University of California at San Diego – Department of Physics – Prof. John McGreevy

# Physics 239/139 Fall 2019
# Assignment 4

**Due 11:00am Tuesday, February 7, 2023**

1. **Error correcting code brain-warmer.**

   (a) For the [7,4] Hamming code discussed in lecture, check that $Ht = 0$ where $t$ is any codeword, and $H$ is the given parity check matrix.

   (b) If you ($B$) are communicating with someone ($A$) through a channel with $H(A|B) = 1/7$ using this code and you receive the string

   $$r = (0, 1, 0, 0, 0, 0, 1)^t,$$

   what is the most likely intended message string?

2. **Huffman code.** Make the Huffman code for the probability distribution $p(x) = \{.5, .2, .16, .1, .04\}$.

   Compare the average word length to the Shannon entropy.

   Bonus: what property of the distribution determines the deviation from optimality?

3. **Huffman code decryption problem.** [Optional, but fun.]

```
010110100001101010010010000111101101101000011100100100000111000,
010101010010110 110000000001000011 101010100101100011000111 1111100011
1111001011000010001011010010111.  001010101 010101010010110
1011010000110001 1 0000 000110010101011010101110010110011111
111011000100100111 1010101001001 010101010010110 1001
110101111111110111001
101011001011000100101110110011001101011001001110 00, 1000010010111011
010010101 1111100011 1101011111111011100 11000 10100100101011011101
0101101001011101 1 010010110111 101001100101000011010100.
```

   You might want to use `Mathematica` to do this problem!

   Hint: I used the letter frequencies from *The Origin of Species*, which is built into `Mathematica`.

4. **Analogy with strong-disorder RG.** [open ended, more optional question]

   Test or decide the following consequence suggested by the analogy between Huffman coding and strong-disorder RG: The optimality of the Huffman code is better when the distribution is broader. A special case is the claim that the Huffman code is worst when all the probabilities are the same. Note that the outcome of the Huffman algorithm in this case depends on the number of elements of the alphabet.

   Measure the optimality by $\langle \ell \rangle - H[p]$ (or maybe $\frac{\langle \ell \rangle - H[p]}{H[p]}$?).

5. **Maximum entropy with fixed average.** Suppose we have a probability distribution $\pi_\ell$ on a positive integer $\ell$ with fixed average:

$$\langle \ell \rangle \equiv \sum_{\ell=1}^{\infty} \pi_\ell \ell = a. \tag{1}$$

   Bound the Shannon entropy of $\pi$ from above:

$$H[\pi] \leq a \log a - (a-1)\log(a-1) \ . \tag{2}$$

   Show that this goes like $\log a$ at large $a$.

   [Hint: use Lagrange multipliers to maximize the entropy subject to the normalization condition and the condition on the average.]

6. **Binary symmetric channel.**

   For the binary symmetric channel $ABE$ defined in lecture, with $a, b, e \in \{0, 1\}$, and

$$p(a) = (p, 1-p)_a, \ p(e) = (1-q, q)_e, \quad \text{and} \quad b = (a+e)_2,$$

   find all the quantities $p(a, b), p(b), p(b|a), p(a|b)$ and $H(B), H(B|A), I(B:A), I(B:A|E)$. Find the channel capacity.

7. **Chain rule for mutual information.** [optional]

   Show from the definitions that the mutual information satisfies the following chain rule:

$$I(X:YZ) = I(X:Y) + I(X:Z|Y) = I(X:Z) + I(X:Y|Z).$$

   More generally,

$$I(X_1 \cdots X_n : Y) = \sum_{i=1}^{n} I(X_i Y | X_{i-1} \cdots X_1). \tag{3}$$

8. **Measuring entropy.** [optional] Build a Lempel-Ziv encoder (and a decoder, so you can check that it works), and use it to estimate the entropy of some stream of data.

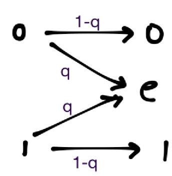
   Where to get the data: you could use a string of iid bits (in which case it's easy to check your answer), or you could use the output of a model of some physical process. A good example is the one in section IIA of this paper:

   > $N$ particles are distributed on a chain of $L > N$ sites with no multiple occupancy allowed. At each time step, represent the occupation numbers of particles as a string of $L$ 0s and 1s. A site is considered *active* if one of its neighbors is also occupied. At each time step, randomly select an active particle and move it to an unoccupied site. A useful order parameter is the fraction $f$ of active sites. If $f \to 0$, the dynamics stops. There is a critical value of the density $\rho = N/L$ above which the system never reaches an $f = 0$ state. Can you see the critical value of $f$ in the entropy?

   This reference also explains an implementation of the LZ77 algorithm.

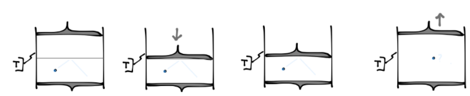
9. **Binary erasure channel.**

   Find the channel capacity of this channel:




10. **Mechanical engineering problem.** [optional]

In lecture I claimed that the expansion of an ideal gas against a piston, as in the figure at right, could be used to lift a weight.



   Design a plausible system of strings and pulleys to make this happen.

11. **Test of Landauer's Principle.** [optional]

   Consider logical bits which are stored in the magnetization (up or down) of little magnets. Show that copying a known bit (say 0) onto an *unkown* bit by the method described in lecture costs energy at least $k_B T \ln 2$.