8. Übung aus Informations- und Codierungstheorie SS 2016

1. Ist (X_n) eine DMS mit Verteilung $P := PX_i^{-1} = (p_1, \dots, p_m)$ und ist $Q := (q_1, \dots, q_m)$ eine weitere Verteilung auf $X(\Omega)$, so bestimme man
$$\lim_n -\frac{1}{n} \log Q(\mathbf{X}_1^n) \text{ und } \lim_n \frac{1}{n} \log \frac{P(\mathbf{X}_1^n)}{Q(\mathbf{X}_1^n)}.$$
Warum wird sich das Ergebnis nicht ohne weiteres auf ergodische Quellen übertragen lassen?
Lösung: Aus dem Gesetz der großen Zahlen folgt
$$\lim_n -\frac{1}{n} \log Q(\mathbf{X}_1^n) = -\frac{1}{n} \sum_{i=1}^n \log Q(X_i)$$
$$= \mathbb{E} \log Q(X_1) = -\sum_{j=1}^m p_j \log q_j = H(P) + D(P|Q).$$
Analog $\lim\limits_n \frac{1}{n} \log \frac{P(\mathbf{X}_1^n)}{Q(\mathbf{X}_1^n)} = D(P|Q)$.
Wir wissen $H(X) = \lim\limits_n H(X_n|\mathbf{X}_0^{n-1})$. Ohne Zusatzvoraussetzungen kann aber kein Konvergenzverhalten für $D(P|Q)$ garantiert werden.
2. Man zeige, dass für eine Quelle $(X_n)_{n\in\mathbb{Z}}$ mit $|X(\Omega)| < \infty$ die Folge $Y_0 := P^{-1}(X_0),\ Y_n := P^{-1}(X_0|\mathbf{X}_{-n}^{-1}),\ n \geq 1$ zusammen mit der Filtration $\mathfrak{S}_n := \mathfrak{S}(\mathbf{X}_{-n}^0)$ ein Martingal bildet.
Beweis: Die $\mathfrak{S}_n$ sind endlich und werden von den Atomen $\{\mathbf{X}_{-n}^0 = \mathbf{x}_{-n}^0\}$ mit $\mathbf{x}_{-n}^0 \in X(\Omega)^{n+1}$ erzeugt. Daher gilt mit $p(\mathbf{x}_i^j) := P(\mathbf{X}_i^j = \mathbf{x}_i^j)$
$$\mathbb{E}(Y_{n+1}|\mathbf{X}_{-n}^0 = \mathbf{x}_{-n}^0) = \frac{1}{p(\mathbf{x}_{-n}^0)} \int\limits_{[\mathbf{X}_{-n}^0 = \mathbf{x}_{-n}^0]} \frac{1}{P(X_0|\mathbf{X}_{-n-1}^{-1})} dP$$
$$= \frac{1}{p(\mathbf{x}_{-n}^0)} \int\limits_{\{\mathbf{x}_{-n}^0\}} \frac{1}{p(x_0|\mathbf{x}_{-n-1}^{-1})} P(\mathbf{X}_{-n-1}^0)^{-1}(dx_0, \dots dx_{-n-1})$$
$$= \sum_{x_{-n-1}\in X(\Omega)} \frac{p(\mathbf{x}_{-n-1}^{-1})\, p(\mathbf{x}_{-n-1}^0)}{p(\mathbf{x}_{-n}^0)\, p(\mathbf{x}_{-n-1}^0)} = \sum_{x_{-n-1}\in X(\Omega)} \frac{p(\mathbf{x}_{-n-1}^{-1})}{p(\mathbf{x}_{-n}^0)} = \frac{p(\mathbf{x}_{-n}^{-1})}{p(\mathbf{x}_{-n}^0)}$$
$$= \frac{1}{p(x_0|\mathbf{x}_{-n}^{-1})}\ \forall\, \mathbf{x}_{-n}^0 \in X(\Omega)^{n+1} \ \Rightarrow\ \mathbb{E}(Y_{n+1}|\mathbf{X}_{-n}^0) = \frac{1}{P(X_0|\mathbf{X}_{-n}^{-1})} = Y_n.$$
3. Man zeige, dass für eine Quelle $(X_n)_{n\in\mathbb{Z}}$ mit $|X(\Omega)| < \infty$ die Folgen $Z_0 := P(X_0),\ Z_n := P(X_0|\mathbf{X}_{-n}^{-1}),\ n \geq 1$ bzw. $V_0 := \log P(X_0)$, $V_n := \log P(X_0|\mathbf{X}_{-n}^{-1}),\ n \geq 1$ zusammen mit der Filtration $\mathfrak{S}_n := \mathfrak{S}(\mathbf{X}_{-n}^0)$

Submartingale bilden. Weiters zeige man, dass (Z_n) gleichmäßig integrierbar ist und dass es ein integrierbares V gibt mit $\lim_n V_n = V$ P–fs.

