MATH 4023-1 **Final Exam Review** **Spring 2014**

The final exam will take place on Friday, May 9, 7:30–9:30 am, in our usual room. In addition to the material covered on the two in–class exams, the final exam will also cover material we've discussed from Chapters IV and V in Lax (and related material discussed in class). The material is cumulative, and to some extent depends on previous material, which you should still know. A (likely non-comprehensive) list of concepts etc. regarding the latter material you should know and be ready to work with is below (references refer to pages, sections, results, and so on in Lax). Refer to the reviews for the in-class exams for analogous lists pertaining to material discussed earlier in the course. If you have questions regarding any of the material, make use of my office hours, and/or ask by email. I anticipate being in my office on Tuesday and Thursday of finals week.

You may use a calculator on the final, but be sure to show all your work to receive full (or partial) credit.

Problems on the final will likely be of a similar nature to those you've encountered on the in-class exams and in the homework (both collected and uncollected) - some proofs, some calculations based on results we've obtained and methods derived from them. Some review problems for the material discussed after Exam 2 are included on the next page. This is **not** a comprehensive list. Additional problems may be found in the Exercises of the sections and supplements we've covered, and the reviews for the in-class exams.

Some concepts, results, definitions. . . you should know:

Know the basic definitions and properties concerning *rings*, *fields* etc. discussed in §IV.1. Numerous examples ($(\mathbb{Z}, +, \cdot)$, $(\mathbb{Z}_n, +, \cdot)$ . . .), notions (*unit*, *zero divisor* . . .) should be familiar.

$\mathbb{Z}_n$ is a field if and only if n is prime (Corollary IV.2.3). If $n = p$ is prime, the field $\mathbb{Z}_p$ is often called $GF(p)$, the Galois field with p elements. Many properties of (the rings) $\mathbb{Z}$ and $\mathbb{Z}_n$ arose earlier in the course (and are in §IV.2).

Know the definition of a *binary linear code*: a subspace of the vector space $GF(2)^n$. (Definition V.2.1, page 215)
(More generally, a *q-ary linear code* is a subspace of $GF(q)^n$. Our focus is on binary codes.)
One way of denoting elements of $GF(2)^2$ is by strings $\hat{x} = x_1x_2 \cdots x_n$, where each x_i is either 0 or 1.

C is an (n, k) binary code if the codewords in C are of length n, that is, $C \subset GF(2)^n$, and the dimension of C is k. $(C, +)$ is a subgroup of the abelian group $(GF(2)^n, +)$ (page 215)

A *generator matrix* for the (n, k) code C is a $k \times n$ matrix G whose rows form a basis for C. G is in standard form if $G = [I_k \colon A]$, where I_k is the $k \times k$ identity matrix and A is a $k \times n - k$ matrix with entries in $GF(2)^n$.
For an (n, k) binary code C with generator matrix G, the codewords in C (i.e., elements of C) are linear combinations of the rows of G, $C = \{\hat{x}G \colon \hat{x} \in GF(2)^k\}$. There are 2^n elements in $GF(2)^n$, and there are 2^k elements in C. (pages 216–218)

Know the definition of a *parity check matrix* for a binary linear code. (Definition V.2.11, page 220)

Know the relationship between codewords in C and a parity check matrix H for C: $\hat{c} \in C \iff H\hat{c} = \hat{0}$. (Proposition V.2.14, page 221 and Proposition V.3.1, page 226)

If $G = [I_k \colon A]$ is a standard form generator matrix for C, then $H = [-A^\top \colon I_{n-k}]$ is a parity check matrix for C.
If $H = [B \colon I_{n-k}]$ is a parity check matrix for C, then $G = [I_k \colon -B^\top]$ is a generator matrix for C. (pages 222-223)

Know the definitions of the *Hamming distance* between codewords, the *weight* of a codeword, and the *minimum distance* $d(C)$ of a code C. If $d(C) = d$, C is said to be an (n, k, d) code. (pages 226–227)

Know the relationship between minimum distance and weight: $d(C) = \min\{w(\hat{c}) \colon \hat{c} \in C, \hat{c} \neq \hat{0}\}$.

Know the significance of the minimum distance of C for error detection and correction using *nearest neighbor decoding*: If $d(C) = d$, then (1) C can detect t errors if $d \geq t + 1$; (2) C can correct t errors if $d \geq 2t + 1$. In particular, if $d(C) = 3$, then C is a 1-error-correcting code. (Theorem V.3.8, page 228, pages 227–228 for discussion of nearest neighbor decoding and underlying assumptions)

Know how to determine the minimum distance of a code from a parity check matrix for the code. (Theorem V.3.10, page 228)

Be able to construct a *standard array* for a code, and use it for decoding. (Algorithm V.4.1, pages 236–237)

Know the definition of a *syndrome*, be able to construct a *syndrome table*, and use it for decoding. (pages 239–240)

Know the definition of the *Hamming binary code* $\mathrm{Ham}(r, 2)$ and its parity check matrix H_r. Syndrome decoding in this context can be done quickly. (pages 244–246)

Some review problems:

1. Write out the addition and multiplication tables for the rings $\mathbb{Z}_7$ and $\mathbb{Z}_8$. Record properties of these rings (ring with identity? commutative ring? etc.).

