Gerhard Keller

Ergodic Theory

Foundations of measurable dynamics

Lecture Notes
Summer semester 2014
Department Mathematik
Universität Erlangen-Nürnberg

July 15, 2014

Contents

1 Ergodicity and ergodic theorems

A huge part of this course is devoted to a study of ergodic decompositions of very general dynamical systems. The main result is the Theorem of Raugi [15] that gives an ergodic decomposition in the context of a countable family of Borel-automorphisms acting non-singularly on a standard probability space (also Lebesgue-Rohlin space).[1] Along the way a number of more sophisticated tools from measure and ergodic theory must be introduced. We start nevertheless with some basic examples.

1.1 Introduction

1.1.1 Examples of dynamical systems

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space. We consider measurable maps $T : \Omega \to \Omega$ that leave the measure class of μ invariant, i.e. for which $\mu \circ T^{-1} \approx \mu$. In many cases we will even have $\mu \circ T^{-1} = \mu$ - then we call $(\Omega, \mathcal{A}, \mu, T)$ a measure preserving dynamical system (mpds).

1.1.1 Example $\Omega = \mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$, $\mathcal{A}$ the σ-algebra of Borel or Lebesgue sets, μ Lebesgue measure. All arithmetic operations are of course taken $\mod \mathbb{Z}^d$.

a) $T(x) = x + a$ for some $a \in \mathbb{R}^d$. Then $\mu \circ T^{-1} = \mu$. In this case $T^{-1}(x) = x - a$.

b) $T(x) = A\,x$ for some $d \times d$ integer matrix A with $\det(A) \neq 0$. In this case T is invertible if and only if $|\det(A)| = 1$, and then $T^{-1}(x) = A^{-1}x$.

In both cases μ is T-invariant.

1.1.2 Example $\Omega = \{0,1\}^{\mathbb{N}}$ or $\Omega = \{0,1\}^{\mathbb{Z}}$ equipped with the product topology, $\mathcal{A}$ the Borel-σ-algebra, $T : \Omega \to \Omega$ the left shift, i.e. $(T(\omega))_n = \omega_{n+1}$ for all n. This transformation is continuous. In case of $\Omega = \{0,1\}^{\mathbb{N}}$, T is not invertible; in case of $\Omega = \{0,1\}^{\mathbb{Z}}$, T is a homeomorphism.

a) μ is the p-Bernoulli measure μ_p for some $p \in (0,1)$, that means μ is the infinite product of the marginal measure $p\delta_1 + (1-p)\delta_0$.

b) $\mu = \int_0^1 \mu_p \, dm(p)$ is a mixture of Bernoulli measures (m is any Borel probability measure on $(0,1)$).

In both cases μ is T-invariant. (For some background on Bernoulli measures see e.g. [10, items 1.3-1.5, 1.9, 1.11, 2.3, 2.16, 2.17, 2.28].)

1.1.3 Example $\Omega = \{0,1\}^{\mathbb{Z}}$, $\mathcal{A}$ and T as before, but $\mu = \left(\mu_p \circ (\pi_{-\infty}^{-1})^{-1}\right) \otimes \left(\mu_{p'} \circ (\pi_0^{\infty})^{-1}\right)$ for $p \neq p'$. This μ is obviously not shift-invariant, but $\mu \circ T^{-1} \approx \mu$.

[1] The ergodic decomposition result is claimed even for uncountable families in [15], but we will see in Example 3.2.12 that at least the proof of this far reaching claim cannot be correct.

1.1.4 Example For a domain $D \subseteq \mathbb{R}^{2d}$ of finite volume let $H \in C^2(D)$. Consider the system of differential equations (Hamiltonian system)

$$\begin{aligned} \dot{q}_i &= \frac{\partial}{\partial p_i} H(q_1, \ldots, q_d, p_1, \ldots, p_d) \\ \dot{p}_i &= -\frac{\partial}{\partial q_i} H(q_1, \ldots, q_d, p_1, \ldots, p_d) \end{aligned} \tag{1.1.1}$$

$(i = 1, \ldots, d)$ and suppose that for initial conditions in D the solution $\Phi_t(q, p)$ exists for all positive and negative times t. Denote by μ the normalized Lebesgue measure on D and by $\mathcal{A}$ the σ-algebra of Lebesgue measurable sets. By *Liouville's theorem* each $T = \Phi_t : D \to D$ preserves μ.

