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

Physics 213 Winter 2023
Assignment 9 – Solutions

Due 11:00am Tuesday March 14, 2023

1. **Brainwarmers.**

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

   A counterexample is:

   $$|\psi\rangle_{ABC} = |0\rangle_A \otimes \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)_{BC}$$

   which has $S(A) = S(BC) = 0, S(B) = S(C) = S(AB) = S(AC) = 1$ and hence $S(A|C) \equiv S(AC) - S(C) = 0, S(B|C) \equiv S(BC) - S(C) = 0 - 1 = -1$ so

   $$S(A|C) + S(B|C) = -1.$$

   In contrast, SSA says *e.g.*

   $$0 \leq S(C|A) + S(C|B) = S(AC) - S(A) + S(AB) - S(B) = 1 - 0 + 1 - 1 = 1.$$

   Notice that SSA in this form is a manifestation of monogamy of entanglement: $S(C|A)$ and $S(C|B)$ can each be negative, precisely when AC or CB are entangled, respectively. But SSA (in the form $0 \leq S(C|A) + S(C|B)$) says that making AC more entangled constrains how entangled BC can be.

   (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:

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

$$\lim_{\alpha\to 1}\frac{\mathrm{sgn}(\alpha)}{1-\alpha}\log\sum_a p_a^\alpha = \lim_{\alpha\to 1}\frac{1}{1-\alpha}\log\sum_a p_a e^{(\alpha-1)\ln p_a} \tag{1}$$

$$= \lim_{\alpha\to 1}\frac{1}{1-\alpha}\log\left(\sum_a p_a\left(1+(\alpha-1)\ln p_a+\mathcal{O}(\alpha-1)^2\right)\right) \tag{2}$$

$$= \lim_{\alpha\to 1}\frac{1}{1-\alpha}\log\left(\sum_a p_a\left(1+(\alpha-1)\ln p_a+\mathcal{O}(\alpha-1)^2\right)\right) \tag{3}$$

$$= \lim_{\alpha\to 1}\frac{1}{1-\alpha}\log\left(\left(1+(\alpha-1)\sum_a p_a\ln p_a+\mathcal{O}(\alpha-1)^2\right)\right) \tag{4}$$

$\log(1+x) = \frac{x}{\ln 2}+\mathcal{O}(x^2)$

$$= \lim_{\alpha\to 1}\frac{1}{1-\alpha}\left(\frac{\alpha-1}{\ln 2}\sum_a p_a\ln p_a+\mathcal{O}(\alpha-1)^2\right) \tag{5}$$

$$= -\frac{1}{\ln 2}\sum_a p_a\ln p_a = \frac{\ln 2}{\ln 2}(-\sum_a p_a\log p_a) = S(\boldsymbol{\rho}). \tag{6}$$

Alternatively, following Jin-Long Huang, we can subtract $0 = \log\sum_a p_a$ from $\log\sum_a p_a^\alpha$ and immediately recognize the expression for the derivative

$$\lim_{\alpha\to 1} S_\alpha(\rho) = -\lim_{\alpha\to 1}\frac{\log\sum_a p_a^\alpha - \log\sum_a p_a^1}{\alpha-1} \tag{7}$$

$$= -\partial_\alpha \log\sum_a p_a^\alpha|_{\alpha=1} \tag{8}$$

$$= -\frac{\sum_a p_a^\alpha\log p_a}{\sum_a p_a^\alpha}|_{\alpha=1} = -\sum_a p_a\log p_a = S(\rho). \tag{9}$$

2. **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.

3. **Holevo quantity and channel capacity.** [optional] 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: $\rho_a = \sum_b \lambda_a^b |b\rangle\langle b|$.

So suppose that $\rho_a = \sum_b p_b^{(a)} |b\rangle\langle b|$ are all simultaneously diagonal. First, notice that

$$p(b|a) = \langle b|\rho_a|b\rangle$$

is the conditional probability for outcome b given signal a. Then

$$S(\rho_{av}) = -\sum_{ab} p_a p_a^{(b)} \log\left(\sum_{a'} p_{a'} p_{a'}^{(b)}\right) \tag{10}$$

$$= -\sum_{ab} p(b|a) p_a \log\left(\sum_{a'} p(b|a) p_{a'}\right) \tag{11}$$

$$= -\sum_b p(b) \log p(b) = S(B). \tag{12}$$

Next, notice that under the assumption that the ρ_a are all diagonal in the $|b\rangle$ basis,

$$S(\rho_a) = -\sum_b \langle b| \rho_a \log \rho_a |b\rangle = -\sum_b \langle b| \sum_{b'} |b'\rangle\langle b'| p_{b'}^{(a)} \log p_{b'}^{(a)} |b\rangle = -\sum_b p_b^{(a)} \log p_b^{(a)} = H(B|A=a).$$

And therefore

$$\sum_a p_a S(\rho_a) = \sum_a p_a H(B|A=a) = H(B|A) = S(AB) - S(A)$$

is the conditional entropy.

