MPRI **Year 2018–19**

Module 2.13.2 : Error correcting codes and applications to cryptography

---

Mid-term exam, November 26

You have 2 hours. Any document including personal lecture notes is authorized.
The exercises are independent.
You can answer either in French or in English.

Exercise 1. (1) (a) Give the list of minimal 2–cyclotomic cosets modulo 9 which permit to classify cyclic codes of length 9 over $\mathbb{F}_2$.

Answer : $\{0\}$, $\{1,2,4,8,7,5\}$, $\{3,6\}$.

(b) How many cyclic codes (including trivial ones) of length 9 over $\mathbb{F}_2$ does there exists ?

Answer : There are 3 minimal cyclotomic cosets so $2^3 = 8$ cyclotomic cosets which gives 8 cyclic codes.

(2) (a) Give the list of minimal 3–cyclotomic cosets modulo 13.

Answer : $\{0\}$, $\{1,3,9\}$, $\{2,6,5\}$, $\{4,12,10\}$, $\{7,8,11\}$.

(b) How many cyclic codes (including trivial ones) of length 13 over $\mathbb{F}_3$ does there exists ?

Answer : 32.

(c) Prove the existence of a $[13, 4, \geqslant 7]_3$ cyclic code and a $[13, 7, \geqslant 5]_3$ cyclic code.

Answer : Using the BCH bound, the code associated to the class $\{1,3,9\} \cup \{2,6,5\} \cup \{4,12,10\}$ contains the consecutive numbers $1,2,3,4,5,6$, hence has minimum distance $\geqslant 7$. Since the class has cardinality 9, the code has dimension $13 - 9 = 4$.
The second code is obtained from the class : $\{2,6,5\} \cup \{7,8,11\}$ which contains $5,6,7,8$ and hence has minimum distance $\geqslant 5$ and dimension 7.

Exercise 2. A code $C \subseteq \mathbb{F}_q^n$ is said to be *non degenerate*, if for any $i \in \{1, \ldots, n\}$, there exists $\mathbf{c} \in C$ such that $c_i \neq 0$.

(1) Reformulate the notion of being *non degenerate* in terms of a generator matrix of C.

Answer : One can reformulate as : *A generator matrix of C has no zero column.*

(2) Reformulate the notion of being *non degenerate* in terms of the minimum distance of $C^\perp$. Justify why this reformulation is equivalent.

Answer : One can reformulate as : *The minimum distance of* $C^\perp$ *is* > 1. Indeed, a result from the course asserts that the minimum distance of a code is the least number of linearly linked columns in a parity check matrix. Since a generator matrix of C is a parity–check matrix of $C^\perp$, the assumption of non degeneracy of C is equivalent to the fact that a generator matrix of C has no zero column, which entails that its dual distance cannot be less than or equal to 1.

Given a non degenerate code $C \subseteq \mathbb{F}_q^n$ and a position $i \in \{1, \dots, n\}$, the *locality of* C *at* i is defined as

$$\mathbf{Loc}(C, i) := \min\{w_H(\mathbf{c}) \mid c \in C^\perp,\ c_i \neq 0\} - 1,$$

where $w_H(\mathbf{x})$ denotes the Hamming weight of $\mathbf{x}$. Next, the *locality* of C is defined as

$$\mathbf{Loc}(C) = \max_{i=1,\dots,n} \{\mathbf{Loc}(C, i)\}.$$

(3) Prove that $\mathbf{Loc}(C) \geqslant d_{\min}(C^\perp) - 1$, where $d_{\min}(\cdot)$ denotes the minimum distance.

Answer : By definition of the locality, for any i, $\mathbf{Loc}(C, i) \geqslant d_{\min}(C^\perp) - 1$. Then, its maximum when i ranges over $\{1, \dots, n\}$ should also be larger than or equal to $d_{\min}(C^\perp) - 1$.

(4) Prove that $\mathbf{Loc}(C) \leqslant \dim(C)$.

Answer : Denote by k the dimension of C. Let $\mathbf{G}$ be a generator matrix of C. Let $i \in \{1, \dots, n\}$. Since $\mathbf{G}$ has k rows, its i–th column is linearly linked to k other ones, which proves the existence of a word of weight $\leqslant k + 1$ in $C^\perp$ whose support contains i. This proves that for any position $i \in \{1, \dots, n\}$, we have $\mathbf{Loc}(C, i) \leqslant k$. Therefore, the code has locality less than or equal to $\dim C$.

(5) Prove that C is MDS if and only if, $\forall i \in \{1, \dots, n\},\ \mathbf{Loc}(C, i) = \dim(C)$.