1.1.5 Remark The action of an invertible T can be interpreted as a measure preserving action of the group $(\mathbb{Z}, +)$ on $(\Omega, \mathcal{A}, \mu)$: $n \cdot \omega := T^n(\omega)$. It satisfies the flow equation

$$(n + m) \cdot \omega = n \cdot (m \cdot \omega) \quad \text{for all } n, m \in \mathbb{Z}\,. \tag{1.1.2}$$

Similarly the group $(\mathbb{R}, +)$ acts by the family $(\Phi_t)_{t \in \mathbb{R}}$ and satisfies the corresponding flow equation for $t \cdot \omega := \Phi_t(\omega)$:

$$(s + t) \cdot \omega = s \cdot (t \cdot \omega)\,.$$

In the same way one can define more general group actions, for example for the action of $(\mathbb{Z}^2, +)$ on $\{0, 1\}^{\mathbb{Z}^2}$ by the left shift $(T_1(\omega))_{(i_1,i_2)} = \omega_{(i_1+1,i_2)}$ and the down shift $(T_2(\omega))_{(i_1,i_2)} = \omega_{(i_1,i_2+1)}$ (note that $T_1 \circ T_2 = T_2 \circ T_1$):

$$(n_1, n_2) \cdot \omega = T_1^{n_1}(T_2^{n_2}(\omega))\,.$$

1.1.6 Remark Each mpds $(\Omega, \mathcal{A}, \mu, T)$ defines a linear operator on the Hilbert space $L^2(\Omega, \mathcal{A}, \mu)$ by

$$U_T f := f \circ T$$

that preserves the scalar product: $\langle U_T f, U_T g \rangle = \int (f \circ T) \cdot (\overline{g} \circ T)\, d\mu = \int f \cdot \overline{g}\, d\mu = \langle f, g \rangle$. Indeed, if T is invertible, U_T is unitary.

A basic question is whether such systems can be decomposed into subsystems that can be studied separately. Here "decomposition" does not necessarily mean a decomposition of the space Ω but rather a decomposition of the measure μ, although this sometimes yields a decomposition of the space as a by-product. We tackle this question in the next subsection.

1.1.2 Ergodicity

Let us say that a mpds $(\Omega, \mathcal{A}, \mu, T)$ is decomposable, if there exists $A \in \mathcal{A}$ with $0 < \mu(A) < 1$ and $T^{-1}(A) = A$. Then also $T^{-1}(A^c) = A^c$, and one can study the two subsystems $T_{|A}$ and $T_{|A^c}$ independently from each other. For each $\omega \in \Omega$, then $A := \{T^n \omega : n \in \mathbb{Z}\}$ is T-invariant, and naïvely one might think that in a uncountable space there should always be many sets made up of complete orbits. But, as we shall see, measurability of A can be a strong restriction to the construction of such sets.

1.1.7 Definition *Denote by $\mathcal{I}(T) := \{A \in \mathcal{A} : T^{-1}(A) = A\}$ the σ-algebra of T invariant measurable sets, and by $\mathcal{I}_\mu(T) := \{A \in \mathcal{A} : T^{-1}(A) = A \bmod \mu\}$. (They are both σ-algebras, indeed).*

A mpds $(\Omega, \mathcal{A}, \mu, T)$ is ergodic, if $\forall A \in \mathcal{I}(T) : \mu(A) \in \{0, 1\}$.

