EXERCISES ON LDS

ANTONIO AGRESTI

1. Exercise 1.

Prove that

$$\lim_{b\to\infty} \overline{\lim_{n\to\infty}} \frac{1}{r_n} \log\left(\int_{f\geq b} e^{r_n f}\, d\mu_n\right) = -\infty\,,$$

is satisfied if there exists $\alpha > 1$, such that

$$C := \sup_n \left(\int e^{\alpha r_n f}\, d\mu_n\right)^{\frac{1}{r_n}} < \infty\,. \tag{1.1}$$

Solution: Using Hölder inequality with exponent $\alpha > 1$ and $\alpha' := \alpha/(\alpha-1) \in (1,\infty)$, we obtain

$$\int_{f\geq b} e^{r_n f}\, d\mu_n \leq \left(\int_{f\geq b} e^{\alpha r_n f}\, d\mu_n\right)^{\frac{1}{\alpha}} (\mu_n(f\geq b))^{\frac{1}{\alpha'}}\,.$$

Applying $1/r_n \log(\cdot)$ to both sides, since the map $(0,\infty) \ni t \mapsto \log t$ maintains the monotonicity, we obtain

$$\begin{aligned}
\frac{1}{r_n}\log \int_{f\geq b} e^{r_n f}\, d\mu_n &\leq \frac{1}{r_n}\log\left[\left(\int_{f\geq b} e^{\alpha r_n f}\, d\mu_n\right)^{\frac{1}{\alpha}}\right] + \frac{1}{r_n}\log\left[(\mu_n(f\geq b))^{\frac{1}{\alpha'}}\right]\\
&= \frac{1}{\alpha}\log\left[\left(\int_{f\geq b} e^{\alpha r_n f}\, d\mu_n\right)^{\frac{1}{r_n}}\right] + \frac{1}{\alpha'}\log\left[(\mu_n(f\geq b))^{\frac{1}{r_n}}\right]\\
&\leq \frac{1}{\alpha}\log C + \frac{1}{\alpha'}\log\left[(\mu_n(f\geq b))^{\frac{1}{r_n}}\right]\,;
\end{aligned}$$

in the last inequality we have used the bound (1.1). Now, observe that (sometimes called Markov type inequalities)

$$\int_{f\geq b} e^{\alpha r_n f}\, d\mu_n \geq e^{\alpha r_n b}\int_{f\geq b} d\mu_n = e^{\alpha r_n b}\mu(f\geq b)\,.$$

This implies

$$\begin{aligned}
\log\left[(\mu(f\geq b))^{\frac{1}{r_n}}\right] &\leq \log\left[e^{-\alpha b}\left(\int_{f\geq b} e^{\alpha r_n f}\, d\mu_n\right)^{\frac{1}{r_n}}\right]\\
&= -\alpha b + \log\left[\left(\int_{f\geq b} e^{\alpha r_n f}\, d\mu_n\right)^{\frac{1}{r_n}}\right]\\
&\leq -\alpha b + \log C\,;
\end{aligned}$$

another time by (1.1). The previous estimates imply

$$\overline{\lim_{n\to\infty}}\frac{1}{r_n}\log\left(\int_{f\geq b} e^{r_n f}d\mu_n\right)\leq \frac{1}{\alpha}\log C+\frac{1}{\alpha'}(-\alpha b+\log C)$$
$$=\log C-\frac{\alpha}{\alpha'}b\to-\infty\,,$$

for $b\to\infty$.

2. Exercise 2.

Consider the following game. Let S_n be the number of tails you get when you flip n times a fair coin. Determine the asymptotic of $\mathbb{E}[3^{S_n}]$.

Solution: Observe that

$$\mathbb{E}[3^{S_n}]=\mathbb{E}[e^{S_n\log 3}]=\int_\Omega e^{n\frac{S_n}{n}\log 3}\,d\mathbb{P}$$
$$=\int_{\mathbb{R}} e^{nt\log 3}d\mu_n(t)\,.$$

Here $(\Omega,\mathfrak{F},\mathbb{P})$ is the probability space and μ_n is the law of S_n/n. Since $S_n/n\in\{0,1/n,\dots,1\}$ then the support of μ_n is contained in $[0,1]\subset\mathbb{R}$ for each $n\in\mathbb{N}$. More precisely, we have

$$\mu_n=\frac{1}{2^n}\sum_{j=0}^n\binom{n}{j}\delta_{\frac{j}{n}}\,,$$

since $\mathbb{P}(S_n/n=j/n)=\binom{n}{j}(1/2)^n$ for $j=0,\dots,n$. Furthermore, define $\chi=\mathbb{R}$ and

$$f(x):=x\log 3\,,\qquad x\in\mathbb{R}\,.$$

Then the previous identity becomes

$$\mathbb{E}[3^{S_n}]=\int_\chi e^{nf(t)}d\mu_n(t)\,.$$

Since f is not bounded by above, we have to use the extended version of Varadhan's Lemma as quoted in the exercises.
We may use Exercise 1 and shows that, for some $\alpha>1$, we have

$$C:=\sup_n\left(\int_\chi e^{n\alpha f}d\mu_n\right)^{\frac{1}{n}}<\infty\,. \tag{2.1}$$

Now, with easy computations

$$\int_\chi e^{n\alpha f}d\mu_n=\frac{1}{2^n}\sum_{j=0}^n 3^{n\,\alpha\,\frac{j}{n}}\binom{j}{n}=\frac{1}{2^n}\sum_{j=0}^n 1^{n-j}3^{\alpha\,j}\binom{j}{n}$$
$$=\frac{1}{2^n}(3^\alpha+1)^n\,,$$

in particular (2.1) follows with $C=(3^\alpha+1)/2$ for all $\alpha>1$.
Combining the extended Varadhan's Theorem with Theorem $I.3$ in F. den Hollander, we obtain

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[3^{S_n}]=\sup\{f(x)-\log 2-x\log x-(1-x)\log(1-x)\,:\,x\in[0,1]\}\,.$$

since $I(x) = \infty$ for all $x \notin [0,1]$ (we use the same notation of Theorem $I.3$ in den Hollander).
It is easy to see that, the function $g(x) := f(x) - \log 2 - x \log x - (1-x)\log(1-x)$ has a local maximum in $x = 3/4$. Simple computations, show that $x = 3/4$ is indeed a maximum on $[0,1]$ for g, so

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[3^{S_n}] = g\left(\frac{3}{4}\right) = \log 2\,.$$

Therefore, $\mathbb{E}[3^{S_n}] = 2^{n(1+o(1))}$.

