

CORRIGÉ DE L'EXAMEN

I L'information mutuelle

1.

$$S_{Y|X} = - \sum_x p(x) \sum_y p(y|x) \log \frac{p(x,y)}{p(x)} = - \sum_{x,y} p(x,y) \log p(x,y) + \sum_x p(x) \log p(x) = S_{X,Y} - S_X .$$

2. $S_{X,Y} = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{4} \log_2 \frac{1}{4} - 2 \times \frac{1}{8} \log_2 \frac{1}{8} = \frac{7}{4} = 1,75$.

La loi marginale $p(x)$ de la première pièce est $p(X = P) = \frac{3}{4}$, $p(X = F) = \frac{1}{4}$, son entropie est donc $S_X = -\frac{3}{4} \log_2 \frac{3}{4} - \frac{1}{4} \log_2 \frac{1}{4} = 2 - \frac{3}{4} \log_2 3 \approx 0,8$.

Les lois de Y conditionnée par X sont d'une part $p(Y = P|X = P) = \frac{2}{3}$, $p(Y = F|X = P) = \frac{1}{3}$, l'entropie de Y quand $X = P$ est donc $-\frac{1}{3} \log_2 \frac{1}{3} - \frac{2}{3} \log_2 \frac{2}{3} = \log_2(3) - \frac{2}{3}$,

d'autre part $p(Y = P|X = F) = p(Y = F|X = F) = \frac{1}{2}$, l'entropie de Y quand $X = F$ est donc 1 bit. En combinant ces deux cas $S_{Y|X} = \frac{3}{4} (\log_2(3) - \frac{2}{3}) + \frac{1}{4} = \frac{3}{4} \log_2(3) - \frac{1}{4} \approx 0,95$.

On a bien $S_{X,Y} = S_X + S_{Y|X}$.

3. Si les variables sont indépendantes $p(y|x) = p(y)$ indépendamment de la valeur de x , on a donc $S_{Y|X} = S_Y$.

4.

$$S_Y - S_{Y|X} = \sum_{x,y} p(x,y) \log \left(\frac{p(y|x)}{p(y)} \right) = \sum_{x,y} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right) . \quad (1)$$

On reconnaît la distance de Kullback-Leibler entre $p(x,y)$ et la loi factorisée $p(x)p(y)$, cette quantité est donc positive, par concavité du logarithme. L'inégalité $S_{Y|X} \leq S_Y$ traduit le fait que le conditionnement diminue l'entropie, en effet en conditionnant selon la valeur de X on gagne en général de l'information sur Y .

5. D'après la réponse à la question 1, $I_{X,Y} = S_Y - S_{Y|X}$; par symétrie entre X et Y on peut aussi écrire $I_{X,Y} = S_X - S_{X|Y}$. Cette quantité est donc positive (cf. la question 4), et mesure le gain d'information (la diminution d'entropie) moyenne que la donnée de X procure sur la valeur de Y (ou réciproquement). En d'autres termes elle quantifie les corrélations entre X et Y . Elle est bien nulle si les deux variables sont indépendantes (cf. question 3).

6. Dégradation de l'information

(a) Conditionnée par la valeur de $X = x$ la variable aléatoire $f(X)$ est déterministe, elle vaut nécessairement $f(x)$, on a donc $S_{f(X)|X} = 0$. En tant qu'entropie on a toujours $S_{X|f(X)} \geq 0$. On peut écrire $S_{X,f(X)} = S_X + S_{f(X)|X} = S_X = S_{f(X)} + S_{X|f(X)}$, ce qui prouve l'inégalité. L'égalité est vraie si et seulement si f est bijective, dans ce cas-là X et $f(X)$ ont le même contenu informatif. Si au contraire $f(x) = f_0$ pour plusieurs valeurs de x , la loi $p(x|f(x) = f_0)$ est uniforme sur les antécédents de f_0 , son entropie est donc strictement positive.

(b) On s'attend à $I_{X,Y} \geq I_{X,f(Y)}$: $f(Y)$ ne contient pas plus d'information sur X que Y , et en contient strictement moins si f n'est pas bijective.