1.1.8 Lemma *Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds and let $A \in \mathcal{A}$. Then*

$$B(A) := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} T^{-k}(A) \in \mathcal{I}(T) \quad \textit{and} \quad \mu\left(\bigcup_{k=n}^{\infty} T^{-k}(A)\right) = \mu(B(A)) \; \forall n \in \mathbb{N} .$$

Proof: Let $B_n := \bigcup_{k=n}^{\infty} T^{-k}(A)$. Then $T^{-1}(B_n) = B_{n+1} \subseteq B_n$. As $\mu(T^{-1}(B_n)) = \mu(B_n)$, this implies $\mu(B_n \backslash B_{n+1}) = 0$. Then $B(A) = \bigcap_{n=1}^{\infty} B_n$, $T^{-1}B(A) = B(A)$ and $\mu(B_n) = \mu(B(A))$ $(n \in \mathbb{N})$. □

1.1.9 Corollary *Let Ω be a topological space with a countable basis, $\mathcal{A}$ its Borel σ-algebra and $(\Omega, \mathcal{A}, \mu, T)$ an ergodic mpds. If $\mu(U) > 0$ for each open set U, then $(T^n \omega)_{n \in \mathbb{N}}$ is dense in Ω for μ-a.e. ω.*

Proof: Let $U_1, U_2, \ldots$ be a basis for the topology. By Lemma 1.1.8, $\mu\left(\bigcap_{i=1}^{\infty} B(U_i)\right) = 1$, so that almost every trajectory visits each U_i infinitely often. □

1.1.10 Theorem *The following conditions are equivalent to the ergodicity of $(\Omega, \mathcal{A}, \mu, T)$:*

1) $\forall A \in \mathcal{I}_\mu(T) : \mu(A) \in \{0, 1\}$.

2) For all $A, B \in \mathcal{A}$ with $\mu(A), \mu(B) > 0$ there is $n \in \mathbb{N}$ such that $\mu(T^{-n}(A) \cap B) > 0$.

3) For all μ-integrable $f : \Omega \to \mathbb{R}$ holds: if $f \circ T = f$ μ-a.e., then $f = const$ μ-a.e.

4) If $f \in L^2(\Omega, \mathcal{A}, \mu)$ satisfies $U_T f = f$, then $f = const$ in L^2.

Proof: The equivalence of 1) and of 2) follow from the foregoing lemma. The other equivalences are an easy exercise, see e.g. [17, Theorems 1.6] □

We check the examples from subsection 1.1.1 for ergodicity.

1.1.11 Theorem *Let $\Omega = \mathbb{T}^d$ and let $\mathcal{A}$ and μ be as in Example 1.1.1.*

a) The mpds with $T(x) = x + a$, $a \in \mathbb{R}^d$, is ergodic if and only if $\forall n \in \mathbb{Z}^d \backslash \{0\} : \langle n, a \rangle \notin \mathbb{Z}$.

b) The mpds with $T(x) = A\,x$ is ergodic if and only if A has no root of unity as an eigenvalue.

Proof: We make use of the fact that each $f \in L^2(\mathbb{T}^d, \mathcal{A}, \mu)$ has a unique Fourier series expansion $f(x) = \sum_{n \in \mathbb{Z}^d} c_n\, e^{2\pi i \langle n, x \rangle}$ with coefficients $c_n \in \mathbb{C}$ (convergence in L^2).

a) For $f \in L^2$ we have

$$f(x) = \sum_{n \in \mathbb{Z}^d} c_n\, e^{2\pi i \langle n, x \rangle} \quad \text{and}$$

$$U_T f(x) = f(x + a) = \sum_{n \in \mathbb{Z}^d} c_n\, e^{2\pi i \langle n, x+a \rangle} = \sum_{n \in \mathbb{Z}^d} c_n\, e^{2\pi i \langle n, a \rangle}\, e^{2\pi i \langle n, x \rangle},$$

so that $U_T f = f$ if and only if $c_n = c_n\, e^{2\pi i \langle n, a \rangle}$ for all $n \in \mathbb{Z}^d$. If $\langle n, a \rangle \notin \mathbb{Z}$ for all $0 \neq n \in \mathbb{Z}^d$, then $c_n = 0$ for all $n \neq 0$ and $f =$ const in L^2. If, on the other hand, there is $0 \neq n \in \mathbb{Z}^d$ such that $\langle n, a \rangle \in \mathbb{Z}$, then $f(x) = e^{2\pi i \langle n, x \rangle}$ is a U_T-invariant non-constant L^2-function. The claim follows now from Theorem 1.1.10.

