Combinatória e Teoria de Códigos
Test - 16/06/2020

1. Instructions

You are to work on this examination by yourself. Any hint of collaborative work will be considered as evidence of academic dishonesty. You are not to have any outside contacts concerning this subject, except myself.
This being a take-home examination, you are expected to hand back a legible document, in terms of the presentation of your answers. Results proven in the course notes may be quoted indicating simply the file and result number (e.g. Notes I, Theorem 10); if you use theorems or other results from other sources, you must identify this source and state them in full, including the proof.
Justify all the steps and include the results of computations.
The questions are written in English but the answers may be written either in Portuguese or in English.
Important: The grade of each question is displayed after the question number. A complete answer consists of parts a) and b) of 1.,2.,3. and 4., together with one of 1.c) or 4.c), and one of 2.c) or 3.c). If you answer both questions in one of these pairs (or in both), the best grade of the two will be chosen.

About the solutions: The solutions presented here refer to one of the versions of the Test. Depending on the problem, some versions may need slightly different approaches.
Although many relevant comments were omited, the solutions include detailed explanations or examples that were not strictly demanded.

1. - Let $f(x), h(x) \in \mathbb{F}_q[x]$, where q is a prime and $f(x)$ is monic and has degree n.

a) (2.0) Show that $h^q(x) - h(x) = \prod_{a \in \mathbb{F}_q}(h(x) - a)$.

b) (2.0) Prove that if $f(x) \mid (h^q(x) - h(x))$ then

$$f(x) = \prod_{a \in \mathbb{F}_q} \gcd(f(x), h(x) - a).$$

c) (1.5) Consider the matrix $A = [a_{ij}]$, where $\sum_{j=0}^{n-1} a_{ij}x^j$ is the remainder from the division of x^{iq} by $f(x)$, for $0 \le i < n$.
Prove that a polynomial $h(x) = \sum_{i=0}^{n-1} h_i x^i$ satisfies

$$f(x) \mid (h^q(x) - h(x))$$

if and only if the vector $\mathbf{h} = (h_0, \cdots, h_{n-1})$ satisfies $\mathbf{h}A = \mathbf{h}$.

Solution: Fermat's Theorem implies $x^q - x = \prod_{a \in \mathbb{F}_q}(x-a)$: dividing the lefthand side by the righthand side gives a remainder $r(x)$ with degree less than q satisfying $r(a) = 0$ for all $a \in \mathbb{F}_q$, ie, $r(x)$ is the zero polynomial.
Denoting that polynomial by $l(x)$, we have $l \circ h(x) = h^q(x) - h(x)$ and also $l \circ h(x) = \prod_{a \in \mathbb{F}_q}(h(x) - a)$.
Each factor $g_a(x) = \gcd(f(x), h(x) - a)$ certainly divides $f(x)$; they are also mutually prime: if $v(x)$ divides both $g_a(x)$ and $g_b(x)$, with $a \neq b$, then

$$v(x) \mid (h(x) - a) \wedge v(x) \mid (h(x) - b) \implies v(x) \mid (b - a),$$

and so $v(x)$ is a nonzero constant, ie, $g_a(x)$ and $g_b(x)$ are co-prime. This implies, by a well-known application of the Euclidean algorithm, that the product $\prod_a g_a(x)$ divides $f(x)$.
Again by the Euclidean algorithm,

$$g_a(x) = s_a(x)f(x) + t_a(x)(h(x) - a),$$

and so

$$\prod_a g_a(x) = S(x)f(x) + T(x)\prod_a(h(x) - a),$$

for a polynomial $S(x)$ and $T(x) = \prod_a t_a(x)$; if $f(x) \mid (h^q(x) - h(x)) \Leftrightarrow f(x) \mid \prod_{a \in \mathbb{F}_q}(h(x) - a)$, then

$$f(x) \mid \prod_a g_a(x).$$

By definition

$$x^{iq} = f(x)v_i(x) + \sum_{j=0}^{n-1} a_{ij}x^j,$$

and so

$$h^q(x) = \left(\sum_{i=0}^{n-1} h_i x^i\right)^q = \sum_{i=0} n-1 h_i x^{iq} =$$

