

Hints or solutions for selected exercises

- 1.1.2. Use Exercise A.3.5 to approximate characteristic functions by continuous functions.
- 1.2.5. Show that if $N > 1/\mu(A)$, then there exists $j \in V_A$ with $0 \leq j \leq N$. Adapting the proof of the previous statement, conclude that if K is a set of non-negative integers with $\#K > 1/\mu(A)$, then we may find $k_1, k_2 \in K$ and $n \in V_A$ such that $n = k_1 - k_2$. That is, the set $K - K = \{k_1 - k_2; k_1, k_2 \in K\}$ intersects V_A . To conclude that S is syndetic assume, by contradiction, that for every $n \in \mathbb{N}$ there is some number l_n such that $\{l_n, l_n + 1, \dots, l_n + n\} \cap V_A = \emptyset$. Consider an element $k_1 \notin V_A$ and construct, recursively, a sequence $k_{j+1} = l_{k_j} + k_j$. Prove that the set $K = \{k_1, \dots, k_N\}$ is such that $(K - K) \cap V_A = \emptyset$.
- 1.2.6. Otherwise, there exists $k \geq 1$ and $b > 1$ such that the set $B = \{x \in [0, 1] : n|f^n(x) - x| > b \text{ for every } n \geq k\}$ has positive measure. Let $a \in B$ be a density point of B . Consider $E = B \cap B(a, r)$, for r small. Get a lower estimate for the return time to E of any point of $x \in E$ and use the Kač theorem to reach a contradiction.
- 1.3.5. Consider the sequence $\log_{10} a_n$, where \log_{10} denotes the base 10 logarithm, and observe that $\log_{10} 2$ is an irrational number.
- 1.3.12. Consider orthonormal bases $\{v_1, \dots, v_d\}$, at x , and $\{w_1, \dots, w_d\}$, at $f(x)$, such that v_1 and w_1 are orthogonal to H_c . Check that $\text{grad}H(f(x)) \cdot Df(x)v = \text{grad}H(x) \cdot v$ for every v . Deduce that the matrix of $Df(x)$ with respect to those bases has the form

$$Df(x) = \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ \beta_2 & \gamma_{2,2} & \cdots & \gamma_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_d & \gamma_{d,2} & \cdots & \gamma_{d,d} \end{pmatrix},$$

with $\|\text{grad}H(f(x))\| |\alpha| = \|\text{grad}H(x)\|$. Note that $\Gamma = (\gamma_{ij})_{ij}$ is the matrix of $D(f|H_c)$ and observe that $|\det \Gamma| = \|\text{grad}H(x)\|/\|\text{grad}H(f(x))\|$. Using the formula of change of variables, conclude that $f|H_c$ preserves the measure $ds/\|\text{grad}H\|$.

- 1.4.4. Choose a set $E \subset M$ with measure less than ε/n and, for each $k \geq 1$, let E_k be the set of points $x \in E$ that return to E in exactly k iterates. Take for B the union of the sets E_k , with $k \geq n$, of the n -th iterates of the sets E_k with $k \geq 2n$, and so

- on. For the second part, observe that if (f, μ) is aperiodic then μ cannot have atoms.
- 1.4.5. By assumption, $f^\tau(y) \in H_{n-\tau(y)}$ whenever $y \in H_n$ with $n > \tau(y)$. Therefore, $T(y) \in H$ if $y \in H$. Consider $A_n = \{1 \leq j \leq n : x \in H_j\}$ and $B_n = \{l \geq 1 : \sum_{i=0}^l \tau(T^i(x)) \leq n\}$. Show, by induction, that $\#A_n \leq \#B_n$ and deduce that $\limsup_n \#B_n/n \geq \theta$. Now suppose that $\liminf_k (1/k) \sum_{i=0}^{k-1} \tau(T^i(x)) > (1/\theta)$. Show that there exists $\theta_0 < \theta$ such that $\#B_n < \theta_0 n$, for every n sufficiently large. This contradicts the previous conclusion.
- 1.5.5. Observe that the maps f, f^2, \dots, f^k commute with each other and then use the Poincaré multiple recurrence theorem.
- 1.5.6. By definition, the complement of $\Omega(f_1, \dots, f_q)^c$ is an open set. The Birkhoff multiple recurrence theorem ensures that the non-wandering set is non-empty.
- 2.1.6. Consider the image $V_*\mu$ of the measure μ under V . Check that $V_*\mu((a, b]) = F(b) - F(a)$ for every $a < b$. Consequently, $V_*\mu(\{b\}) = F(b) - \lim_{a \rightarrow b} F(a)$. Therefore, $(-\infty, b]$ is a continuity set for $V_*\mu$ if and only if b is a continuity point for F . Using Theorem 2.1.2, it follows that if $(V_{k*}\mu)_k$ converges to $V_*\mu$ in the weak* topology then $(V_k)_k$ converges to V in distribution. Conversely, if $(V_k)_k$ converges to V in distribution then $V_{k*}\mu((a, b]) = F_k(b) - F_k(a)$ converges to $F(b) - F(a) = V_*\mu((a, b])$, for any continuity points $a < b$ of F . Observing that such intervals $(a, b]$ generate the Borel σ -algebra of the real lines, conclude that $(V_{k*}\mu)_k$ converges to $V_*\mu$ in the weak* topology.
- 2.1.8. (Billingsley [Bil68]) Use the hypothesis to show that if $(U_n)_n$ is an increasing sequence of open subsets of M such that $\bigcup_n U_n = M$ then, for every $\varepsilon > 0$ there exists n such that $\mu(U_n) \geq 1 - \varepsilon$ for every $\mu \in \mathcal{K}$. Next, imitate the proof of Proposition A.3.7.
- 2.2.2. For the first part of the statement use induction in q . The case $q = 1$ corresponds to Theorem 2.1. Consider continuous transformations $f_i : M \rightarrow M$, $1 \leq i \leq q + 1$ commuting with each other. By the induction hypothesis, there exists a probability ν invariant under f_i for $1 \leq i \leq q$. Define $\mu_n = (1/n) \sum_{j=0}^{n-1} (f_{q+1})_*^j(\nu)$. Note that $(f_i)_*\mu_n = \mu_n$ for every $1 \leq i \leq q$ and every n . Hence, every accumulation point of $(\mu_n)_n$ is invariant under every f_i , $1 \leq i \leq q$. By compactness, there exists some accumulation point $\mu \in \mathcal{M}_1(M)$. Check that μ is invariant under f_{q+1} . For the second part, denote by $M_q \subset \mathcal{M}_1(M)$ the set of probability measures invariant under f_i , $1 \leq i \leq q$. Then, $(M_q)_q$ is a non-increasing sequence of closed non-empty subsets of $\mathcal{M}_1(M)$. By compactness, the intersection $\bigcap_q M_q$ is non-empty.
- 2.2.6. Define μ in each iterate $f^j(W)$, $j \in \mathbb{Z}$ by letting $\mu(A) = m(f^{-j}(A))$ for each measurable set $A \subset f^j(W)$.
- 2.3.2. Clearly, convergence in norm implies weak convergence. To prove the converse, assume that $(x^k)_k$ converges to zero in the weak topology but not in the norm topology. The first condition implies that, for every fixed N , the sum $\sum_{n=0}^N |x_n^k|$ converges to zero when $k \rightarrow \infty$. The second condition means that, up to restricting to a subsequence, there exists $\delta > 0$ such that $\|x^k\| > \delta$

for every k . Then, there exists some increasing sequence $(l_k)_k$ such that

$$\sum_{n=0}^{l_{k-1}} |x_n^k| \leq \frac{1}{k} \quad \text{but} \quad \sum_{n=0}^{l_k} |x_n^k| \geq \|x^k\| - \frac{1}{k} \geq \delta - \frac{1}{k} \quad \text{for every } k.$$

Take $a_n = \overline{x_n^k}/|x_n^k|$ for each $l_{k-1} < n \leq l_k$. Then, for every k ,

$$\left| \sum_{n=0}^{\infty} a_n x_n^k \right| \geq \sum_{l_{k-1} < n \leq l_k} |x_n^k| - \sum_{n \leq l_{k-1}} |x_n^k| - \sum_{n > l_k} |x_n^k| \geq \|x^k\| - \frac{4}{k} \geq \delta - \frac{4}{k}.$$

This contradicts the hypotheses. Now take $x_n^k = 1$ if $k = n$ and $x_n^k = 0$ otherwise. Given any $(a_n)_n \in c_0$, we have that $\sum_n a_n x_n^k = a_k$ converges to zero when $k \rightarrow \infty$. Therefore, $(x^k)_k$ converges to zero in the weak* topology. But $\|x^k\| = 1$ for every k , hence $(x^k)_k$ does not converge to zero in the norm topology.

- 2.3.6. Take $W = U(H)^\perp$ and $V = (\bigoplus_{n=0}^\infty U^n(W))^\perp$.
- 2.3.7. Suppose that there exist tangent functionals T_1 and T_2 with $T_1(v) > T_2(v)$ for some $v \in E$. Show that $\phi(u + tv) + \phi(u - tv) - 2\phi(u) \geq t(T_1(u) - T_2(u))$ for every t and deduce that ϕ is not differentiable in the direction of v .
- 2.4.1. Consider the set \mathcal{P} of all probability measures on $X \times M$ of the form $\nu^{\mathbb{Z}} \times \eta$. Note that \mathcal{P} is compact in the weak* topology and is invariant under the operator F_* .
- 2.4.2. The condition $\hat{p} \circ g = \hat{f} \circ \hat{p}$ entails $\hat{f}^n \circ \hat{p} = \hat{p} \circ g^n$ for every $n \in \mathbb{Z}$. Using $\pi \circ \hat{p} = p$, it follows that $\pi \circ \hat{f}^n \circ \hat{p} = p \circ g^n$ for every $n \leq 0$. Therefore, $\hat{p}(y) = (p(g^n(y)))_{n \leq 0}$. This proves the existence and uniqueness of \hat{p} . Now suppose that p is surjective. The hypotheses of compactness and continuity ensure that

$$(g^{-n}(p^{-1}(\{x_n\})))_{n \leq 0}$$

is a nested sequence of compact sets, for every $(x_n)_{n \leq 0} \in \hat{M}$. Take y in the intersection and note that $\hat{p}(y) = (x_n)_{n \leq 0}$.

- 2.5.2. Fix q and l . Assume that for every $n \geq 1$ there exists a partition $\{S_1^n, \dots, S_l^n\}$ of the set $\{1, \dots, n\}$ such that no subset of S_j^n contains an arithmetic progression of length q . Consider the function $\phi_n : \mathbb{N} \rightarrow \{1, \dots, l\}$ given by $\phi_n(i) = j$ if $i \in S_j$ and $\phi_n(i) = l$ if $i > n$. Take $(n_k)_k \rightarrow \infty$ such that the subsequence $(\phi_{n_k})_k$ converges at every point to some function $\phi : \mathbb{N} \rightarrow \{1, \dots, l\}$. Consider $S_j = \phi^{-1}(j)$ for $j = 1, \dots, l$. Some S_j contains some arithmetic progression of length q . Then $S_j^{n_k}$ contains that arithmetic progression for every k sufficiently large.
- 2.5.4. Consider $\Sigma = \{1, \dots, l\}^{\mathbb{N}^k}$ with the distance $d(\omega, \omega') = 2^{-N}$ where $N \geq 0$ is largest such that $\omega(i_1, \dots, i_k) = \omega'(i_1, \dots, i_k)$ for every $i_1, \dots, i_k < N$. Note that Σ is a compact metric space. Given $q \geq 1$, let $F_q = \{(a_1, \dots, a_k) : 1 \leq a_i \leq q \text{ and } 1 \leq i \leq k\}$. Let e_1, \dots, e_m be an enumeration of the elements of F_q . For each $j = 1, \dots, m$, consider the shift map $\sigma_j : \Sigma \rightarrow \Sigma$ given by $(\sigma_j \omega)(n) = \omega(n + e_j)$ for $n \in \mathbb{N}^k$. Consider the point $\omega \in \Sigma$ defined by $\omega(n) = i \Leftrightarrow n \in S_i$. Let Z be the closure of $\{\sigma_1^{l_1} \cdots \sigma_m^{l_m}(\omega) : l_1, \dots, l_m \in \mathbb{N}\}$. Note that Z is invariant under the shift maps σ_j . By the Birkhoff multiple recurrence theorem, there exist $\zeta \in Z$ and $s \geq 1$ such that $d(\sigma_j^s(\zeta), \zeta) < 1$ for every $j = 1, \dots, m$. Let $e = (1, \dots, 1) \in \mathbb{N}^k$. Then $\zeta(e) = \zeta(e + se_1) = \dots = \zeta(e + se_m)$. Consider $\sigma_1^{l_1} \cdots \sigma_m^{l_m}(\omega)$ close enough to ζ that $\omega(b) = \omega(b + se_1) = \dots = \omega(b + se_m)$, where $b = e + l_1 e_1 +$

$\dots + l_m e_m$. It follows that if $i = \omega(b)$, then $b + sF_q \subset S_i$. Given that there are only finitely many sets S_i , some of them must contain infinitely many sets of the type $b + sF_q$, with q arbitrarily large.