(c) En soustrayant les deux derniers membres de la règle de composition avec $X_1 = X$, $X_2 = Y$, et $X_3 = f(Y)$, il vient $I_{X,Y} - I_{X,f(Y)} = I_{X,Y|f(Y)} - I_{X,f(Y)|Y}$. En transformant cette expression avec la définition de l'information mutuelle conditionnée on obtient $I_{X,Y} - I_{X,f(Y)} = S_{Y|f(Y)} - S_{Y|(X,f(Y))} - S_{f(Y)|Y} + S_{f(Y)|(X,Y)}$. Les deux derniers termes de cette expression sont

nuls : connaissant Y (*a fortiori* Y et X) la variable aléatoire $f(Y)$ est déterministe. Par ailleurs $S_{Y|f(Y)} \geq S_{Y|(X,f(Y))}$: augmenter le conditionnement (ici en ajoutant la connaissance de X) diminue l'entropie, selon la réponse à la question 4.

7. Propagation de l'information

- (a) Comme les deux valeurs possibles de X_0 sont choisies avec la même probabilité, et que les erreurs de transmission sont les mêmes pour les deux symboles, tout le processus est symétrique sous l'échange $P \leftrightarrow F$, notamment $p(X_n = P) = p(X_n = F) = 1/2$. Les entropies des lois marginales sont donc $S_{X_0} = S_{X_n} = 1$ bit, et $I_n = 2 - S_{X_0, X_n}$. La loi jointe de X_0, X_n contient quatre événements de probabilité respectives $\frac{p_n}{2}, \frac{p_n}{2}, \frac{1-p_n}{2}, \frac{1-p_n}{2}$, son entropie est donc $S_{X_0, X_n} = 1 + H(p_n)$, d'où finalement $I_n = 1 - H(p_n)$.
- (b) $X_n \neq X_0$ si $X_{n-1} \neq X_0$ et $X_n = X_{n-1}$, ou si $X_{n-1} = X_0$ et $X_n \neq X_{n-1}$, soit $p_n = p_{n-1}(1-p) + (1-p_{n-1})p$. On doit donc résoudre l'équation de récurrence $p_n = (1-2p)p_{n-1} + p$, avec la condition aux bords $p_1 = p$. En cherchant la solution sous la forme $p_n = \alpha(1-2p)^n + \beta$, on détermine les constantes α, β et on obtient

$$p_n = \frac{1 - (1-2p)^n}{2} . \quad (2)$$

Dès que $p > 0$, on a $p_n \rightarrow \frac{1}{2}$ pour n grand, comme si les deux pièces étaient tirées indépendamment : les erreurs s'accumulent et font disparaître les corrélations entre X_0 et X_n .

- (c) Le développement de l'entropie $H(x)$ autour de son maximum $1/2$ donne $H\left(\frac{1+\epsilon}{2}\right) \sim 1 - \frac{\epsilon^2}{2 \log 2}$, soit $I_n \sim \frac{1}{2 \log 2}(1-2p)^{2n}$.
- (d) Les configurations sont tirées avec la loi de probabilité de Gibbs-Boltzmann $e^{-\beta H}/Z$; l'Hamiltonien respecte la symétrie d'inversion globale $\sigma_i \leftrightarrow -\sigma_i$, donc toutes les lois marginales sont triviales, $p(\sigma_i = +1) = p(\sigma_i = -1) = 1/2$. Par ailleurs les lois conditionnelles pour deux spins voisins sont

$$p(\sigma_i | \sigma_{i-1}) = \frac{e^{\beta J \sigma_i \sigma_{i-1}}}{2 \cosh \beta J} . \quad (3)$$

La probabilité que $\sigma_i \neq \sigma_{i-1}$ est donc $(1 - \tanh(\beta J))/2$, ce qui correspond au paramètre p dans la description précédente. p_n se comporte, à des constantes près, comme $(1-2p)^n = (\tanh(\beta J))^n$, ce qui est bien le comportement de la fonction de corrélation du modèle d'Ising unidimensionnel. La longueur de corrélation est $-1/\log(1-2p) = -1/\log(\tanh(\beta J))$.