b) Consider again $f \in L^2$. Now

$$U_T f(x) = f(A\,x) = \sum_{n \in \mathbb{Z}^d} c_n\, e^{2\pi i \langle n, A\,x \rangle} = \sum_{n \in \mathbb{Z}^d} c_n\, e^{2\pi i \langle A^t n, x \rangle} .$$

Therefore, $U_T f = f$ if and only if $c_{A^t n} = c_n$ for all $n \in \mathbb{Z}^d$, and is equivalent to

$$\forall n \in \mathbb{Z}^d\ \forall k \in \mathbb{N} : \ c_{(A^t)^k n} = c_n .$$

As $\sum_{n \in \mathbb{Z}^d} |c_n|^2 = \|f\|_2^2 < \infty$, $c_n \neq 0$ implies that $(A^t)^k n = n$ for some $k \in \mathbb{N}\backslash\{0\}$. If A has no eigenvalue which is a root of unity, it follows that $c_n = 0$ for all $n \in \mathbb{Z}\backslash\{0\}$, i.e. $f = \text{const}$ in L^2. In view of Theorem 1.1.10 this proves the ergodicity of T. Conversely, if there A (and hence A^t) has a k-th root of unity as eigenvalue, then $(A^t)k$ has eigenvalue 1. As $(A^t)^k$ has only integer coefficients, the it has an eigenvector $v \in \mathbb{Q}^d\backslash\{0\}$, and multiplying it with the product of all its denominators, we can even assume that $v \in \mathbb{Z}^d\backslash\{0\}$. Define $f(x) = \sum_{j=0}^{k-1} e^{2\pi i \langle (A^t)^j v, x \rangle}$. Then $f \neq \text{const}$ but $U_T f = f$, so T is not ergodic.

□

Analogous characterizations of ergodicity are known for translations and endomorphisms of compact metric groups equipped with Haar measure, see [17, §1.5]. Here we are more interested in understanding the non-ergodic examples.

1.1.12 Remark Let $\Omega = \mathbb{T}^d$, $\mathcal{A}$ and μ be as in Example 1.1.1 and the previous theorem. We focus on the non-ergodic cases.

a) Let $T(x) = x + a$. Consider first the case $d = 1$. Then non-ergodicity is equivalent to a being rational, $a = p/q$. Then $T^q(x) = x$ for all x, i.e. each set $J \cup T(J) \cup \cdots \cup T^{q-1}(J)$ is invariant such that there are plenty non-trivial measurable invariant sets. Denote by μ_x the equi-distribution on the closed orbit $O_T(x) := \{x, T(x), \ldots, T^{q-1}(x)\}$. Then $\mu = \int_\Omega \mu_x\, d\mu(x)$ in the sense that

$$\mu(A) = \int_\Omega \mu_x(A)\, d\mu(x) \quad \text{for all } A \in \mathcal{A} . \tag{1.1.3}$$

(This statement includes the assertion that the functions $x \mapsto \mu_x(A)$ are measurable.) As each $(O_T(x), \mathcal{A}_{|O_T(x)}, \mu_x, T_{|O_T(x)})$ is an ergodic mpds, (1.1.3) is an ergodic decomposition of the original mpds. A more convincing way to write this down would be to consider the space $\mathbb{O}_T := \{O_T(x) : x \in \mathbb{T}^1\}$ of all orbits of T. (In this algebraic setting $\mathbb{O}_T$ can be interpreted as the factor group $\mathbb{T}^1/[\frac{1}{q}]$ where $[\frac{1}{q}]$ denotes the subgroup of $\mathbb{T}^1$ generated by $\frac{1}{q}$.[2]) So one might write

$$\mu = \int_{\mathbb{O}_T} \mu_\omega\, d(\mu \circ \pi^{-1})(\omega)$$

where μ_ω denotes the equidistribution on $\omega \in \mathbb{O}_T$ and $\pi : \mathbb{T}^1 \to \mathbb{O}_T$ maps x to $\omega = O_T(x)$.