- 3.1.1. Mimic the proof of Theorem 3.1.6.
- 3.1.2. Suppose that for every $k \in \mathbb{N}$ there exists $n_k \in \mathbb{N}$ such that $\mu(A \cap f^{-j}(A)) = 0$ for every $n_k + 1 \leq j \leq n_k + k$. It is no restriction to assume that $(n_k)_k \rightarrow \infty$. Take $\varphi = \chi_A$. By Exercise 3.1.1, $(1/k) \sum_{j=n_k+1}^{n_k+k} \varphi \cdot \varphi \circ f^j \rightarrow \varphi \cdot P(\varphi)$. The left-hand side is identically zero and the right-hand side is equal to $\|P(\varphi)\|^2$. Hence, the time average $P(\varphi) = 0$ and so $\mu(A) = \int P(\varphi) d\mu = 0$.
- 3.2.3. (a) Consider $\varepsilon = 1$ and let $C = \sup\{|\varphi(l)| : |l| \leq L(1)\}$. Given $n \in \mathbb{Z}$, fix $s \in \mathbb{Z}$ such that $sL(1) < n \leq (s+1)L(1)$. By hypothesis, there exists $\tau \in \{sL(1) + 1, \dots, (s+1)L(1)\}$ such that $|\varphi(k + \tau) - \varphi(k)| < 1$ for every $k \in \mathbb{Z}$. Take $k = n - \tau$ and observe that $|k| \leq L(1)$. It follows that $|\varphi(n)| < 1 + C$. (b) Take $\rho\varepsilon > 2L(\varepsilon) \sup|\varphi|$. For every $n \in \mathbb{Z}$ there exists some ε -quasi-period $\tau = n\rho + r$ with $1 \leq r \leq L(\varepsilon)$. Then,

$$\left| \sum_{j=n\rho+1}^{(n+1)\rho} \varphi(j) - \sum_{j=1-r}^{\rho-r} \varphi(j) \right| < \rho\varepsilon \quad \text{and} \quad \left| \sum_{j=1-r}^{\rho-r} \varphi(j) - \sum_{j=1}^{\rho} \varphi(j) \right| \leq 2r \sup|\varphi| < \rho\varepsilon.$$

(c) Given $\varepsilon > 0$, take ρ as in part (b). For each $n \geq 1$, write $n = s\rho + r$, with $1 \leq r \leq \rho$. Then,

$$\frac{1}{n} \sum_{j=1}^n \varphi(j) = \frac{\rho}{s\rho + r} \sum_{i=0}^{s-1} \frac{1}{\rho} \sum_{l=i\rho+1}^{(i+1)\rho} \varphi(l) + \frac{1}{n} \sum_{l=s\rho+1}^{s\rho+r} \varphi(l).$$

For s large, the first term on the right-hand side is close to $(1/\rho) \sum_{j=0}^{\rho-1} \varphi(j)$ (by part (b)) and the last term is close to zero (by part (a)). Conclude that the left-hand side of the identity is a Cauchy sequence. (d) Observe that

$$\left| \frac{1}{n} \sum_{j=1}^n \varphi(x+k) - \frac{1}{n} \sum_{j=1}^n \varphi(j) \right| \leq \frac{2|x|}{n} \sup|\varphi|$$

and use parts (a) and (c).

- 3.3.3. Let μ be a probability measure invariant under a flow $f^t : M \rightarrow M$, $t \in \mathbb{R}$ and let $(\varphi_s)_{s>0}$ be a family of functions, indexed by the positive real numbers, such that $\varphi_{s+t} \leq \varphi_t + \varphi_s \circ f^t$ and the function $\Phi = \sup_{0 < s < 1} \varphi_s^+$ is in $L^1(\mu)$. Then, $(1/T)\varphi_T$ converges at μ -almost every point to a function φ such that $\varphi^+ \in L^1(\mu)$ and $\int \varphi d\mu = \lim_{T \rightarrow \infty} (1/T) \int \varphi_T d\mu$. To prove this, take $\varphi = \lim_n (1/n)\varphi_n$ (Theorem 3.3.3). For $T > 0$ non-integer, write $T = n + s$ with $N \in \mathbb{N}$ and $s \in (0, 1)$. Then,

$$\varphi_T \leq \varphi_n + \varphi_s \circ f^n \leq \varphi_n + \Phi \circ f^n \quad \text{and} \quad \varphi_T \geq \varphi_{n+1} - \varphi_{1-s} \circ f^T \geq \varphi_n - \Phi \circ f^T.$$

Using Lemma 3.2.5, the first inequality shows that $\limsup_{T \rightarrow \infty} (1/T)\varphi_T \leq \varphi$. Analogously, using the version of Lemma 3.2.5 for continuous time, the second inequality above gives that $\liminf_{T \rightarrow \infty} (1/T)\varphi_T \geq \varphi$. It also follows that $\lim_{T \rightarrow \infty} (1/T) \int \varphi_T d\mu$ coincides with $\lim_n (1/n) \int \varphi_n d\mu$. By Theorem 3.3.3, this last limit is equal to $\int \varphi d\mu$.

- 3.3.6. Since $\log^+ \|\phi\| \in L^1(\mu)$, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\mu(B) < \delta$ implies $\int_B \log^+ \|\theta\| d\mu < \varepsilon$. Using that $\log^+ \|\phi^n\| \leq \sum_{j=0}^{n-1} \log^+ \|\theta\| \circ f^j$, one gets that

$$\mu(E) < \delta \Rightarrow \frac{1}{n} \int_E \log^+ \|\phi^n\| d\mu \leq \frac{1}{n} \sum_{j=0}^{n-1} \int_{f^{-j}(E)} \log^+ \|\theta\| d\mu \leq \varepsilon.$$

- 3.4.3. Consider local coordinates $x = (x_1, x_2, \dots, x_d)$ such that Σ is contained in $\{x_1 = 0\}$. Write $\nu = \psi(x) dx_1 dx_2 \dots dx_d$. Then $\nu_\Sigma = \psi(y) dx_2 \dots dx_d$ with $y = (0, x_2, \dots, x_d)$. Given $A \subset \Sigma$ and $\delta > 0$, the map $\zeta : (t, y) \mapsto g^t(y)$ is a diffeomorphism from $[0, \delta] \times A$ to A_δ . Therefore, $\nu(A_\delta) = \int_{[0, \delta] \times A} (\psi \circ \zeta) |\det D\zeta| dt dx_2 \dots dx_d$ and, consequently,

$$\lim_{\delta \rightarrow 0} \frac{\nu(A_\delta)}{\delta} = \int_A \psi(y) |\det D\zeta|(y) dx_2 \dots dx_d.$$

Next, note that $|\det D\zeta|(y) = |X(y) \cdot (\partial/\partial t)| = \phi(y)$ for every $y \in \Sigma$. It follows that the flux of ν coincides with the measure $\eta = \phi\nu_\Sigma$. In particular, η is invariant under the Poincaré map.

- 4.1.2. Use the theorem of Birkhoff and the dominated convergence theorem.
- 4.1.8. Assume that $U_f \varphi = \lambda \varphi$. Since U_f is an isometry, $|\lambda| = 1$. If $\lambda^n = 1$ for some n then $\varphi \circ f^n = \varphi$ and, by ergodicity, φ is constant almost everywhere. Otherwise, given any $c \neq 0$, the sets $\varphi^{-1}(\lambda^{-k}c)$, $k \geq 0$ are pairwise disjoint. Since they all have the same measure, this measure must be zero. Finally, the set $\varphi^{-1}(c)$ is invariant under f and, consequently, its measure is either zero or total.
- 4.2.4. Let K be such a set. We may assume that K contains an infinite sequence of periodic orbits $(O_n)_n$ with period going to infinity. Let $Y \subset K$ be the set of accumulation points of that sequence. Show that Y cannot consist of a single point. Let $p \neq q$ be periodic points in Y and z be a *heteroclinic point*, that is, such that $\sigma^n(z)$ converges to the orbit of p when $n \rightarrow -\infty$ and to the orbit of q when $n \rightarrow +\infty$. Show that $z \in Y$ and deduce the conclusion of the exercise.
- 4.2.10. Let $J_k = (0, 1/k)$, for each $k \geq 1$. Check that the continued fraction expansion of x is of bounded type if and only if there exists $k \geq 1$ such that $G^n(x) \notin J_k$ for every n . Observe that $\mu(J_k) > 0$ for every k . Deduce that for every k and μ -almost every x there exists $n \geq 1$ such that $G^n(x) \in J_k$. Conclude that \mathcal{L} has zero Lebesgue measure.
- 4.2.11. For each $L \in \mathbb{N}$, consider $\varphi_L(x) = \min\{\phi(x), L\}$. Then, $\varphi_L \in L^1(\mu)$ and, by ergodicity, $\tilde{\varphi}_L = \int \varphi_L d\mu$ at μ -almost every point. To conclude, observe that $\tilde{\phi} \geq \tilde{\varphi}_L$ for every L and $\int \varphi_L d\mu \rightarrow +\infty$.
- 4.3.7. Let $M = \{0, 1\}^{\mathbb{N}}$ and, for each n , let μ_n be the invariant measure supported on the periodic orbit $\alpha^n = (\alpha_k^n)_k$, with period $2n$, defined by $\alpha_k^n = 0$ if $0 \leq k < n$ and $\alpha_k^n = 1$ if $n \leq k < 2n$. Show that $(\mu_n)_n$ converges to $(\delta_0 + \delta_1)/2$, where 0 and 1 are the fixed points of the shift map.
- 4.3.9. (a) Take $k \geq 1$ such that every cylinder of length k has diameter less than δ . Take $y = (y_j)$ defined by $y_{j+n_i} = x_j^i$ for each $0 \leq j < m_i + k$. (b) Take $\delta > 0$ such that $d(z, w) < \delta$ implies $|\varphi(z) - \varphi(w)| < \varepsilon$ and consider $k \geq 1$ given by part (a). Choose m_i , $i = 1, \dots, s$ such that $m_i/n_s \approx \alpha_i$ for every i . Then take y as in part (a). (c) By the ergodic theorem, $\int \varphi d\mu = \int \tilde{\varphi} d\mu$. Take $x^1, \dots, x^s \in \Sigma$ and

$\alpha^1, \dots, \alpha^s$ such that $\int \tilde{\varphi} d\mu \approx \sum_i \alpha_i \tilde{\varphi}(x^i)$. Note that $\tilde{\varphi}(y) = \int \varphi dv_y$, where v_y is the invariant measure supported on the orbit of y . Recall Exercise 4.1.1.

- 4.4.3. On each side of the triangle, consider the foot of the corresponding height, that is, the orthogonal projection of the opposite vertex. Show that the trajectory defined by those three points is a periodic orbit of the billiard.
- 4.4.5. Using (4.4.10) and the twist condition, we get that for each $\theta \in \mathbb{R}$ there exists exactly one number $\rho_\theta \in (a, b)$ such that $\Theta(\theta, \rho_\theta) = \theta$. The function $\theta \mapsto \rho_\theta$ is continuous and periodic, with period 1. Consider its graph $\Gamma = \{(\theta, \rho_\theta) : \theta \in S^1\}$. Every point in $\Gamma \cap f(\Gamma)$ is fixed under f : if $(\theta, \rho_\theta) = f(\gamma, \rho_\gamma) = (\Theta(\gamma, \rho_\gamma), R(\gamma, \rho_\gamma))$ then, since $\Theta(\gamma, \rho_\gamma) = \gamma$, it follows that $\theta = \gamma$ and so $\rho_\theta = \rho_\gamma$. Since f preserves the area measure, none of the connected components of $A \setminus \Gamma$ may be mapped inside itself. This implies that $f(\Gamma)$ intersect Γ at no less than two points.
- 4.4.7. Taking inspiration from Example 4.4.12, show that the billiard map in Ω extends to a *Dehn twist* in the annulus $A = S^1 \times [-\pi/2, \pi/2]$, that is, a homeomorphism $f : A \rightarrow A$ that coincides with the identity on both boundary components but is homotopically non-trivial: actually, f admits a lift $F : \mathbb{R} \times [-\pi/2, \pi/2] \rightarrow \mathbb{R} \times [-\pi/2, \pi/2]$ such that $F(s, -\pi/2) = (s - 2\pi, -\pi/2)$ and $F(s, \pi/2) = (s, \pi/2)$ for every s . Consider rational numbers $p_n/q_n \in (-2\pi, 0)$ with $q_n \rightarrow \infty$. Use Exercise 4.4.6 to show that g has periodic points of period q_n . One way to ensure that these periodic points are all distinct is to take the q_n mutually prime.
- 5.1.7. The statement does not depend on the choice of the ergodic decomposition, since the latter is essentially unique. Consider the construction in Exercise 5.1.6. The set M_0 is saturated by the partition \mathcal{W}^s , that is, if $x \in M_0$ then $\mathcal{W}^s(x) \subset M_0$. Moreover, the map $y \mapsto \mu_y$ is constant on each $\mathcal{W}^s(x)$. Since the partition \mathcal{P} is characterized by $\mathcal{P}(x) = \mathcal{P}(y) \Leftrightarrow \mu_x = \mu_y$, it follows that $\mathcal{P} \prec \mathcal{W}^s$ restricted to M_0 .
- 5.2.1. Consider the canonical projections $\pi_{\mathcal{P}} : M \rightarrow \mathcal{P}$ and $\pi_{\mathcal{Q}} : M \rightarrow \mathcal{Q}$, the quotient measures $\hat{\mu}_{\mathcal{P}} = (\pi_{\mathcal{P}})_* \mu$ and $\hat{\mu}_{\mathcal{Q}} = (\pi_{\mathcal{Q}})_* \mu$ and the disintegrations $\mu = \int \mu_P d\hat{\mu}_{\mathcal{P}}(P)$ and $\mu = \int \mu_Q d\hat{\mu}_{\mathcal{Q}}(Q)$. Moreover, for each $P \in \mathcal{P}$, consider $\hat{\mu}_{P, \mathcal{Q}} = (\pi_{\mathcal{Q}})_* \mu_P$ and the disintegration $\mu_P = \int \mu_{P, Q} d\hat{\mu}_{P, \mathcal{Q}}(Q)$. Observe that $\int \hat{\mu}_{P, \mathcal{Q}} d\hat{\mu}_{\mathcal{P}}(P) = \hat{\mu}_{\mathcal{Q}}$: given any $B \subset \mathcal{Q}$,

$$\int \hat{\mu}_{P, \mathcal{Q}}(B) d\hat{\mu}_{\mathcal{P}}(P) = \int \mu_P(\pi_{\mathcal{Q}}^{-1}(B)) d\hat{\mu}_{\mathcal{P}}(P) = \mu(\pi_{\mathcal{Q}}^{-1}(B)) = \hat{\mu}_{\mathcal{Q}}(B).$$

To check that $\mu_{\pi(Q), Q}$ is a disintegration of μ with respect to \mathcal{Q} : (a) $\mu_{P, Q}(Q) = 1$ for $\hat{\mu}_{P, \mathcal{Q}}$ -almost every Q and $\hat{\mu}_{\mathcal{P}}$ -almost every P . Moreover, $\mu_{P, Q} = \mu_{\pi(Q), Q}$ for $\hat{\mu}_{P, \mathcal{Q}}$ -almost every Q and $\hat{\mu}_{\mathcal{P}}$ -almost every P , because $\mu_P(P) = 1$ for $\hat{\mu}_{\mathcal{P}}$ -almost every P . By the previous observation, it follows that $\mu_{\pi(Q), Q}(Q) = 1$ for $\hat{\mu}_{\mathcal{Q}}$ -almost every Q . (b) $P \mapsto \mu_P(E)$ is measurable, up to measure zero, for every Borel set $E \subset M$. By construction (Section 5.2.3), there exists a countable generating algebra \mathcal{A} such that $\mu_{P, Q}(E) = \lim_n \mu_P(E \cap Q_n) / \mu_P(Q_n)$ for every $E \in \mathcal{A}$ (where Q_n is the element of \mathcal{Q}_n that contains Q). Deduce that $P \mapsto \mu_{\pi(Q), Q}(E)$ is measurable, up to measure zero, for every $E \in \mathcal{A}$. Extend this conclusion to every Borel set E , using the monotone class argument in

Section 5.2.3. (c) Note that $\mu = \int \mu_P d\hat{\mu}_{\mathcal{P}}(P) = \int \int \mu_{P,Q} d\hat{\mu}_{P,Q}(Q) d\hat{\mu}_{\mathcal{P}}(P) = \int \int \mu_{\pi(Q),Q} d\hat{\mu}_{P,Q}(Q) d\hat{\mu}_{\mathcal{P}}(P) = \int \mu_{\pi(Q),Q} d\hat{\mu}_{\mathcal{Q}}(Q)$.