8. En appliquant deux fois la relation de Bayes on peut écrire

$$p(x_1, x_2, x_3) = p(x_1)p(x_2, x_3|x_1) = p(x_1)p(x_2|x_1)p(x_3|x_1, x_2) .$$

L'entropie de la loi jointe est donc

$$S_{X_1, X_2, X_3} = - \sum_{x_1, x_2, x_3} p(x_1, x_2, x_3) \log(p(x_1)p(x_2|x_1)p(x_3|x_1, x_2)) = S_{X_1} + S_{X_2|X_1} + S_{X_3|(X_1, X_2)} . \quad (4)$$

On en déduit donc

$$I_{X_1, (X_2, X_3)} = S_{X_1} + S_{X_2, X_3} - S_{X_1, X_2, X_3} = S_{X_2, X_3} - S_{X_2|X_1} - S_{X_3|(X_1, X_2)} . \quad (5)$$

Par définition

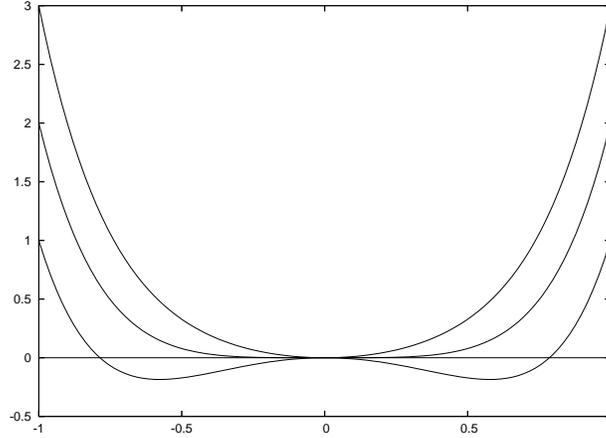
$$I_{X_1, X_2} + I_{X_1, X_3|X_2} = S_{X_2} - S_{X_2|X_1} + S_{X_3|X_2} - S_{X_3|(X_1, X_2)} . \quad (6)$$

En remplaçant dans cette dernière égalité $S_{X_2} + S_{X_3|X_2}$ par S_{X_2, X_3} on obtient bien $I_{X_1, (X_2, X_3)} = I_{X_1, X_2} + I_{X_1, X_3|X_2}$. L'autre égalité est équivalente en remplaçant $X_2 \leftrightarrow X_3$.

II Théorie de Landau et tricriticalité

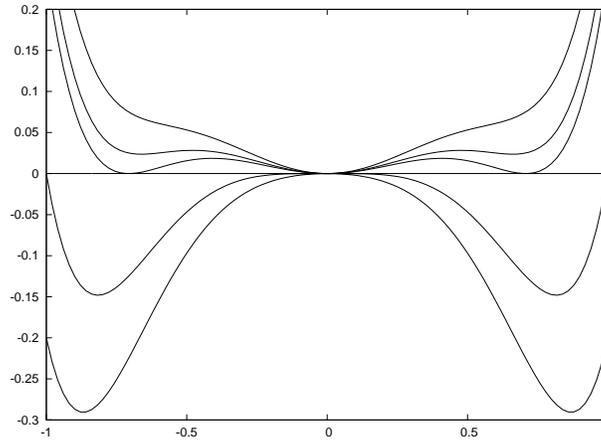
A Ordre des transitions et signe des coefficients

1. Ce développement s'appliquerait par exemple à un système de spins qui respecte la symétrie $\phi \leftrightarrow -\phi$.
2. On doit avoir $a_6 > 0$, sinon le système est instable, le minimum de l'énergie libre étant atteint pour $\phi \rightarrow \pm\infty$ si $a_6 < 0$.
3. Pour $a_4 > 0$ la transition est du deuxième ordre, l'allure des courbes $R(\phi)$ est :



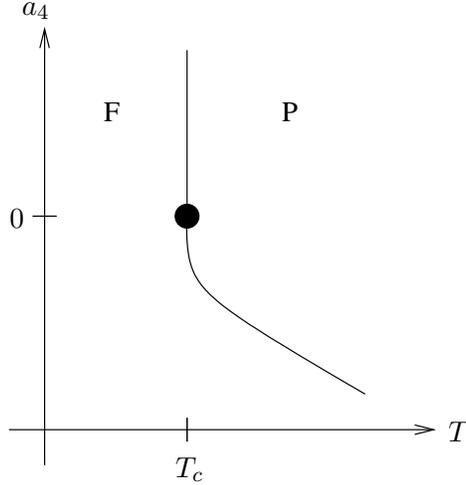
où de haut en bas $T > T_c$, $T = T_c$, $T < T_c$.

Pour $a_4 < 0$ la transition est du premier ordre, l'allure des courbes $R(\phi)$ est :



où de haut en bas : deux valeurs de $T > T^* > T_c$, $T = T^*$, $T = T_c$, $T < T_c$.