$$= \sum_{i=0}^{n-1} h_i \left(f(x)v_i(x) + \sum_{j=0}^{n-1} a_{ij}x^j\right) = f(x)v(x) + \sum_{j=0}^{n-1}\left(\sum_{i=0}^{n-1} h_i a_{ij}\right) x^j,$$

showing that the second summand is the remainder of the division of $h^q(x)$ by $f(x)$.
If $f(x) \mid (h^q(x) - h(x))$, as $\deg(h(x)) < n$, this remainder equals $h(x)$, ie,

$$h(x) = \sum_{j=0}^{n-1}\left(\sum_{i=0}^{n-1} h_i a_{ij}\right) x^j,$$

which is equivalent to $\mathbf{h}A = \mathbf{h}$.
And conversely, this equality implies that $h^q(x) = f(x)v(x) + h(x)$, ie, $f(x) \mid (h^q(x) - h(x))$.

2. - Let C be the cyclic binary code of length 15, with generator polynomial

$$g(x) = 1 + x^4 + x^6 + x^7 + x^8.$$

a) (2.0) Determine the defining set of the code, with respect to the primitive 15-root of unity α satisfying $\alpha^4 = \alpha + 1$, and prove that the minimal distance of C is 5.

b) (2.5) Decode the output

$$u = (1\,0\,0\,1\,0\,1\,1\,0\,0\,1\,1\,0\,0\,1\,0),$$

under the assumption that the possible error is a burst of length at most 3. Find the codeword at minimal distance from u.

c) (1.5) Confirm that C is 3-burst error correcting: justify that it is enough to show that the coset containing $b_0(x) = 1+x+x^2$ does not contain another nonzero burst $b(x)$ of length less or equal than 3; compare the values of $b_0(x)$ and $b(x)$ at appropriate powers of α.

Solution: The cyclotomic cosets modulo 15, with respect to 2, are

$C_0 = (0), C_1 = (1,2,4,8), C_3 = (3,6,9,12), C_5 = (5,10), C_7 = (7,11,13,14);$

if $g(x)$ is a generator polynomial for a cyclic code of length 15 is has to be the product of some of the irreducible factors of $x^{15}-1$; each of these irreducible factors is of the form $p_i(x) = \prod_{j \in C_i}(x-\alpha^j)$. Comparing degrees, we conclude that $g(x)$ is the product of two of the three irreducible polynomials of degree 4. A way to identify them is to divide $g(x)$ by the irreducible factor we know beforehand $p_1(x) = x^4+x+1$; we find that

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

and so $p_1(x)$ is one of the factors. We may identify the cyclotomic coset associated to the other factor without using a table of powers of α noting that

$$p_3(x) = (x+\alpha^3)(x+\alpha^6)(x+\alpha^9)(x+\alpha^{12}) = x^4+Sx^3+Sx^2+Sx+1,$$

where $S = (\alpha^3+\alpha^6+\alpha^9+\alpha^{12})$, which immediately shows that

$$p_3(x) = x^4+x^3+x^2+x+1:$$

the coefficient S is either 1 or 0 and it can not be 0, as x^4+1 is not irreducible.
So the defining set T is the union $C_1 \cup C_3$ and by the BCH bound the minimal distance of the code is at least 5; as the weight of $g(x)$ is 5 we conclude that $d=5$.
The received output corresponds to the polynomial

$$u(x) = 1+x^3+x^5+x^6+x^9+x^{10}+x^{13};$$

following the error-trapping algorithm, we compute the syndromes S_j of $x^j u(x)$ and the remainders $r_j(x)$ of the division of $x^{15-j}S_j(x)$, which are the candidates to error patterns; we find that

$$S_6 = 1+x+x^2,$$

corresponding to the error pattern $x^9+x^{10}+x^{11}$, which is the only burst of length less or equal to 3 that appears in the sequence of $r_j(x)$. This leads to the codeword

