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

Physics 239/139 Fall 2019
Assignment 9

Due 12:30pm Wednesday, December 4, 2019

1. **Brainwarmers.**

   (a) Is it true that $0 \leq S(A|C) + S(B|C)$? Prove or give a counterexample.

   (b) Show that the von Neumann entropy is the special case $S(\boldsymbol{\rho}) = \lim_{\alpha \to 1} S_\alpha(\boldsymbol{\rho})$ of the Renyi entropies:

   $$S_\alpha(\boldsymbol{\rho}) \equiv \frac{\mathrm{sgn}(\alpha)}{1-\alpha} \log \mathrm{tr} \boldsymbol{\rho}^\alpha = \frac{\mathrm{sgn}(\alpha)}{1-\alpha} \log \sum_a p_a^\alpha \ .$$

2. **Majorization questions.**

   (a) Show that if a doubly stochastic map is reversible (invertible and the inverse is also doubly stochastic) then it is a permutation.

   (b) Show that the set of doubly stochastic maps is convex (that is: a convex combination $\sum_a p_a D_a$ of doubly stochastic maps is doubly stochastic). What are the extreme points of this set? (This is the easier direction of the Birkhoff theorem.)

   (c) Show that a pure state and uniform state satisfy $(1, 0, 0 \cdots) \succ p \succ (1/L, 1/L \cdots)$ for any p on an L-item space.

   (d) A useful visualization of majorization relations is called the 'Lorenz curve': this is just a plot of the cumulative probability $P_p(K) = \sum_{k=1}^K p_k^\downarrow$ as a function of K. What does $p \succ q$ mean for the Lorenz curves of p and q? Draw the Lorenz curves for the uniform distribution and for a pure state.

   (e) Show that the set of probability vectors majorized by a fixed vector x is convex. That is: if $x \succ y$ and $x \succ z$ then $x \succ ty + (1-t)z, t \in [0, 1]$. Hints: (1) the analogous relation is true if we replace x, y, z with real numbers and $\succ$ with $\geq$. (2) Show that $P_{p^\downarrow}(K) \geq P_{\pi p^\downarrow}(K)$ (where $\pi p^\downarrow$ indicates any other ordering of the distribution).

   (f) For the case of a 3-item sample space we can draw some useful pictures of the whole space of distributions. The space of probability distributions on three elements is the triangle $x_1 + x_2 + x_3 = 1, x_i \geq 0$, which can be drawn in the plane. We can simplify the picture further by ordering the elements $x_1 \geq x_2 \geq x_3$, since majorization does not care about the order. Pick some distribution x with $x_1 \neq x_2 \neq x_3$ and draw the set of distributions which x majorizes, the set of distributions majorized by x, and the set of distributions with which x does not participate in a majorization relation ('not comparable to x').

3. **Work and the Holevo bound.** [optional]

   (a) Show that the Holevo quantity $\chi(p_a, \rho_a) \equiv S(\rho_{av}) - \sum_a p_a S(\rho_a)$ (with $\rho_{av} \equiv \sum_a p_a \rho_a$) can be written as $\chi(p_a, \rho_a) = \sum_a p_a D(\rho_a || \rho_{av})$.

   (b) Show that
   $$\sum_a p_a D(\rho_a||\sigma) = \chi(p_a, \rho_a) + D(\rho_{av}||\sigma).$$

   (c) Suppose A labors in contact with a heat bath at temperature T, and is governed by hamiltonian H. Convince yourself that in order to create the signal state ρ_a, the required work A must do is
   $$W_a \geq F_T[\rho_a] - F_T[\rho_T] = (k_B T \ln 2)\, D(\rho_a||\rho_T),$$
   where $F_T[\rho] \equiv \mathrm{tr}\rho H - T S_{vN}[\rho]$ is the free energy functional.

   (d) Show that the average work $\bar{W} \equiv \sum_a p_a W_a$ satisfies
   $$\bar{W} \geq (k_B T \ln 2)\, \chi(p_a, \rho_a).$$
   (hint: $D(\rho||\sigma) \geq 0$).

   (e) Apply the Holevo bound to conclude
   $$\bar{W} \geq (k_B T \ln 2)\, I(A:B),$$
   so that that every bit of information A can convey to B requires average work at least $k_B T \ln 2$. Yay, Landauer.

   (f) [optional] Estimate the amount of work done per bit sent to your cellular telephone.