4. A la température de transition T^* il existe des minima locaux en $\pm\phi^*$, avec $\phi^* > 0$, qui ont la même énergie libre que le minimum en $\phi = 0$. Les deux inconnues T^* et ϕ^* sont donc fixées par les deux équations $\mathcal{R}(\phi^*, T^*) = \frac{\partial \mathcal{R}}{\partial \phi}(\phi^*, T^*) = 0$. En résolvant ce système il vient la forme de l'énoncé avec $\alpha = 3/16$, et $\phi^* = \sqrt{\frac{-3a_4}{4a_6}}$.
5. Pour $a_4 > 0$ la transition entre les phases paramagnétique (P) et ferromagnétique (F) est du deuxième ordre et se produit sur la ligne $T = T_c$; pour $a_4 < 0$ la température de transition du premier ordre T_* dépend de a_4 :



Ces lignes se rencontrent au point tricritique • de coordonnées $(T = T_c, a_4 = 0)$.

B Étude du point tricritique

1. La transition est du deuxième ordre, le paramètre d'ordre croît continument en fonction de la température à la transition, l'énergie libre et ses dérivées premières sont continues.
2. Les solutions de $\mathcal{R}'(\phi) = 0$ sont $\phi = (\tilde{a}_2(T_c - T)/a_6)^{1/4}$, i.e. $\beta = 1/4$ au lieu de $1/2$ pour $a_4 > 0$.
3. Le champ h apparaît sous la forme de $-h\phi$ dans l'énergie libre. En minimisant $a_6\phi^6/6 - h\phi$ on obtient $\delta = 5$, au lieu de $\delta = 3$ pour $a_4 > 0$.

C Dimension critique supérieure pour le point tricritique

1. Quand $T \neq T_c$ la longueur de corrélation ξ est finie, $\Gamma(r) \sim e^{-r/\xi}$ quand $r \rightarrow \infty$. A $T = T_c$ ξ est infinie, la décroissance de r est plus lente, $\Gamma(r) \sim 1/r^{d-2}$.
2. Notons $M = \int_{\xi^d} \phi(\vec{r}) d\vec{r}$ l'aimantation totale dans un volume de corrélation. Sa moyenne est $\langle M \rangle = \xi^d \phi_{\text{eq}}$, et ses fluctuations sont données par $\langle M^2 \rangle - \langle M \rangle^2 = \xi^d \int_{\xi^d} \Gamma(\vec{r}) d\vec{r}$. Le rapport de Ginzburg peut s'exprimer comme $R_G = (\langle M^2 \rangle - \langle M \rangle^2) / \langle M \rangle^2$, il permet de vérifier si la description de champ moyen qui néglige les fluctuations est cohérente, il doit alors être petit.
3. En notant $\epsilon = T_c - T$ l'écart à la température critique, on a $\phi_{\text{eq}} \propto \epsilon^\beta$ et $\xi \propto \epsilon^{-\nu}$. Ainsi,

$$R_G \propto \epsilon^{-2\beta} \epsilon^{d\nu} \int_0^\xi \frac{r^{d-1} dr}{\xi^{d-2}} F\left(\frac{r}{\xi}\right) \propto \epsilon^{-2\beta} \epsilon^{d\nu} \xi^2 \int_0^1 du F(u) \propto \epsilon^{(d-2)\nu-2\beta}. \quad (7)$$

La dimension critique supérieure d_{sup} est telle que R_G ne diverge pas quand $\epsilon \rightarrow 0$ pour $d \geq d_{\text{sup}}$, soit $d_{\text{sup}} = 2 + 2\frac{\beta}{\nu}$. Au point tricritique $\beta = 1/4$, $\nu = 1/2$, soit $d_{\text{sup}} = 3$, alors que pour $a_4 > 0$ on retrouve la valeur $d_{\text{sup}} = 4$ puisque $\beta = \nu = 1/2$ dans ce cas-là.

4. On ajoute un terme de champ magnétique et un terme en gradient qui pénalise les inhomogénéités du champ,

$$\mathcal{R} = \int d\vec{r} \left[\frac{1}{2} a_2 \phi(\vec{r})^2 + \frac{1}{4} a_4 \phi(\vec{r})^4 + \frac{1}{6} a_6 \phi(\vec{r})^6 - h(\vec{r}) \phi(\vec{r}) + \lambda (\vec{\nabla} \phi)^2 \right]. \quad (8)$$

Dans la suite on reprendra $a_4 = 0$.