5.2.2. Argue that the partition \mathcal{Q} of the space $\mathcal{M}_1(M)$ into points is measurable. Given any disintegration $\{\mu_P : P \in \mathcal{P}\}$, consider the measurable map $M \mapsto \mathcal{M}_1(M)$, $x \mapsto \mu_{P(x)}$. The pre-image of \mathcal{Q} under this map is a measurable partition. Check that this pre-image coincides with \mathcal{P} on a subset with full measure.

6.1.3. The function φ is invariant.

6.2.5. Denote by X the closure of the orbit of x . If X is minimal, for each $y \in X$ there exists $n(y) \geq 1$ such that $d(f^{n(y)}(y), x) < \varepsilon$. Then, by continuity, y admits an open neighborhood $V(y)$ such that $d(f^{n(y)}(z), x) < \varepsilon$ for every $z \in V(y)$. Take y_1, \dots, y_s such that $X \subset \bigcup_i V(y_i)$ and let $m = \max_i n(y_i)$. Given any $k \geq 1$, take i such that $f^k(x) \in V(y_i)$. Then, $d(f^{k+n_i}(x), x) < \varepsilon$, that is, $k + n_i \in R_\varepsilon$. This proves that, given any $m + 1$ consecutive integers, at least one of them is in R_ε . Hence, R_ε is syndetic. Now assume that X is not minimal. Then, there exists a non-empty, closed invariant set F properly contained in X . Note that $x \notin F$ and so, for every ε sufficiently small, there exists an open set U that contains F and does not intersect $B(x, \varepsilon)$. On the other hand, since R_ε is syndetic, there exists $m \geq 1$ such that for any $k \geq 1$ there exists $n \in \{k, \dots, k + m\}$ satisfying $f^n(x) \in B(x, \varepsilon)$. Take k such that $f^k(x) \in U_1$, where $U_1 = U \cap f^{-1}(U) \cap \dots \cap f^{-m}(U)$, and find a contradiction.

6.2.6. By Exercise 6.2.5, the set $R_\varepsilon = \{n \in \mathbb{N} : d(x, f^n(x)) < \varepsilon\}$ is syndetic for every $\varepsilon > 0$. If y is close to x then $\{n \in \mathbb{N} : d(f^n(x), f^n(y)) < \varepsilon\}$ contains blocks of consecutive integers with arbitrary length, no matter the choice of $\varepsilon > 0$. Let U_1 be any neighborhood of x . It follows from the previous observations that there exist infinitely many values of $n \in \mathbb{N}$ such that $f^n(x), f^n(y)$ are in U_1 . Fix n_1 with this property. Next, consider $U_2 = U_1 \cap f^{-n_1}(U_1)$. By the previous step, there exists $n_2 > n_1$ such that $f^{n_2}(x), f^{n_2}(y) \in U_2$. Continuing in this way, construct a non-increasing sequence of open sets U_k and an increasing sequence of natural numbers n_k such that $f^{n_k}(U_{k+1}) \subset U_k$ and $f^{n_k}(x), f^{n_k}(y) \in U_k$. Check that $f^{n_{i_1} + \dots + n_{i_k}}(x)$ and $f^{n_{i_1} + \dots + n_{i_k}}(y)$ are in U_1 for any $i_1 < \dots < i_k, k \geq 1$.

6.2.7. Consider the shift map $\sigma : \Sigma \rightarrow \Sigma$ in $\Sigma = \{1, 2, \dots, q\}^{\mathbb{N}}$. The partition $\mathbb{N} = S_1 \cup \dots \cup S_q$ defines a certain element $\alpha = (\alpha_n) \in \Sigma$, given by $\alpha_n = i$ if and only if $n \in S_i$. Consider β in the closure of the orbit of α such that α and β are near and the closure of the orbit of β is a minimal set. Apply Exercise 6.2.6 with $x = \beta, y = \alpha$ and $U = [0; \alpha_0]$ to obtain the result.

6.3.6. Write $g = (a_{11}, a_{12}, a_{21}, a_{22})$. Then,

$$E_g(x_{11}, x_{12}, x_{21}, x_{22}) = (a_{11}x_{11} + a_{12}x_{21}, a_{11}x_{12} + a_{12}x_{22}, a_{21}x_{11} + a_{22}x_{21}, a_{21}x_{12} + a_{22}x_{22}).$$

Write the right-hand side as $(y_{11}, y_{12}, y_{21}, y_{22})$. Use the formula of change of variables, observing that $\det(y_{11}, y_{12}, y_{21}, y_{22}) = (\det g) \det(x_{11}, x_{12}, x_{21}, x_{22})$ and

$$dy_{11} dy_{12} dy_{21} dy_{22} = (\det g)^2 dx_{11} dx_{12} dx_{21} dx_{22}.$$

In the complex case, take

$$\int_{GL(2, \mathbb{R})} \varphi d\mu = \int \frac{\varphi(z_{11}, z_{12}, z_{21}, z_{22})}{|\det(z_{11}, z_{12}, z_{21}, z_{22})|^4} dx_{11} dy_{11} dx_{12} dy_{12} dx_{21} dy_{21} dx_{22} dy_{22},$$

where $z_{jk} = x_{jk} + y_{jk}i$. [Observation: Generalize these constructions to any dimension!]

- 6.3.9. Given $x \in M$, there exists a unique number $0 \leq r < 10^k$ such that $f^r(x) \in [b_0, \dots, b_{k-1}]$. Moreover, $f^n(x) \in [b_0, \dots, b_{k-1}]$ if and only if $n - r$ is a multiple of 10^k . Use this observation to conclude that

$$\tau([b_0, \dots, b_{k-1}], x) = 10^{-k} \quad \text{for every } x \in M.$$

Conclude that if f admits an ergodic probability measure μ then $\mu([b_0, \dots, b_{k-1}]) = 10^{-k}$ for every b_0, \dots, b_{k-1} . This determines μ uniquely. To conclude, show that μ is well defined and invariant.

- 6.3.11. Consider the sequence of words w_n defined inductively by $w_1 = \alpha$ and $s(w_{n+1}) = w_n$ for $n \geq 1$. Decompose the word $s(\alpha) = w_2 = \alpha r_1$ and prove, by induction, that w_{n+1} may be decomposed as $w_{n+1} = w_n r_n$, for some word r_n with length greater than or equal to n , such that $s(r_n) = r_{n+1}$. Define $w = \alpha r_1 r_2 \dots$ and note that $s(w) = s(\alpha)s(r_1)s(r_2)\dots = \alpha r_1 r_2 r_3 \dots = w$. This proves existence. To prove uniqueness, let $\gamma \in \Sigma$ be a sequence starting with α and such that $S(\gamma) = \gamma$. Decompose γ as $\gamma = \alpha \gamma_1 \gamma_2 \gamma_3 \dots$, in such a way that γ_i and r_i have the same length. Note that $S(\alpha) = \alpha \gamma_1 = \alpha r_1$, and so $\gamma_1 = r_1$. Conclude by induction.

- 6.4.2. Given any $0 \leq \alpha < \beta \leq 1$, we have that $\sqrt{n} \in (\alpha, \beta)$ in the circle if and only if there exists some integer $k \geq 1$ such that $k^2 + 2k\alpha + \alpha^2 < n < k^2 + 2k\beta + \beta^2$. For each k the number of values of n that satisfy this inequality is equal to the integer part of $2k(\beta - \alpha) + (\beta^2 - \alpha^2)$. Therefore,

$$\#\{1 \leq n < N^2 : \sqrt{n} \in (\alpha, \beta)\} \leq \sum_{k=1}^{N-1} 2k(\beta - \alpha) + (\beta^2 - \alpha^2)$$

and the difference between the term on the right and the one on the left is less than N . Hence,

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \#\{1 \leq n < N^2 : \sqrt{n} \in (\alpha, \beta)\} = \beta - \alpha.$$

A similar calculation shows that the sequence $(\log n \pmod{\mathbb{Z}})_n$ is not equidistributed in the circle. [Observation: But it does admit a continuous (non-constant) limit density. Calculate that density!]

- 6.4.3. Define $\phi_n = a^n + (-1/a)^n$. Check that $(\phi_n)_n$ is the Fibonacci sequence and, in particular, $\phi_n \in \mathbb{N}$ for every $n \geq 1$. Now observe that $(-1/a)^n$ converges to zero. Hence, $\{n \geq 1 : a^n \pmod{\mathbb{Z}} \in I\}$ is finite, for any interval $I \subset S^1$ whose closure does not contain zero.

- 7.1.1. It is clear that the condition is necessary. To see that it is sufficient: Given A , consider the closed subspace \mathcal{V} of $L^2(\mu)$ generated by the functions 1 and $\mathcal{X}_{f^{-k}(A)}$, $k \in \mathbb{N}$. The hypothesis ensures that $\lim_n U_f^n(\mathcal{X}_A) \cdot \mathcal{X}_{f^{-k}(A)} = (\mathcal{X}_A \cdot 1)(\mathcal{X}_{f^{-k}(A)} \cdot 1)$ for every k . Conclude that $\lim_n U_f^n(\mathcal{X}_A) \cdot \phi = (\mathcal{X}_A \cdot 1)(\phi \cdot 1)$ for every $\phi \in \mathcal{V}$. Given a measurable set B , write $\mathcal{X}_B = \phi + \phi^\perp$ with $\phi \in \mathcal{V}$ and $\phi^\perp \in \mathcal{V}^\perp$ to conclude that $\lim_n U_f^n(\mathcal{X}_A) \cdot \mathcal{X}_B = (\mathcal{X}_A \cdot 1)(\mathcal{X}_B \cdot 1)$.

- 7.1.2. Assuming that E exists, decompose $(1/n) \sum_{j=0}^{n-1} |a_j|$ into two terms, one over $j \in E$ and the other over $j \notin E$. The hypotheses imply that the two terms converge to zero. Conversely, assume that $(1/n) \sum_{j=0}^{n-1} |a_j|$ converges to zero. Define $E_n =$

$\{j \geq 0: |a_j| \geq (1/m)\}$ for each $m \geq 1$. The sequence $(E_m)_m$ is increasing and each E_m has density zero; in particular, there exists $\ell_m \geq 1$ such that $(1/n)\#\{0, \dots, n-1\} \cap E_m < (1/m)$ for every $n \geq \ell_m$. Choose $(\ell_m)_m$ increasing and define $E = \bigcup_m (E_m \cap \{\ell_m, \dots, \ell_{m+1} - 1\})$. For the second part of the exercise, apply the first part to both sequences, $(a_n)_n$ and $(a_n^2)_n$.

7.1.6. (Pollicott and Yuri [PY98]) It is enough to treat the case when $\int \varphi_j d\mu = 0$ for every j . Use induction on the number k of functions. The case $k = 1$ is contained in Theorem 3.1.6. Use the inequalities

$$\frac{1}{N} \sum_{n=1}^N a_n \leq \frac{1}{N} \sum_{n=1}^{N-m+1} \left(\frac{1}{m} \sum_{j=0}^{m-1} a_{n+j} \right) + \frac{m}{N} \left(\max_{1 \leq i \leq m} |a_i| + \max_{N-m \leq i \leq N} |a_i| \right)$$

$$\left(\frac{1}{N} \sum_{n=1}^N b_n \right)^2 \leq (1/N) \sum_{n=1}^N |b_n|^2$$

to conclude that $\int |(1/N) \sum_{j=0}^{N-1} (\varphi_1 \circ f^n) \cdots (\varphi_k \circ f^{kn})|^2 d\mu$ is bounded above by

$$\frac{1}{N} \sum_{n=1}^N \left(\int \left| \frac{1}{m} \sum_{j=0}^{m-1} (\varphi_1 \circ f^{n+j}) \cdots (\varphi_k \circ f^{k(n+j)}) \right|^2 d\mu + \left(\frac{2m}{N} + \frac{m^2}{N^2} \right) \left(\max_{1 \leq i \leq k} \sup_{j \geq 0} |\varphi_i| \right)^2 \right)$$

The integral is equal to

$$\sum_{i=0}^{m-1} \sum_{j=0}^{m-1} \int \prod_{l=1}^k \left(\varphi_l (\varphi_l \circ f^{l(j-i)}) \right) \circ f^{l(n+i)} d\mu.$$

By the induction hypothesis,

$$\frac{1}{N} \sum_{n=1}^N \prod_{l=2}^k \left(\varphi_l (\varphi_l \circ f^{l(j-i)}) \right) \circ f^{l(n+i)} \rightarrow \prod_{l=2}^k \int \varphi_l (\varphi_l \circ f^{l(j-i)}) d\mu$$

in $L^2(\mu)$, when $N \rightarrow \infty$. Therefore,

$$\frac{1}{N} \sum_{n=1}^N \int \prod_{l=1}^k \left(\varphi_l (\varphi_l \circ f^{l(j-i)}) \right) \circ f^{l(n+i)} d\mu \rightarrow \prod_{l=1}^k \int \varphi_l (\varphi_l \circ f^{l(j-i)}) d\mu$$

in $L^2(\mu)$, when $N \rightarrow \infty$. Combining these estimates,

$$\limsup_N \int \left| \frac{1}{N} \sum_{n=1}^N (\varphi_1 \circ f^n) \cdots (\varphi_k \circ f^{kn}) \right|^2 d\mu \leq \frac{1}{m^2} \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} \prod_{l=1}^k \int \varphi_l (\varphi_l \circ f^{l(j-i)}) d\mu.$$

Since (f, μ) is weak mixing, $\int \varphi_l (\varphi_l \circ f^{lr}) d\mu$ converges to 0 when $r \rightarrow \infty$, restricted to a set of values of l with density 1 at infinity (recall Exercise 7.1.2). Therefore, the expression on the right-hand side is close to zero when m is large.