4. **Holevo quantity and channel capacity.** Consider a collection of mutually-commuting density matrices $\{\rho_a\}$. Show that in this case, the Holevo quantity
$$\chi(p_a, \rho_a) \equiv S(\rho_{av}) - \sum_a p_a S(\rho_a) = \sum_a p_a D(\rho_a||\rho_{av}), \quad \rho_{av} \equiv \sum_a p_a \rho_a$$

is the mutual information $I(A:B)$, where the random variable B is the variable b labelling the mutual eigenvectors of the ρ_a: $\rho_a = \sum_b \lambda_a^b |b\rangle\langle b|$.

This suggests that a good definition of the capacity of a quantum channel for sending classical information (let's call it classical capacity) is determined by the Holevo quantity as

$$C = \chi(p_a, \rho_a)/\mathcal{T}$$

(where $\mathcal{T}$ is how long the information takes to go down the channel). And indeed, recall the Holevo bound, which says that $I(A:B) \leq \chi(p_a, \rho_a)$ where B is the outcomes of *any* measurement done on $\sum_a p_a \rho_a$.

5. **Channel capacity of the radiation field.**

   Suppose (crazy idea) we wanted to send signals using the electromagnetic field.

   The radiation field is a collection of quantum harmonic oscillators labelled by frequency, ω. For simplicity, let's consider a one-dimensional field with only one polarization, so there is one oscillator for each value of ω. In the first part of the problem, we'll put the system in a box, so that the allowed frequencies are integer multiples of some fundamental frequency, and the energy of a state with n_j photons in mode j is $E(\{n\}) = \sum_j j n_j h \equiv Nh$ for some constant h.

   The signal information could be stored for example in the number of photons $\bar{n}(\omega)$ with a given frequency. As in other examples, to send message a, A puts the field in the state ρ_a. And the message can be extracted by measurements on the resulting radiation field, for example by counting photons.

   For practical reasons, we will fix the power P of the signal. There are several ways to implement this constraint; we'll consider two below.

   At first we ignore the presence of noise.

   (a) Show that the Holevo quantity χ (and hence the channel capacity, no matter what measurement we do) is bounded by the entropy of the average signal $\sum_a p_a \rho_a$.

   (b) What is the ρ_{av} which maximizes the entropy, subject to the constraint of fixed energy $E(\{n\}) = P\mathcal{T}$ (where $\mathcal{T}$ is the duration of the signal)?

   (c) As a useful intermediate step, show that the entropy for a single harmonic oscillator in thermal equilibrium can be written in terms of the average occupation number $\bar{n}$ as $S_B(\bar{n})$ where

   $$S_B(n) \equiv (n+1)\log(n+1) - n\log n.$$

(d) Using the definition of classical capacity in the previous problem, determine the classical capacity of the channel in part 5b at large $\mathcal{T}$.

You may use the Hardy-Ramanujan formula, which counts partitions of N at large N:

$$\mathcal{N}(N) = \frac{1}{4\sqrt{3}N} e^{\pi\sqrt{\frac{2}{3}N}} + \mathcal{O}\left(e^{\frac{\pi}{2}\sqrt{\frac{2}{3}N}}\right).$$

(e) Alternatively, we may impose the condition of fixed power as a condition on the *average* energy. The state which maximizes entropy at fixed average energy is a thermal state. The temperature is determined by the average energy, which is in turn related to the power carried by the signal. Find the relation between T and P. Find a bound on the channel capacity at fixed average energy. (In this part of the problem you may take the infinite-volume limit.)

Inevitably there will be noise, represented by an additional number of photons $\bar{n}(\omega)$ at each frequency which are out of our control. Assume the noise is thermal, in equilibrium at temperature T_N. Suppose the power of the *signal* P (which is some amount of extra photons on top of the noise) is still fixed.

(f) Convince yourself that the upper bound on the channel capacity is now reduced by the entropy of the noise:

$$C\mathcal{T} \leq S(\rho_{T_{S+N}}) - S(\rho_{T_N})$$

where ρ_T is the thermal density matrix with temperature T, T_N is the noise temperature, and T_{S+N} is the temperature at an average energy which includes both the noise and the signal. Find T_{S+N} in terms of T_N and P.

(g) Do the integral over frequency. Study the high- and low-temperature limits of your answer. Confirm Landauer's principle in the former case in the following sense: compute the minimum power required to send a single bit.