EXERCISES 8

Excercise 1. Using the data-processing inequality of the relative entropy, show the following statements:

(1) Show that the relative entropy is joint convex, i.e., that for quantum states $\rho_1, \rho_2, \sigma_1, \sigma_2 \in D(\mathcal{H})$ and any $\lambda \in [0,1]$, we have

$$D((1-\lambda)\rho_1 + \lambda\rho_2 \| (1-\lambda)\sigma_1 + \lambda\sigma_2) \leqslant (1-\lambda)D(\rho_1\|\sigma_1) + \lambda D(\rho_2\|\sigma_2).$$

(2) Show that for quantum states $\rho, \sigma \in D(\mathcal{H})$, we have

$$D(\rho\|\sigma) \geqslant 0.$$

This is sometimes called *Klein's inequality*.

Excercise 2. In this exercise, we will prove three important properties of the von Neumann entropy. They are called *concavity*, *subadditivity* and *strong subadditivity*:

(1) Show that for any $\rho, \sigma \in D(H)$ and any $\lambda \in [0,1]$ we have

$$H((1-\lambda)\rho + \lambda\sigma) \geqslant (1-\lambda)H(\rho) + \lambda H(\sigma).$$

Hint: Express the von Neumann entropy using the relative entropy.

(2) For a quantum state ρ_{AB} show that

$$H(\rho_{AB}) \leqslant H(\rho_A) + H(\rho_B),$$

where $\rho_A = \mathrm{Tr}_B[\rho_{AB}]$ and $\rho_B = \mathrm{Tr}_A[\rho_{AB}]$.
Hint: Compute the relative entropy

$$D(\rho_{AB}\|\rho_A \otimes \rho_B).$$

(3) For a quantum state $\rho_{ABC} \in D(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ show that

$$H(\rho_{ABC}) + H(\rho_B) \leqslant H(\rho_{AB}) + H(\rho_{BC}),$$

where $\rho_{AB} = \mathrm{Tr}_C[\rho_{ABC}]$, $\rho_{BC} = \mathrm{Tr}_A[\rho_{ABC}]$, and $\rho_B = \mathrm{Tr}_{AC}[\rho_{ABC}]$.
Hint: Compute the relative entropies

$$D(\rho_{ABC}\|\frac{\mathbb{1}_{\mathcal{H}_A}}{\dim(\mathcal{H}_A)} \otimes \rho_{BC}) \quad \text{and} \quad D(\rho_{AB}\|\frac{\mathbb{1}_{\mathcal{H}_A}}{\dim(\mathcal{H}_A)} \otimes \rho_B),$$

and then think about how to apply the data-processing inequality.

Excercise 3 (Conditional entropy)**.** For $\rho_{AB} \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ define the *conditional entropy* as

$$H(A|B)_{\rho_{AB}} = H(\rho_{AB}) - H(\rho_B),$$

For quantum states $\rho_{ABC} \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ we use the notation $H(A|B,C)_{\rho_{ABC}}$ to refer to the conditional entropy where the conditioning is with respect to the joint system 'BC' and we use the notation $H(A|B)_{\rho_{ABC}}$ instead of $H(A|B)_{\rho_{AB}}$ where $\rho_{AB} = \mathrm{Tr}_C[\rho_{ABC}]$. This notation extends to any number of systems in the obvious way.

(1) Show that in the classical case, we have

$$H(A|B)_{p_{AB}} = H(p_{AB}) - H(p_B) \geqslant 0$$

for any joint probability distribution $p_{AB} \in \mathcal{P}(\Sigma_A \times \Sigma_B)$.
Hint: Use the product rule for joint probability distributions.

(2) Find a quantum state $\rho_{AB} \in D(\mathbb{C}^d \otimes \mathbb{C}^d)$ for which

$$H(A|B)_{\rho_{AB}} < 0.$$

(3) Show that

$$H(A|B,C)_{\rho_{ABC}} \leqslant H(A|B)_{\rho_{ABC}},$$

i.e., conditioning does not increase entropy.

(4) For $t \in \mathbb{N}$ consider a family $\mathcal{F} = \{F_1, \ldots, F_K\}$ of subsets of $\{1, 2, \ldots, n\}$ such that

$$|\{k \in \{1, \ldots, K\} \ : \ i \in F_k\}| = t$$

for any $i \in \{1, 2, \ldots, n\}$, i.e., every $i \in \{1, 2, \ldots, n\}$ is contained in exactly t of the sets in $\mathcal{F}$. Show that

$$H(\rho_{1,2,\ldots,n}) \leqslant \frac{1}{t} \sum_{F \in \mathcal{F}} H(\rho_F),$$

for any $\rho_{1,2,\ldots,n} \in D(\mathcal{H}^{\otimes n})$ and where ρ_F denotes the reduction of $\rho_{1,2,\ldots,n}$ to the systems with labels in F, e.g., $\rho_{1,2} = \mathrm{Tr}_{3,4,\ldots,n}\left[\rho_{1,2,\ldots,n}\right]$.

Hint: For a subset $F \subseteq \{1, 2, \ldots, n\}$ denote its elements by $(i_1, \ldots, i_l)$ ordered increasingly. Start by showing that

$$\sum_{j=1}^{|F|} H(i_j | i_1, \ldots, i_{j-1})_{\rho_{1,2,\ldots,n}} = H(\rho_F).$$

Use this in the proof together with the result from the previous subexercise.

Excercise 4 (Fannes' inequality)**.** For any $\rho, \sigma \in D(\mathcal{H})$ satisfying $\|\rho - \sigma\|_1 \leqslant 1/e$, we will show that

$$|H(\rho) - H(\sigma)| \leqslant \|\rho - \sigma\|_1 \log(d) + \eta\left(\|\rho - \sigma\|_1\right),$$

where $\eta(x) = -x \log(x)$ and $d = \dim(\mathcal{H})$. Consider the spectral decompositions

$$\rho = \sum_{i=1}^{d} \lambda_i |v_i\rangle\langle v_i| \quad \text{and} \quad \sigma = \sum_{i=1}^{d} \mu_i |w_i\rangle\langle w_i|,$$

such that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$ and $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_d$ and follow the steps below:

(1) Start by writing $\rho - \sigma = P - Q$ for $P, Q \in B(\mathcal{H})^+$ satisfying $PQ = 0$ and define an operator $X = P + \rho = Q + \sigma$. If $\tau_1 \geqslant \tau_2 \geqslant \ldots \geqslant \tau_d$ are the eigenvalues of X show that $2\tau_i \geqslant \lambda_i + \mu_i + |\lambda_i - \mu_i|$ and use this to show that

$$\|\rho - \sigma\|_1 \geqslant \sum_{i=1}^{d} |\lambda_i - \mu_i|.$$

(2) Show that $|\eta(\lambda_i) - \eta(\mu_i)| \leqslant \eta(|\lambda_i - \mu_i|)$ for all $i \in \{1, \ldots, d\}$ and conclude that

$$|H(\rho) - H(\sigma)| \leqslant \sum_{i=1}^{d} \eta\left(|\lambda_i - \mu_i|\right).$$

(3) Finally, show that $\eta(|\lambda_i - \mu_i|) = S\eta\left(|\lambda_i - \mu_i|/S\right) - |\lambda_i - \mu_i| \log(S)$ for $S = \sum_{i=1}^{d} |\lambda_i - \mu_i|$. Use this observation to prove

$$|H(\rho) - H(\sigma)| \leqslant \|\rho - \sigma\|_1 \log(d) + \eta\left(\|\rho - \sigma\|_1\right),$$

whenever $\|\rho - \sigma\|_1 \leqslant 1/e$.