7.2.5. The first statement is analogous to Exercise 7.2.1. The definition ensures that μ_k has memory k . Given $\varepsilon > 0$ and any (uniformly) continuous function $\varphi : \Sigma \rightarrow \mathbb{R}$, there exists $\kappa \geq 1$ such that $|\int_C \varphi d\eta - \varphi(x)\eta(C)| \leq \varepsilon \eta(C)$ for every $x \in C$, every cylinder C of length $l \geq \kappa$ and every probability measure η . Since $\mu = \mu_k$ for cylinders of length k , it follows that $|\int \varphi d\mu_k - \int \varphi d\mu| \leq \varepsilon$ for every $k \geq \kappa$. This proves that $(\mu_k)_k$ converges to μ in the weak* topology.

- 7.2.6. (a) Use that $P_{i,i}^{n_1+n_2} = \sum_j P_{i,j}^{n_1} P_{j,i}^{n_2}$. All the terms in this expression are non-negative and the term corresponding to $j = i$ is positive. (b) Up to replacing R by R/κ , we may suppose that $\kappa = 1$. Start by showing that if $S \subset \mathbb{Z}$ is closed under addition and subtraction then $S = a\mathbb{Z}$, where a is the smallest positive element of S . Use that fact to show that if a_1, \dots, a_s are positive integers with greatest common divisor equal to 1 then there exist integers b_1, \dots, b_s such that $b_1 a_1 + \dots + b_s a_s = 1$. Now take $a_1, \dots, a_s \in R$ such that their greatest common divisor is equal to 1. Using the previous observation, and the hypothesis that R is closed under addition, conclude that there exists $p, q \in R$ such that $p - q = 1$. To finish, show that R contains every integer $n \geq pq$. (c) Consider any $i, j \in X$ and let κ_i, κ_j be the greatest common divisors of $R(i), R(j)$, respectively. By irreducibility, there exist $k, l \geq 1$ such that $P_{i,j}^k > 0$ and $P_{j,i}^l > 0$. Deduce that if $n \in R(i)$ then $n + k + l \in R(j)$. In view of (b), this is possible only if $\kappa_i \geq \kappa_j$. Exchanging the roles of i and j , it also follows that $\kappa_i \leq \kappa_j$. If $\kappa \geq 2$ then, given any i , we have $P_{i,i}^n = 0$ for n arbitrarily large and so P cannot be aperiodic. Now suppose that $\kappa = 1$. Then, using (b) and the hypothesis that X is finite, there exists $m \geq 1$ such that $P_{i,i}^m > 0$ for every $i \in X$ and every $n \geq m$. Then, since P is irreducible and X is finite, there exists $k \geq 1$ such that for any i, j there exists $l \leq k$ such that $P_{i,j}^l > 0$. Deduce that $P_{i,j}^{m+k} > 0$ for every i, j and so P is aperiodic. (d) Fix any $i \in X$ and, for each $r \in \{0, \dots, \kappa - 1\}$, define $X_r = \{j \in X : \text{there exists } n \equiv r \pmod{\kappa} \text{ such that } P_{i,j}^n > 0\}$. Check that these sets X_r cover X and are pairwise disjoint. Show that the restriction of P^κ to each of them is aperiodic.
- 7.3.1. By the theorem of Darboux, there exist coordinates (x_1, x_2) in the neighborhood of any point of S such that $\omega = dx_1 \wedge dx_2$. Consider the expression of the vector field in those coordinates: $X = X_1(\partial/\partial x_1) + X_2(\partial/\partial x_2)$. Show that $\beta = X_1 dx - 2 - X_2 dx_1$ and so $d\beta = (\text{div } X) dx_1 \wedge dx_2$. Hence, β is closed if and only if the divergent of X vanishes.
- 7.3.5. Observe that f is invertible and if A is a d -adic interval of level $r \geq 1$ (that is, an interval of the form $A = [id^{-r}, (i+1)d^{-r}]$), then there exists $s \geq r$ such that $f(A)$ consists of d^{s-r} d -adic intervals of level s . Deduce that f preserves the Lebesgue measure. Show also that if A and B are d -adic intervals then, since σ has no periodic points, $m(f^k(A) \cap B) = m(A)m(B)$ for every large k .
- 7.4.2. (a) Given $y^1, y^2 \in M$, write $f^{-1}(y^j) = \{x_1^j, \dots, x_d^j\}$ with $d(x_1^j, x_2^j) \leq \sigma^{-1}d(y^1, y^2)$. Then,

$$|\mathcal{L}\varphi(y^1) - \mathcal{L}\varphi(y^2)| = \frac{1}{d} \sum_{j=1}^d |\varphi(x_1^j) - \varphi(x_2^j)| \leq K_\theta(\varphi) \sigma^{-\theta} d(y^1, y^2)^\theta.$$

- (b) It follows that $\|\mathcal{L}\varphi\| \leq \sup|\varphi| + \sigma^{-\theta} K_\theta(\varphi) \leq \|\varphi\|$ for every $\varphi \in \mathcal{E}$, and the identity holds if and only if φ is constant. Hence, $\|\mathcal{L}\| = 1$. (c) Let $J_n = [\inf \mathcal{L}^n \varphi, \sup \mathcal{L}^n \varphi]$. By part (a), the sequence $(J_n)_n$ is decreasing and the diameter of J_n converges to zero exponentially fast. Take v_φ to be the point in the intersection and note that $\|\mathcal{L}^n \varphi - v_\varphi\| = \sup|\mathcal{L}^n \varphi - v_\varphi| + K_\theta(\mathcal{L}^n \varphi)$. (d) The constant functions are eigenvectors of \mathcal{L} , associated with the eigenvalue $\lambda = 1$. It follows that $v_{\varphi+c} = v_\varphi + c$ for every $\varphi \in \mathcal{E}$ and every $c \in \mathbb{R}$. Then, $H = \{\varphi : v_\varphi = 0\}$ is a hyperplane of \mathcal{E} transverse to the line of constant functions. This

hyperplane is invariant under \mathcal{L} and, by part (c), the spectral radius of $\mathcal{L} \mid H$ is less or equal than $\sigma^{-\theta} < 1$. (e) By part (b), $\|\mathcal{L}^n \varphi - \mathcal{L}^n \psi\| \leq \|\mathcal{L}^k \varphi - \mathcal{L}^k \psi\|$ for every $n \geq k \geq 1$. Making $n \rightarrow \infty$, we get that $|\nu_\varphi - \nu_\psi| \leq \|\mathcal{L}^k \varphi - \mathcal{L}^k \psi\|$ for every $k \geq 1$. Using part (a) and making $k \rightarrow \infty$, we get that $|\nu_\varphi - \nu_\psi| \leq \sup |\varphi - \psi|$. Therefore, the linear operator $\psi \mapsto \nu_\psi$ is continuous, relative to the norm in the space $C^0(M)$.

8.1.2. Denote $X_i = X \cap [0; i]$ and $p_i = \mu([0; i])$, for $i = 1, \dots, k$. Since μ is a Bernoulli measure, $\mu(X_i) = p_i \mu(f(X_i))$. Hence, $\sum_i p_i \mu(f(X_i)) = 1$. Since $\sum_i p_i = 1$, it follows that $\mu(f(X_i)) = 1$ for every i . Consequently, $\bigcap_i f(X_i)$ has full measure. Take x in that intersection. If (f, μ) and (g, ν) are ergodically equivalent, there exists a bijection $\phi : X \rightarrow Y$ between full measure invariant subsets such that $\phi \circ f = g \circ \phi$. Take $x \in X$ with k pre-images x_1, \dots, x_k in X . The points $\phi(x_i)$ are pre-images of $\phi(x)$ for the transformation g . Hence, $k \leq l$; by symmetry, we also have that $l \leq k$.

8.2.5. Assume that (f, μ) is not weak mixing. By Theorem 8.2.1, there exists a non-constant function φ such that $U_f \varphi = \lambda \varphi$ for some $\lambda = e^{2\pi i \theta}$. By ergodicity, the absolute value of φ is constant μ -almost everywhere. Using that f^n is ergodic for every n (Exercise 4.1.8), θ is irrational and any set where φ is constant has measure zero. Given $\alpha < \beta$ in $[0, 2\pi]$, consider $A = \{x \in \mathbb{C} : \alpha \leq \arg(\varphi(x)) \leq \beta\}$. Show that for every $\varepsilon > 0$ there exists n such that $\mu(f^{-n}(A) \setminus A) < \varepsilon$. Show that, by choosing $|\beta - \alpha|$ sufficiently small, one gets to contradict the inequality in the statement.

8.2.7. Note that $f_{n+1}(x) = f_n(x)$ for every $x \in J_n$ that is not on the top of \mathcal{S}_n . Hence, (for example, arguing as in Exercise 6.3.10), $f(x) = f_n(x)$ for every $x \in [0, 1]$ and every n sufficiently large; moreover, f preserves the Lebesgue measure. Let $a_n = \#\mathcal{S}_n$ be the height of each pile \mathcal{S}_n . Denote by $\{I^e, I^c, I^d\}$ the partition of each $I \in \mathcal{S}_n$ into subintervals of equal length, ordered from left to right. (a) If A is a set with $m(A) > 0$ then for every $\varepsilon > 0$ there exists $n \geq 1$ and some interval $I \in \mathcal{S}_n$ such that $m(A \cap I) \geq (1 - \varepsilon)m(I)$. If A is invariant, it follows that $m(A \cap J) \geq (1 - \varepsilon)m(J)$ for every $J \in \mathcal{S}_n$. (b) Assume that $U_f \varphi = \lambda \varphi$. Since U_f is an isometry, $|\lambda| = 1$. By ergodicity, $|\varphi|$ is constant almost everywhere; we may suppose that $|\varphi| \equiv 1$. Initially, assume that there exists n and some interval $I \in \mathcal{S}_n$ such that the restriction of φ to I is constant. Take $x \in I^e$ and $y \in I^c$ and $z \in I^d$. Then, $\varphi(x) = \varphi(y) = \varphi(z)$ and $\varphi(y) = \lambda^{a_n} \varphi(x)$ and $\varphi(z) = \lambda^{a_n+1} \varphi(y)$. Hence, $\lambda = 1$ and, by ergodicity, φ is constant. In general, use the theorem of Lusin (Theorems A.3.5–A.3.9) to reach the same conclusion. (c) A is a union of intervals I_j in the pile \mathcal{S}_n for each $n \geq 2$. Then, $f^{a_n}(I_j^e) = I_j^c$ for every j . Hence, $m(f^{a_n}(A) \cap A) \geq m(A)/3 = 2/27$.

8.3.1. Let $\{v_j : j \in \mathcal{I}\}$ be a basis of H formed by eigenvectors with norm 1 and λ_j be the eigenvalue associated with each eigenvector v_j . The hypothesis ensures that we may consider $\mathcal{I} = \mathbb{N}$. Show that for every $\delta > 0$ and every $k \geq 1$ there exists $n \geq 1$ such that $|\lambda_j^n - 1| \leq \delta$ for every $j \in \{1, \dots, k\}$ (use the pigeonhole principle). Decompose $\varphi = \sum_j c_j v_j$, with $c_j \in \mathbb{C}$. Observe that $U_f^n \varphi = \sum_{j \in \mathbb{N}} c_j \lambda_j^n v_j$, and so

$$\|U_f^n \varphi - \varphi\|_2^2 \leq \sum_{j=1}^k |c_j(\lambda_j^n - 1)|^2 + \sum_{j=k+1}^\infty 2|c_j|^2 \leq \delta^2 \|\varphi\|_2^2 + \sum_{j=k+1}^\infty 2|c_j|^2.$$

Given $\varepsilon > 0$, we may choose δ and k in such a way that each one of the terms on the right-hand side is less than $\varepsilon/2$.