Answer : One can use the lecture notes and use the fact that C is MDS if and only if $C^\perp$ is MDS, or we can prove it again. Suppose C is MDS and let $\mathbf{G}$ be a generator matrix of C. We claim that any k columns of C are independent. Indeed, if some k–tuple of columns was linked, then one could construct by Gaussian elimination a nonzero codeword vanishing at these k positions which would have weight $< n - k + 1$ which is a contradiction. Therefore any k columns of $\mathbf{G}$ are independent and hence the minimum distance of $C^\perp$ is larger than or equal to $k + 1$. We proved that the dual of an MDS code is MDS.

Next, suppose that C is MDS, then combining the results of questions 3 and 4, we get :

$$\dim C \geqslant \mathrm{Loc}(C, i) \geqslant d_{\min}(C^\perp) - 1$$

But if C (and hence $C^\perp$) is MDS, then the right hand side equals $n - \dim(C^\perp) = \dim C$. Conversely, suppose that $\mathbf{Loc}(C, i) \geqslant \dim C$ for any possible i. Then, the minimum distance of $C^\perp$ is larger than or equal to $\dim C + 1$. Thus, $C^\perp$ is MDS and hence so is C.

Given $I \subseteq \{1, \dots, n\}$ the *puncturing* and *shortening* of a code A at I are defined as

$$\mathcal{P}_I(A) := \{(a_i)_{i \in \{1,\dots,n\} \setminus I} \mid \mathbf{a} \in A\} \quad \text{and} \quad \mathcal{S}_I(A) := \{(a_i)_{i \in \{1,\dots,n\} \setminus I} \mid \mathbf{a} \in A \text{ and } \forall i \in I,\ a_i = 0\}.$$

We admit the following statement : *for any code* $A \subseteq \mathbb{F}_q$, $\mathcal{S}_I(A)^\perp = \mathcal{P}_I(A^\perp)$.

(6) Let C be a non degenerate code and $I \subseteq \{1, \ldots, n\}$. Prove that $\mathbf{Loc}(\mathcal{S}_I(C)) \leqslant \mathbf{Loc}(C)$.

Answer : Let $j \in \{1, \ldots, n\} \setminus I$. By definition

$$\begin{aligned}\mathbf{Loc}(\mathcal{S}_I(C), j) &= \min\{w_H(\mathbf{c}) \mid \mathbf{c} \in \mathcal{S}_I(C)^\perp,\ c_j \neq 0\} \\ &= \min\{w_H(\mathbf{c}) \mid \mathbf{c} \in \mathcal{P}_I(C^\perp),\ c_j \neq 0\} \\ &\leqslant \min\{w_H(\mathbf{c}) \mid \mathbf{c} \in \mathcal{C}^\perp,\ c_j \neq 0\} = \mathbf{Loc}(C, j).\end{aligned}$$

Thus, $\mathbf{Loc}(\mathcal{S}_I(C)) \leqslant \mathbf{Loc}(C)$.

(7) Let $\mathbf{c} \in C^\perp$ with $c_1 \neq 0$, $w_H(\mathbf{c}) = \mathbf{Loc}(C, 1) + 1$ and $I \subseteq \{1, \ldots, n\}$ be the *support of* $\mathbf{c}$, i.e.

$$I := \{i \mid c_i \neq 0\}.$$

Prove that $\mathcal{S}_I(C)$ is an $[n - \mathbf{Loc}(C, 1) - 1, k - \mathbf{Loc}(C, 1)]_q$–code.

Answer : The assertion on the length is obvious, we only have to prove that the dimension equals $k - \mathbf{Loc}(C, 1)$. Consider the projection map $C^\perp \to \mathcal{P}_I(C^\perp)$. Its kernel contains the words of $C^\perp$ whose support are in I. The subcode of such words has dimension 1 and spanned by $\mathbf{c}$, indeed, if this subcode had a larger dimension, then, by elimination one could construct other codewords in $C^\perp$ whose support contains 1 and which is strictly included in I. This would be a contradiction with the definition of the locality at 1. Therefore, the kernel of the projection, $C^\perp \to \mathcal{P}_I(C^\perp)$ has dimension 1, thus $\dim \mathcal{P}_I(C^\perp) = n - k - 1$ and hence the dimension of its dual

$$\begin{aligned}\dim \mathcal{S}_I(C) &= n - |I| - (n - k - 1) \\ &= k - |I| + 1 \\ &= k - \mathbf{Loc}(C, 1).\end{aligned}$$