$$c(x) = u(x)-(x^9+x^{10}+x^{11}) = 1+x^3+x^5+x^6+x^{11}+x^{13}.$$

On the other hand, $S_0(x) = r_0(x) = x^3+x^7$ is a error pattern with weight 2 and so the unique codeword at minimal distance from $u(x)$ is

$$u(x)-(x^3+x^7) = 1+x^5+x^6+x^7+x^9+x^{10}+x^{13}.$$

A code is 3-burst error correcting if and only if each coset contains at most one burst of length less or equal to 3. Two bursts $b(x)$ and $b'(x)$ are in the same coset iff $b(x) - b'(x) \in C$. Given that the minimal distance of C is 5, we only need to consider the possible coexistence in the same coset of two bursts of weight 3 or one of weight 3 and the other of weight 2. Because C is cyclic, given a burst of weight 3, $x^j + x^{j+1} + x^{j+2}$ and another burst $b(x)$,

$$x^j + x^{j+1} + x^{j+2} - b(x) \in C \Leftrightarrow 1 + x + x^2 - x^{-j}b(x) \in C.$$

So we need only to consider the possibility that $b_0(x) = 1 + x + x^2$ and another burst $b(x)$ (with weight 2 or 3) are in the same coset.
Now we recall that $s(x) = b_0(x) - b(x) \in C$ iff $s(\alpha^i) = 0$ for all $i \in T = \{1, 2, 3, 4, 6, 8, 9, 12\}$. The case where $b(x)$ has also weight 3 is easier: then

$$b_0(x) - b(x) = 1 + x + x^2 + x^j + x^{j+1} + x^{j+2} = (1 + x^j)(1 + x + x^2),$$

where $2 < j < 13$ (otherwise the weight would be less than 5); $1+x+x^2$ has roots α^5 and α^{10} (the elements of the field with order 3); on the other hand, the roots of $1 + x^j$ are

- $1, \alpha^5, \alpha^{10}$ if $j = 3$;
- $1, \alpha^3, \alpha^6, \alpha^9, \alpha^{12}$ if $j = 5$;
- 1 otherwise.

In any case, α is never a root.
If $b(x)$ has weight 2, some more computations are needed. There are two cases:

- $b_0(x) - b(x) = 1 + x + x^2 + x^j + x^{j+1}$, with $2 < j < 14$,
- $b_0(x) - b(x) = 1 + x + x^2 + x^j + x^{j+2}$, with $2 < j < 13$,

and in each one of these cases we must show that the equalities

$$b_0(\alpha) - b(\alpha) = 0 \qquad b_0(\alpha^3) - b(\alpha^3) = 0$$

can not occur simultaneously.
For these computations we need to know the expressions

$$\alpha^t = a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3, \qquad a_i \in \mathbb{F}_2,$$

for $0 \leq t < 15$. We have then

$$1+\alpha+\alpha^2+\alpha^j+\alpha^{j+1} = 0 \Leftrightarrow \alpha^{10}+\alpha^j(1+\alpha) = 0 \Leftrightarrow \alpha^{10}+\alpha^{j+4} = 0 \Leftrightarrow j = 6;$$

on the other hand,

$$1 + \alpha^3 + \alpha^6 + \alpha^{3j}(1 + \alpha^3) = 0 \Leftrightarrow \alpha^8 + \alpha^{3j+14} = 0 \Leftrightarrow j = 3.$$

The second case is similar.