Putting these together, the Holevo quantity is

$$\chi(p_a, \rho_a) = S(\rho_{av}) - \sum_a p_a S(\rho_a) = S_B - (S_{AB} - S_A) = S_A + S_B - S_{AB} = I(A:B),$$

the mutual information.

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 $\sum_a p_a \rho_a$.

4. **Channel capacity of the radiation field.** [optional but highly encouraged]

   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 $\sum_a p_a \rho_a$.

$S(\rho_a) \geq 0$, so $\chi \leq S(\rho_{av})$.

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

It is the uniform state on the set of energy eigenstates with $E(\{n\}) = P\mathcal{T} \equiv E_N$,

$$\rho_{max} = \frac{\mathbb{1}_{E_N}}{\mathcal{N}_N},$$

where $\mathcal{N}_N$ is the dimension of this space. $\mathcal{N}_N$ is equal to the number of partitions of N. This is like the microcanonical ensemble.

N is related to P by

$$P\mathcal{T} \equiv E_N = Nh.$$

(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

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

The partition function for a single SHO mode with frequency ω is

$$Z = \text{tr} e^{-\beta H} = \sum_{n=0}^{\infty} e^{-\beta\omega(n+\frac{1}{2})} = \frac{e^{\beta\omega/2}}{e^{\beta\omega}-1}.$$

The average occupation number is the Bose function $\bar{n} = \frac{1}{e^{\beta\omega}-1}$, which satisfies

$$\bar{n}+1 = \frac{e^{\beta\omega}}{e^{\beta\omega}-1}, \quad \frac{\bar{n}+1}{\bar{n}} = e^{\beta\omega}.$$

Therefore

$$S = -\partial_T F = -\ln\left(e^{\beta\omega}-1\right) + \beta\omega\frac{e^{\beta\omega}}{e^{\beta\omega}-1} \tag{13}$$

$$= \ln\bar{n} + \ln\frac{\bar{n}+1}{\bar{n}}\bar{n} + 1 = -\bar{n}\ln\bar{n} + (\bar{n}+1)\ln(\bar{n}+1). \tag{14}$$

Alternatively, we can write $p(n) = e^{-\beta\omega(n+\frac{1}{2})}/Z$ in terms of $\bar{n}$ and $\bar{n}+1$ as

$$p(n) = e^{-\beta n\omega}(1-e^{-\beta\omega}) = \frac{\bar{n}^n}{(\bar{n}+1)^{n+1}}$$

and use

$$S(p) = -\sum_{n=0}^{\infty} p(n) \log p(n) \tag{15}$$

$$= \sum_n (n+1) \frac{\bar{n}^n}{(\bar{n}+1)^{n+1}} \log(\bar{n}+1) - \sum_n n \frac{\bar{n}^n}{(\bar{n}+1)^{n+1}} \log \bar{n} \tag{16}$$

$$= -\bar{n} \ln \bar{n} + (\bar{n}+1) \ln(\bar{n}+1). \tag{17}$$

(d) Using the definition of classical capacity in the previous problem, determine the classical capacity of the channel in part 4b at large $\mathcal{T}$.

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

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

We can use the H-R formula because N is large when $P\mathcal{T}$ is large.

The channel capacity is a *rate*:

$$C = \frac{S_{max}}{\mathcal{T}} = \frac{\log \mathcal{N}(N)}{\mathcal{T}} = \frac{1}{\mathcal{T}} \pi \sqrt{\frac{2}{3}N} = \frac{1}{\mathcal{T}} \pi \sqrt{\frac{2}{3} \frac{P\mathcal{T}}{h}} = \pi \sqrt{\frac{2P}{3h\mathcal{T}}}.$$

(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.)

The energy *density* in the radiation field at temperature $T = 1/\beta$ (in one infinite dimension) is

$$\langle \mathcal{E} \rangle = \int d\omega \, \langle n(\omega) \rangle \, \omega = \int d\omega \omega \frac{1}{e^{\beta\omega} - 1} = \frac{\pi T^2}{12\hbar}.$$

If we wait a time $\mathcal{T}$ a chunk of radiation of length $c\mathcal{T}$ will pass us; its energy is $E = \langle \mathcal{E} \rangle c\mathcal{T}$. We must equate this with $E = P\mathcal{T}$ giving

$$P = c \langle \mathcal{E} \rangle = \frac{c\pi T^2}{12\hbar}.$$

Similarly, when counting the rate of communication, the signal is moving at the speed of light, so in time $\mathcal{T}$, a chunk of length $c\mathcal{T}$ passes B. So the bound on the rate is given by cS where S is the entropy *density*.

$$C = cS_{max} = c\int d\omega S_B(\bar{n}(\omega)).$$

Here $S_B(n) = (n+1)\log(n+1) - n\log n$ is the entropy of a bosonic mode with average occupation number n. In thermal equilibrium at temperature T, $\bar{n}(\omega) = \frac{1}{e^{\omega/T}-1}$ is the Bose distribution.