(8) Let $t = \left\lceil \frac{k}{\ell} \right\rceil - 1$. **Until the end of the exercise, we suppose that** $n > (\ell + 1)t$. Prove that there exists a finite sequence of distinct indexes $i_1, \ldots, i_t \in \{1, \ldots, n\}$ and a sequence $\mathbf{c}_1, \ldots, \mathbf{c}_t \in C^\perp$ such that :

(i) for any $j \in \{2, \ldots, t\}$, i_j is not contained in the supports of $\mathbf{c}_1, \ldots, \mathbf{c}_{j-1}$;

(ii) for any $j \in \{1, \ldots, t\}$, $w_H(\mathbf{c}_j) = \mathbf{Loc}(C, j) + 1$.

Answer : Take $\mathbf{c}_1$ to be the vector $\mathbf{c}$ of the previous question. We iteratively choose i_j out of the union of the supports of $\mathbf{c}_1, \ldots, \mathbf{c}_{j-1}$ and $\mathbf{c}_j$ to be a codeword in $C^\perp$ whose support contains i_j and whose weight equals the locality of the code at i_j. By definition, these supports have cardinality at most $\ell + 1$, hence, one can repeat this process at least t times.

(9) Let $s \in \{1, \ldots, t\}$ (where t has been defined in Question 8). Let I_s be the union of the supports of $\mathbf{c}_1, \ldots, \mathbf{c}_s$ and $[n_s, k_s, d_s]$ be the parameters of $\mathcal{S}_{I_s}(C)$. Prove that $d_s \geqslant d$ and $n_s - k_s \leqslant n - k - s$.

Hint. Use Question 7 and proceed by induction on s.

Answer : The shortening is constructed from a subcode of C by removing zero positions. Hence, its minimum distance is at least that of C. Therefore $d_s \geqslant d$.
From question 7, we have $n_1 - k_1 \leqslant n - k - 1$. Applying this result iteratively we get

$$n_s - k_s \leqslant n - k - s.$$

(10) Let ℓ be the locality of C. Prove that the parameters $[n, k, d]$ of C satisfy

$$d \leqslant n - k - \left\lceil \frac{k}{\ell} \right\rceil + 2.$$

Hint. Consider the shortening of C at the union of the supports of the words $\mathbf{c}_1, \dots, \mathbf{c}_t$.

Answer : Applying Singleton bound to $\mathcal{S}_{I_t}(C)$. This code satisfies

$$d_s \leqslant n_s - k_s + 1$$

Using the previous questions, we deduce :

$$d \leqslant n - k - t + 1.$$

This yields the result.

Exercise 3. Let n be a positive integer, σ be a permutation on n elements and ϕ_σ be the linear map :

$$\phi_\sigma : \begin{cases} \mathbb{F}_q^n & \longrightarrow & \mathbb{F}_q^n \\ (x_1, \dots, x_n) & \longmapsto & (x_{\sigma(1)}, \dots, x_{\sigma(n)}) \end{cases}.$$

(1) Show that if $C \subseteq \mathbb{F}_q^n$ is a code, then C and $\phi_\sigma(C)$ have the same weight distribution.

Answer : The map σ preserves the weights, hence for any $a \in \{0, \dots, n\}$ it induces a bijection between the set of words of weight a of C and the set of words of weight a in $\sigma(C)$.

We aim at solving the following problem :

Problem : *Given two codes C, D, is there a permutation σ such that $D = \phi_\sigma(C)$?*

(2) Propose a naive brute force algorithm to solve the problem and compute its complexity.

Answer : Let $\mathbf{G}$ be a generator matrix of C and $\mathbf{H}$ a parity–check matrix of D. Enumerate any permutation $\sigma \in \mathfrak{S}_n$. For any such permutation σ, denote by $\mathbf{G}^\sigma$ the matrix $\mathbf{G}$ whose columns have been permuted using the permutation σ. Then, compute

$$\mathbf{H} \cdot \mathbf{G}^\sigma.$$

If the above matrix is zero, then $\phi_\sigma(C) = D$.
The complexity of one iteration is the complexity of a product of matrices, i.e. $O(n^3)$ and hence the overall complexity is in $O(n!n^3)$ (say $\widetilde{O}(n!)$).