5. Pour un choix de $h(\vec{r})$ donné le paramètre d'ordre $\phi(\vec{r})$ minimise la fonctionnelle d'énergie libre, soit

$$0 = a_2 \phi(\vec{r}) + a_6 \phi(\vec{r})^5 - h(\vec{r}) - \lambda \Delta \phi. \quad (9)$$

La fonction de corrélation à deux points est donnée par $\Gamma(\vec{r}, \vec{r}') = \frac{1}{\beta} \frac{\delta \phi(\vec{r})}{\delta h(\vec{r}')}$, soit

$$0 = (a_2 + 5a_6 \phi(\vec{r})^4) \beta \Gamma(\vec{r}, \vec{r}') - \delta(\vec{r} - \vec{r}') - \lambda \Delta (\beta \Gamma). \quad (10)$$

Pour un système homogène, Γ ne dépend que de $\vec{r}-\vec{r}'$ que l'on notera \vec{r} , et $\phi(\vec{r})$ devient ϕ_{eq} . L'équation est alors

$$\beta(-\lambda\Delta + a_2 + 5a_6\phi_{\text{eq}}^4)\Gamma(\vec{r}) = \delta(\vec{r}) . \quad (11)$$

6. La transformée de Fourier de Γ est donc

$$\Gamma(\vec{q}) = \frac{1}{\beta(\lambda q^2 + a_2 + 5a_6\phi_{\text{eq}}^4)} = \frac{1}{\beta\lambda} \frac{1}{q^2 + \xi^{-2}} , \quad (12)$$

avec la longueur de corrélation ξ donnée par

$$\xi = \sqrt{\frac{\lambda}{a_2 + 5a_6\phi_{\text{eq}}^4}} \propto (T_c - T)^{-1/2} , \quad (13)$$

ce qui confirme l'exposant $\nu = 1/2$.

7. La tension de surface s'annule à T_c . Plus précisément, un profil unidimensionnel $\phi(z)$ passant de $-\phi_{\text{eq}}$ à $+\phi_{\text{eq}}$ sur une longueur d'ordre ξ contribue à l'énergie libre par unité de surface transverse à z selon

$$\lambda \int dz \phi'(z)^2 \propto \xi \left(\frac{\phi_{\text{eq}}}{\xi} \right)^2 \propto \epsilon^{2\beta+\nu} . \quad (14)$$

Pour le point tricritique $\beta = 1/4$, $\nu = 1/2$, l'exposant d'annulation de la tension de surface est donc 1 ; pour $a_4 > 0$ on retrouve la valeur 3/2 obtenue en TD.

III Une approche microscopique aux relations d'Onsager

1. En l'absence des termes aléatoires le système évolue dans la direction opposée au gradient de H , il cherche donc à minimiser son énergie.
2. P_{eq} sera un état stationnaire si $WP_{\text{eq}} = 0$. En notant que $\frac{\partial P_{\text{eq}}}{\partial x_i} = -\beta \frac{\partial H}{\partial x_i} P_{\text{eq}}$, on voit que c'est bien le cas si $\beta \frac{\Gamma}{2} = 1$, autrement dit si $\Gamma = 2k_B T$. Remarquons que si la température du thermostat tend vers 0 on retrouve la situation de la question précédente : tout le transfert d'énergie se produit alors du système vers le thermostat.
3. A \underline{x}' fixé l'évolution de la probabilité d'être en \underline{x} obéit à l'équation de Fokker-Planck, soit $\frac{dQ}{d\tau} = WQ$. Par ailleurs la condition initiale est $Q_{\underline{x},\underline{x}'}(\tau = 0) = \delta(\underline{x} - \underline{x}')$, la solution de l'équation différentielle est donc $Q = \exp[\tau W]$.

