

EXAMEN

DURÉE : 3H ; DOCUMENTS ET CALCULATRICES INTERDITS, DICTIONNAIRES AUTORISÉS

Barème indicatif : I sur 20/3 points, II sur 20/3 points, III sur 20/3 points

I L'information mutuelle

Lorsque deux variables aléatoires X et Y ne sont pas indépendantes, il est intéressant de quantifier l'information que l'on acquiert sur l'une d'elle en connaissant l'autre. C'est ce que nous nous proposons de faire dans cet exercice, où les notions d'entropie conditionnelle et d'information mutuelle vont être introduites.

On note $p(x, y)$ la loi jointe du couple (X, Y) , $p(y|x)$ la loi conditionnelle de Y sachant X , et $p(x)$ la loi marginale de X . Les variables étant supposées discrètes, on a ainsi $p(x) = \sum_y p(x, y)$. On rappelle la formule de Bayes entre ces trois lois : $p(x, y) = p(y|x)p(x)$, et on définit l'entropie conditionnelle $S_{Y|X}$ comme l'entropie de la loi conditionnelle $p(y|x)$, moyennée sur x :

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

Dans toutes les définitions d'entropie on prendra le logarithme en base 2.

1. Établir le lien entre l'entropie S_X de la variable X , l'entropie $S_{X,Y}$ de la loi jointe, et l'entropie conditionnelle $S_{Y|X}$.
2. Supposons par exemple, dans cette question seulement, que X et Y représentent le résultat (Pile ou Face) du tirage de deux pièces (corrélées), avec

$$p(P, P) = \frac{1}{2}, \quad p(P, F) = \frac{1}{4}, \quad p(F, P) = \frac{1}{8}, \quad p(F, F) = \frac{1}{8}.$$

Calculer les entropies $S_{X,Y}$, S_X et $S_{Y|X}$; on donne $3 \log_2 3 \approx 4,8$.

3. Quelle est la relation entre $S_{Y|X}$ et S_Y si les variables X et Y sont indépendantes ?
4. Mettre la différence d'entropie $S_Y - S_{Y|X}$ sous la forme d'une distance de Kullback-Leibler entre $p(x, y)$ et une autre loi de probabilité. En déduire qu'en général $S_{Y|X} \leq S_Y$, et commenter cette inégalité.
5. On définit l'information mutuelle par

$$I_{X,Y} = S_X + S_Y - S_{X,Y} . \quad (2)$$

L'exprimer d'une part en fonction de S_Y et $S_{Y|X}$, d'autre part en fonction de S_X et $S_{X|Y}$. Que mesure l'information mutuelle et quel est son signe ? Quelle est la valeur de $I_{X,Y}$ dans le cas où les variables X et Y sont indépendantes ?

6. Dégradation de l'information

- (a) On considère une variable aléatoire X et on note $f(X)$ la variable aléatoire obtenue par l'application d'une fonction f (arbitraire mais déterministe) sur le résultat du tirage de X . Que vaut $S_{f(X)|X}$? Que peut-on dire de $S_{X|f(X)}$? En déduire que $S_{f(X)} \leq S_X$, et donner une condition sur f pour qu'il y ait égalité.
- (b) Quelle inégalité est-on en droit d'attendre entre $I_{X,Y}$ et $I_{X,f(Y)}$, où f est une fonction déterministe quelconque (aucune preuve n'est demandée ici, seulement un argument de plausibilité) ?

- (c) L'information mutuelle entre deux variables (X et Y), conditionnées par une troisième (Z), est définie selon $I_{X,Y|Z} = S_{X|Z} - S_{X|(Y,Z)} = S_{Y|Z} - S_{Y|(X,Z)}$. En admettant la règle de composition suivante

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

justifier votre réponse intuitive à la question précédente [on pourra prendre $X_1 = X$, $X_2 = Y$, et $X_3 = f(Y)$].

7. Propagation de l'information

On considère des variables aléatoires $X_0, X_1, \dots, X_n$ qui peuvent chacune prendre deux valeurs, P et F par exemple. Une personne choisit un de ces deux “messages” avec probabilité $1/2$, ce qui fixe la valeur de X_0 . Elle transmet ensuite cette information à son voisin X_1 , mais à cause du bruit ambiant la probabilité que X_1 entende la bonne valeur de X_0 n'est que de $1 - p$, il y a donc une probabilité p que la transmission soit erronée, i.e. que $X_1 \neq X_0$. X_1 transmet alors à X_2 , et de proche en proche l'information est ainsi transférée jusqu'à X_n , avec à chaque fois la même probabilité d'erreur p entre X_i et X_{i+1} . On va calculer l'information mutuelle $I_n \equiv I_{X_0,X_n}$ entre l'émetteur et le récepteur final, en particulier quand la distance n entre eux devient très grande.