2. An element r in a ring R is *nilpotent* if $r^n = 0$ for some natural number n. Show that a nilpotent element is a zero divisor. Are there any nilpotent elements in $\mathbb{Z}_7$? In $\mathbb{Z}_8$?

3. Recall that an integral domain is a commutative ring with identity 1 ($1 \neq 0$) that has no zero divisors except 0. For which values of n is $\mathbb{Z}_n$ an integral domain?

4. Find the minimum distance between any two of the codewords in each of the following lists.

   (a) $\{0000, 1010, 0101, 1111\}$

   (b) $\{11110, 11011, 01111, 10111, 11101\}$

   (c) $\{100001, 010100, 001010, 011110, 101011, 110101\}$

5. Let $G = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$ be a generator matrix for a binary linear code C

   (a) Find all the codewords in C.

   (b) Find a parity check matrix for the C.

   (c) Determine the minimum distance of C. How many errors can C detect? How many can it correct?

   (d) Construct a syndrome table for C. Use it to decode each of the following messages:
   110010, 101010, 111000, 001110, 111111.

6. Consider the Hamming binary code Ham$(4, 2)$.

   (a) Write the parity check matrix H_4 for this code.

   (b) If G is a generator matrix for this code, what is the size of G? How many codewords are in this code?

   (c) Use the parity check matrix H_4 to decode the messages 111000000000111 and 000111000111000.

7. Hamming's square code encodes a 4 digit binary message by writing the 4 bits in a square, computing the sums of rows and columns modulo 2, and forming a 9 digit word from the 3 rows. For instance, the message 1101 is encoded as 110011101.

$$1101 \longrightarrow \begin{array}{|cc} 1 & 1 \\ 0 & 1 \\ \hline \end{array} \longrightarrow \begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 1 \\ \hline 1 & 0 & 1 \end{array} \longrightarrow 110011101$$

   A general message $x_1x_2x_3x_4$ is encoded as $c_1c_2c_3c_4c_5c_6c_7c_8c_9$

$$x_1x_2x_3x_4 \longrightarrow \begin{array}{|cc} x_1 & x_2 \\ x_3 & x_4 \\ \hline \end{array} \longrightarrow \begin{array}{cc|c} c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \\ \hline c_7 & c_8 & c_9 \end{array} \longrightarrow c_1c_2c_3c_4c_5c_6c_7c_8c_9, \quad \text{where}$$

   $c_1 = x_1$, $c_2 = x_2$, $c_3 - x_1 + x_2$, $c_4 = x_3$, $c_5 = x_4$, $c_6 = x_3 + x_4$, $c_7 = x_1 + x_3$, $c_8 = x_2 + x_4$, $c_9 = x_1 + x_2 + x_3 + x_4$.

   (a) Determine how many codewords are in this code.

   (b) If $\hat{c}$ is a codeword in this code, express $\hat{c}$ as a linear combination of vectors (with coefficients given by the digits of the message $\hat{x}$), and use this to find a generator matrix for this code.

   (c) Detect and correct the errors in the messages 111110011, 011111110, 100110011

Answers to the review problems:

1. For $\mathbb{Z}_7$:

$+$	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

$\cdot$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

For $\mathbb{Z}_8$:

$+$	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7	0
2	2	3	4	5	6	7	0	1
3	3	4	5	6	7	0	1	2
4	4	5	6	7	0	1	2	3
5	5	6	7	0	1	2	3	4
6	6	7	0	1	2	3	4	5
7	7	0	1	2	3	4	5	6

$\cdot$	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	0	2	4	6
3	0	3	6	1	4	7	2	5
4	0	4	0	4	0	4	0	4
5	0	5	2	7	4	1	6	3
6	0	6	4	2	0	6	4	2
7	0	7	6	5	4	3	2	1

Both are commutative rings with identity. $\mathbb{Z}_7$ is a field (and is an integral domain), $\mathbb{Z}_8$ is not.

2. If $r^n = 0$, then $r \cdot r^{n-1} = 0$ and $r^{n-1} \cdot r = 0$. Since $r^{n-1} \neq 0$, r is a zero divisor.
Every nonzero element in $\mathbb{Z}_7$ is a unit, so there are no zero divisors in $\mathbb{Z}_7$, and no nilpotent elements in $\mathbb{Z}_7$.
Since $0^1 \equiv 0$, $2^3 \equiv 0$, $4^2 \equiv 0$, $6^3 \equiv 0$, the congruence classes of $0, 2, 4, 6$ are nilpotent in $\mathbb{Z}_8$ (the other elements in $\mathbb{Z}_8$ are units, so cannot be nilpotent).