Comments 2.1. One can avoid the use of extended Varadhan's Theorem, using contraction principle, with $\mathcal{Y} = [-1,2]$ and $T : \chi \to \mathcal{Y}$, defined as follows: $T(x) = x$ on $[-1,2]$, $T(x) = 2$ for $x > 2$ and $T(x) = -1$ for $x < -1$.

3. Exercise 3.

Let Z_k be i.i.d. with $P(Z_k = 1) = P(Z_k = 2) = 1/2$. Let $W_n := Z_1 Z_2 \cdot \dots \cdot Z_n$.
(a) Show that $\mathbb{E}[W_n] = (3/2)^n$.
(b) Show that, given $\varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}((2-\varepsilon)^n < W_n < (2+\varepsilon)^n)) = 1$.

Solution:
(a) It is easy to see

$$\mathbb{E}[Z_i] = \int_{\{Z_i=1\}} Z_i d\mathbb{P} + \int_{\{Z_i=2\}} Z_i d\mathbb{P} = 1\frac{1}{2} + 2\frac{1}{2} = \frac{3}{2}\,,$$

for all $i \in \mathbb{N}$. By independence of Z_i, we have

$$\mathbb{E}[W_n] = \Pi_{i=1}^n \mathbb{E}[Z_i] = \left(\frac{3}{2}\right)^n\,.$$

(b) For all $i \in \mathbb{N}$, define $Y_i := \log Z_i$. Then Y_i's are i.i.d since Z_i's are so, furthermore

$$\mathbb{E}[Y_i] = \int_{\{Y_i=1\}} Y_i d\mathbb{P} + \int_{\{Y_i=2\}} Y_i d\mathbb{P} = \frac{1}{2}\log 2 = \log\sqrt{2}\,.$$

Then by the weak law of large numbers, for all $\varepsilon' > 0$, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\Big|\frac{1}{n}\sum_{i=1}^n Y_i - \mathbb{E}[Y_i]\Big| < \varepsilon'\right) = 1\,. \tag{3.1}$$

Note that

$$\frac{1}{n}\sum_{i=1}^n Y_i = \log(Z_1 \cdot \dots \cdot Z_n)^{\frac{1}{n}} = \log W_n^{\frac{1}{n}}\,. \tag{3.2}$$

Furthermore

$$\Big|\frac{1}{n}\sum_{i=1}^n Y_i - \mathbb{E}[Y_i]\Big| = |\log W_n^{\frac{1}{n}} - \log\sqrt{2}| < \varepsilon'\,, \tag{3.3}$$

or equivalently

$$\log\sqrt{2} - \varepsilon' < \log W_n^{\frac{1}{n}} < \log\sqrt{2} + \varepsilon' \Leftrightarrow (\sqrt{2}e^{-\varepsilon'})^n < W_n < (\sqrt{2}e^{\varepsilon'})^n\,. \tag{3.4}$$

For any $\varepsilon > 0$, one can choose a $\varepsilon' > 0$ sufficiently small, such that $e^{-\varepsilon'} > 1-(\varepsilon/\sqrt{2})$ and $e^{\varepsilon'} < 1+(\varepsilon/\sqrt{2})$. With this choice (3.2)-(3.4) imply

$$\{(\sqrt{2}-\varepsilon)^n < W_n < (\sqrt{2}+\varepsilon)^n\} \supset \left\{\left|\frac{1}{n}\sum_{i=1}^n Y_i - \mathbb{E}[Y_i]\right| < \varepsilon'\right\}.$$

Combining the previous observations with (3.1), the claim follows.

4. Exercise 4.

Let $\mathcal{S}$ be a Polish space. Let X_n be a sequence of i.i.d. $\mathcal{S}$-valued random variables with common distribution λ and let $\phi : \mathcal{S} \to \mathbb{R}^d$ be a continuous function. Assume that

$$\mathbb{E}[e^{a|\phi(X_1)|}] < \infty\,, \qquad \forall a > 0\,.$$

Prove by contraction from Sanov's theorem that

$$\frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^n \phi(X_i)\,,$$

satisfies a LDP with speed n and good and convex rate function $I : \mathbb{R}^d \to [0,\infty]$ given by

$$I(z) := \inf\{H(\nu|\lambda) : \nu(\phi) = z\} = \sup_{\theta\in\mathbb{R}^d}\{z\cdot\theta \log\lambda(e^{\theta\cdot\phi})\}\,.$$

For simplicity suppose that ϕ is bounded.

Solution: a) Set $L_n^X = 1/n\sum_{i=1}^n \delta_{X_i}$; then $L_n^X : \Omega \to \mathcal{M}_1(S)$ is an $\mathcal{M}_1(S)$-valued random variable.
By Sanov's Theorem, the law $\mathcal{L}(L_n^X)$ of L_n^X satisfies a LDP on $\mathcal{M}_1(S)$ with rate function $I(\nu) = H(\nu|\lambda)$.
If $\phi : S \to \mathbb{R}^d$ is bounded and continuous, then automatically satisfies $\mathbb{E}[e^{a|\phi(X_1)|}] < \infty$ for all $a > 0$. Furthermore, consider the following continuous map

$$\begin{aligned} T : \mathcal{M}_1(S) &\to \mathbb{R}^d\,,\\ \nu &\mapsto (\nu(\phi^1),\dots,\nu(\phi^d))\,, \end{aligned}$$

where ϕ^i are the component of ϕ; further we set $\nu(\phi) := (\nu(\phi^1),\dots,\nu(\phi^d))$.
To see the continuity, recall that by definition of product topology, a map f with value $\mathbb{R}^d$ is continuous if and only if $p_i(f)$ is continuous for all $i = 1,\dots,d$; here p_i is the projection on the i-th component. So T is continuous if and only if the maps

$$\nu \mapsto \nu(\phi^i)$$

are continuous for all $i = 1,\dots,d$; since $\phi^i \in C_b(S)$ by hypothesis, this clear follows by definition of weak topology on $\mathcal{M}_1(S)$.
By Contraction principle, we have that $\mathcal{L}(L_n^X)\circ T^{-1}$ satisfy a LDP with speed n and rate function

$$I(x) = \inf_{\nu\in\mathcal{M}_1(S)\,,\,T(\nu)=x} H(\nu|\lambda) = \inf_{\nu\in\mathcal{M}_1(S)\,,\,\nu(\phi)=x} H(\nu|\lambda)\,.$$