Beweis: Da die Folge $Y_n := P^{-1}(X_0|\mathbf{X}_{-n}^{-1})$, wie in Bsp. 2 gezeigt, ein Martingal ist, und die Funktionen $\varphi(x) := \frac{1}{x}$ bzw. $\psi(x) := -\log x$ konvex sind, sind $(Z_n, \mathfrak{S}_n) = (\varphi(Y_n), \mathfrak{S}_n)$ und $(V_n, \mathfrak{S}_n) = (\psi(Y_n), \mathfrak{S}_n)$ Submartingale. (Z_n) ist gleichmäßig integrierbar, da $0 \le Z_n = P(X_0|\mathbf{X}_{-n}^{-1}) \le 1 \quad \forall\, n \in \mathbb{N}$. Aus $P(X_0|\mathbf{X}_{-n}^{-1}) \le 1$ folgt $V_n = \log P(X_0|\mathbf{X}_{-n}^{-1}) \le 0 \quad \forall\, n \in \mathbb{N}$. Zudem gilt $-\infty < H(X_0) = \mathbb{E}V_0 \le \mathbb{E}V_n \le 0 \quad \forall\, n \;\Rightarrow\; \sup_n \mathbb{E}|V_n| \le \infty$. Daher gibt es nach dem Konvergenzsatz von Doob ein $V \in \mathbf{L}_1$ mit $\lim_n V_n = V$ P–fs.

direkter Beweis: Da $\varphi(x) := \frac{1}{x}$ konvex ist, folgt aus der Jensen-Ungleichung

$$\mathbb{E}(\, P(X_0|\mathbf{X}_{-n-1}^{-1})|\mathbf{X}_{-n}^{0} = \mathbf{x}_{-n}^{0}) = \sum_{x_{-n-1}} \frac{1}{\frac{p(\mathbf{x}_{-n-1}^{-1})}{p(\mathbf{x}_{-n-1}^{0})}}\, p(x_{-n-1}|\mathbf{x}_{-n}^{0})$$

$$\ge \frac{1}{\sum\limits_{x_{-n-1}} \frac{p(\mathbf{x}_{-n-1}^{-1})\, p(\mathbf{x}_{-n-1}^{0})}{p(\mathbf{x}_{-n-1}^{0})\, p(\mathbf{x}_{-n}^{0})}} = \frac{p(\mathbf{x}_{-n}^{0})}{\sum\limits_{x_{-n-1}} p(\mathbf{x}_{-n-1}^{-1})} = \frac{p(\mathbf{x}_{-n}^{0})}{p(\mathbf{x}_{-n}^{-1})} = p(x_0|\mathbf{x}_{-n}^{-1}).$$

4. Geben Sie ein Beispiel für eine nichtergodische Quelle, für die trotzdem gilt $\lim_n -\frac{1}{n}\, \log P(\mathbf{X}_0^{n-1}) = const.$

   Lösung: Die Markoff-Kette mit $P_0 := \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ und $\Pi := \begin{pmatrix} \frac{1}{2}, & \frac{1}{2}, & 0, & 0 \\ \frac{1}{2}, & \frac{1}{2}, & 0, & 0 \\ 0, & 0, & \frac{1}{2}, & \frac{1}{2} \\ 0, & 0, & \frac{1}{2}, & \frac{1}{2} \end{pmatrix}$ ist stationär und nicht ergodisch ($A := \{\omega_1, \omega_2\}$ und $B := \{\omega_3, \omega_4\}$ sind invariant mit $P(A) = P(B) = \frac{1}{2}$). Aber aus $P(\mathbf{X}_0^{n-1}) = \frac{1}{4}\,\left(\frac{1}{2}\right)^{n-1}$ folgt $\lim_n -\frac{1}{n}\, \log P(\mathbf{X}_0^{n-1}) = \lim_n \frac{n+1}{n} = 1$.