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.

If N is the power in the noise

$$(P+N)\mathcal{T} = \frac{\pi}{12}T_{S+N}^2$$

So

$$T_{S+N} = \sqrt{\frac{12}{\pi}Pc + T_N^2}.$$

(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.

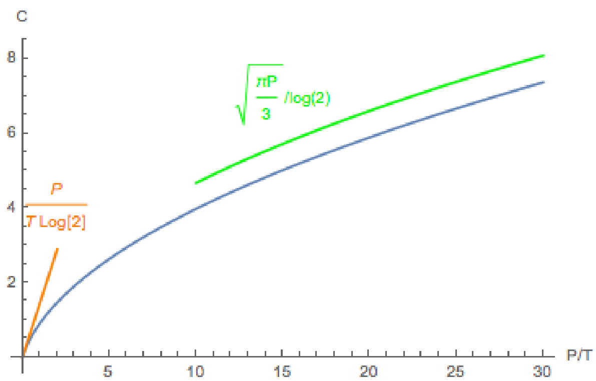
$$\int_{-\infty}^{\infty} d\omega S_B(\langle n\rangle_T) = T\int_{-\infty}^{\infty} d\theta S_B\left(\frac{1}{e^\theta - 1}\right) = T\frac{\pi^2}{3},$$

as Mathematica can tell you.

Then

$$C \le S(\rho_{T_{S+N}}) - S(\rho_{T_N}) \tag{18}$$
$$= \frac{\pi^2}{3}\left(T_{S+N} - T_N\right) \tag{19}$$
$$= \frac{\pi^2}{3}\left(\sqrt{\frac{12}{\pi}Pc + T_N^2} - T_N\right) \tag{20}$$

This function looks like this (in units where $T = 1$):



When $P \gg T_N^2$, we can ignore the noise and we reproduce the answer from the first part of the problem. In the high-temperature limit we find

$$C \le \frac{Pc}{k_B T_N \log 2}$$

which says that the condition to send a single bit is precisely: the power must exceed the Landauer bound, $k_B T_N \log 2$.

This problem is loosely based on the discussion in Vedral, quant-ph/0102094, which I found incredibly confusing. For example: there is a 3 missing in eqn 45, the expressions for the SHO thermal density matrices should read $\rho(\omega) = \sum_n \frac{1-e^{-\beta\omega}}{e^{n\beta\omega}} |n\rangle\langle n|$, there is a minus sign missing in the expression for S_B, and the factors of c are missing, while all the factors of $\hbar, k_B$ are present. For more on this subject see this review by Caves and Drummond or this beautiful PRL by Yuen and Ozawa.

5. **Direct application of Lieb's theorem.**

   We only used a very special case of Lieb's theorem to prove monotonicity of the relative entropy. Surely there is more to learn from it.

   Consider an ensemble of states $\boldsymbol{\rho} = \sum_i p_i \boldsymbol{\rho}_i$, and a unitary operator $\mathbf{U}$ (for example, it may be closed-system time evolution).

Show that the relative entropy between $\boldsymbol{\rho}(t) \equiv \mathbf{U}\boldsymbol{\rho}\mathbf{U}^\dagger$ and $\boldsymbol{\rho}$ is convex in $\boldsymbol{\rho}$:

$$D(\boldsymbol{\rho}(t)||\boldsymbol{\rho}) \leq \sum_i p_i D(\boldsymbol{\rho}_i(t)||\boldsymbol{\rho}_i).$$

In the notation we used for Lieb's theorem, $f_{s,X}(\rho, \sigma) \equiv \mathrm{tr} X^\dagger \rho^{1-s} X \sigma^s$,

$$D(\boldsymbol{\rho}(t)||\boldsymbol{\rho}) = -\partial_s|_{s=0} \mathrm{tr} \mathbf{U}\boldsymbol{\rho}^{1-s}\mathbf{U}^\dagger \boldsymbol{\rho}^s = -\partial_s|_{s=0} f_{s,\mathbf{U}^\dagger}(\boldsymbol{\rho}, \boldsymbol{\rho}).$$

Since Lieb's theorem says $f_{\mathbf{U},s}(\boldsymbol{\rho}, \boldsymbol{\rho})$ is jointly concave, this function is jointly convex:

$$D(\boldsymbol{\rho}(t)||\boldsymbol{\rho}) \leq \sum_i p_i \left(-\partial_s|_{s=0} f_{s,\mathbf{U}^\dagger}(\boldsymbol{\rho}_i, \boldsymbol{\rho}_i)\right) = \sum_i p_i D(\boldsymbol{\rho}_i(t)||\boldsymbol{\rho}_i)$$

as requested.