It remains to show that $\mathcal{L}(L_n^X)\circ T^{-1} = \mathcal{L}(S_n/n)$.
Indeed, let $B \in \mathcal{B}(\mathbb{R}^d)$ be a Borel set, then it is enough to show that

$$\{L_n^X \in T^{-1}(B)\} = \{S_n/n \in B\}\,. \tag{4.1}$$

If $\omega \in \{L_n^X \in T^{-1}(B)\}$, then $L_n^X(\omega) = \nu$ for some $\nu \in T^{-1}(B)$; in other words $T(L_n^X(\omega)) \in B$. A straightforward computation shows that

$$T(L_n^X(\omega)) = \frac{1}{n}\sum_{i=1}^{n} \phi(X_i(\omega)) = \frac{S_n(\omega)}{n}\,;$$

since $\delta_{X_i(\omega)}(\phi^j) = \phi^j(X_i(\omega))$ for all $i \in \mathbb{N}$, $\omega \in \Omega$ and $j = 1, \dots, d$. This proves that $\{L_n^X \in T^{-1}(B)\} \subset \{S_n/n \in B\}$, reversing the argument just performed one obtain the opposite inclusion, this implies the equality in (4.1).

2) *Convexity.* Fix $\varepsilon > 0$ small and $x, y \in \mathbb{R}^d$, then there exists $\nu_1^\varepsilon, \nu_2^\varepsilon$ such that

$$I(x) + \varepsilon \geq H(\nu_1^\varepsilon|\lambda)\,, \qquad \nu(\nu_1^\varepsilon) = x\,, \tag{4.2}$$

$$I(y) + \varepsilon \geq H(\nu_2^\varepsilon|\lambda)\,, \qquad \nu(\nu_2^\varepsilon) = y\,. \tag{4.3}$$

Note that $\nu^\varepsilon = t\nu_1^\varepsilon + (1-t)\nu_2^\varepsilon$ verifies $\nu^\varepsilon(\phi) = tx + (1-t)y$. So,

$$\begin{aligned} I(tx + (1-t)y) &\leq \int_S \frac{d\nu^\varepsilon}{d\lambda} \log \frac{d\nu^\varepsilon}{d\lambda}\, d\lambda \\ &= H(\nu^\varepsilon|\lambda) \\ &\leq tH(\nu_1^\varepsilon|\lambda) + (1-t)H(\nu_2^\varepsilon|\lambda) \\ &\leq t(I(x)+\varepsilon) + (1-t)(I(x)+\varepsilon)\,, \end{aligned}$$

where we have used the convexity of the relative entropy $H(\cdot|\lambda)$ and (4.2)-(4.3); sending $\varepsilon \searrow 0$ one obtains the claim.

3) *Legendre Transform.* Note that

$$\begin{aligned} I^*(\theta) &= \sup_{z\in\mathbb{R}^d} \Big\{ z\cdot\theta - \inf_{\nu\in\mathcal{M}_1(S)\,,\,\nu(\phi)=z} H(\nu|\lambda) \Big\} \\ &= \sup_{z\in\mathbb{R}^d} \left(\sup_{\nu\in\mathcal{M}_1(S)\,,\,\nu(\phi)=z} \{ z\cdot\theta - H(\nu|\lambda)\} \right) \\ &= \sup_{z\in\mathbb{R}^d} \left(\sup_{\nu\in\mathcal{M}_1(S)\,,\,\nu(\phi)=z} \{ \nu(\phi\cdot\theta) - H(\nu|\lambda)\} \right) \\ &= \sup_{\nu\in\mathcal{M}_1(S)} \Big(\nu(\phi\cdot\theta) - H(\nu|\lambda) \Big)\,. \end{aligned}$$

4) We know that $H(\nu|\lambda) = \sup_{f\in C_b(S)}\{\nu(f) - \log(\lambda(e^f))\}$. Then

$$\begin{aligned} I^*(\theta) &= (H(\nu|\lambda))^*(\phi\cdot\theta) \\ &= (p^*)^*(\phi\cdot\theta) = p(\phi\cdot\theta)\,; \end{aligned}$$

where in the last equality we have used the Fenchel-Moreau Theorem.

5) Note that

$$\begin{aligned} I(z) &= (I^*)^*(z) \\ &= \sup_{\theta\in\mathbb{R}^d}\{z\cdot\theta - I^*(\theta)\} \\ &= \sup_{\theta\in\mathbb{R}^d}\{z\cdot\theta - p(\phi\cdot\theta)\} \\ &= \sup_{\theta\in\mathbb{R}^d}\{z\cdot\theta - \log\lambda(e^{\phi\cdot\theta})\}\,. \end{aligned}$$

5. Exercise 5.

Solve Exercise 2.16 in [2].

Solution:
a) Let G be an open subset of $\mathcal{E}$, then for $\delta > 0$ define

$$\begin{aligned} G_\delta &= \{x\in G \,:\, \mathrm{dist}(x,\partial G) > \delta\} \\ &= \{x\in\mathcal{E} \,:\, \mathrm{dist}(x,\partial G) > \delta\}\cap G\,. \end{aligned}$$

Then $G_\delta \subset G$ is an open subset of $\mathcal{E}$.
Fix $\delta > 0$, then by hypothesis there exists $n_0\in\mathbb{N}$ such that $d(\xi_n(\omega),\eta_n(\omega)) < \delta$ for all $n > n_0$ and $\omega\in\Omega$. Then

$$\{\eta_n\in G\}\supset\{\xi_n\in G_\delta\}\,,\qquad \forall n > n_0\,. \tag{5.1}$$

This implies

$$\liminf_{n\to\infty}\frac{1}{r_n}\log\mathbb{P}(\eta_n\in G)\geq\liminf_{n\to\infty}\frac{1}{r_n}\log\mathbb{P}(\xi_n\in G_\delta)\geq -\inf_{x\in G_\delta} I(x)\,;$$

where the first inequality follows by (5.1) and the second by the hypothesis on ξ_n.
In particular, we have

$$\liminf_{n\to\infty}\frac{1}{r_n}\log\mathbb{P}(\eta_n\in G)\geq -\liminf_{n\to\infty}\inf_{x\in G_{\delta_n}} I(x)\,;$$

where δ_n is any sequence of positive real numbers such that $\delta_n\searrow 0$.
By this, it is enough to show that

$$\liminf_{n\to\infty}\inf_{x\in G_{\delta_n}} I(x)\leq\inf_{x\in G} I(x)\,.$$

To prove this, consider $(x_n)_n$ a sequence such that $I(x_n)\searrow\inf_{x\in G} I(x)$. Then, for δ_n sufficiently small, $x\in G_{\delta_n}$ and

$$\inf_{x\in G_{\delta_n}} I(x)\leq I(x_n)\,. \tag{5.2}$$

Taking the $\liminf_{n\to\infty}$ to both sides of (5.2), we obtain the claim.