4.

$$\langle A(t) \rangle = \int d\underline{x} A(\underline{x}) P(\underline{x}, t) = \int d\underline{x}_1 d\underline{x}_0 A(\underline{x}_1) Q_{\underline{x}_1, \underline{x}_0}(t) P_0(\underline{x}_0) . \quad (15)$$

5.

$$\langle A(t)B(t') \rangle = \int d\underline{x}_2 d\underline{x}_1 d\underline{x}_0 A(\underline{x}_2) B(\underline{x}_1) Q_{\underline{x}_2, \underline{x}_1}(t - t') Q_{\underline{x}_1, \underline{x}_0}(t') P_0(\underline{x}_0) . \quad (16)$$

6. La distribution P_{eq} étant stationnaire, on a $\int d\underline{x}_0 Q_{\underline{x}, \underline{x}_0}(\tau) P_{\text{eq}}(\underline{x}_0) = P_{\text{eq}}(\underline{x})$. Les formules précédentes se simplifient en

$$\langle A(t) \rangle = \int d\underline{x} A(\underline{x}) P_{\text{eq}}(\underline{x}) , \quad (17)$$

i.e. à tous les temps les moyennes des observables sont celles de l'équilibre canonique. De plus

$$\langle A(t)B(t') \rangle = \int d\underline{x}_2 d\underline{x}_1 A(\underline{x}_2) B(\underline{x}_1) Q_{\underline{x}_2, \underline{x}_1}(t - t') P_{\text{eq}}(\underline{x}_1) \quad (18)$$

ne dépend que de la différence $t - t'$ entre les deux temps d'observation du système.

7. En transformant cette dernière expression avec l'identité de l'énoncé on obtient

$$\langle A(t)B(t') \rangle = \int d\underline{x}_2 d\underline{x}_1 A(\underline{x}_2) B(\underline{x}_1) Q_{\underline{x}_1, \underline{x}_2}(t-t') P_{\text{eq}}(\underline{x}_2) \quad (19)$$

$$= \int d\underline{x}_2 d\underline{x}_1 A(\underline{x}_1) B(\underline{x}_2) Q_{\underline{x}_2, \underline{x}_1}(t-t') P_{\text{eq}}(\underline{x}_1) = \langle B(t)A(t') \rangle, \quad (20)$$

en échangeant les indices muets de la première à la deuxième ligne.

8. Cette invariance par renversement temporel est l'étape cruciale dans la preuve de la symétrie de la matrice des coefficients cinétiques, dans le cadre de la théorie de la réponse linéaire.

9. En terme de multiplication d'opérateurs,

$$\frac{\partial}{\partial x_i} P_{\text{eq}} = P_{\text{eq}} \frac{\partial}{\partial x_i} - \beta \left(\frac{\partial H}{\partial x_i} \right) P_{\text{eq}}, \quad (21)$$

et

$$\frac{\partial^2}{\partial x_i^2} P_{\text{eq}} = P_{\text{eq}} \frac{\partial^2}{\partial x_i^2} - \beta \left(\frac{\partial H}{\partial x_i} \right) P_{\text{eq}} \frac{\partial}{\partial x_i} - \beta \left(\frac{\partial^2 H}{\partial x_i^2} \right) P_{\text{eq}} + \beta^2 \left(\frac{\partial H}{\partial x_i} \right)^2 P_{\text{eq}} - \beta \left(\frac{\partial H}{\partial x_i} \right) P_{\text{eq}} \frac{\partial}{\partial x_i}. \quad (22)$$

En reportant dans l'expression de W et compte-tenu de $\beta\Gamma = 2$, il vient

$$WP_{\text{eq}} = P_{\text{eq}} \sum_{i=1}^N \left[- \left(\frac{\partial H}{\partial x_i} \right) \frac{\partial}{\partial x_i} + \frac{\Gamma}{2} \frac{\partial^2}{\partial x_i^2} \right]. \quad (23)$$

De plus

$$\left[\left(\frac{\partial H}{\partial x_i} \right) \frac{\partial}{\partial x_i} \right]^\dagger = \frac{\partial}{\partial x_i} \dagger \left(\frac{\partial H}{\partial x_i} \right) = - \frac{\partial}{\partial x_i} \left(\frac{\partial H}{\partial x_i} \right) = - \left(\frac{\partial^2 H}{\partial x_i^2} \right) - \left(\frac{\partial H}{\partial x_i} \right) \frac{\partial}{\partial x_i}, \quad (24)$$

ce qui permet de montrer $WP_{\text{eq}} = P_{\text{eq}}W^\dagger$. Comme $Q = \exp[\tau W]$, ceci implique $QP_{\text{eq}} = P_{\text{eq}}Q^\dagger$, qui était l'identité utilisée à la question 7.