- 8.4.3. Let $U : H \rightarrow H$ be a non-invertible isometry. Recalling Exercise 2.3.6, show that there exist closed subspaces V and W of H such that $U : H \rightarrow H$ is unitarily conjugate to the operator $U_1 : V \oplus W^{\mathbb{N}} \rightarrow V \oplus W^{\mathbb{N}}$ given by $U_1|_V = U|_V$ and $U_1|_{W^{\mathbb{N}}} = \text{id}$. Let $U_2 : V \oplus W^{\mathbb{Z}} \rightarrow V \oplus W^{\mathbb{Z}}$ be the linear operator defined by $U_2|_V = U|_V$ and $U_2|_{W^{\mathbb{Z}}} = \text{id}$. Check that U_2 is a unitary operator such that $U_2 \circ j = j \circ U_1$, where $j : V \oplus W^{\mathbb{N}} \rightarrow V \oplus W^{\mathbb{Z}}$ is the natural inclusion. Show that if $E \subset V \oplus W^{\mathbb{N}}$ satisfies the conditions in the definition of Lebesgue spectrum for U_1 then $j(E)$ satisfies those same conditions for U_2 . Conclude that the rank of U_1 is well defined.
- 8.4.6. The lemma of Riemann–Lebesgue ensures that F takes values in c_0 . The operator F is continuous: $\|F(\varphi)\| \leq \|\varphi\|$ for every $\varphi \in L^1(\lambda)$. Moreover, F is injective: if $F(\varphi) = 0$ then $\int \varphi(z)\psi(z)d\lambda(z) = 0$ for every linear combination $\psi(z) = \sum_{|j| \leq l} a_j z^j$, $a_j \in \mathbb{C}$. Given any interval $I \subset S^1$, the sequence $\psi_N = \sum_{|n| \leq N} c_n z^n$, $c_n = \int_I z^{-n} d\lambda(z)$ of partial sums of the Fourier series of the characteristic function χ_I is bounded (see [Zyg68, page 90]). Using the dominated convergence theorem, it follows that $F(\varphi) = 0$ implies $\int_I \varphi(z) d\lambda(z) = 0$, for any interval I . Hence, $\varphi = 0$. If F were bijective then, by the open mapping theorem, its inverse would be a continuous linear operator. Then, there would be $c > 0$ such that $\|F(\varphi)\| \geq c\|\varphi\|$ for every $\varphi \in L^1(\lambda)$. But that is false: consider $D_N(z) = \sum_{|n| \leq N} z^n$ for $N \geq 0$. Check that $F(D_N) = (a_n^N)_n$ with $a_n^N = 1$ if $|n| \leq N$ and $a_n^N = 0$ otherwise. Hence, $\|F(D_N)\| = 1$ for every N . Writing $z = e^{2\pi it}$, check that $D_N(z) = \sin((2N+1)\pi t)/\sin(\pi t)$. Conclude that $\|D_N\| = \int |D_N(z)| d\lambda(z)$ converges to infinity when $N \rightarrow \infty$. [Observation: One can also give explicit examples. For instance, if $(a_n)_n$ converges to zero and satisfies $\sum_{n=1}^{\infty} a_n/n = \infty$ then the sequence $(a_n)_n$ given by $a_n = a_n/(2i)$ for $n \geq 1$ and $a_n + a_{-n} = 0$ for every $n \geq 0$ may not be written in the form $a_n = \int z^n d\nu(z)$. See Section 7.3.4 of Edwards [Edw79].]
- 8.5.3. By Exercise 8.5.2, \tilde{f} is always injective. Conclude that if \tilde{f} is surjective then it is invertible: there exists a homomorphism of measure algebras $h : \tilde{\mathcal{B}} \rightarrow \tilde{\mathcal{B}}$ such that $h \circ \tilde{f} = \tilde{f} \circ h = \text{id}$. Use Proposition 8.5.6 to find $g : M \rightarrow M$ such that $g \circ f = f \circ g$ at μ -almost every point. The converse is easy: if (f, μ) is invertible at almost every point then the homomorphism of measure algebras \tilde{g} associated with $g = f^{-1}$ satisfies $\tilde{g} \circ \tilde{f} = \tilde{f} \circ \tilde{g} = \text{id}$; in particular, \tilde{f} is surjective.
- 8.5.6. Check that the unions of elements of $\bigcup_n \mathcal{P}_n$ are pre-images, under the inclusion ι , of an open subset of K . Use that fact to show that if the chains have measure zero then for each $\delta > 0$ there exists an open set $A \subset K$ such that $m(A) < \delta$ and every point outside A is in the image of the inclusion: in other words, $K \setminus \iota(M_{\mathcal{P}}) \subset A$. Conclude that $\iota(M_{\mathcal{P}})$ is a Lebesgue measurable set and its complement in K has measure zero. For the converse, use the fact that (a) implies (c) in Exercise A.1.13.
- 9.1.1. $H_{\mu}(\mathcal{P}/\mathcal{R}) \leq H_{\mu}(\mathcal{P} \vee \mathcal{Q}/\mathcal{R}) = H_{\mu}(\mathcal{Q}/\mathcal{R}) + H_{\mu}(\mathcal{P}/\mathcal{Q} \vee \mathcal{R}) \leq H_{\mu}(\mathcal{Q}/\mathcal{R}) + H_{\mu}(\mathcal{P}/\mathcal{Q})$.
- 9.1.3. Let $g = f^k$. Then $H_{\mu}(\bigvee_{i=0}^{k-1} f^{-i}(\mathcal{P})/\bigvee_{j=k}^n f^{-j}(\mathcal{P})) = H_{\mu}(\mathcal{P}^k/\bigvee_{i=1}^{n-k} g^{-i}(\mathcal{P}^k))$. By Lemma 9.1.12, this expression converges to $h_{\mu}(g, \mathcal{P}_k)$. Now use Lemma 9.1.13.

- 9.2.5. Write $\mathcal{Q}^n = \bigvee_{j=0}^{n-1} f^{-j}(\mathcal{Q})$ for each n and let \mathcal{A} be the σ -algebra generated by $\bigcup_n \mathcal{Q}^n$. Check that f is measurable with respect to the σ -algebra \mathcal{A} . Show that the hypothesis implies that $\mathcal{P} \subset \mathcal{A}$. By Corollary 9.2.4, it follows that $H_\mu(\mathcal{P}/\mathcal{Q}^n)$ converges to zero. By Lemmas 9.1.11 and 9.1.13, we have that $h_\mu(f, \mathcal{P}) \leq h_\mu(f, \mathcal{Q}) + H_\mu(\mathcal{P}/\mathcal{Q}^n)$ for every n .
- 9.2.7. The set \mathcal{A} of all finite disjoint of rectangles $A_i \times B_i$, with $A_i \subset M$ and $B_i \subset N$, is an algebra that generates the σ -algebra of $M \times N$. Given partitions \mathcal{P} and \mathcal{Q} of M and N , respectively, the family $\mathcal{P} \times \mathcal{Q} = \{P \times Q : P \in \mathcal{P} \text{ and } Q \in \mathcal{Q}\}$ is a partition of $M \times N$ contained in \mathcal{A} and such that $h_{\mu \times \nu}(f \times g, \mathcal{P} \times \mathcal{Q}) = h_\mu(f, \mathcal{P}) + h_\nu(g, \mathcal{Q})$. Conversely, given any partition $\mathcal{R} \subset \mathcal{A}$ of $M \times N$, there exist partitions \mathcal{P} and \mathcal{Q} such that $\mathcal{R} \prec \mathcal{P} \times \mathcal{Q}$ and so $h_{\mu \times \nu}(f \times g, \mathcal{R}) \leq h_\mu(f, \mathcal{P}) + h_\nu(g, \mathcal{Q})$. Conclude using Exercise 9.2.6.
- 9.3.3. It is clear that $B(x, n, \varepsilon) \subset B(f(x), n - 1, \varepsilon)$. Hence, $h_\mu(f, x) \geq h_\mu(f, f(x))$ for μ -almost every x . On the other hand, $\int h_\mu(f, x) d\mu(x) = \int h_\mu(f, f(x)) d\mu(x)$ since the measure μ is invariant under f .
- 9.4.2. Use the following consequence of the Jordan canonical form: there exist numbers $\rho_1, \dots, \rho_l > 0$, there exists an A -invariant decomposition $\mathbb{R}^d = E_1 \oplus \dots \oplus E_l$ and, given $\alpha > 0$, there exists an inner product in \mathbb{R}^d relative to which the subspaces E_j are orthogonal and satisfy $e^{-\alpha} \rho_j \|v\| \leq \|Av\| \leq e^\alpha \rho_j \|v\|$ for every $v \in E_j$. Moreover, the ρ_i are the absolute values of the eigenvalues of A and they satisfy $\sum_{i=1}^d \log^+ |\lambda_i| = \sum_{j=1}^l \dim E_j \log^+ \rho_j$.
- 9.4.3. Consider any countable partition \mathcal{P} with $\{B, B^c\} \prec \mathcal{P}$. Let \mathcal{Q} be the restriction of \mathcal{P} to the set B . Write $\mathcal{P}^n = \bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})$ and $\mathcal{Q}^k = \bigvee_{j=0}^{k-1} g^{-j}(\mathcal{Q})$. Check that, for every $x \in B$ and $k \geq 1$, there exists $n_k \geq 1$ such that $\mathcal{Q}^k(x) = \mathcal{P}^{n_k}(x)$. Moreover, by ergodicity, $\lim_k k/n_k = \tau(B, x) = \mu(B)$ for almost every x . By the theorem of Shannon–McMillan–Breiman,

$$h_\nu(g, \mathcal{Q}, x) = \lim_k -\frac{1}{k} \log \nu(\mathcal{Q}^k(x)) \quad \text{and} \quad h_\mu(f, \mathcal{P}, x) = \lim_k -\frac{1}{n_k} \log \mu(\mathcal{P}^{n_k}(x)).$$

Conclude that $h_\nu(f, \mathcal{Q}, x) = \mu(B)h_\nu(g, \mathcal{Q}, x)$ for almost every $x \in B$. Varying \mathcal{P} , deduce that $h_\nu(f) = \mu(B)h_\nu(g)$.

- 9.5.2. Consider $A \in \bigcap_n f^{-n}(\mathcal{B})$ with $m(A) > 0$. Then, for each n there exists $A_n \in \mathcal{B}$ such that $A = f^{-n}(A_n)$. Consider the intervals $I_{j,n} = ((j - 1)/10^n, j/10^n)$. Then,

$$\frac{m(A \cap I_{j,n})}{m(I_{j,n})} = \frac{m(A_n)}{m((0, 1))} = m(A) \quad \text{for every } 1 \leq j \leq 10^n.$$

Making $n \rightarrow \infty$, conclude that A^c has no points of density. Hence, $m(A^c) = 0$.

- 9.5.6. Assume that $h_\mu(f, \mathcal{P}) = 0$. Use Lemma 9.5.4 to show that $\mathcal{P} \prec \bigvee_{j=1}^\infty f^{-jk}(\mathcal{P})$ for every $k \geq 1$. Deduce that, up to measure zero, \mathcal{P} is contained in $f^{-k}(\mathcal{B})$ for every $k \geq 1$. Conclude that the partition \mathcal{P} is trivial.
- 9.6.1. Uniqueness is immediate. To prove the existence, consider the functional Ψ defined by $\Psi(\psi) = \int (\int \psi d\eta) dW(\eta)$ in the space of bounded measurable functions $\psi : M \rightarrow \mathbb{R}$. Note that Ψ is linear and non-negative and satisfies $\Psi(1) = 1$. Use the monotone convergence theorem to show that if B_n , $n \geq 1$ are pairwise disjoint measurable subsets of M then $\Psi(\mathcal{X}_{\bigcup_n B_n}) = \sum_n \Psi(\mathcal{X}_{B_n})$. Conclude that $\zeta(B) = \Psi(\mathcal{X}_B)$ defines a probability measure in the σ -algebra

of measurable subsets of M . Show that $\int \psi d\xi = \Psi(\psi)$ for every bounded measurable function. Take $\text{bar}(W) = \xi$.

- 9.6.5. For the penultimate identity one would need to know that $n^{-1} \log \mu_P(\mathcal{Q}^n(x))$ is a dominated sequence (for example).
 9.7.4. By Exercise 9.7.3, given any bounded measurable function ψ ,

$$\int (\psi \circ f) d\eta = \int \psi(x) \left(\sum_{z \in f^{-1}(x)} \frac{1}{J_{\eta f}(z)} \right) d\eta(x).$$

Deduce the first part of the statement. For the second part, note that if η is invariant then $\eta(f(A)) = \eta(f^{-1}(f(A))) \geq \eta(A)$ for every domain of invertibility A .

- 9.7.7. The “if” part of the statement is easy: we may exhibit the ergodic equivalence explicitly. Assume that the two systems are ergodically equivalent. The fact that $k = l$ follows from Exercise 8.1.2. To prove that p and q are permutations of one another, use the fact that the Jacobian is invariant under ergodic equivalence (Exercise 9.7.6), together with the expressions of the Jacobians given by Example 9.7.1.
 10.1.6. Note that $\psi(M)$ is compact and the inverse $\psi^{-1} : \psi(M) \rightarrow M$ is (uniformly) continuous. Hence, given $\varepsilon > 0$ there exists $\delta > 0$ such that if $E \subset M$ is (n, ε) -separated for f then $\psi(E) \subset N$ is (n, δ) -separated for g . Conclude that $s(f, \varepsilon, M) \leq s(g, \delta, N)$ and deduce that $h(f) \leq h(g)$. [Observation: The statement remains valid in the non-compact case, as long as we assume the inverse $\psi^{-1} : \psi(M) \rightarrow M$ to be uniformly continuous.]

For the second part, consider the distance defined in Σ by

$$d((x_n)_n, (y_n)_n) = \sum_{n \in \mathbb{Z}} 2^{-|n|} |x_n - y_n|.$$

Consider a discrete set $A \subset [0, 1]$ with n elements. Check that the restriction to $A^{\mathbb{Z}}$ of the distance of $[0, 1]^{\mathbb{Z}}$ is uniformly equivalent to the distance defined in (9.2.15). Using Example 10.1.2, conclude that the topological entropy of σ is at least $\log n$, for any n .

- 10.1.10. (Carlos Gustavo Moreira) Let $\theta_1 = 0, \theta_2 = 01$ and, for $n \geq 2$, $\theta_{n+1} = \theta_n \theta_{n-1}$. We claim that, for every $n \geq 1$, there exists a word τ_n such that $\theta_n \theta_{n+1} = \tau_n \alpha_n$ and $\theta_{n+1} \theta_n = \tau_n \beta_n$, where $\alpha_n = 10$ and $\beta_n = 01$ if n is even and $\alpha_n = 01$ and $\beta_n = 10$ if n is odd. That holds for $n = 1$ with $\tau_1 = 0$ and for $n = 2$ with $\tau_2 = 010$. If it holds for a given n , then $\theta_{n+1} \theta_{n+2} = \theta_{n+1} \theta_{n+1} \theta_n = \theta_{n+1} \tau_n \beta_n = \tau_{n+1} \alpha_{n+1}$ and also $\theta_{n+2} \theta_{n+1} = \theta_{n+1} \theta_n \theta_{n+1} = \theta_{n+1} \tau_n \alpha_n = \tau_{n+1} \beta_{n+1}$, as long as we take $\tau_{n+1} = \theta_{n+1} \tau_n$. This proves the claim. It follows that the last letters of θ_n and θ_{n+1} are distinct.