3. - Let A be the Reed-Solomon code over $\mathbb{F}_{3^2}$ defined by

$$C = \{(f(1), f(\lambda), \cdots, f(\lambda^7)) : f(x) \in P_4 \subset \mathbb{F}_{3^2}[x]\}$$

where λ is a primitive 8-root of unity satisfying $\lambda^2 = \lambda + 1$, and P_4 denotes as usual the vector space of polynomials with degree less than 4.

a) (2.0) Determine the dimension and minimal distance of the trace code $Tr(A) \subset \mathbb{F}_3^8$.

b) (2.5) Let B be the $[4, 2, 3]$ code over F_3 with generator matrix

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{bmatrix}.$$

Decode the following output from the concatenated code $A[B]$, using syndrome decoding for B and Peterson's algorithm for A:

$$(1211 \quad 1102 \quad 1010 \quad 0100 \quad 2110 \quad 0000 \quad 0111 \quad 2022).$$

c) (1.5) Prove that the concatenated decoding algorithm used in b) corrects all errors with weight $w < 6$, but that there are errors with weight 6 that are not corrected. Prove also that the algorithm succefully decodes any burst with length $l < 8$.

Solution: To facilitate following the solution, we start by computing a table for the powers of λ:

j	0	1	2	3	4	5	6	7
λ^j	1	λ	$1+\lambda$	$1+2\lambda$	2	2λ	$2+2\lambda$	$2+\lambda$

A has generator matrix

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \lambda & \lambda^2 & \lambda^3 & \lambda^4 & \lambda^5 & \lambda^6 & \lambda^7 \\ 1 & \lambda^2 & \lambda^4 & \lambda^6 & 1 & \lambda^2 & \lambda^4 & \lambda^6 \\ 1 & \lambda^3 & \lambda^6 & \lambda & \lambda^4 & \lambda^7 & \lambda^2 & \lambda^5 \end{bmatrix}$$

The values of the trace function $Tr(a) = a + a^3$ are $Tr(0) = 0$ and

j	0	1	2	3	4	5	6	7
$Tr(\lambda^j)$	2	1	0	1	1	2	0	2

If we compute componentwise the trace of each row of G and of each row multiplied by λ, we obtain

$$\begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 & 1 & 2 & 0 & 2 \\ 1 & 0 & 1 & 1 & 2 & 0 & 2 & 2 \\ 2 & 0 & 1 & 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 1 & 0 & 1 & 1 & 2 & 0 & 2 \\ 1 & 1 & 2 & 0 & 2 & 2 & 1 & 0 \end{bmatrix}$$

Reducing to standard form (which includes eliminating row linear dependences) we obtain

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 2 \end{bmatrix}$$

showing that the trace code $Tr(A)$ has dimension 5; the minimal distance $d = 3$ may be read directly from this generator matrix or from the parity-check matrix.

The syndromes of the blocks of the output, with respect to the parity-check matrix for B

$$]H = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{bmatrix}$$

are

$$21 \quad 00 \quad 02 \quad 12 \quad 00 \quad 00 \quad 20 \quad 00$$

with corresponding coset leaders

$$0200 \quad 0000 \quad 0002 \quad 0100 \quad 0000 \quad 0000 \quad 0020 \quad 0000$$

we obtain the sequence of codewords for B

$$1011 \quad 1102 \quad 1011 \quad 0000 \quad 2110 \quad 0000 \quad 0121 \quad 2022$$

which are in turn the encoding of the messages

$$10 \quad 11 \quad 10 \quad 00 \quad 21 \quad 00 \quad 01 \quad 20$$

The corresponding output of the outercode is

$$r = (1 \quad \lambda^2 \quad 1 \quad 0 \quad \lambda^7 \quad 0 \quad \lambda \quad 2)$$

The syndrome of r is $(\lambda^3 \quad \lambda^7 \quad 2 \quad \lambda^2)$.
According to Peterson's algorithm we construct the matrix

$$D(r) = \begin{bmatrix} \lambda^3 & \lambda^7 & 2 \\ \lambda^7 & 2 & \lambda^2 \end{bmatrix}$$

and compute a vector in the kernel whose entries are the coefficients of the locator polynomial. It is easy to see that the locator polynomial will have degree 2, as the first two columns of $D(r)$ are linearly independent. Solving the system of equations we obtain