In case $d \geqslant 2$ the situation is a bit more complex. For $n \in \mathbb{Z}^d$ denote

$$G_n := \{x \in \mathbb{T}^d : \langle n, x \rangle \in \mathbb{Z}\} .$$

[2] In more general situations this space may be rather awkward, and to deal with it properly is one of the things we will have to learn.

G_n is well defined, because for $x \in \mathbb{R}^d$ and $m \in \mathbb{Z}^d$ holds: $\langle n, x\rangle \in \mathbb{Z} \Leftrightarrow \langle n, x+m\rangle \in \mathbb{Z}$. G_n is a subgroup of $\mathbb{T}^d$. If T is not ergodic, i.e. if $\langle n, a\rangle \in \mathbb{Z}$, then $\langle n, T(x)\rangle = \langle n, x\rangle + \langle n, a\rangle \in \langle n, x\rangle + \mathbb{Z}$. In particular, $T(G_n) = G_n$, and also all cosets of G_n are invariant so that each subset $A \subseteq \mathbb{T}^d$ made up of cosets of G_n is T-invariant. Let $x \in G_n$ and $v \in \mathbb{R}^d$ with $0 < |\langle n, v\rangle| < 1$. Then $\langle n, x+v\rangle = \langle n, v\rangle + \langle n, x\rangle \notin \mathbb{Z}$ such that $x+v \notin G_n$. Therefore G_n is a finite collection of parallel hyperplanes separated by a distance at least $1/\|n\|_2$. This implies that G_n and also each coset is a finite union of parallel closed $d-1$-dimensional submanifolds. In many cases this is already the ergodic decomposition, but that need not be the case: consider $d = 2$ and $a \in \mathbb{Q}^2$, say $a = (p_1/n_1, p_2/n_2)$. Then $\langle n, a\rangle = p_1 + p_2 \in \mathbb{Z}$ and $T^{|n_1 n_2|}(x) = x + |n_1 n_2|\, a = x + (n_2 p_1, n_1 p_2) = x$ in $\mathbb{T}^2$. Hence T acts like a rational rotation inside each coset.

b) Let $T(x) = A\,x$ and assume that $|\det(A)| = 1$, i.e. T is a group isomorphism of $\mathbb{T}^d$. If T is not ergodic, then A has an eigenvalue which is a root of unity, and the proof of Theorem 1.1.11 shows that there are indeed $k \in \mathbb{N}\backslash\{0\}$ and $n \in \mathbb{Z}^d\backslash\{0\}$ such that $(A^t)^k n = n$. Hence, for all $x \in \mathbb{T}^d$ we have $\langle n, x\rangle = \langle (A^t)^k n, x\rangle = \langle n, A^k x\rangle = \langle n, T^k(x)\rangle$, in particular $x \in G_n \Leftrightarrow T^k(x) \in G_n$. The identity shows even more: T^k leaves all cosets of G_n invariant, but again the ergodic decomposition may be finer, just think of the case $T = \mathrm{id}_{\mathbb{T}^d}$ where the ergodic decomposition is into single points (point masses on the level of measures).