b) Let F be a closed subset of $\mathcal{E}$, then for $\delta > 0$ define

$$F^\delta = \{x\in\mathcal{E} \,:\, \mathrm{dist}(x,F)\leq\delta\}\,.$$

Then $F^\delta\supset F$ is a closed subset of $\mathcal{E}$.
Fix $\delta > 0$, then by hypothesis there exists $n_0\in\mathbb{N}$ such that $d(\xi_n(\omega),\eta_n(\omega)) < \delta$ for all $n > n_0$ and $\omega\in\Omega$. Then

$$\{\eta_n\in F\}\subset\{\xi_n\in F^\delta\}\,,\qquad \forall\ n > n_0\,. \tag{5.3}$$

This implies

$$\limsup_{n\to\infty} \frac{1}{r_n} \log \mathbb{P}(\eta_n \in F) \le \limsup_{n\to\infty} \frac{1}{r_n} \log \mathbb{P}(\xi_n \in F^\delta) \le - \inf_{x\in F^\delta} I(x)\,;$$

where the first inequality follows by (5.3) and the second by the hypothesis on ξ_n. As before, it is sufficient to prove that

$$\liminf_{n\to\infty} \inf_{F^{\delta_n}} I(x) \ge \inf_{x\in F} I(x)\,; \tag{5.4}$$

for some $(\delta_n)_n$, such that $\delta_n \searrow 0$. If $I \not\equiv +\infty$ on F^δ (otherwise the inequality in (5.4) is trivial), then we can assume F^δ is compact; since we can replace F^δ with $F^\delta \cap \{I \le I(x)\}$ for some $x \in F^\delta$ such that $I(x) < \infty$ (recall that I is good by assumption).
Note that, for each $n \in \mathbb{N}$ there exists $x_n \in F^{\delta_n}$ such that

$$I(x_n) - \frac{1}{n} \le \inf_{x\in F^{\delta_n}} I(x)\,. \tag{5.5}$$

By compactness of F^δ, there exists a subsequence of $(x_n)_n$ which converges in $\mathcal{E}$ to an element $\tilde{x} \in F$; we still denote $(x_n)_n$ this subsequence.
So $\inf_{x\in F} I(x) \le I(\tilde{x})$, but for the lower semicontinuity of the rate function I, we have

$$I(\tilde{x}) \le \liminf_{n\to\infty} I(x_n) \le \liminf_{n\to\infty} \inf_{x\in F^{\delta_n}} I(x)\,;$$

where in the last inequality follows by (5.5). This conclude the proof of (5.4).

6. Exercise 6.

Solve Exercise 5.20 in [2].

Solution:
a) Since $p^*(z) < \infty$ there exists $\rho \in \mathcal{M}_1(\mathcal{S})$ such that $c := H(\rho|\lambda) < \infty$. Furthermore, the map

$$\mathcal{M}_1(\mathcal{S}) \ni \nu \mapsto \mathbb{E}^\nu[\mathcal{H}]\,,$$

is continuous for the weak topology, so for each $z \in \mathbb{R}$, the set $\{\nu \in \mathcal{M}_1(\mathcal{S}) : E^\nu[\mathcal{H}] = z\}$ is closed.
Since $\{\nu \in \mathcal{M}_1(\mathcal{S}) : H(\nu|\lambda) \le c\}$ is compact, then the set

$$Z := \{\nu \in \mathcal{M}_1(\mathcal{S}) : H(\nu|\lambda) \le c\} \cap \{\nu \in \mathcal{M}_1(\mathcal{S}) : E^\nu[\mathcal{H}] = z\}\,,$$

is compact. It is easy to see that

$$\inf\{H(\nu|\lambda) : E^\nu[\mathcal{H}] = z\,,\, \nu \in \mathcal{M}_1(\mathcal{S})\} = \inf\{H(\nu|\lambda) : \nu \in Z\}\,.$$

Let $\{\nu_k\}_k \subset Z$ be a minimizing sequence, i.e.

$$H(\nu_k|\lambda) \searrow \inf\{H(\nu|\lambda) : \nu \in Z\}\,, \qquad \text{as } k \nearrow \infty\,. \tag{6.1}$$

By the compactness of Z, there exists a subsequence $\{\nu_{k_j}\}_j \subset Z$ and $\nu_z \in \mathcal{M}_1(\mathcal{S})$ such that $\nu_{k_j} \to \nu_z \in \mathcal{M}_1(\mathcal{S})$. By lower semicontinuity of the map $\nu \mapsto H(\nu|\lambda)$, we obtain

$$H(\nu_z|\lambda) \le \liminf_{j\to\infty} H(\nu_{k_j}|\lambda) = \inf\{H(\nu|\lambda) : \nu \in Z\} = p^*(z)\,;$$

where we have used (6.1). Observe that $\mathbb{E}^{\nu_z}[\mathcal{H}] = \lim_{j\to\infty} \mathbb{E}^{\nu_{k_j}}[\mathcal{H}] = z$, then $\nu_z \in Z$ and this concludes the proof of the existence.
For the uniqueness, suppose that there exist $\nu_1, \nu_2 \in \mathcal{M}_1(\mathcal{S})$ such that $H(\nu_1|\lambda) =$

$H(\nu_2|\lambda) = \inf\{H(\nu|\lambda) \, : \, \nu \in Z\}$, then by the strict convexity of $H(\cdot|\lambda)$ (see [2] Exercise 5.5) we have

$$H(\frac{1}{2}\nu_1 + \frac{1}{2}\nu_2|\lambda) < \frac{1}{2}H(\nu_1|\lambda) + \frac{1}{2}H(\nu|\lambda) \leq \inf\{H(\nu|\lambda) \, : \, \nu \in Z\}\,,$$

so $\mathbb{E}^{\frac{1}{2}\nu_1+\frac{1}{2}\nu_2}[\mathcal{H}] = z$ and $H(\frac{1}{2}\nu_1 + \frac{1}{2}\nu_2|\lambda) < \inf\{H(\nu|\lambda) \, : \, E^{\nu}[\mathcal{H}] = z\,, \nu \in \mathcal{M}_1\}$; which is clearly an absurd.