5. Für eine binäre Quelle $(X_n)_{n\in\mathbb{N}_0}$ wird durch einen fairen Münzwurf entschieden, ob gilt $X_{2\,n} = X_{2\,n+1} \quad \forall\, n \in \mathbb{N}_0$ oder $X_{2\,n} = X_{2\,n-1} \quad \forall\, n \in \mathbb{N}_0$. Die geraden Glieder der Folge sind unabhängig und $B_{\frac{1}{2}}$-verteilt. Die Quelle wird so codiert, dass jedem Tupel $(x_{2\,n}, x_{2\,n+1})$ ein Codewort $c(x_{2\,n}, x_{2\,n+1})$ auf folgende Art zugeordnet wird: $c(0,0) := 0, c(1,1) := 10$, $c(1,0) := 110, c(0,1) := 111$. Man zeige, dasss die Quelle stationär und ergodisch ist, dass aber die durchschnittliche Anzahl von Codebuchstaben pro Nachrichtensymbol mit Wahrscheinlichkeit $\frac{1}{2}$ gegen $\frac{3}{4}$ und mit Wahrscheinlichkeit $\frac{1}{2}$ gegen $\frac{9}{8}$ konvergiert.

   Lösung: $M_0:\ X_{2\,n} = X_{2\,n+1}$, d.h. $P(X_{2\,n+1} = x|X_{2\,n} = x, M_0) = 1$ für $x = 0, 1$ und $P(X_{2\,n} = x|X_{2\,n-1} = y, M_0) = \frac{1}{4},\ x, y \in \{0,1\}$. Unter M_1 : $X_{2\,n} = X_{2\,n-1}$ gilt $P(X_{2\,n+1} = x|X_{2\,n} = y, M_1) = \frac{1}{4},\ x, y \in \{0,1\}$ und $P(X_{2\,n-1} = x|X_{2\,n} = x, M_1) = 1 \;\Rightarrow\; P(X_{2\,n} = x|X_{2\,n-1} = x, M_1) = 1$ für $x = 0, 1$. Daher gilt mit $\mathrm{sh}(\mathbf{x}_n) := \mathbf{x}_{n+1}$ und alle $m < n$ offensichtlich $P(\mathbf{X}_m^n = \mathbf{x}_m^n|M_i) = P(\mathrm{sh}(\mathbf{X}_m^n) = \mathbf{x}_m^n|M_{1-i}),\ i = 0, 1$. Das impliziert

$$P(\mathbf{X}_m^n = \mathbf{x}_m^n) = P(M_0)\,P(\mathbf{X}_m^n = \mathbf{x}_m^n | M_0) + P(M_1)\,P(\mathbf{X}_m^n = \mathbf{x}_m^n | M_1)$$
$$= \frac{1}{2} P(\mathrm{sh}(\mathbf{X}_m^n) = \mathbf{x}_m^n | M_1) + \frac{1}{2} P(\mathrm{sh}(\mathbf{X}_m^n) = \mathbf{x}_m^n | M_0) = P(\mathrm{sh}(\mathbf{X}_m^n) = \mathbf{x}_m^n).$$

Ist A invariant, so gilt $A = \mathrm{sh}^{-n}(A) \in \mathfrak{S}(\mathbf{X}_n^\infty) \quad \forall\ n$, d.h. A ist terminal. Aber $\mathfrak{S}(\mathbf{X}_0^n)$ und $\mathfrak{S}(\mathbf{X}_{n+2}^\infty)$ sind unabhängig. Somit ist jede terminale Menge von sich selbst unabhängig, d.h. $P(A) = P(A)^2$. Daher ist die Quelle (X_n) ergodisch.
Unter M_0 treten die Tupel $(x_{2\,n}, x_{2\,n+1}) = (0,0)$ und $(x_{2\,n}, x_{2\,n+1}) = (1,1)$ mit Wahrscheinlichkeit $\frac{1}{2}$ auf. Daher treten nur die Codewortlängen 1 oder 2 gleichwahrscheinlich auf. Das ergibt $\frac{3}{4}$ Codebuchstaben pro Symbol.
Unter M_1 treten alle Tupel $(x_{2\,n}, x_{2\,n+1})$ gleichwahrscheinlich auf. Somit ist die mittlere Wortlänge $\frac{1}{4}\,(1+2+3+3) = \frac{9}{4}$. Das entspricht $\frac{9}{8}$ Codebuchstaben pro Symbol.