Exam

3 hours ; documents and pocket calculators not allowed ; dictionaries allowed

Tentative scale : 20/3 points for each of the three problems

I Mutual information

Whenever two random variables X and Y are not independent, it is interesting to quantify the information obtained on one from the knowledge of the other. This is the purpose of the present exercise, where the notions of conditional entropy and mutual information will be introduced.

We denote by $p(x, y)$ the joint probability of (X, Y), $p(y|x)$ is the conditional probability of variable Y given X, and $p(x)$ is the probability of X (marginal law). The variables are supposed discrete, and we then have $p(x) = \sum_y p(x, y)$. We recall that Bayes formula connects the three above probabilities through : $p(x, y) = p(y|x)p(x)$, and we then define the conditional entropy $S_{Y|X}$ as the entropy of the conditional law $p(y|x)$, subsequently averaged over x :

$$S_{Y|X} \;=\; -\sum_x p(x) \sum_y p(y|x) \log_2 p(y|x) \; . \tag{1}$$

In all these definitions, the log in base 2 shall be considered.

1. Establish the link between the entropy S_X of variable X, the entropy $S_{X,Y}$ of the joint distribution, and the conditional entropy $S_{Y|X}$.
2. For concreteness, we consider, in this question only, that X and Y represent the result (heads or tails) of tossing two correlated coins, with

$$p(H, H) = \frac{1}{2} \; , \quad p(H, T) = \frac{1}{4} \; , \quad p(T, H) = \frac{1}{8} \; , \quad p(T, T) = \frac{1}{8} \; .$$

   Compute the entropies $S_{X,Y}$, S_X and $S_{Y|X}$; we give $3 \log_2 3 \approx 4,8$.
3. What is the relation between $S_{Y|X}$ and S_Y if the variables X and Y are independent ?
4. Write the entropy difference $S_Y - S_{Y|X}$ in the form of the Kullback-Leibler distance between $p(x, y)$ and another probability distribution. Show then that $S_{Y|X} \le S_Y$, and comment this inequality.
5. We define the mutual information by

$$I_{X,Y} \;=\; S_X + S_Y - S_{X,Y} \; . \tag{2}$$

   Express the above quantity as a function of S_Y and $S_{Y|X}$ on the one hand, and as function of S_X and $S_{X|Y}$ on the other hand. What does mutual information measure, and what is its sign ? What is the value of $I_{X,Y}$ in the case where X and Y are independent ?
6. Information degradation
   (a) We consider a random variable X and the random variable $f(X)$ obtained by applying an arbitrary (but deterministic) function to X. What is the value of $S_{f(X)|X}$? What can we say about $S_{X|f(X)}$? Proceed to show that $S_{f(X)} \le S_X$, and provide a condition on f for turning the inequality into an equality.
   (b) Which inequality can we expect between $I_{X,Y}$ and $I_{X,f(Y)}$, where f is again a deterministic function (no proof is asked, only a plausibility argument) ?

(c) The mutual information between two variables (X and Y), conditioned by a third (Z), is defined as $I_{X,Y|Z} = S_{X|Z} - S_{X|(Y,Z)} = S_{Y|Z} - S_{Y|(X,Z)}$. Admitting the following composition rule

$$I_{X_1,(X_2,X_3)} = I_{X_1,X_2} + I_{X_1,X_3|X_2} = I_{X_1,X_3} + I_{X_1,X_2|X_3} \,, \tag{3}$$

recover your intuitive answer to the previous question [hint : it proves convenient to take $X_1 = X$, $X_2 = Y$, and $X_3 = f(Y)$].

7. Propagation of information

We consider random variables $X_0, X_1, \ldots, X_n$ that may each take two values, P and F for instance. A given person chooses one of the above "messages" with probability 1/2, which fixes the value of X_0. This person then transmits the information to its neighbour X_1, but because of ambient noise, the probability that X_1 hears the right value of X_0 is only $1-p$. There is therefore a probability p that the transmission is erroneous, i.e. that $X_1 \neq X_0$. X_1 further transmits to X_2, and so forth. Information is thereby transferred to X_n, with the same probability p of transmission error between X_i and X_{i+1}. We shall compute the mutual information $I_n \equiv I_{X_0,X_n}$ between the emitter and the final receiver, in particular when the distance n between both is large.

(a) What is the marginal law of X_n ? Show then that $I_n = 1 - H(p_n)$, where $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the entropy of a binary variable with probability x, and p_n is the probability that $X_0 \neq X_n$.

(b) Determine p_n, from a recurrence relation. Comment on the limiting behaviour of p_n when $n \to \infty$.

(c) Obtain the asymptotic formula for I_n in the form $I_n \underset{n\to\infty}{\sim} a\, b^n$. What are the values of the constants a and b ?