$b)$ Recall that $p(t) = \log \mathbb{E}^{\lambda}[e^{t\mathcal{H}}]$, then by the chain rule and the interchange of differentiation and integration theorem, we obtain

$$\begin{aligned} p'(t) = \frac{1}{\mathbb{E}^{\lambda}[e^{t\mathcal{H}}]}\mathbb{E}^{\lambda}[\mathcal{H}e^{t\mathcal{H}}] &= \int \mathcal{H}\frac{e^{t\mathcal{H}}}{\mathbb{E}^{\lambda}[e^{t\mathcal{H}}]}d\lambda \\ &= \int \mathcal{H} d\mu_{-t} = \mathbb{E}^{\mu_{-t}}[\mathcal{H}]\,; \end{aligned}$$

where the last two equalities follows by the definition of the Gibbs measures.
For the interchange between integral and derivation one can easily see that the hypothesis $i) - iii)$ of Theorem 3.6.3 in [1] are satisfied and so the interchange is valid.
The same applies to the second derivative

$$\begin{aligned} p''(t) &= \frac{d}{dt}\left[\int \mathcal{H}\frac{e^{t\mathcal{H}}}{\mathbb{E}^{\lambda}[e^{t\mathcal{H}}]}d\lambda\right] \\ &= \int \mathcal{H}\,\frac{\mathcal{H}e^{t\mathcal{H}}\mathbb{E}^{\lambda}[e^{t\mathcal{H}}] - e^{t\mathcal{H}}\mathbb{E}^{\lambda}[\mathcal{H}e^{t\mathcal{H}}]}{(\mathbb{E}^{\lambda}[e^{t\mathcal{H}}])^2}d\lambda \\ &= \frac{1}{\mathbb{E}^{\lambda}[e^{t\mathcal{H}}]}\int \mathcal{H}^2 e^{t\mathcal{H}}d\lambda - \left(\frac{\mathbb{E}^{\lambda}[\mathcal{H}e^{t\mathcal{H}}]}{\mathbb{E}^{\lambda}[e^{t\mathcal{H}}]}\right)^2 \\ &= \mathbb{E}^{\mu_{-t}}[\mathcal{H}^2] - (\mathbb{E}^{\mu_{-t}}[\mathcal{H}])^2\,. \end{aligned}$$

Now, we prove that $\lim_{t\to 0} p'(t) = \mathbb{E}^{\lambda}[\mathcal{H}]$. Indeed, by Lebesgue dominated convergence (recall that $\mathcal{H} \in L^{\infty}(S)$) we have

$$\int e^{t\mathcal{H}}\mathcal{H}d\lambda \to \mathbb{E}^{\lambda}[\mathcal{H}]\,, \qquad \int e^{t\mathcal{H}}d\lambda \to \mathbb{E}^{\lambda}[1] = 1\,;$$

as $t \to 0$, then the claim follows.

Fix $\varepsilon > 0$, by definition of ess sup, we have $\lambda(\{B - \varepsilon \leq \mathcal{H} \leq B\}) > 0$ and

$$(B - \varepsilon) \leq \frac{\int_{\{B-\varepsilon\leq\mathcal{H}\leq B\}} \mathcal{H}e^{t\mathcal{H}}\, d\lambda}{\int_{\{B-\varepsilon\leq\mathcal{H}\leq B\}} e^{t\mathcal{H}}\, d\lambda} \leq B\,.$$

Furthermore,

$$\int_{\{B-\varepsilon\leq\mathcal{H}\leq B\}} e^{t\mathcal{H}}\, d\lambda \sim \int e^{t\mathcal{H}}d\lambda\,, \qquad \text{as } t \nearrow \infty\,,$$

$$\int_{\{B-\varepsilon\leq\mathcal{H}\leq B\}} \mathcal{H}e^{t\mathcal{H}}\, d\lambda \sim \int \mathcal{H}e^{t\mathcal{H}}d\lambda\,, \qquad \text{as } t \nearrow \infty\,;$$

(here $f(t) \sim g(t)$ as $t \to t_0$ iff $\lim_{t\to t_0}(f(t)/g(t)) = 1$). We prove the first, the other follows identically.

Since $\int e^{t\mathcal{H}} d\lambda = \int_{\{B-\varepsilon \le \mathcal{H} \le B\}} e^{t\mathcal{H}} d\lambda + \int_{\{\mathcal{H} < B-\varepsilon\}} e^{t\mathcal{H}} d\lambda$, it is enough to show that

$$\frac{\int_{\{\mathcal{H} < B-\varepsilon\}} e^{t\mathcal{H}} d\lambda}{\int_{\{B-\varepsilon \le \mathcal{H} \le B\}} e^{t\mathcal{H}} d\lambda} \to 0\,, \qquad t \nearrow \infty\,.$$

Since

$$\begin{aligned} e^{t\varepsilon/2}\lambda(\{B-\varepsilon/2 \le \mathcal{H} \le B\}) &\le \int_{\{B-\varepsilon/2 \le \mathcal{H} \le B\}} e^{t(\mathcal{H}-B+\varepsilon)} d\lambda \\ &\le \int_{\{B-\varepsilon \le \mathcal{H} \le B\}} e^{t(\mathcal{H}-B+\varepsilon)} d\lambda\,, \end{aligned}$$

then

$$0 \le \frac{\int_{\{\mathcal{H} < B-\varepsilon\}} e^{t(\mathcal{H}-B+\varepsilon)} d\lambda}{\int_{\{B-\varepsilon \le \mathcal{H} \le B\}} e^{t(\mathcal{H}-B+\varepsilon)} d\lambda} \le \frac{\int_{\{\mathcal{H} < B-\varepsilon\}} e^{t(\mathcal{H}-B+\varepsilon)} d\lambda}{e^{t\varepsilon/2}\lambda(\{B-\varepsilon/2 \le \mathcal{H} \le B\})} \searrow 0\,,$$

as $t \nearrow \infty$; since $\int_{\{\mathcal{H} < B-\varepsilon\}} e^{t(\mathcal{H}-B+\varepsilon)} d\lambda \searrow 0$ by Lebesgue dominated convergence and $\varepsilon > 0$.