- Quelle est la loi marginale de X_n ? En déduire que $I_n = 1 - H(p_n)$, où $H(x) = -x \log_2 x - (1-x) \log_2 (1-x)$ est l'entropie d'une variable binaire de probabilité x , et p_n est la probabilité que $X_0 \neq X_n$.
- Déterminer p_n , en raisonnant par récurrence. Commenter la limite de p_n quand $n \rightarrow \infty$.
- Obtenir un équivalent de I_n de la forme $I_n \underset{n \rightarrow \infty}{\sim} a b^n$, où l'on précisera les valeurs des constantes a et b .
- On considère le modèle d'Ising unidimensionnel avec conditions aux bords libres, i.e. un système de N spins d'Ising $\sigma_0, \dots, \sigma_{N-1} \in \{+1, -1\}$. L'énergie d'une configuration est $H = -J \sum_{i=1}^{N-1} \sigma_{i-1} \sigma_i$, et l'on suppose le système à l'équilibre avec un thermostat de température inverse β . Montrer que ce modèle réalise exactement le processus défini ci-dessus; on donnera un lien quantitatif entre p et les paramètres du modèle d'Ising, et l'on retrouvera la longueur de corrélation établie en TD.

8. Question subsidiaire : démontrer la relation (3). Pour ce faire, on pourra remarquer –après justification– que

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

II Théorie de Landau et tricriticalité

Dans une approche à la Landau, on effectue le développement de l'énergie libre en puissances du paramètre d'ordre ϕ , sous la forme

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

où le coefficient a_2 dépend linéairement de la température T au voisinage de T_c : $a_2 = (T - T_c) \tilde{a}_2$ (avec $\tilde{a}_2 > 0$). On suppose pour simplifier a_4 et a_6 indépendants de la température.

A Ordre des transitions et signe des coefficients

- Donner un exemple de système physique où un tel développement pourrait apparaître.
- Quel doit être le signe de a_6 ?
- Donner l'allure des profils d'énergie libre $\mathcal{R}(\phi)$ pour différentes températures, en distinguant les cas $a_4 > 0$ et $a_4 < 0$. Pour chacun de ces deux cas, quel est l'ordre de la transition de phase que décrit la relation (5)?

4. Dans le cas où la transition est du premier ordre, on note T^* sa température critique. Quelles sont les conditions sur $\mathcal{R}$ déterminant cette température ? A $T = T^*$, donner les valeurs de ϕ minimisant l'énergie libre, et montrer que

$$T^* = T_c + \alpha \frac{a_4^2}{a_2 a_6},$$

où l'on précisera la valeur de α .

5. Proposer un diagramme de phase dans le plan (T, a_4) , pour des valeurs fixées de $\tilde{a}_2$ et de a_6 , où seront indiquées les lignes de transition du premier et deuxième ordre. En quel point de ce diagramme les lignes associées à ces deux types de transition se rencontrent-elles ?

B Étude du point tricritique

On suppose désormais $a_4 = 0$, ce qui définit pour ce modèle un point dit tricritique.

1. Quel est l'ordre de la transition ?
2. Comment le paramètre d'ordre dépend-il de la température au voisinage de T_c ? En déduire la valeur de l'exposant β défini par $\phi \propto (T_c - T)^\beta$ en dessous du point critique. Quelle est la valeur de β lorsque $a_4 > 0$?
3. En présence d'un champ extérieur h , quel terme additionnel apparaît-il dans l'énergie libre ? Préciser l'exposant δ qui mesure, à $T = T_c$, la réponse au champ appliqué via $\phi \propto h^{1/\delta}$. Même question lorsque $a_4 > 0$.

C Dimension critique supérieure pour le point tricritique

Nous admettons dans un premier temps que la fonction de corrélation spatiale du paramètre d'ordre, $\Gamma(\vec{r})$, est isotrope et de la forme

$$\Gamma(\vec{r}) = \frac{1}{\xi^{d-2}} F\left(\frac{r}{\xi}\right), \quad (6)$$

où $r = |\vec{r}|$, d désigne la dimension d'espace, ξ la longueur de corrélation, et F est une fonction non précisée. On peut de manière équivalente supposer que $\Gamma(\vec{r}) = \frac{1}{r^{d-2}} \tilde{F}\left(\frac{r}{\xi}\right)$, les fonctions d'échelle F et $\tilde{F}$ étant reliées l'une à l'autre de manière simple.

Nous traitons ici de nouveau le cas $a_4 = 0$.

1. Discuter qualitativement le comportement de Γ en fonction de r à T_c et à $T \neq T_c$ (aucun calcul n'est demandé).
2. On considère la phase de basse température du système, et l'on définit le rapport de Ginzburg par

$$R_G = \frac{1}{\phi_{\text{eq}}^2} \frac{1}{\xi^d} \int_{\xi^d} \Gamma(\vec{r}) d\vec{r}, \quad (7)$$

où ϕ_{eq} est la valeur d'équilibre du paramètre d'ordre. Quel est l'intérêt de la quantité R_G ?