Open ended bonus problem: see if you can find a better result by directly applying Lieb's joint concavity theorem to a problem in many body physics.

6. **Random singlets.** [optional]

   Consider qbits arranged on a chain. Suppose that the groundstate is made of random singlets, in the following sense: for a given site i, with probability $f(|i - j|a)$ (a is the lattice spacing), the spins at i and j are in the state $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$. Every spin is paired with some other spin. Consider in turn the case of short-range singlets where $f(x) \propto e^{-x/\xi}$, and long-range singlets where $f(x) \propto \frac{1}{x^2+\delta^2}$.

   (a) Consider a region A which is an interval $\left[-\frac{R-\epsilon}{2}, \frac{R-\epsilon}{2}\right]$ ($\epsilon \ll R$) and B is what we called $\bar{A}^-$ (nearly the complement), more precisely: $B \equiv [-\infty, -\frac{R}{2}] \cup [\frac{R}{2}, \infty]$. Let $I_\epsilon(R) \equiv I(A : B) = S(A) + S(B) - S(AB)$ be their mutual information.

   Find $\overline{\left\langle \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_j \right\rangle}$ (where $\vec{\mathbf{S}} = \frac{1}{2}(\sigma^x, \sigma^y, \sigma^z)$) and $\overline{I_\epsilon(R)}$. In both cases assume the regions are big enough that you can average over regions and use a continuum approximation ($\xi, \delta \gg$ lattice spacing).

   Check that the answer is consistent with the mutual information bound on correlations.

   If two spins ij are paired, $\langle singlet|_{ij} \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_j |singlet\rangle_{ij} = -\frac{3}{4}$. We assume that otherwise $\left\langle \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_j \right\rangle = 0$. Therefore $\overline{\left\langle \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_j \right\rangle} = -\frac{3}{4} f(|i - j|)$.

   The mutual information is equal to $2\log 2$ times the number of singlet bonds connecting the two regions. On average, this is

$$\overline{I(A:B)} = 2\log 2 \int_B dy \int_A dx f(x - y).$$

In the case of short-range singlets this gives

$$\overline{I(A:B)}_{\text{short}} = 4\int_{-\infty}^{R/2} dy e^{y/\xi} \int_{-R/2}^{R/2} dx e^{-x/\xi} = 8\xi^2 e^{-\frac{R}{2\xi}} \sinh\frac{R}{2\xi},$$

which approaches $4\xi^2$ for $R \gg \xi$.

For long range singlets, Mathematica says

$$\overline{I(A:B)}_{\text{long}} = 2\frac{R}{\delta}\left(\pi - 2\tan^{-1}\left(\frac{R}{\delta}\right)\right) + \ln\left(1 + \frac{R^2}{\delta^2}\right)$$

which is unbounded as $R \gg \delta$.

Actually the bound we proved does not quite apply to the operator $\vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_j$ because this is not of the form $\mathcal{O}_A\mathcal{O}_B$ (rather it is a sum of three such operators). Let's instead check X_iX_i, for which

$$\langle singlet|_{ij} X_iX_j |singlet\rangle_{ij} = -1, \overline{\langle X_iX_j\rangle} = -f(|i-j|), \| X_i \| = 1.$$

Then

$$\frac{\overline{\langle X_iX_j\rangle}^2}{2\| X \|^2} = \frac{1}{2}f(|i-j|)^2.$$

In both cases, I is larger than $f(\epsilon)$, where ϵ is the maximum separation between A and $\bar{A}^-$.

(b) Consider instead the case where $B = [-\infty, -\frac{R}{2} - L] \cup [\frac{R}{2} + L, \infty]$, so that A and B are separated by a distance L. Show that: for short-range singlets, (i) all (averaged) correlation functions decay exponentially in L (ii) $I(A : B) \sim e^{-L/\xi}$ for large L (and hence the mutual information satisfies an area law). For long-range singlets (i) (averaged) correlation functions have power law decay (ii) $I(A : B) \sim \log(2R - L)$ for large L, and there is no area law.

Clearly the averaged correlation functions are simply proportional to $f(i-j)$, which (since f is monotonically decreasing in both cases) is less than its value at the minimum separation between A and B, namely $f(L)$.

Now

$$\overline{I(A:B)}_{\text{short}} = 2\xi^2 e^{-\frac{2L+R}{2\xi}} \sinh\frac{R}{2\xi} \overset{L\gg\xi}{\sim} 4\xi^2 e^{-L/\xi},$$

exponential decay.

This problem is from this paper.