By the asymptoptic behaviour, we have

$$\begin{aligned} (B-\varepsilon) &\le \liminf_{t\to\infty} \frac{\int \mathcal{H} e^{t\mathcal{H}}\, d\lambda}{\int e^{t\mathcal{H}}\, d\lambda} \le B\,, \\ (B-\varepsilon) &\le \limsup_{t\to\infty} \frac{\int \mathcal{H} e^{t\mathcal{H}}\, d\lambda}{\int e^{t\mathcal{H}}\, d\lambda} \le B\,; \end{aligned}$$

by the arbitrariness of $\varepsilon > 0$, one can easily conclude the proof.
Similarly, one can prove that $p'(t) = \lim_{t\to-\infty} = A$; we omit the details.

c) Since $p'(t)$ is a smooth function of $t \in \mathbb{R}$, then $p''(t) > 0$ for every $t \in \mathbb{R}$, implies that $p'(t)$ is strictly increasing. So, for each $z \in (A, B)$ there exists an unique $\beta = \beta_z$ such that $z = p'(-\beta)$.
Furthermore, by the definition of p^* we have

$$p^*(z) = \sup_{t\in\mathbb{R}}\{tz - p(t)\}\,.$$

We will show that the sup on the above RHS is indeed a maximum and it is attained at $t = -\beta$. This would imply that $p^*(z) = z(-\beta) - p(\beta) = -z\beta - p(-\beta)$ and this concludes the first part of the point *c*).
To prove this, for a fixed $z \in \mathbb{R}$, set $f(t) = tz - p(t)$. Then

$$f'(t) = z - p'(t)\,, \qquad f''(t) = -p''(t) < 0\,.$$

Since the equation $f'(t) = z - p'(t) = 0$ admits an *unique* solution, which we called $-\beta$, then the condition $f''(t) < 0$ trivially implies the value $f(-\beta)$ is a local maximum and since there are not other critical point of f, then $-\beta$ is indeed a global maximum.

For the last part of the point c), note that

$$\begin{aligned} H(\mu_\beta|\lambda) &= \int \frac{d\mu_\beta}{d\lambda}\,\log\frac{d\mu_\beta}{d\lambda}\,d\lambda \\ &= \int \frac{e^{-\beta\mathcal{H}}}{\mathbb{E}^\lambda[e^{-\beta\mathcal{H}}]}\log\frac{e^{\beta\mathcal{H}}}{\mathbb{E}^\lambda[e^{-\beta\mathcal{H}}]}\,d\lambda \\ &= \frac{1}{\mathbb{E}^\lambda[e^{-\beta\mathcal{H}}]}\int e^{-\beta\mathcal{H}}\log e^{-\beta\mathcal{H}}d\lambda - \frac{\log\mathbb{E}^\lambda[e^{-\beta\mathcal{H}}]}{\mathbb{E}^\lambda[e^{-\beta\mathcal{H}}]}\int e^{-\beta\mathcal{H}}d\lambda \\ &= -\beta\frac{\mathbb{E}^\lambda[\mathcal{H}e^{-\beta\mathcal{H}}]}{\mathbb{E}^\lambda[e^{-\beta\mathcal{H}}]} - \log\mathbb{E}^\lambda[e^{-\beta\mathcal{H}}] \\ &= -\beta z - p(t) = p^*(z)\,; \end{aligned}$$

where $z = p'(-\beta)$ and the last equality follows by the first part of point c) just proved. Recall that, by part a) we have that for each z then there exists an unique ν_z such that $p^*(z) = H(\nu_z|\lambda)$. The uniqueness implies that $\nu_z = \mu_\beta$.

7. Exercise 7.

Solve Exercise 6.5 in [2].

Solution:
For clarity, we divide the proof in two steps.
Step 1. We assume that $g(\omega) = g(\omega_k)$ for a fixed $k \in \mathbb{Z}^d$, i.e. the local function f depends only on $\omega_k \in S$ for $\omega = (\omega_i)_{i\in\mathbb{Z}^d} \in \Omega := S^{\mathbb{Z}^d}$, (here $d \in \mathbb{N}$).
Let be $V_n := (-n,n)^d \cap \mathbb{Z}^d$, define

$$\tilde{V}_n^k := \{i \in V_n \,:\, k-i \in V_n\}\,. \tag{7.1}$$

Note that for each $i \in \tilde{V}_n^k$ we have

$$(\theta_i\omega^{(n)})_k = \omega^{(n)}_{k-i} = \omega_{k-i} = (\theta_i\omega)_k\,;$$

then $g((\theta_i\omega^{(n)})_k) - g((\theta_i\omega)_k) = 0$ for each $i \in \tilde{V}_n^k$.
This implies that

$$\begin{aligned} |R(\omega,g) - \tilde{R}(\omega,g)| &= \frac{1}{|V_n|}\Big|\sum_{i\in V_n} g((\theta_i\omega^{(n)})_k) - g((\theta_i\omega)_k)\Big| \\ &= \frac{1}{|V_n|}\Big|\sum_{i\in V_n\setminus\tilde{V}_n^k} g((\theta_i\omega^{(n)})_k) - g((\theta_i\omega)_k)\Big| \\ &\le 2|g|_{L^\infty(S)}\frac{|V_n\setminus\tilde{V}_n^k|}{|V_n|}\,; \end{aligned}$$

this implies

$$\sup_{\omega\in\Omega}|R(\omega,g) - \tilde{R}(\omega,g)| \le 2|g|_{L^\infty(\Omega)}\frac{|V_n\setminus\tilde{V}_n^k|}{|V_n|}\,. \tag{7.2}$$

Since, $|V_n\setminus\tilde{V}_n^k| \sim c_n n^{d-1}$ and $|V_n| \sim c_n' n^d$ as $n \nearrow \infty$ for suitable $c_n, c_n' > 0$, the claim follows by (7.2).