3. Déterminer la dépendance en température de R_G au voisinage de la température critique, en admettant que ξ diverge à T_c avec un exposant $\nu = 1/2$, et en utilisant l'exposant β obtenu dans la partie B (ou en le traitant comme un paramètre libre le cas échéant). En déduire la dimension critique supérieure du modèle, et la commenter.
4. Écrire une fonctionnelle de Ginzburg-Landau qui généralise l'expression (5) dans le cas où ϕ varie avec la position $\vec{r}$, et qui inclue un terme dû à un champ extérieur $h(\vec{r})$.
5. En déduire l'équation différentielle vérifiée par la fonction de corrélation. Que devient cette équation lorsque le système est spatialement homogène ?
6. Déduire de ce qui précède que l'on a bien $\nu = 1/2$.
7. Quelle prédiction simple peut-on faire concernant le comportement de la tension de surface au voisinage de T_c ?

III Une approche microscopique pour les relations d'Onsager

L'objectif de l'exercice est de justifier, à partir d'une modélisation microscopique de la dynamique d'un système, les propriétés de symétrie exprimées par les relations d'Onsager.

On suppose que la configuration microscopique d'un système est donnée par N variables réelles $x_1, \dots, x_N$. On notera $\underline{x} = (x_1, \dots, x_N)$ la configuration globale du système, et $H(\underline{x})$ son énergie. Au contact d'un thermostat, sa configuration évolue selon les équations de Langevin

$$\frac{dx_i}{dt} = -\frac{\partial H}{\partial x_i}(x_1(t), \dots, x_N(t)) + \xi_i(t) \quad \text{pour } i = 1, \dots, N, \quad (8)$$

où les $\xi_i(t)$ sont des fonctions aléatoires modélisant le couplage au thermostat, avec

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

On admettra alors que la probabilité $P(\underline{x}, t)$ d'observer le système dans la configuration $\underline{x}$ à l'instant t évolue selon l'équation de Fokker-Planck

$$\frac{\partial P}{\partial t} = WP, \quad (10)$$

où l'action de l'opérateur de Fokker-Planck W sur les fonctions $f(\underline{x})$ de l'espace des configurations est défini par

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

1. Interpréter qualitativement les équations de Langevin en l'absence des termes aléatoires ξ_i .
2. Donner une condition suffisante sur la valeur de Γ pour que la distribution de Gibbs-Boltzmann $P_{\text{eq}}(\underline{x}) = e^{-\beta H(\underline{x})}/Z$ soit stationnaire. On supposera cette condition vérifiée dans toute la suite.

On suppose qu'à $t = 0$ la condition initiale $\underline{x}(t = 0)$ est tirée aléatoirement selon une loi de probabilité $P_0(\underline{x})$, et que le système évolue pour $t \geq 0$ selon la modélisation expliquée ci-dessus. On note $\langle \bullet \rangle$ les moyennes sur la condition initiale et sur l'évolution aléatoire ultérieure. Pour deux observables A, B (i.e. deux fonctions $A(\underline{x}), B(\underline{x})$) on note $\langle A(t) \rangle = \langle A(\underline{x}(t)) \rangle$ la moyenne au temps t et $\langle A(t)B(t') \rangle = \langle A(\underline{x}(t))B(\underline{x}(t')) \rangle$ leur fonction de corrélation à deux temps.

3. On note $Q_{\underline{x}, \underline{x}'}(\tau)$ la probabilité que le système soit dans la configuration $\underline{x}$ à un certain instant t , sachant qu'il était dans la configuration $\underline{x}'$ au temps $t - \tau$. Exprimer Q en fonction de W (on traitera formellement les opérateurs comme des matrices).
4. Donner une expression de $\langle A(t) \rangle$ en fonction de P_0 , Q et de la fonction $A(\underline{x})$.
5. Exprimer de même $\langle A(t)B(t') \rangle$, en supposant $t > t'$.
6. Simplifiez vos réponses aux deux dernières questions en supposant le système à l'équilibre au temps initial, i.e. avec $P_0 = P_{\text{eq}}$.
7. En admettant que

$$Q_{\underline{x}, \underline{x}}(\tau) P_{\text{eq}}(\underline{x}) = Q_{\underline{x}, \underline{x}'}(\tau) P_{\text{eq}}(\underline{x}'), \quad (12)$$

montrer que pour un système à l'équilibre, $\langle A(t)B(t') \rangle = \langle B(t)A(t') \rangle$.

8. Rappeler brièvement les conséquences de cette invariance par renversement temporel dans le cadre de la thermodynamique des systèmes faiblement hors d'équilibre.
9. Question subsidiaire : démontrer la relation (12).

On pourra commencer par démontrer que $WP_{\text{eq}} = P_{\text{eq}}W^\dagger$, où P_{eq} est l'opérateur (diagonal) multiplication par $P_{\text{eq}}(\underline{x})$ et $\dagger$ dénote la transposée. On notera que

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