(d) We consider the one dimensional Ising model with free boundary conditions, that is a system of N Ising spins $\sigma_0, \ldots, \sigma_{N-1} \in \{+1,-1\}^N$. The energy of a configuration reads $H = -J\sum_{i=1}^{N-1} \sigma_{i-1}\sigma_i$, and we suppose that the system is in thermal equilibrium at inverse temperature β. Show that this Ising model realises exactly the afore-defined process. What is the connection between p and Ising model parameters ? Show that the correlation length derived in class is recovered.

8. Subsidiary question : derive relation (3). To this end, one can use after justification that

$$S_{X_1,(X_2,X_3)} = S_{X_1,X_2,X_3} = S_{X_1} + S_{X_2|X_1} + S_{X_3|(X_1,X_2)} \,. \tag{4}$$

II Landau theory and tricriticality

Within a Landau approach, the free energy is expanded in powers of the order parameter ϕ, in the form

$$\mathcal{R} = \frac{1}{2}a_2\,\phi^2 + \frac{1}{4}a_4\,\phi^4 + \frac{1}{6}a_6\,\phi^6 \,, \tag{5}$$

where in the vicinity of T_c, the coefficient a_2 linearly depends on temperature T : $a_2 = (T - T_c)\,\widetilde{a}_2$ (with $\widetilde{a}_2 > 0$). For simplicity, we suppose that a_4 and a_6 are independent on temperature.

A Order of transitions and sign of coefficients

1. Provide an example of a physical system where such an expansion may be relevant.
2. What should the sign of a_6 be ?
3. Sketch the free energy profiles $\mathcal{R}(\phi)$ for different temperatures, treating separately the cases $a_4 > 0$ and $a_4 < 0$. For each case, what is the order of the corresponding phase transition embodied in Eq. (5) ?

4. In the first order transition case, we denote T^* the critical temperature. What are the conditions on $\mathcal{R}$ which determine this temperature? At $T = T^*$, give the values of ϕ that minimize the free energy and show that
$$T^* = T_c + \alpha \frac{a_4^2}{\widetilde{a}_2\, a_6} .$$
What is the value of α?
5. Draw a schematic phase diagram in the plane (T, a_4), for fixed values of $\widetilde{a}_2$ and a_6. Indicate the phase transition lines of first and second orders. In which point of the diagram do they meet?

B Study of the tricritical point

We will now assume that $a_4 = 0$, which defines for our model a so-called tricritical point.

1. What is the order of the phase transition?
2. How does the order parameter depend on temperature, in the vicinity of T_c? Infer from this behaviour the value of the β exponent, defined by $\phi \propto (T_c - T)^\beta$ below the critical point. What is the value of β when $a_4 > 0$?
3. Under the action of an applied external field h, which additional term should appear in the free energy? What is the exponent δ which measures, at $T = T_c$, the response to h through $\phi \propto h^{1/\delta}$? Same question when $a_4 > 0$.

C Upper critical dimension for the tricrical point

We first admit that the spatial autocorrelation function of the order parameter, $\Gamma(\vec{r})$, is isotropic and of the form
$$\Gamma(\vec{r}) = \frac{1}{\xi^{d-2}} F\left(\frac{r}{\xi}\right), \tag{6}$$
with $r = |\vec{r}|$, d the space dimension, ξ the correlation length, and F an unspecified function. We may equivalently assume that $\Gamma(\vec{r}) = \frac{1}{r^{d-2}} \widetilde{F}\left(\frac{r}{\xi}\right)$, the scaling functions F and $\widetilde{F}$ being simply related.

We again address the case $a_4 = 0$.

1. Discuss qualitatively the behaviour of Γ as a function of r at T_c and at $T \neq T_c$ (no calculation asked).
2. We consider the low temperature phase and we define the Ginzburg ratio as
$$R_G = \frac{1}{\phi_{\text{eq}}^2} \frac{1}{\xi^d} \int_{\xi^d} \Gamma(\vec{r}) d\vec{r} , \tag{7}$$
where ϕ_{eq} is the equilibrium value of the order parameter. Why is the quantity R_G interesting?
3. Determine the temperature dependence of R_G in the vicinity of the critical temperature, assuming that ξ diverges at T_c with an exponent $\nu = 1/2$, and making use of the β exponent derived in part B (or treating β as a free parameter if needed). Give the upper critical dimension of the model, and comment upon it.
4. Write a free energy functional of the Ginzburg-Landau type, that generalizes expression (5) to situations where ϕ changes with position $\vec{r}$, and which includes a term stemming from an external field $h(\vec{r})$.
5. What is the differential equation fulfilled by the correlation function? What form does it take when the system is spatially homogeneous?
6. Show then that one indeed has $\nu = 1/2$.
7. Which simple prediction can be made for the behaviour of the surface tension when T approaches T_c?