Now, we claim that $\theta = \lim_n \theta_n$ is not pre-periodic. Indeed, suppose that θ were pre-periodic and let m be its period. Since the length of θ_n is F_{n+1} (where F_k is the k -th Fibonacci number), we may take n large such that m divides F_{n+1} and the pre-period (that is, the length of the non-periodic part) of θ is less than F_{n+2} . Then, θ starts with $\theta_{n+3} = \theta_{n+2} \theta_{n+1} = \theta_{n+1} \theta_n \theta_{n+1}$. However, since the length F_{n+1} of θ_n is a multiple of the period m , the F_{m+2} -th letter of θ , which is the last letter of θ_{n+1} , must coincide with the $(F_{m+2} + F_{n+1})$ -th letter of θ , which is the last letter of θ_n . This would contradict the conclusion of the previous paragraph.

Next, we claim that $c_{k+1}(\theta) > c_k(\theta)$ for every k . Indeed, suppose that $c_{k+1}(\theta) = c_k(\theta)$ for some k . Then, every subword of length k can have only one continuation of length $k + 1$. Hence, we have a transformation in the set of subwords of length k , assigning to each subword its unique continuation, without the first letter. Since the domain is finite, all the orbits of this transformation are pre-periodic. In particular, θ is also pre-periodic, which contradicts the conclusion in the previous paragraph.

Since $c_1(\theta) = 2$, it follows that $c_k(\theta) \geq k + 1$ for every k . We claim that $c_{F_{n+1}}(\theta) \leq F_{n+1} + 1$ for every $n > 1$. To prove that fact, note that θ may be written as a concatenation of words belonging to $\{\theta_n, \theta_{n+1}\}$ because (by induction) every θ_r with $r \geq n$ may be written as a concatenation of words belonging to $\{\theta_n, \theta_{n+1}\}$. Thus, any subword of θ of length F_{n+1} (which is the length of θ_n) is a subword of $\theta_n\theta_{n+1}$ or $\theta_{n+1}\theta_n$. Since $\theta_n\theta_{n+1} = \theta_n\theta_n\theta_{n-1}$, is a subword of $\theta_n\theta_n\theta_{n-1}\theta_{n-2} = \theta_n\theta_n\theta_n$, there are at most $|\theta_n| = F_{n+1}$ subwords of length $|\theta_n| = F_{n+1}$ of $\theta_n\theta_n\theta_n$ and, hence, of $\theta_n\theta_{n+1}$. Since $\theta_n\theta_{n+1} = \tau_n\alpha_n$ and $\theta_{n+1}\theta_n = \tau_n\beta_n$, and $\theta_{n+1}\theta_n$ ends with θ_n and $|\beta_n| = 2$, the unique subword of $\theta_{n+1}\theta_n$ of length $|\theta_n| = F_{n+1}$ that may not be a subword of $\theta_n\theta_{n+1}$ is the subword that ends with the first letter of β_n (that is, one position before the end of $\theta_{n+1}\theta_n$). Hence, $c_{F_{n+1}}(\theta) \leq F_{n+1} + 1$ as stated.

We are ready to obtain the statement of the exercise. Assume that $c_k(\theta) > k + 1$ for some k . Taking n such that $F_{n+1} > k$, we would have $c_{F_{n+1}}(\theta) - c_k(\theta) < F_{n+1} + 1 - (k + 1) = F_{n+1} - k$ and that would imply that $c_{m+1}(\theta) \leq c_m(\theta)$ for some m with $k \leq m < F_{n+1}$. This would contradict the conclusion in the previous paragraph.

- 10.2.4. By Proposition 10.2.1, $h(f) = g(f, \delta, M)$ whenever f is ε -expansive and $\delta < \varepsilon/2$. Show that if $d(f, h) < \delta/3$ then $g(h, \delta/3, M) \leq g(f, \delta, M)$. Deduce that if $(f_k)_k$ converges to f then $\limsup_k h(f_k) = \limsup_k g(f_k, \delta/3, M) \leq g(f, \delta, M) = h(f)$.
- 10.2.8. (Bowen [Bow72]) Write $a = g_*(f, \varepsilon)$. Observe that if E is an (n, δ) -generating set of M , with $\delta < \varepsilon$, then $M = \bigcup_{x \in E} B(x, n, \varepsilon)$. Combining this fact with the result of Bowen, show that $g_n(f, \delta, M) \leq \#Ee^{c+(a+b)n}$. Take $b \rightarrow 0$ to conclude the inequality.
- 10.3.3. (a) The hypothesis implies that for every n and every subcover δ of β^n there exists a subcover γ of α^n such that $\gamma < \delta$. Taking γ minimal, $\#\gamma \leq \#\delta$ and $\sum_{U \in \gamma} \inf_{x \in U} e^{\phi_n(x)} \leq \sum_{V \in \delta} \inf_{y \in V} e^{\phi_n(y)}$. It follows that $Q_n(f, \phi, \alpha) \leq Q_n(f, \phi, \beta)$ for every n . (b) Lemma 10.1.11 gives that α^{n+k-1} is a subcover of $(\alpha^k)^n$. A variation of the argument in part (a) gives that $Q_n(f, \phi, \alpha^k) \leq e^{(k-1) \sup |\phi|} Q_{n+k-1}(f, \phi, \alpha)$ for every n . Hence, $Q^\pm(f, \phi, \alpha^k) \leq Q^\pm(f, \phi, \alpha)$. [Observation: Analogously, $P(f, \phi, \alpha^k) \leq P(f, \phi, \alpha)$.] By the second part of Lemma 10.1.11, for every subcover β of $(\alpha^k)^n$ there exists a subcover γ of α^{n+k-1} such that $\gamma < \beta$, $\#\gamma \leq \#\beta$ and $\sum_{U \in \gamma} \inf_{x \in U} e^{\phi_{n+k-1}(x)} \leq e^{(k-1) \sup |\phi|} \sum_{V \in \beta} \inf_{y \in V} e^{\phi_n(y)}$ (taking γ minimal). Deduce that $Q_{n+k-1}(f, \phi, \alpha) \leq e^{(k-1) \sup |\phi|} Q_n(f, \phi, \alpha^k)$. Hence, $Q^\pm(f, \phi, \alpha) \leq Q^\pm(f, \phi, \alpha^k)$. (c) Follows from part (b) and Corollary 10.3.3. (d) If the elements of α are disjoint then $(\alpha^k)^n = \alpha^{n+k-1}$ and so

$$P_n(f, \phi, \alpha^k) = \inf \left\{ \sum_{U \in \gamma} \sup_{x \in U} e^{\phi_n(x)} : \gamma \subset (\alpha^k)^n \right\} = \inf \left\{ \sum_{U \in \gamma} \sup_{x \in U} e^{\phi_{n+k-1}(x)} : \gamma \subset \alpha^{n+k-1} \right\}.$$

It follows that $e^{-(k-1)\sup|\phi|} P_n(f, \phi, \alpha^k) \leq P_{n+k-1}(f, \phi, \alpha) \leq e^{(k-1)\sup|\phi|} P_n(f, \phi, \alpha^k)$.

(e) Follows from part (d) and the definition of pressure (Lemma 10.3.1). (f)

Note that $\alpha^{\pm k} = f^k(\alpha^{2k})$ and use Exercise 10.3.2.

- 10.3.8. Show that given $\varepsilon > 0$ there exists $\kappa \geq 1$ such that every dynamical ball $B(x, n, \varepsilon)$ has diameter equal to $\varepsilon 2^{-n}$ and contains some periodic point p_x^n of period $n + \kappa$. Show that given $C, \theta > 0$ there exists $K > 0$ such that $|\phi_n(y) - \phi_n(p_x^n)| \leq K$ for every $y \in B(x, n, \varepsilon)$, every $n \geq 1$ and every (C, θ) -Hölder function $\phi : S^1 \rightarrow \mathbb{R}$. Use this fact to replace generating (or separated) sets by sets of periodic points in the definition of pressure.

- 10.3.10. Since $\xi | \Lambda^c = \eta | \Lambda^c$,

$$\begin{aligned} |\mathcal{E}(\xi, \eta) - E_\Lambda(\xi | \Lambda) + E_\Lambda(\eta | \Lambda)| &= \left| \sum_{(k,l) \in \Lambda \times \Lambda^c \cup \Lambda^c \times \Lambda} \Psi(k-l, \xi_k, \xi_l) \right. \\ &\quad \left. - \Psi(k-l, \eta_k, \eta_l) \right| \\ &\leq \sum_{(k,l) \in \Lambda \times \Lambda^c \cup \Lambda^c \times \Lambda} 2Ke^{-\theta|k-l|}. \end{aligned}$$

Recalling that Λ is an interval, the cardinal of $\{(k, l) \in \Lambda \times \Lambda^c \cup \Lambda^c \times \Lambda : |k - l| = n\}$ is less than or equal to $4n$ for every $n \geq 1$. Hence,

$$|\mathcal{E}(\xi, \eta) - E_\Lambda(\xi | \Lambda) + E_\Lambda(\eta | \Lambda)| \leq \sum_{n=1}^{\infty} 8Kne^{-\theta n} < \infty.$$

The second part of the statement is an immediate consequence of the first one.

- 10.4.4. Consider the shift map σ in the space $\Sigma = \{0, 1\}^{\mathbb{N}}$. Consider the function $\phi : \Sigma \rightarrow \mathbb{R}$ defined by $\phi(x) = 0$ if $x_0 = 0$ and $\phi(x) = 1$ if $x_0 = 1$. Let N be the set of points $x \in \Sigma$ such that the time average in the orbit of x does not converge. Check that N is invariant under σ and is non-empty: for each finite sequence (z_0, \dots, z_k) one can find $x \in N$ with $x_i = z_i$ for $i = 0, \dots, k$. Deduce that the topological entropy of the restriction $f | N_\phi$ is equal to $\log 2$. Justify that N does not support any probability measure invariant under f .
- 10.4.5. Consider the open cover ξ of K whose elements are $K \cap [0, \alpha]$ and $K \cap [1 - \beta, 1]$. Check that $P(f, \phi) = P(f, \phi, \xi)$ for every potential ϕ . Moreover,

$$P_n(f, -t \log g', \xi) = \sum_{U \in \alpha^n} [(g^n)']^{-t}(U) = (\alpha^t + \beta^t)^n.$$

Conclude that $\psi(t) = \log(\alpha^t + \beta^t)$. Check that $\psi' < 0$ and $\psi'' > 0$ (convexity also follows from Proposition 10.3.7). Moreover, $\psi(0) > 0 > \psi(1)$. By the variational principle, the last inequality implies that $h_\mu(f) - \int \log g' d\mu < 0$.

- 10.5.3. The Gibbs property gives that $\lim_n (1/n) \log \mu(C^n(x)) = \bar{\phi}(x) - P$, where $C^n(x)$ is the cylinder of length n that contains x . Combine this identity with the theorem of Brin–Katok (Theorem 9.3.3) and the theorem of Birkhoff to get the first claim. Now assume that μ_1 and μ_2 are two ergodic Gibbs states with the same constant P . Observe that there exists C such that $C^{-1}\mu_1(A) \leq \mu_2(A) \leq C\mu_1(A)$ for every A in the algebra formed by the finite disjoint unions of cylinders. Using the monotone class theorem (Theorem A.1.18), deduce that $C^{-1}\mu_1(A) \leq \mu_2(A) \leq C\mu_1(A)$ for any measurable set A . This implies

that μ_1 and μ_2 are equivalent measures. Using Lemma 4.3.1, it follows that $\mu_1 = \mu_2$.

- 10.5.5. By Proposition 10.3.7, the pressure function is convex. By Exercise A.5.1, it follows that it is also continuous. By the smoothness theorem of Mazur (recall Exercise 2.3.7), there exists a residual subset $\mathcal{R} \subset C^0(M)$ such that the pressure function is differentiable at every $\varphi \in \mathcal{R}$. Apply Exercise 10.5.4.
- 11.1.3. Adapt the arguments in Section 9.4.2, as follows. Start by checking that the iterates of f have bounded distortion: there exists $K > 1$ such that

$$\frac{1}{K} \leq \frac{|Df^n(x)|}{|Df^n(y)|} \leq K,$$

for every $n \geq 1$ and any points x, y with $\mathcal{P}^n(x) = \mathcal{P}^n(y)$. Consider the sequence $\mu_n = (1/n) \sum_{j=0}^{n-1} f_*^j m$ of averages of iterates of the Lebesgue measure m . Show that the Radon–Nikodym derivatives $d\mu_n/dm$ are uniformly bounded and are Hölder, with uniform Hölder constants. Deduce that every accumulation point μ of that sequence is an invariant probability measure absolutely continuous with respect to the Lebesgue measure. Show that the Radon–Nikodym derivative $\rho = d\mu/dm$ is bounded from zero and infinity (in other words, $\log \rho$ is bounded). Show that ρ and $\log \rho$ are Hölder.