3. If n is prime, then $(a, n) = 1$ for every integer a with $1 \leq a \leq n-1$. Consequently, the congruence class of a modulo n is not a zero divisor in $\mathbb{Z}_n$ (it is in fact a unit in the field $\mathbb{Z}_n$).
If $n = rs$ is not prime, then, for instance, the congruence classes of r and s are zero divisors in $\mathbb{Z}_n$.
Thus, $\mathbb{Z}_n$ is an integral domain if and only if n is prime (in which case, $\mathbb{Z}_n$ is a field).

4. The minimum distance is 2 in each of parts (a), (b), and (c).

5. (a) There are 2^3 codewords in C. If $\hat{r}_1, \hat{r}_2, \hat{r}_3$ are the rows of the generator matrix G, these codewords are $\hat{0} = 000000$, $\hat{r}_1 = 100101$, $\hat{r}_2 = 010011$, $\hat{r}_3 = 001110$, $\hat{r}_1 + \hat{r}_2 = 110110$, $\hat{r}_1 + \hat{r}_3 = 101011$, $\hat{r}_2 + \hat{r}_3 = 011101$, $\hat{r}_1 + \hat{r}_2 + \hat{r}_3 = 111000$.

   (b) $H = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$

   (c) $d(C) = 3$. C can detect 2 errors, and correct 1 error.

   (d) Below, syndromes are expressed as strings, as opposed to column vectors, e.g., 101 corresponds to $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

coset leader	syndrome
000000	000
100000	101
010000	011
001000	110
000100	100
000010	010
000001	001
100010	111

The syndrome of 110010 is 100. So we decode 110010 as 110010+000100=110110
The syndrome of 101010 is 001. So we decode 101010 as 101010+000001=101011
The syndrome of 111000 is 000. So we decode 111000 as 111000+000000=111000
The syndrome of 001110 is 000. So we decode 001110 as 001110+000000=001110
The syndrome of 111111 is 111. So we decode 111111 as 111111+100010=011101

6. (a) $H_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$

The columns of H_4 correspond to the binary expansions of $1, 2, \ldots 15$ (in this order).

(b) Since H_4 is 4×15, and is $n - k \times n$, we have $n = 15$ and $n - k = 4$, so $k = 11$. A generator matrix G is $k \times n$, so is 11×15. Since the rows of G form a basis for $\mathrm{Ham}(4, 2)$, this code has 2^{11} codewords.

(c) The syndrome of 111000000000111 is 1100, which is column 12 of H_4. So we decode 111000000000111 as 111000000000111+000000000001000=111000000001111

The syndrome of 000111000111000 is 1010, which is column 10 of H_4. So we decode 000111000111000 as 000111000111000 + 000000000100000=000111000011000

7. (a) There are 2^4 codewords.

(b) If $\hat{c} = (c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9)$, with $c_1 = x_1$, $c_2 = x_2$, $c_3 - x_1 + x_2$, $c_4 = x_3$, $c_5 = x_4$, $c_6 = x_3 + x_4$, $c_7 = x_1 + x_3$, $c_8 = x_2 + x_4$, $c_9 = x_1 + x_2 + x_3 + x_4$, then

$$\begin{aligned} \hat{c} &= (x_1, x_2, x_1 + x_2, x_3, x_4, x_3 + x_4, x_1 + x_3, x_2 + x_4, x_1 + x_2 + x_3 + x_4) \\ &= x_1(1, 0, 1, 0, 0, 0, 1, 0, 1) + x_2(0, 1, 1, 0, 0, 0, 0, 1, 1) + x_3(0, 0, 0, 1, 0, 1, 1, 0, 1) + x_4(0, 0, 0, 0, 1, 1, 0, 1, 1) \end{aligned}$$

So a generator matrix for C is $G = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}$

(c) Using the generator matrix G found in part (b), one can find a parity check matrix H (by finding a basis for the nullspace of G). Then one can use syndrome decoding as in the previous problems.

Alternatively, one can work directly with the code as follows:

Express the message 111110011 as an array $\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 1 & 0 \\ \hline 0 & 1 & 1 \end{array}$. From the construction of the code, there are errors in (row 1, column 3) and (row 3, column 2). That is, the first row sum and second column sum are wrong. Consequently, the entry in (row 1, column 2) should be changed from a 1 to a 0. This gives the corrected array $\begin{array}{cc|c} 1 & 0 & 1 \\ 1 & 1 & 0 \\ \hline 0 & 1 & 1 \end{array}$ (corresponding to the codeword 101110011), which is the result of the string $x_1x_2x_3x_4 = 1011$.

Similarly, 011111110 $\longrightarrow$ $\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 1 & 1 \\ \hline 1 & 1 & 0 \end{array}$ $\xrightarrow{\text{correct}}$ $\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & 1 \\ \hline 1 & 1 & 0 \end{array}$, which is the result of the string 1110.

Similarly, 100110011 $\longrightarrow$ $\begin{array}{cc|c} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \hline 0 & 1 & 1 \end{array}$ $\xrightarrow{\text{correct}}$ $\begin{array}{cc|c} 1 & 0 & 1 \\ 1 & 1 & 0 \\ \hline 0 & 1 & 1 \end{array}$, which is the result of the string 1011.