1.1.13 Theorem *The shift dynamical system* $(\{0,1\}^{\mathbb{Z}}, \mathcal{A}, \mu_p, T)$ *from Example 1.1.2a is ergodic.*

Proof: Suppose $A = T^{-1}(A) \in \mathcal{A}$ and let $\epsilon > 0$. As the measure μ is determined by the algebra of sets that are specified on finitely many indices $i \in \mathbb{Z}$, there are $s \in \mathbb{N}$ and $B \subseteq \Omega$ specified by the indices from $\{-s, \dots, s\}$ and such that $\mu(A \triangle B) < \epsilon$. As μ is T-invariant, it follows easily that

$$\mu(T^{-2s}(A) \triangle T^{-2s}(B)) < \epsilon \text{ and}$$
$$\mu((A \cap T^{-2s}(A)) \triangle (B \cap T^{-2s}(B))) < 2\epsilon\,.$$

Hence,

$$\begin{aligned}|\mu(A) - \mu(A)^2| &= |\mu(A \cap T^{-2s}(A)) - \mu(A)\mu(T^{-2s}(A))| \\ &\leqslant |\mu(B \cap T^{-2s}(B)) - \mu(B)\mu(T^{-2s}(B))| + 4\epsilon\,.\end{aligned}$$

But as μ is a product measure and as B is specified by indices from $\{-s, \dots, s\}$, the last difference is zero so that $|\mu(A) - \mu(A)^2| < 4\epsilon$. As ϵ was arbitrary, this implies $\mu(A) = \mu(A)^2$ and hence $\mu(A) \in \{0, 1\}$. □

1.1.14 Remark What about the measure $\mu = \int_0^1 \mu_p\, dm(p)$ from Example 1.1.2b? It is explicitly described as a mixture of ergodic invariant measures, so its integral representation should be already the ergodic decomposition. This decomposition can also be realized as a decomposition of the underlying space $\Omega = \{0,1\}^{\mathbb{Z}}$: For $p \in [0,1]$ let

$$\Omega_p^{\pm} := \left\{\omega \in \Omega : \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \omega_{\pm k} = p\right\}$$

Obviously, $\Omega = \biguplus_{p\in[0,1]} \Omega_p^+$ and also $\Omega = \biguplus_{p\in[0,1]} \Omega_p^-$. Furthermore, by the law of large numbers, $\mu_p(\Omega_p^+) = \mu_p(\Omega_p^-) = 1$. This shows a problem with the decomposition of the space: It is far from unique, because $\Omega_p^+ \cap \Omega_{p'}^- \neq \emptyset$ for all $p, p' \in [0,1]$. Observe also that each $\Omega_p^\pm$ is dense in Ω by Corollary 1.1.9. Indeed, for each $p \in (0,1)$ the orbit $(T^n(\omega))_{n\in\mathbb{N}}$ is dense in Ω for μ_p-a.e. ω.

It remains to study the non-shift invariant measure $\mu = \big(\mu_p \circ (\pi_{-\infty}^{-1})^{-1}\big) \otimes \big(\mu_{p'} \circ (\pi_0^\infty)^{-1}\big)$ from Example 1.1.3. I do not know whether it is ergodic.

1.1.15 Remark For Hamiltonian flows as in Example 1.1.4 the Hamiltonian H itself is a constant of motion:

$$\frac{d}{dt}(H \circ \Phi_t) = \sum_{i=1}^d \left(\frac{\partial H}{\partial q_i} \circ \Phi_t \cdot \dot{q}_i(t) + \frac{\partial H}{\partial p_i} \circ \Phi_t \cdot \dot{p}_i(t) \right) = \sum_{i=1}^d \left(-\dot{p}_i(t)\, \dot{q}_i(t) + \dot{q}_i(t)\, \dot{p}_i(t) \right) = 0$$

It follows that $U_{\Phi_T} H = H$ for each $t \in \mathbb{R}$. If H is non-constant (otherwise there is no movement), Hamiltonian flows and their time-t-maps are not ergodic. Whether the "energy shells" $\{H = E\}$ for $E \in \mathbb{R}$ yield the ergodic decomposition or whether there are additional integrals of motion can be a mathematically very deep problem – depending on the concrete example.