(3) Prove that if two codes C, D satisfy $D = \phi_\sigma(C)$, then,
(i) $D^\perp = \phi_\sigma(C^\perp)$;

Answer : Let $\mathbf{d} \in D$ and $\mathbf{c} \in C^\perp$. Then,

$$\langle \phi_\sigma(\mathbf{c}), \mathbf{d} \rangle = \langle \mathbf{c}, \phi_{\sigma^{-1}}(\mathbf{d}) \rangle$$

Since $D = \phi_\sigma(C)$, then there exists $\mathbf{c}_0 \in C$ such that $\mathbf{d} = \phi_\sigma(\mathbf{c}_0)$. Thus,

$$\langle \phi_\sigma(\mathbf{c}), \mathbf{d} \rangle = \langle \mathbf{c}, \phi_{\sigma^{-1}} \circ \phi_\sigma(\mathbf{d}) \rangle = \langle \mathbf{c}, \mathbf{d} \rangle = 0.$$

Thus, $\phi_\sigma(C^\perp) \subseteq D^\perp$ and since these codes have the same dimensions, the inclusion is an equality.

(ii) $D \cap D^{\perp} = \phi_{\sigma}(C \cap C^{\perp})$.

Answer : It is a direct consequence of the previous question.

(4) Consider the following algorithm.
- **if** $C \cap C^{\perp}$ and $D \cap D^{\perp}$ do not have the same weight distribution, **return false**.
- **else return true**

(a) Does this algorithm always solve the problem ?

Answer : If the algorithm returns false, then the codes are not permutation–equivalent. If it returns true, the codes many not be equivalent, for instance, it may happen that $C \cap C^{\perp}$ and $D \cap D^{\perp}$ and, the codes may not be permutation–equivalent.

(b) Express the complexity of this algorithm in function of the dimension s of $C \cap C^{\perp}$. We suppose that the computation of the weight of a word costs $O(n)$ and that the best manner to compute the weight distribution is to enumerate all the codewords.

Answer : $O(nq^{dim C \cap C^{\perp}})$.

(c) Explain the advantages and possible drawbacks of comparing the weight distributions of $C \cap C^{\perp}$ and $D \cap D^{\perp}$ instead of comparing those of C, D ?

Answer : Unless the codes are contained in their dual, in general $C \cap C^{\perp}$ is strictly contained in C and hence the computation of its weight distribution will be much less expensive.

(5) Given a code C and $i \in \{1, \ldots, n\}$, we denote by C_i the code obtained by removing the i–th entry of any codeword of C. Namely :

$$C_i = \{(c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n) \mid (c_1, \ldots, c_n) \in C\} \subseteq \mathbb{F}_q^{n-1}$$

Using these codes C_i the algorithm can be refined as follows : if $C \cap C^{\perp}$ and $D \cap D^{\perp}$ have the same weight distributions, then compute the weight distributions of $C_i \cap C_i^{\perp}$ and $D_i \cap D_i^{\perp}$ for all $i \in \{1, \ldots, n\}$.

(a) If the weight distributions of the codes $C_i \cap C_i^{\perp}$ for $i \in \{1, \ldots, n\}$ are distinct, explain why is it possible to solve the problem.

Answer : Compute the weight distribution of $C_i \cap C_i^{\perp}$ and $D_i \cap D_i^{\perp}$ for any $i \in \{1, \ldots, n\}$. If for any i there exists $j_i \in \{1, \ldots, n\}$ such that $C_i \cap C_i^{\perp}$ and $D_{j_i} \cap D_{j_i}^{\perp}$ have the same weight distribution, then consider the permutation $\sigma : i \mapsto j_i$ and check whether $D = \phi_{\sigma}(C)$. If it does, you found the permutation. If not, or if there was no j_i for at least on i then the codes are not permutation equivalent.

(b) If not, what kind of information on σ (if exists) can we get ?

Answer : You can consider a partition $U_1 \cup \cdots \cup U_r$ of $\{1, \ldots, n\}$ such that the weight distribution of $C_i \cap C_i^{\perp}$ is the same for any $i \in U_j$. You can compute the same partition for D and compare the sequence of cardinalities of these partitions. If they differ, then the codes are non equivalent.

(c) Suppose that there exists a **cyclic** code E and permutations σ_1, σ_2 such that $C = \phi_{\sigma_1}(E)$ and $D = \phi_{\sigma_2}(E)$. Show that in this situation, the previous refinement will not be helpful.

Answer : If the codes are cyclic, then the weight distribution of $C_i \cap C_i^\perp$ will be the same for any i.

(d) In the case of a cyclic code as described in Question (5c), propose an improvement of the refinement which may solve the problem.

Answer : One can for instance consider the weight distributions of $C_{1i} \cap C_{1i}^\perp$ and $D_{1j} \cap D_{1j}^\perp$ for $i, j \in \{2, \ldots, n\}$.