III A microscopic approach for Onsager relations

The goal is here to show, starting from a microscopic model of a system's dynamics, that Onsager's symmetry relations hold.

We assume that the microscopic configuration of the system is specified by a set of N real variables $x_1, \ldots, x_N$. We denote $\underline{x} = (x_1, \ldots, x_N)$ the global configuration of the system, and $H(\underline{x})$ its energy. When put in contact with a thermostat, the configuration evolves according to Langevin's equations

$$\frac{dx_i}{dt} = -\frac{\partial H}{\partial x_i}(x_1(t), \ldots, x_N(t)) + \xi_i(t) \qquad \text{for } i = 1, \ldots, N \ , \tag{8}$$

where the $\xi_i(t)$ are random functions accounting for the coupling to the thermostat, with

$$\langle \xi_i(t) \rangle = 0 \ , \qquad \langle \xi_i(t) \xi_j(t') \rangle = \Gamma \, \delta_{i,j} \, \delta(t - t') \ . \tag{9}$$

We admit that the probability $P(\underline{x}, t)$ to observe the system in configuration $\underline{x}$ at time t evolves following the Fokker-Planck equation

$$\frac{\partial P}{\partial t} = WP \ . \tag{10}$$

The Fokker-Planck operator W acts on functions $f(\underline{x})$ of phase space configurations following

$$(Wf)(\underline{x}) = \sum_{i=1}^{N} \frac{\partial}{\partial x_i} \left[\frac{\partial H}{\partial x_i} f(\underline{x}) + \frac{\Gamma}{2} \frac{\partial f}{\partial x_i} \right] \ . \tag{11}$$

1. Provide a qualitative interpretation of Langevin equations when no random terms are at work ($\xi_i = 0$).
2. Give a sufficient condition on Γ such that Gibbs-Boltzmann distribution $P_{\text{eq}}(\underline{x}) = e^{-\beta H(\underline{x})}/Z$ be stationary. We shall suppose this condition fulfilled in the remainder.

   We assume that at $t = 0$, the initial condition $\underline{x}(t = 0)$ is randomly chosen from the probability distribution function $P_0(\underline{x})$, and that the system evolves, for $t \geq 0$, according to the above Langevin dynamics. We denote by $\langle \bullet \rangle$ averages over the initial condition and over the stochastic subsequent evolution. For two observables A, B (i.e. two functions $A(\underline{x}), B(\underline{x})$), we denote $\langle A(t) \rangle = \langle A(\underline{x}(t)) \rangle$ the average at time t and $\langle A(t) B(t') \rangle = \langle A(\underline{x}(t)) B(\underline{x}(t')) \rangle$ the two-time correlation function.

3. We denote by $Q_{\underline{x},\underline{x}'}(\tau)$ the probability that the system be in configuration $\underline{x}$ at a given time t, knowing that it was in configuration $\underline{x}'$ at time $t - \tau$. Write Q as a function of W (operators will be formally treated as matrices).
4. Write the expression of $\langle A(t) \rangle$ as a function of P_0, Q and $A(\underline{x})$.
5. Same question for $\langle A(t) B(t') \rangle$, assuming that $t > t'$.
6. Simplify your answers to the two previous questions assuming that the system is initially in equilibrium, i.e. with $P_0 = P_{\text{eq}}$.
7. Admitting that
   $$Q_{\underline{x}',\underline{x}}(\tau) P_{\text{eq}}(\underline{x}) = Q_{\underline{x},\underline{x}'}(\tau) P_{\text{eq}}(\underline{x}') \ , \tag{12}$$
   show that for an equilibrium system, $\langle A(t) B(t') \rangle = \langle B(t) A(t') \rangle$.
8. Briefly recall the consequences of this invariance by time reversal, in the framework of weakly out of equilibrium thermodynamics.
9. Subsidiary question : prove the relation (12).
   Hint : one can start by showing that $W P_{\text{eq}} = P_{\text{eq}} W^\dagger$, where P_{eq} is the diagonal multiplication by $P_{\text{eq}}(\underline{x})$ operator and $\dagger$ denotes transposition. To this end, it can be remarked that

$$W = \sum_{i=1}^{N} \left[\left(\frac{\partial^2 H}{\partial x_i^2} \right) + \left(\frac{\partial H}{\partial x_i} \right) \frac{\partial}{\partial x_i} + \frac{\Gamma}{2} \frac{\partial^2}{\partial x_i^2} \right] \ , \qquad \text{and that} \quad \left(\frac{\partial}{\partial x_i} \right)^\dagger = -\frac{\partial}{\partial x_i} \ . \tag{13}$$