate-wise multiplication of c by α.

Recall that we did not enforce the second condition for binary linear codes. That is because this condition is redundant over $\mathbb{F}_2$.

It is not hard to show that every linear subspace $\mathcal{C}$ of $\mathbb{F}_q^n$ can be written as

$$\mathcal{C} = \text{span}(v_1, \ldots, v_k)$$

for some linearly independent vectors $v_1, \ldots, v_k \in \mathbb{F}_q^n$. Then, $|\mathcal{C}| = q^k$ and we call k the *dimension* of $\mathcal{C}$. We can also write $\mathcal{C}$ as the column span of the $n \times k$ matrix

$$G = \begin{pmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_k \\ | & | & \cdots & | \end{pmatrix}.$$

We call G the *generator matrix* of $\mathcal{C}$.

The *rank* of a matrix is the largest number of linearly independent columns. If an $n \times k$ matrix M has rank $\min(k, n)$, then we say that M is *full rank*. Note that the generator matrix G above is full rank.

3.2 Minimum distance of linear codes

As was the case for binary linear codes, the minimum distance of a q-ary linear code is connected to its minimum Hamming weight.

Theorem 6 *A linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ has minimum distance d if and only if*[1]

$$\min_{c \in \mathcal{C} \setminus \{0\}} w_H(c) = d,$$

where, as before, $w_H(c) = |\{i : c_i \neq 0\}|$.

We say that a linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of dimension k and minimum distance at least d is an $[n, k, d]_q$-code.

3.3 Parity-check matrix and dual codes

As we did for binary linear codes, we can also take the dual perspective. Every linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of dimension k can be seen as the kernel of some $(n-k) \times n$ full rank matrix H. That is,

$$\mathcal{C} = \ker(H) = \{v : Hv = 0\}.$$

Given a linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$, we define its dual code $\mathcal{C}^\perp$ as

$$\mathcal{C}^\perp = \{v : \langle v, c \rangle = 0 \text{ for all } c \in \mathcal{C}\},$$

[1] When the context is clear we abuse notation and also use 0 to denote the zero vector $(0, \ldots, 0) \in \mathbb{F}_q^n$.

where $\langle v, c \rangle = \sum_{i=1}^{n} v_i \cdot c_i$ and all operations are done in $\mathbb{F}_q$. Note that $\mathcal{C}^\perp$ is also a q-ary linear code. If $\mathcal{C}$ has dimension k, then $\mathcal{C}^\perp$ has dimension $n - k$ and its generator matrix is H^T, with H the parity-check matrix of $\mathcal{C}$.

Analogously to the binary case, we can determine the minimum distance of $\mathcal{C}$ by inspecting the columns of its parity-check matrix H.

Theorem 7 *Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a linear code with parity-check matrix H. Then, the minimum distance of $\mathcal{C}$ equals the size of the smallest collection of columns of H that are linearly dependent.*