- 11.1.5. Check that $J_\mu f = (\rho \circ f)|f'|/\rho$ and use the Rokhlin formula (Theorem 9.7.3).
- 11.2.4. Take $\Lambda = \{2^{-n} : n \geq 0\} \pmod{\mathbb{Z}}$. The restriction $f : \Lambda \rightarrow \Lambda$ cannot be an expanding map because $1/2$ is an isolated point in Λ but $1 = f(1/2)$ is not. [Observation: Note that $\Lambda = S^1 \setminus \bigcup_{n=0}^\infty f^{-n}(I)$, where $I = (1/2, 1) \pmod{\mathbb{Z}}$. Modifying suitably the choice of I , one finds many other examples, possibly with Λ uncountable.]
- 11.3.3. Let $a = \int \varphi d\mu_1$ and $b = \int \varphi d\mu_2$. Assume that $a < b$ and write $r = (b - a)/5$. By the ergodic decomposition theorem, we may assume that μ_1 and μ_2 are ergodic. Then, there exist x_1 and x_2 such that $\tilde{\varphi}(x_1) = a$ and $\tilde{\varphi}(x_2) = b$. Using the hypothesis that f is topologically exact, construct a pseudo-orbit $(z_n)_{n \geq 0}$ alternating (long) segments of the orbits of x_1 and x_2 in such a way that the sequence of time averages of φ along the pseudo-orbit $(z_n)_n$ oscillates from $a + r$ to $b - r$ (meaning that $\liminf \leq a + r$ and $\limsup \geq b - r$). Next, use the shadowing lemma to find $x \in M$ whose orbit shadows this pseudo-orbit. Using that φ is uniformly continuous, conclude that the sequence of time averages of φ along the orbit of x oscillates from $a + 2r$ to $b - 2r$.
- 12.1.2. Theorem 2.1.5 gives that $\mathcal{M}_1(M)$ is weak* compact and it is clear that it is convex. Check that the operator $\mathcal{L} : C^0(M) \rightarrow C^0(M)$ is continuous and deduce that its dual $\mathcal{L}^* : \mathcal{M}(M) \rightarrow \mathcal{M}(M)$ is also continuous. If $(\eta_n)_n \rightarrow \eta$ in the weak* topology then $(\int \mathcal{L}1 d\eta_n)_n \rightarrow \int \mathcal{L}1 d\eta$. Conclude that the operator $G : \mathcal{M}_1(M) \rightarrow \mathcal{M}_1(M)$ is continuous. Hence, by the Tychonoff–Schauder theorem, G has some fixed point ν . This means that $\mathcal{L}^*\nu = \lambda\nu$, where $\lambda = \int \mathcal{L}1 d\nu$. Since $\lambda > 0$, this proves that ν is a reference measure. Using Corollary 12.1.9, check that $\lambda = \limsup_n \sqrt[n]{\|\mathcal{L}^n 1\|}$ and deduce that λ is the spectral radius of \mathcal{L} .
- 12.2.4. Fix in S^1 the orientation induced by \mathbb{R} . Consider the fixed point $p_0 = 0$ of f and let p_1, \dots, p_d be its pre-images, ordered cyclically, with $p_d = p_0$. Analogously, let q_0 be a fixed point of g and q_1, \dots, q_d be its pre-images, ordered cyclically,

with $q_d = q_0$. Note that f maps each $[p_{i-1}, p_i]$ and g maps each $[q_{i-1}, q_i]$ onto S^1 . Then, for each sequence $(i_n)_n \in \{1, \dots, d\}^{\mathbb{N}}$ there exists exactly one point $x \in S^1$ and one point $y \in S^1$ such that $f^n(x) \in [p_{i_n-1}, p_{i_n}]$ and $g^n(y) \in [q_{i_n-1}, q_{i_n}]$ for every n . Clearly, the maps $(i_n)_n \mapsto x$ and $(i_n)_n \mapsto y$ are surjective. Consider two sequences $(i_n)_n$ and $(j_n)_n$ to be equivalent if there exists $N \in \mathbb{N} \cup \{\infty\}$ such that (1) $i_n = j_n$ for every $n \leq N$ and either (2a) $i_n = 1$ and $j_n = d$ for every $n > N$ or (2b) $i_n = d$ and $j_n = 1$ for every $n > N$. Show that the points x corresponding to $(i_n)_n$ and $(j_n)_n$ coincide if and only if the two sequences are equivalent and a similar fact holds for the points y corresponding to the two sequences. Conclude that the map $\phi : x \mapsto y$ is well defined and is a bijection in S^1 such that $\phi(f(x)) = g(\phi(x))$ for every x . Observe that ϕ preserves the orientation of S^1 and, thus, is a homeomorphism.

- 12.2.5. (a) \Rightarrow (b): Trivial. (b) \Rightarrow (c): Let μ_a be the absolutely continuous invariant probability measure and μ_m be the measure of maximum entropy of f ; let ν_a and ν_m be the corresponding measures for g . Show that $\mu_a = \mu_m$. Let $\phi : S^1 \rightarrow S^1$ be a topological conjugacy. Show that $\nu_m = \phi_* \mu_m$ and $\nu_a = \phi_* \mu_a$ if ϕ is absolutely continuous. Use Corollary 12.2.4 to conclude that in the latter case $|(g^n)'(x)| = k^n$ for every $x \in \text{Fix}(f^n)$. (c) \Rightarrow (a): The hypothesis implies that $\nu_a = \nu_m$ and so $\nu_a = \phi_* \mu_a$. Recall (Proposition 12.1.20) that the densities $d\mu_a/dm$ and $d\nu_a/dm$ are continuous and bounded from zero and infinity. Conclude that ϕ is differentiable, with $\phi' = (d\mu/dm)/(d\nu/dm) \circ \phi$.

- 12.3.2. Consider $A = (a, 1)$, $P = (p, 1)$, $Q = (q, 1)$, $B = (b, 1)$, $O = (0, 0) \in \mathbb{C} \times \mathbb{R}$. Let A' (respectively, B') be the point where the line parallel to OQ (respectively, OP) passing through P (respectively, Q) intersects the boundary of C . Note that all these points belong to the plane determined by P , Q and O ; note also that $A' \in OA$ and $B' \in OB$. By definition, $\alpha(P, Q) = |B'Q|/|OP|$ and $\beta(P, Q) = |OQ|/|A'P|$. Check that $|AP|/|AQ| = |A'P|/|OQ|$ and $|BQ|/|BP| = |B'Q|/|OP|$. Hence,

$$\theta(P, Q) = \log \frac{\beta(P, Q)}{\alpha(P, Q)} = \log \frac{|OQ|/|OP|}{|A'P|/|B'Q|} = \log \frac{|AQ|/|BP|}{|AP|/|BQ|}.$$

In other words, $d(p, q) = \log(|aq||bp|)/(|ap||bq|) = \Delta(p, q)$, for any $p, q \in \mathbb{D}$.

- 12.3.4. Consider the cone C_0 of positive continuous functions in M . The corresponding projective distance θ_0 is given in Example 12.3.5. Check that θ_1 is the restriction of θ_0 to the cone C_1 . Consider a sequence of positive differentiable functions converging uniformly to a (continuous but) non-differentiable function g_0 . Show that $(g_n)_n$ converges to g_0 with respect to the distance θ_0 and, thus, is a Cauchy sequence for θ_0 and θ_1 . Argue that $(g_n)_n$ cannot be convergent for θ_1 .
- 12.3.8. (a) It is clear that $\log g$ is (b, β) -Hölder and $\sup g / \inf g$ is close to 1 if the norm $\|v\|_{\beta, \rho}$ is small; this will be implicit in all that follows. Then, $g \in C(b, \beta, R)$. To estimate $\theta(1, g)$, use the expression given by Lemma 12.3.8. Observe that

$$\beta(1, g) = \sup \left\{ g(x), \frac{\exp(b\delta)g(x) - g(y)}{\exp(b\delta) - 1} : x \neq y, d(x, y) < \rho \right\} \quad \text{where } \delta = d(x, y)^\beta.$$

Clearly, $g(x) \leq 1 + \sup |v|$. Moreover,

$$\begin{aligned} \frac{\exp(b\delta)g(x) - g(y)}{\exp(b\delta) - 1} &\leq \frac{\exp(b\delta)g(y) + \exp(b\delta)H_{\beta,\rho}(v)\delta - g(y)}{\exp(b\delta) - 1} \\ &= g(y) + \frac{\delta \exp(b\delta)}{\exp(b\delta) - 1} H_{\beta,\rho}(v). \end{aligned}$$

Take $K_1 > K_2 > 0$, depending only on b, β, ρ , such that $K_1 \geq \exp(bs)s / (\exp(bs) - 1) \geq K_2$ for every $s \in [0, \rho^\beta]$. Then, the term on the right-hand side of the previous inequality is bounded by $1 + \sup |v| + K_1 H_{\beta,\rho}(v)$. Hence, $\log \beta(1, g) \leq \log(1 + \sup |v| + K_1 H_{\beta,\rho}(v)) \leq K'_1 \|v\|_{\beta,\rho}$, where $K'_1 = \max\{K_1, 1\}$. Varying x and y in the previous arguments, we also find that $\beta(1, g) \geq 1 + \sup |v|$ and $\beta(1, g) \geq 1 - \sup |v| + K_2 H_{\beta,\rho}(v)$. Deduce that

$$\log \beta(1, g) \geq \max \{ \log(1 + \sup |v|), \log(1 - \sup |v| + K_2 H_{\beta,\rho}(v)) \} \geq K'_2 \|v\|_{\beta,\rho},$$

where the constant K'_2 depends only on K_2, β and ρ . Analogously, there exist constants $K'_3 > K'_4 > 0$ such that $-K'_3 \|v\|_{\beta,\rho} \leq \log \alpha(1, g) \leq -K'_4 \|v\|_{\beta,\rho}$. Fixing

$$K \geq \max\{K_1 + K_3, 1/(K_2 + K_4)\},$$

it follows that $K^{-1} \|v\|_{\beta,\rho} \leq \theta(1, g) \leq K \|v\|_{\beta,\rho}$. (b) It is no restriction to assume that $\|v\|_{\beta,\rho} < r$. Note that $\mathcal{P}^n g = 1 + \mathcal{P}^n v$ for every n . Corollary 12.3.12 gives that

$$\theta(\mathcal{P}^{kN} g, 1) \leq \Lambda_0^k \theta(1, g) \quad \text{for every } k,$$

with $\Lambda_0 < 1$. By part (a), it follows that $\|\mathcal{P}^{kN} v\| \leq K^2 \Lambda_0^k$ for every k . This yields the statement, with $\tau = \Lambda_0^{1/k}$ and $C = K^2 \|\mathcal{P}\|^N \Lambda_0^{-1}$.

12.4.4. Consider $0 < \delta \leq \rho$. For every cover \mathcal{U} of A with diameter less than δ , we have

$$\sum_{U \in \mathcal{U}} (\text{diam } U)^d \geq \sum_{U \in \mathcal{U}} K^{-1} \mu(U) \geq K^{-1} \mu(A).$$

Taking the infimum over \mathcal{U} , we get that $m_d(A, \delta) \geq K^{-1} \mu(A)$. Making $\delta \rightarrow 0$, we find that $m_d(A) > K^{-1} \mu(A)$; hence, $d(A) \geq d$.

12.4.7. Consider $\ell = 1$. Then, D, D_1, \dots, D_N (Section 12.4.3) are compact intervals.

It is no restriction to assume that $D = [0, 1]$. Write $D_{i^n} = h_{i_0} \circ \dots \circ h_{i_{n-1}}(D)$ for each $i^n = (i_0, \dots, i_{n-1})$ in $\{1, \dots, N\}^n$. Starting from the bounded distortion property (Proposition 12.4.5), prove that there exists $c > 0$ such that, for every i^n and every n ,

- (i) $c \leq |(f^n)'(x)| \text{diam } D_{i^n} \leq c^{-1}$ for every $x \in D_{i^n}$;
- (ii) $d(D_{i_n}, D_{j^n}) \geq c \text{diam } D_{i^n}$ for every $j^n \neq i^n$;
- (iii) $\text{diam } D_{i^{n+1}} \geq c \text{diam } D_{i^n}$ for every i_n , where $i^{n+1} = (i_0, \dots, i_{n-1}, i_n)$.

Let ν be the reference measure of the potential $\varphi = -d_0 \log |f'|$. Since $P(f, \varphi) = 0$, it follows from Lemma 12.1.3 and Corollary 12.1.15 that $\int_{D_{i^n}} \nu = |\nu|^{d_0}$. Deduce that $c \leq |(f^n)'(x)|^{d_0} \nu(D_{i^n}) \leq c^{-1}$ for any $x \in D_{i^n}$ and, using (i) once more, conclude that

$$c^2 \leq \frac{\text{diam}(D_{i^n})^{d_0}}{\nu(D_{i^n})} \leq c^{-2} \quad \text{for every } i^n \text{ and every } n.$$

It follows that $\sum_{i^n} \text{diam}(D_{i^n})^{d_0} \leq c^{-2} \sum_{i^n} \nu(D_{i^n}) = c^{-2}$. Since the diameter of D_{i^n} converges uniformly to zero when $n \rightarrow \infty$, this implies that $m_{d_0}(\Lambda) \leq c^{-2}$.

For the lower estimate, let us prove that ν satisfies the hypothesis of the mass distribution principle (Exercise 12.4.4). Given any U with $\text{diam } U < c \min\{\text{diam } D_1, \dots, \text{diam } D_N\}$, there exist $n \geq 1$ and i^n such that D_{i^n} intersects U and $c \text{diam } D_{i^n} > \text{diam } U$. By (ii), we have that $\nu(U) \leq \nu(D_{i^n}) \leq c^{-2} \text{diam } D_{i^n}^{d_0}$. Take n maximum. Then, using (iii), $\text{diam } U \geq c \text{diam } D_{i^{n+1}} \geq c^2 \text{diam } D_{i^n}$ for some choice of i_n . Combining the two inequalities, we get $\nu(U) \leq c^{-2-2d_0} (\text{diam } U)^{d_0}$. Then, by the mass distribution principle, $m_{d_0}(\Lambda) \geq c^{2+2d_0}$. Finally, extend these arguments to any dimension $\ell \geq 1$.

