1. In this problem, you are going to design your own (6,3) single error-correcting code (SEC) and do a lot of manipulations with such.

(a) Design a (6,3) single error-correcting linear block code (SEC) by giving $G$, $H$, and the set of codewords $C$ for such. (There are many potential solutions. Do not use the one from class or one that matches your friends’ solutions!)

(b) Find the minimum distance $d_{\text{min}}$ of your code.

(c) Using your number for $d_{\text{min}}$ from (b), identify $d_{\text{min}}$ columns of $H$ that sum to zero.

(d) Find the (full) standard array and identify the coset leaders. (Yes, I know this is going to take a little bit of time.)

(e) Suppose the information bits are (011) and the error pattern is (010000). Do the following:
   - Find the transmitted codeword.
   - Find the received vector.
   - Find the syndrome.
   - List all of the possible error patterns that could have occurred given your syndrome, and, for each, identify the codeword you would output if that were the error pattern. (This should be a table with 8 rows; in each row, you should have an error pattern and the corresponding codeword you would output.)
   - Give the output codeword if you choose the error pattern corresponding to the coset leader.
   - Was the correct codeword identified?

(f) Repeat part (e), except for the fourth bullet, for each of the following sets of information bits and error patterns:
   - (010) and (010000).
   - (110) and (000111).
   - (101) and (010100).
   - (111) and an error pattern corresponding to one of your other non-zero codewords (not the one that is sent; other than that, your choice!)

(g) Suppose your code is used on a binary symmetric channel with crossover probability $p$. What is the probability of error of your code?

(h) Suppose I shorten this code by removing one of the columns of $H$:
   - Give the new $n$ and $k$. How does the rate compare to your (6,3) code?
   - Give the new generator matrix and set of codewords $C$.
   - Find the minimum distance $d_{\text{min}}$ of the new code.
2. (a) Argue that for any binary strings $x, y, z$ of length $n$ that

$$d_H(x, y) \leq d_H(x, z) + d_H(z, y).$$

(b) Use your result from (a) to show that a code correct any set of $t$ errors as long as:

$$t \leq \left\lfloor \frac{d_{\text{min}}(C) - 1}{2} \right\rfloor$$

3. Consider a (5,2) linear block code for the binary symmetric channel with crossover probability $p$. Let the code be given by:

$$C = \{00000, 11010, 01111, 10101\}.$$  

(a) Find the generator matrix $G$ and the parity check matrix $H$ for this code.

(b) Give the standard array, denoting the coset leaders.

(c) Use your solution to (b) to decode the following received vectors: 01100, 11111, 01101, 00010.

(d) What are the correctible error patterns? Using these, find the probability of error of the code as a function of $p$.

4. Suppose we add another circle to the “circle diagram” used in class to introduce the (7,4) Hamming code. Draw the new circle around the entire diagram and place the position 8 in the new circle (but outside the three original circles). Suppose that we add a fourth parity check that checks that the number of ones inside the new circle is even (i.e. it checks for even parity across all 8 positions).

(a) Find the parity check matrix $H$.

(b) Find the minimum distance of the code. By finding the indication in the circle diagram of two errors, argue that this code can be used simultaneously to correct a single error and detect any pattern of two errors.

5. A customer has a (binary) channel that accepts strings of length $n = 7$, and the only error patterns that ever occur are:

$$0000000, 1000000, 1100000, 1110000,$$

$$1111000, 1111100, 1111110, 1111111.$$  

Design a linear block code with as high a rate as possible that will correct all such error patterns.