Step 2. Here we suppose that

$$g(\omega) = g(\omega_{i_1},\omega_{i_2},\dots,\omega_{i_k})\,,$$

for $i_1, \dots, i_k \in \mathbb{Z}^d$ and $k > 1$ fixed.
Now, for $i \in V_n$, we can write

$$\begin{aligned}
&g(\theta_i\omega^{(n)}) - g(\theta_i\omega)\\
&\qquad = [g((\theta_i\omega^{(n)})_{i_1}, (\theta_i\omega^{(n)})_{i_2}, \dots, (\theta_i\omega^{(n)})_{i_k}) - g((\theta_i\omega)_{i_1}, (\theta_i\omega^{(n)})_{i_2}, \dots, (\theta_i\omega^{(n)})_{i_k})]\\
&\qquad + [g((\theta_i\omega)_{i_1}, (\theta_i\omega^{(n)})_{i_2}, \dots, (\theta_i\omega^{(n)})_{i_k}) - g((\theta_i\omega)_{i_1}, (\theta_i\omega)_{i_2}, \dots, (\theta_i\omega^{(n)})_{i_k})]\\
&\qquad \dots\dots\\
&\qquad + [g((\theta_i\omega)_{i_1}, (\theta_i\omega)_{i_2}, \dots, (\theta_i\omega^{(n)})_{i_k}) - g((\theta_i\omega)_{i_1}, (\theta_i\omega)_{i_2}, \dots, (\theta_i\omega)_{i_k})]\\
&\qquad = \sum_{h=1}^{k} \Delta g_h^i(\omega)\,;
\end{aligned}$$

where, for each $h = 1, \dots, k$;

$$\begin{aligned}
\Delta g_h^i(\omega) := g((\theta_i\omega)_{i_1}, \dots, (\theta_i\omega^{(n)})_{i_h}, \dots (\theta_i\omega^{(n)})_{i_k})&\\
- g((\theta_i\omega)_{i_1}, \dots, (\theta_i\omega)_{i_h}, \dots (\theta_i\omega^{(n)})_{i_k})&\,.
\end{aligned}$$

Now, for each $i_j \in \mathbb{Z}^d$ and $j = 1, \dots, k$, with the same notation as in (7.1) we set

$$\tilde V_n^{i_j} = \{i \in V_n \,:\, i_j - i \in V_n\}\,.$$

Arguing as in **Step 1**, we have

$$\begin{aligned}
R(\omega, g) - \tilde R(\omega, g) &= \frac{1}{|V_n|}\Big|\sum_{i\in V_n} g(\theta_i\omega^{(n)}) - g(\theta_i\omega)\Big|\\
&\le \frac{1}{|V_n|}\sum_{h=1}^{k}\Big|\sum_{i\in V_n}\Delta g_h^i(\omega)\Big|\\
&= \frac{1}{|V_n|}\sum_{h=1}^{k}\Big|\sum_{i\in V_n\setminus \tilde V_n^{i_h}}\Delta g_h^i(\omega)\Big|\\
&\le 2|g|_{L^\infty(\Omega)}\sum_{h=1}^{k}\frac{|V_n\setminus \tilde V_n^{i_h}|}{|V_n|}\,.
\end{aligned}$$

Since k is fixed, the same argument performed at the end of **Step 1** completes the proof.

8. Exercise 8.

Solve Exercise 6.10 in [2].

Solution: We will need the following property:

(P) Let $\alpha\,, \mu \in \mathcal{M}_1(\chi)$ and $\beta\,, \nu \in \mathcal{M}_1(\mathcal{Y})$ be probability measures. Then

$$\alpha\otimes\beta \ll \mu\otimes\nu \quad \leftrightarrow \quad \alpha \ll \mu \quad \text{and} \quad \beta \ll \nu\,.$$

Furthermore, if $f := d\alpha/d\mu$ and $g := d\beta/d\nu$, then

$$\frac{d(\alpha\otimes\beta)}{d(\mu\otimes\nu)} = f\,g\,.$$

We postpone the proof of this fact at the end of the exercise.

Solution:
(a) By property (P), we have only to check the equality under the condition $\alpha \otimes \beta \ll \mu \otimes \nu$; otherwise

$$H(\alpha \otimes \beta | \mu \otimes \nu) = \infty \,,$$

and at least one between the quantities $H(\alpha|\mu)$, $H(\beta|\nu)$ is infinite.

Now under the hypothesis $\alpha \otimes \beta \ll \mu \otimes \nu$, we obtain by property (P) (and using that notation for the Radon-Nikodym derivative)

$$\begin{aligned}
H(\alpha \otimes \beta | \mu \otimes \nu) &= \int_{\chi \times \mathcal{Y}} \frac{d(\alpha \otimes \beta)}{d(\mu \otimes \nu)} \log\left(\frac{d(\alpha \otimes \beta)}{d(\mu \otimes \nu)}\right) d(\mu \otimes \nu) \\
&= \int_{\chi \times \mathcal{Y}} f\, g(\log f + \log g) d(\mu \otimes \nu) \\
&= \int_{\chi} f \log f \left(\int_{\mathcal{Y}} g\, d\nu\right) d\mu + \int_{\mathcal{Y}} g \log g \left(\int_{\chi} f\, d\mu\right) d\nu \\
&= H(\alpha|\mu) + H(\beta|\nu) \,,
\end{aligned}$$

since $\int_{\mathcal{Y}} g\, d\nu = \beta(\mathcal{Y}) = 1$ and $\int_{\chi} f\, d\mu = \alpha(\chi) = 1$.

(b) Now, observe that

$$\begin{aligned}
H_n(\mu^{\otimes \mathbb{Z}^d} | \nu^{\otimes \mathbb{Z}^d}) = H(\otimes_{i \in V_n} \mu | \otimes_{i \in V_n} \nu) &= \sum_{i \in V_n} H(\mu|\nu) \\
&= |V_n| H(\mu|\nu) \,.
\end{aligned}$$

Here the second equality follows by part (a) and a simple induction argument. By this, we immediately obtain the claim, since

$$\begin{aligned}
h(\mu^{\otimes \mathbb{Z}^d} | \nu^{\otimes \mathbb{Z}^d}) = \lim_{n \to \infty} \frac{1}{|V_n|} H_n(\mu^{\otimes \mathbb{Z}^d} | \nu^{\otimes \mathbb{Z}^d}) &= \lim_{n \to \infty} \frac{1}{|V_n|} |V_n|\ H(\mu|\nu) \\
&= H(\mu|\nu) \,.
\end{aligned}$$

Proof of Property (P). ($\leftarrow$) as before $f := d\alpha/d\mu$ and $g := d\beta/d\nu$, then for all $A \times B$ such that $A \in \mathcal{B}(\chi)$ and $B \in \mathcal{B}(\mathcal{Y})$, we have

$$(\alpha \otimes \beta)(A \times B) = \int_{A \times B} fg\ d(\mu \otimes \nu) \,.$$

Thus define $\lambda(E) := \int_E fgd(\mu \otimes \nu)$ for each $E \in \mathcal{B}(\chi \times \mathcal{Y})$; it is easy to see that is a measure on the product space $\chi \times \mathcal{Y}$. Furthermore, the set

$$\mathcal{F} = \{E \in \mathcal{B}(\chi \times \mathcal{Y}) \,|\, \lambda(E) = (\alpha \otimes \beta)(E)\} \,,$$

is a sigma algebra containing all Borel set of the form $A \times B$ for $A \in \mathcal{B}(\chi)$ and $B \in \mathcal{B}(\mathcal{Y})$, then $\mathcal{F} \supset \sigma(\{A \times B \,|\, A \in \mathcal{B}(\chi)\,,\, B \in \mathcal{B}(\mathcal{Y})\}) = \mathcal{B}(\chi \times \mathcal{Y})$; where the last equality follows by the definition of product sigma algebra. This implies $\lambda = \alpha \otimes \beta$ on $\mathcal{B}(\chi \times \mathcal{Y})$ and this finishes the first part of the proof.