- A.1.9. Given $A_1 \supset \dots \supset A_i \supset \dots$, take $A = \bigcap_{i=1}^{\infty} A_i$. For $j \geq 1$, consider $A'_j = A_j \setminus A$. By Theorem A.1.14, we have that $\mu(A'_j) \rightarrow 0$ and so $\mu(A_j) \rightarrow \mu(A)$. Given $A_1 \subset \dots \subset A_i \subset \dots$, take $A = \bigcup_{i=1}^{\infty} A_i$. For each j , consider $A'_j = A \setminus A_j$. By Theorem A.1.14, we have that $\mu(A'_j) \rightarrow 0$, that is, $\mu(A_j) \rightarrow \mu(A)$.
- A.1.13. (Royden [Roy63]) (b) \Rightarrow (a) Assume that there exist Borel sets B_1, B_2 such that $B_1 \subset E \subset B_2$ and $m(B_2 \setminus B_1) = 0$. Deduce that $m^*(E \setminus B_1) = 0$, hence $E \setminus B_1$ is a Lebesgue measurable set. Conclude that E is a Lebesgue measurable set. (a) \Rightarrow (c) Let E be a Lebesgue measurable set such that $m^*(E) < \infty$. Given $\varepsilon > 0$, there exists a cover by open rectangles $(R_k)_k$ such that $\sum_k m^*(R_k) < m^*(E) + \varepsilon$. Then, $A = \bigcup_k R_k$ is an open set containing E and such that $m^*(A) - m^*(E) < \varepsilon$. Using that E is a Lebesgue measurable set, deduce that $m^*(A \setminus E) < \varepsilon$. For the general case, write E as a disjoint union of Lebesgue measurable sets with finite exterior measure. (c) \Leftrightarrow (d) It is clear that E is a Lebesgue measurable set if and only if its complement is. (c) and (d) \Rightarrow (b) For each $k \geq 1$, consider a closed set $F_k \subset E$ and an open set $A_k \supset E$ such that $m^*(E \setminus F_k)$ and $m^*(A_k \setminus E)$ are less than $1/k$. Then, $B_1 = \cup F_k$ and $B_2 = \cap A_k$ are Borel sets such that $B_1 \subset E \subset B_2$ and $m^*(E \setminus B_1) = m^*(B_2 \setminus E) = 0$. Conclude that $m(B_2 \setminus B_1) = m^*(B_2 \setminus B_1) = 0$.
- A.1.18. Show that $x \mapsto \frac{1}{n} \#\{0 \leq j \leq n-1 : a_j = 5\}$ is a simple function for each $n \geq 1$. By Proposition A.1.31, it follows that ω_5 is measurable.
- A.2.8. (a) Assume that \mathcal{F} is uniformly integrable. Consider $C > 0$ corresponding to $\alpha = 1$ and take $L = C + 1$. Check that $\int |f| d\mu < L$ for every $f \in \mathcal{F}$. Given $\varepsilon > 0$, consider $C > 0$ corresponding to $\alpha = \varepsilon/2$ and take $\delta = \varepsilon/(2C)$. Check that $\int_A |f| d\mu < \varepsilon$ for every $f \in \mathcal{F}$ and every set with $\mu(A) < \delta$. Conversely, given $\alpha > 0$, take $\delta > 0$ corresponding to $\varepsilon = \alpha$ and let $C = L/\delta$. Show that $\int_{|f|>C} |f| d\mu < \alpha$. (b) Applying Exercise A.2.5 to the function $|g|$, show that \mathcal{F} satisfies the criterion in (a). (c) Let us prove three facts about $f = \lim_n f_n$. (i) f is finite at almost every point: Consider L as in (a). Note that $\mu(\{x : |f_n(x)| \geq k\}) \leq L/k$ for every $n, k \geq 1$ (Exercise A.2.4) and deduce that $\mu(\{x : |f(x)| \geq k\}) \leq L/k$ for every $k \geq 1$. (ii) f is integrable: Fix $K > 0$. Given any $\varepsilon > 0$, take δ as in (a). Take n sufficiently large that $\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) < \delta$. Note that

$$\int_{|f| \leq K} |f| d\mu \leq \int_{|f_n - f| \leq \varepsilon} |f| d\mu + \int_{|f| \leq K, |f_n - f| > \varepsilon} |f| d\mu \leq (L + \varepsilon) + K\delta.$$

Deduce that $\int_{|f| \leq K} |f| d\mu \leq L$ for every K and $\int |f| d\mu \leq L$. (iii) $(f_n)_n$ converges to f in $L^1(\mu)$: Show that given $\varepsilon > 0$ there exists $K > 0$ such that $\int_{|f| > K} |f| d\mu < \varepsilon$ and $\int_{|f| > K} |f_n| d\mu < \varepsilon$ for every n . Take δ as in part (a) and n large enough that

$\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) < \delta$. Then

$$\int_{|f| \leq K} |f_n - f| d\mu \leq \int_{|f_n - f| \leq \varepsilon} |f_n - f| d\mu + \int_{|f_n - f| > \varepsilon} |f_n| d\mu + \int_{|f_n| \leq K, |f_n - f| > \varepsilon} |f| d\mu.$$

The right-hand side is bounded above by $2\varepsilon + K\delta$. Combining these inequalities, $\int |f_n - f| d\mu < 4\varepsilon + K\delta$ for every n sufficiently large.

- A.2.14. It is no restriction to assume that the B_n are pairwise disjoint. For each n , consider the measure η_n defined in B_n by $\eta_n(A) = \eta(f(A))$. Then, $\eta_n \ll (\eta|_{B_n})$ and, by the theorem of Radon–Nikodym, there exists $\rho_n : B_n \rightarrow [0, +\infty]$ such that $\int_{B_n} \phi d\eta_n = \int_{B_n} \phi \rho_n d\eta$ for every bounded measurable function $\phi : B_n \rightarrow \mathbb{R}$. Define $J_\eta|_{B_n} = \rho_n$. The essential uniqueness of J_η is a consequence of the essential uniqueness of the Radon–Nikodym derivative.
- A.3.5. Given any Borel set $B \subset M$, use Proposition A.3.2 and Lemma A.3.4 to construct Lipschitz functions $\psi_n : M \rightarrow [0, 1]$ such that $\mu(\{x \in M : \psi_n(x) \neq \chi_B(x)\}) \leq 2^{-n}$ for every n . Conclude that the claim in the exercise is true for every simple function. Extend the conclusion to every bounded measurable function, using the fact that it is a uniform limit of simple functions. Finally, for any integrable function, use the fact that the positive part and the negative part are monotone pointwise limits of bounded measurable functions. Now consider $M = [0, 1]$ and assume that there exists a sequence of continuous functions $\psi_n : M \rightarrow \mathbb{R}$ converging to the characteristic function ψ of $M \cap \mathbb{Q}$ at every point. Consider the set $R = \bigcap_m \bigcup_{n > m} \{x \in M : \psi_n(x) > 1/2\}$. On the one hand, $R = \mathbb{Q} \cap M$; on the other hand, R is a residual subset of M ; this is a contradiction.
- A.4.6. By the inverse function theorem, for every $x \in M$ there exist neighborhoods $U(x) \subset M$ of x and $V(x) \subset N$ of $f(x)$ such that f maps $U(x)$ diffeomorphically onto $V(x)$. This implies that the function $y \mapsto \#f^{-1}(y)$ is lower semi-continuous. Moreover, this function is bounded. Indeed, if there were $y_n \in N$ with $\#f^{-1}(y_n) \geq n$ for every $n \geq 1$ then, since M is compact, we could find $x_n, x'_n \in f^{-1}(y_n)$ distinct with $d(x_n, x'_n) \rightarrow 0$. Let x be any accumulation point of either sequence. Then f would not be injective in the neighborhood of x , contradicting the hypothesis. Let k be the maximum value of $\#f^{-1}(y)$. The set B_k of points $y \in N$ such that $\#f^{-1}(y) = k$ is open, closed and non-empty. Since N is connected, it follows that $B_k = M$.
- A.4.9. Consider local charts $\varphi_\alpha : U_\alpha \rightarrow X_\alpha, x \mapsto \varphi_\alpha(x)$ of M and $\varphi_\alpha : T_{U_\alpha}M \rightarrow X_\alpha \times \mathbb{R}^d, (x, v) \mapsto (\varphi_\alpha, D\varphi_\alpha(x)v)$ of TM . Note that $\varphi_\alpha \circ \pi \circ D\varphi_\alpha^{-1}$ is the canonical projection $X_\alpha \times \mathbb{R}^d \rightarrow X_\alpha$, which is infinitely differentiable. Since M is of class C^r and TM is of class C^{r-1} , it follows that π is of class C^{r-1} .
- A.5.2. (a) Use the fact that the exponential function is convex. (b) Starting from the Young inequality, show that $\int |f\bar{g}| d\mu \leq 1$ whenever $\|f\|_p = \|g\|_q = 1$. Deduce the general case of the Hölder inequality. (c) Start by noting that $|f + g|^p \leq |f| |f + g|^{p-1} + |g| |f + g|^{p-1}$. Apply the Hölder inequality to each of the terms on the right-hand side of this inequality to obtain the Minkowski inequality.
- A.5.6. (Rudin [Rud87, Theorem 6.16]) Note that $\Phi(g) \in L^p(\mu)^*$ and $\|\Phi(g)\| \leq \|g\|_q$; for $q < \infty$, that follows from the Hölder inequality; the case $q = \infty$ is immediate. It is clear that Φ is linear. To see that it is injective, given g such that $\Phi(g) = 0$, consider a function β with values on the unit circle such that $\beta g = |g|$. Then, $\phi(g)\beta = \int |g| d\mu = 0$, hence $g = 0$. We are left to prove that for every

$\phi \in L^p(\mu)^*$ there exists $g \in L^q(\mu)$ such that $\phi = \Phi(g)$ and $\|g\|_q = \|\phi\|$. For each measurable set $B \subset M$, define $\eta(B) = \phi(\mathcal{X}_B)$. Check that η is a complex measure (to prove σ -additivity one needs $p < \infty$) and observe that $\eta \ll \mu$. Consider the Radon–Nikodym derivative $g = (d\eta/d\mu)$. Then, $\phi(\mathcal{X}_B) = \int_B g d\mu$ for every B ; conclude that $\phi(f) = \int fg d\mu$ for every $f \in L^\infty(\mu)$. In the case $p = 1$, this construction yields $|\int_B g d\mu| \leq \|\phi\| \mu(B)$ for every measurable set. Deduce that $\|g\|_\infty \leq \|\phi\|$. Now suppose that $1 < p < \infty$. Take $f_n = \mathcal{X}_{B_n} |g|^{q-1}$, where $B_n = \{x : |g(x)| \leq n\}$. Observe that $f_n \in L^\infty(\mu)$ and $|f_n|^p = |g|^q$ in the set B_n and

$$\int_{B_n} |g|^q d\mu = \int f_n g d\mu = \phi(f_n) \leq \|\phi\| \left(\int |f_n|^p d\mu \right)^{1/p} \leq \|\phi\| \left(\int_{B_n} |g|^q d\mu \right)^{1/p}.$$

This yields $\int_{B_n} |g|^q d\mu \leq \|\phi\|^q$ for every n and, thus, $\|g\|_q \leq \|\phi\|$. Finally, $\phi(f) = \int fg d\mu$ for every $f \in L^p(\mu)$, since the two sides are continuous functionals and they coincide on the dense subset L^∞ .

- A.6.5. By definition, $u \cdot Lv = L^*u \cdot v$ and $u \cdot L^*v = (L^*)^*u \cdot v$ for any u and v . Hence, $v \cdot (L^*)^*u = L^*v \cdot u$ for any u and v . Reversing the roles of u and v , we see that $L = (L^*)^*$. Note that $\|L^*u \cdot v\| \leq \|L\| \|u\| \|v\|$ for every u and v . Taking $v = L^*u$, it follows that $\|L^*u\| \leq \|L\| \|u\|$ for every u and so $\|L^*\| \leq \|L\|$. Since $L = (L^*)^*$, it follows that $\|L\| \leq \|L^*\|$, hence the two norms coincide. Since the operator norm is submultiplicative, $\|L^*L\| \leq \|L\|^2$. On the other hand, $u \cdot L^*Lu = \|Lu\|^2$ and so $\|L^*L\| \|u\|^2 \geq \|Lu\|^2$, for every u . Deduce that $\|L^*L\| \geq \|L\|^2$ and so the two expressions coincide. Analogously, $\|LL^*\| = \|L\|^2$.
- A.6.8. Assume that $v \in H$ and $(u_n)_n$ is a sequence in E such that $u_n \cdot v \rightarrow u \cdot v$ for every $v \in H$. Considering $v \in E^\perp$, conclude that $v \in (E^\perp)^\perp$. By Exercise A.6.7, it follows that $u \in E$. Therefore, E is closed in the weak topology. Now consider any sequence $(v_n)_n$ in $U(E)$ converging to some $v \in H$. For each n , take $u_n = h^{-1}(v_n) \in E$. Since h is an isometry, $\|u_m - u_n\| = \|v_m - v_n\|$ for any m, n . It follows that $(u_n)_n$ is a Cauchy sequence in E and so it admits a limit $u \in E$. Hence, $v = h(u)$ is in $U(E)$.
- A.7.1. The inverse of $T + H$ is given by the equation $(T + H)(T^{-1} + J) = \text{id}$, which may be rewritten as a fixed point equation $J = -L^{-1}HL^{-1} + L^{-1}HJ$. Use the hypothesis to show that this equation admits a (unique) solution. Hence, $T + H$ is an isomorphism. Deduce that $L - \lambda \text{id}$ whenever $\lambda > \|L\|$. Therefore, the spectrum of L is contained in the disk of radius $\|L\|$. It also follows from the previous observation that if $L - \lambda \text{id}$ is an isomorphism then the same is true for $L - \lambda' \text{id}$ if λ' is sufficiently close to λ .
- A.7.4. (a) Observe that $L - \lambda \text{id} = \int (z - \lambda) dE(z)$ and use Lemma A.7.4. By the continuity from below property (Exercise A.1.9), $E(\{\lambda\}) = \lim_n E(\{z : |z - \lambda| \leq 1/n\})$. It follows that $E(\{\lambda\})v = v$. (b) It follows from Exercise A.7.3 that $E(B)E(\{\lambda\}) = E(\{\lambda\})$ if $\lambda \in B$ and $E(B)E(\{\lambda\}) = E(\emptyset) = 0$ otherwise. Since $L = \int z dE(z)$, we get that $Lv = \lambda E(\{\lambda\})v = \lambda v$.