$$Q_1(x) = x^2 + \lambda^3 x + \lambda^2 = (x-1)(x-\lambda^2).$$

The last factorization may be obtained by trial and error, by noting that $1 + \lambda^2 = -\lambda^3$, or even by the formula for the roots of a quadratic polynomial (which works for fields of charcteristic different from 2).
We see that the errors occur in the first and third coordinates. We could compute the polynomial $Q_0(x)$ or, in alternative, to search for an error pattern of the form

$$(\lambda^t \quad 0 \quad \lambda^s \quad 0 \quad 0 \quad 0 \quad 0 \quad 0)$$

with the given syndrome. This amounts to solve the system

$$\begin{cases} \lambda^t + \lambda^{2+s} = \lambda^3 \\ \lambda^t + \lambda^{4+s} = \lambda^7 \\ \lambda^t + \lambda^{6+s} = 2 \\ \lambda^t + \lambda^s = \lambda^2 \end{cases}$$

and we obtain $t = 1$ and $s = 0$, ie, the codeword is

$$c = r - (\lambda \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0) = (\lambda^3 \quad \lambda^2 \quad 0 \quad 0 \quad \lambda^7 \quad 0 \quad \lambda \quad 2$$

The entries of c are the values $f(\lambda^j)$, with $0 \leq j < 8$, where $f(x)$ is a polynomial with degree less than 4. In this case, it is easy to determine $f(x)$ because, as

$$f(\lambda^2) = f(\lambda^3) = f(\lambda^5) = 0,$$

and $\deg(f(x)) \leq 3$, we must have

$$f(x) = a(x-\lambda^2)(x-\lambda^3)(x-\lambda^5)$$

for some constant $a \in \mathbb{F}_9$. Computing for example $f(1)$, we find that $a = \lambda^3$, and so

$$f(x) = \lambda^3 x^3 + \lambda^5 x^2 + x + \lambda.$$

The corresponding input depends on the generator matrix, ie, on the choice of basis of P_4. If we choose the canonical basis, which corresponds to the generator matrix given above, the input is

$$(\lambda \quad 1 \quad \lambda^5 \quad \lambda^3)$$

represented by the vector from $\mathbb{F}_3^8$

$$u = (0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 2 \quad 1 \quad 2).$$

To discuss the error-correcting capabilities of the algorithm, we notice that the decoding algorithm for the outer code corrects any 2 errors, while syndrome decoding for the inner code decodes any single error. This means that blocks $v_i \in \mathbb{F}_3^4$ containing 1 error will be correctly decoded in the inner step of the decoding. If less than 3 blocks are not decoded correcly in the inner step, they will still be decoded in the outer step.
Suppose that $r \in \mathbb{F}_3^{32} = (\mathbb{F}_3^4)^8$ contains less than 6 errors; then at most 2 of the blocks may contain an error with weight > 1. So the inner decoding takes care of possible single errors and the outer step will correctly decode the two errors that may remain.
If r contains 6 errors, the algorithm may fail to decode correctly: suppose that 3 of the blocks contain 2 errors each, and that syndrome decoding fails to correct them in the inner step. Then, it may happen that the outer decoding fails to correct the 3 errors.
It is easy to construct examples where the locator polynomial can not be constructed or has no zeros, despite the fact that the syndrome is not zero: suppose for example that the zero codeword is sent and the received vector is

$$(1100 \quad 1100 \quad 1100 \quad 0000 \quad 0000 \quad 0000 \quad 0000 \quad 0000).$$

Another example is the following: suppose that the zero codeword is sent and

$$(2100 \quad 1001 \quad 0210 \quad 0000 \quad 0000 \quad 0000 \quad 0000 \quad 0000)$$

is received. The inner step of the decoding algorithm we obtain the syndromes

$$(20 \quad 20 \quad 01 \quad 00 \quad 00 \quad 00 \quad 00 \quad 00)$$

and consequently the codeword

$$(2110 \quad 1011 \quad 0212 \quad 0000 \quad 0000 \quad 0000 \quad 0000 \quad 0000)$$

which is the encoding of

$$(21 \quad 10 \quad 02 \quad 00 \quad 00 \quad 00 \quad 00 \quad 00).$$

So in the outer step of the algorithm we apply Peterson's algorithm to the decoding of

$$(\lambda^7 \quad 1 \quad \lambda^5 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0).$$