($\rightarrow$) Let $A \in \mathcal{B}(\chi)$ be a Borel and set $F := d(\alpha \otimes \beta)/d(\mu \otimes \nu)$, then

$$\begin{aligned} \alpha(A) = \alpha(A)\beta(\mathcal{Y}) &= \int_{A \times Y} F\, d(\mu \otimes \nu) \\ &= \int_A \left(\int_{\mathcal{Y}} F\, d\nu \right) d\mu\,; \end{aligned}$$

so $d(\alpha)/d(\mu) = \int_{\mathcal{Y}} F\, d\nu$. The statement for β follows in the same manner.

9. Exercise 9.

Consider $\mu, \nu \in \mathcal{M}_1(\chi)$ with $\nu \neq \mu$. As a consequence, there exists $f \in C_b(\chi)$ such that $\mu(f) \neq \nu(f)$. Use this observation and Birkhoff's ergodic theorem to deduce that $\mu^{\otimes \mathbb{Z}^d}$ and $\nu^{\otimes \mathbb{Z}^d}$ are mutually singular (in agreement with the fact that different probability measures in $\mathcal{M}_\theta$ are mutually singular).

Solution: Set $\Omega = \chi^{\mathbb{Z}^d}$, and let $f \in C_b(\chi)$ be a bounded continuous function on χ such that $\mu(f) \neq \nu(f)$. We may regard f as a function on Ω,

$$\tilde{f}(\omega) := f(\omega_0)\,.$$

Of course $\tilde{f}$ is continuous on Ω, since $\tilde{f} = f(p_0)$, where p_0 is the projection on the 0-th component of ω.
It is easy to see that

$$\int_\Omega \tilde{f}\, d\mu^{\otimes \mathbb{Z}^d} = \int_\chi f\, d\mu = \mu(f)\,,$$

and $\nu^{\otimes \mathbb{Z}^d}(\tilde{f}) = \nu(f)$. By Exercise B.24 in [2] we have that $\nu^{\otimes \mathbb{Z}^d}, \mu^{\otimes \mathbb{Z}^d}$ are ergodic measures, so Birkhoff's Theorem applies and we obtain

$$\begin{aligned} &R_n(\omega, \tilde{f}) \to \mu^{\otimes \mathbb{Z}^d}(\tilde{f}) = \mu(f)\,, && \forall \omega \in \Omega \setminus A\,, \\ &R_n(\omega, \tilde{f}) \to \nu^{\otimes \mathbb{Z}^d}(\tilde{f}) = \nu(f)\,, && \forall \omega \in \Omega \setminus B\,, \end{aligned}$$

with $\mu^{\otimes \mathbb{Z}^d}(A) = 0$ and $\nu^{\otimes \mathbb{Z}^d}(B) = 0$.
Since by assumption $\nu(f) \neq \mu(f)$, the uniqueness of the limit implies

$$(\Omega \setminus A) \cap (\Omega \setminus B) = \emptyset\,,$$

and the claim follows.

10. Exercise 10.

Solve Exercise 8.5 in [2].

Solution: (a) $h(\cdot|\Phi)$ is affine.
By (8.3) in [2] pag. 122. for each $\nu \in \mathcal{M}_\theta(\Omega)$, we have

$$h(\nu|\Phi) = \mathbb{E}^\nu[f_\Phi] + h(\nu|\lambda) + P(\Phi)\,. \tag{10.1}$$

Since Φ is fixed and the assignment

$$\nu \mapsto \mathbb{E}^\nu[f_\Phi] + h(\nu|\lambda)\,,$$

is linear by Proposition 6.8 in [2], then $h(\cdot|\Phi)$ is manifestly affine.

(b) $h(\cdot|\Phi)$ is lower semi-continuous.
Since on $\mathcal{M}_\theta(\Omega)$ we consider the topology induced by the weak topology according

to the duality $(C_b(\Omega))^* = M_b(\Omega)$, then the map $\nu \mapsto \mathbb{E}^\nu[f_\Phi]$ is continuous since $f_\Phi \in C_b(\Omega)$. Moreover, by Proposition 6.8 in [2] then $\nu \mapsto h(\nu|\lambda)$ is lower semi-continuous, then by (10.1) we obtain the claim.

(c) $h(\cdot|\Phi)$ has compact level.
It turns out (see (6.4) in Theorem 6.7 in [2]) that the following equality holds

$$h(\nu|\lambda) = \sup_{\Lambda\in\mathcal{R}} \frac{1}{|\lambda|} H_\Lambda(\nu|\lambda)\,;$$

where $\mathcal{R}$ is the collection of finite rectangles in $\mathbb{Z}^d$.
By this, for any $c \in \mathbb{R}$, we have

$$\{\nu \in \mathcal{M}_\theta(\Omega)\,:\, h(\nu|\lambda) \leq c\} = \bigcap_{\Lambda\in\mathcal{R}} \{\nu \in \mathcal{M}_\theta(\Omega)\,:\, H_\Lambda(\nu|\lambda) \leq c|\Lambda|\}\,.$$

Since $H_\Lambda(\cdot|\lambda)$ has compact level (see Proposition 6.8 in [2]) and the fact that an intersection of compact subsets is compact (recall that we are in an Hausdorff space), the claim follows.

References

[1] Benedetto, John J and Czaja, Wojciech, *Integration and modern analysis*, Springer Science & Business Media, 2010.

[2] Rassoul-Agha, Firas and Seppäläinen, Timo, *A course on large deviations with an introduction to Gibbs measures*, American Mathematical Soc., 2015.