The syndrome is $(1 \quad 0 \quad \lambda^3 \quad 1)$, and the locator polynomial is

$$Q_1(x) = x^2 + \lambda x + \lambda^7 = (x - \lambda^3)(x - \lambda^4).$$

We may decode as above, finding an error pattern

$$(0 \quad 0 \quad 0 \quad \lambda^t \quad \lambda^s \quad 0 \quad 0 \quad 0)$$

with the given syndrome, or we may determine the polynomial $Q_0(x)$ which is

$$Q_0(x) = 2x^5 + \lambda^2 x^4 + \lambda x^3 + \lambda^2 x^2 + \lambda^5 x + \lambda.$$

The input polynomial obtained is

$$u(x) = -\frac{Q_0(x)}{Q_1(x)} = x^3 + \lambda^7 x^2 + \lambda x + \lambda^6$$

which corrsponds to the ternary input

$$(2 \quad 2 \quad 0 \quad 1 \quad 2 \quad 1 \quad 1 \quad 0)$$

and to the codeword

$$2110 \quad 1011 \quad 0212 \quad 2110 \quad 0212 \quad 0000 \quad 0000 \quad 0000).$$

So in this example the true errors are not detected and the algorithm produces a wrong decoding.

If an output is $r = c + b$ where b is a burst of length $l < 8$, then b affects at most 3 of the blocks, and the error in at least one of these blocks has weight 1 and so is corrected in the inner step of the algorithm, while the remaining two blocks correspond to an error of weight 2 for the outer step of the algorithm and are therefore corrected.

4. -

a) (2.0) Compute the free distance of the binary convolutional code C with generator matrix

$$[1 + D \quad 1 + D^2 \quad 1 + D + D^2].$$

b) (2.0) Decode by Viterbi's algorithm the following output from C, presented in interleaved form,

$$(\ 111 \quad 101 \quad 101 \quad 010 \quad 010 \quad 100 \quad 110 \quad 010 \quad 100 \quad 101 \quad 011\ ).$$

c) (1.5) Determine the internal and external degrees of the matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & 1 + D \\ 0 & 1 & 0 & 1 + D^2 \\ 0 & 0 & 1 & 1 + D + D^2 \end{bmatrix},$$

and find a canonical generator matrix for the code generated by G.

Solution: The free distance of C is 7. This may be confirmed examining the state diagram; the input 1 gives rise to an output

$$(\ 111 \quad 101 \quad 011\ ),$$

which is the unique codeword with minimal weight.
I'm not presenting here the complete solution of ***4.b)****. If needed, I'll include later the trellis diagram of the decoding.*
The codeword is

$$(\ 111 \quad 101 \quad 100 \quad 110 \quad 010 \quad 100 \quad 110 \quad 010 \quad 100 \quad 101 \quad 011\ ),$$

corresponding to the input

$$(1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1).$$

We have

$$\operatorname{intdeg}(G) = 2, \qquad \operatorname{extdeg}(G) = 5;$$

G is basic (the gcd of the 3×3 minors is 1) but not reduced.
We may obtain generator matrices with lower exernal degree, applying row elementary operations: replacing the last row by the sum of the 3 rows we obtain

$$G_1 = \begin{bmatrix} 1 & 0 & 0 & 1+D \\ 0 & 1 & 0 & 1+D^2 \\ 1 & 1 & 1 & 1 \end{bmatrix};$$

G_1 is also basic but not reduced. But adding the first row multiplied by $1+D$ to the second row, we get

$$G_2 = \begin{bmatrix} 1 & 0 & 0 & 1+D \\ 1+D & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix};$$

G_2 is basic, by the same reason as above, and

$$\operatorname{intdeg}(G_2) = \operatorname{extdeg}(G_2) = 2,$$

implying that G_2 is reduced. So G_2 is canonical.