1.1.16 Remark If one does not just study the action of a single measure preserving transformation or one-parameter flow but the action $\mathcal{S} = (S_g)_{g\in G}$ of a more general group on $(\Omega, \mathcal{A}, \mu)$, then ergodicity of this action is defined in terms of the σ-algebra

$$\mathcal{I}(\mathcal{S}) := \{A \in \mathcal{A} : S_g^{-1}(A) = A \ \forall g \in G\} = \bigcap_{g\in G} \{A \in \mathcal{A} : S_g^{-1}(A) = A\},$$

and if this σ-algebra is not trivial, one would like to know its ergodic decomposition.

1.1.3 Exercises

1.1.1 Determine $\frac{d(\mu \circ T^{-1})}{d\mu}$ in Example 1.1.3.

1.1.2 Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds - invertible or not. Prove that $U_T^* U_T = \mathrm{id}_{L^2}$ and that $U_T U_T^*$ is the orthogonal projection onto $U_T(L^2)$.

1.1.3 Suppose that T is invertible on $(\Omega, \mathcal{A}, \mu)$ and that $\mu \circ T \approx \mu$ but that μ is not necessarily T-invariant. Define $U_T f := \sqrt{\frac{d(\mu \circ T)}{d\mu}} \cdot (f \circ T)$ and prove that U_T is a unitary operator on $L^2(\Omega, \mathcal{A}, \mu)$.

1.1.4 Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds. Prove that the system is ergodic if and only if each for T-invariant probability ν on $(\Omega, \mathcal{A})$ holds: $\nu \ll \mu \Rightarrow \nu = \mu$. *Hint:* The proof for invertible systems is rather straightforward; for non-invertible ones conditional expectations can help.

1.1.5 Let T be a rotation or an endomorphism of $\mathbb{T}^d$. Prove that T is ergodic if and only if Lebesgue-almost all $\omega \in \mathbb{T}^d$ have a dense orbit under T.

1.1.6 Let $\Omega = \{0,1\}^{\mathbb{N}}$, $\mu = \mu_{1/2}$ and $\mathcal{A}$ the Borel σ-algebra (or its μ-completion) of Ω. For $\omega \in \Omega$ denote

$$\ell(\omega) = \inf\{\ell \in \mathbb{N} : \omega_\ell = 0\} \qquad (= +\infty \text{ if no such } \ell \text{ exists}).$$

Define $T : \Omega \to \Omega$ as

$$(T(\omega))_i = \begin{cases} 0 & \text{if } i < \ell(\omega) \\ 1 & \text{if } i = \ell(\omega) \\ \omega_i & \text{if } i > \ell(\omega) \end{cases}$$

(This is the so called binary adding machine - addition with carry to the right. It is also called an odometer.) Prove that $\mu \circ T^{-1} = \mu$ and that the system $(\Omega, \mathcal{A}, \mu, T)$ is ergodic.

1.2 Birkhoff's ergodic theorem

1.2.1 Theorem *Let* $(\Omega, \mathcal{A}, \mu, T)$ *be a mpds and* $f \in L^1(\Omega, \mathcal{A}, \mu)$. *Then*

$$\bar{f}(\omega) := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(\omega))$$

exists for μ*-a.e.* $\omega \in \Omega$, *and* $\bar{f}$ *is a version of the conditional expectation* $E_\mu[f \mid \mathcal{I}(T)]$, *i.e.*

(1) $\bar{f}$ *is* $\mathcal{I}(T)$*-measurable* $\bmod \mu$ *and*

(2) $\forall\, A \in \mathcal{I}(T) : \int_A \bar{f}\, d\mu = \int_A f\, d\mu$.

(Observe that $\bar{f}$ *is uniquely characterized (* $\bmod \mu$*) by properties (1) and (2).)*

There are many different *proofs* of this theorem. The one in [10, Satz 15.1] assumes that the concept of a conditional expectation is known in the sense that a function $\bar{f}$ with properties (1) and (2) exists. We recall briefly the main steps to the construction of this object.

Let μ and ν be σ-finite measures on a measurable space $(\Omega, \mathcal{A})$.

- ν has density f w.r.t. μ, in short $\nu = f\mu$, if $\forall A \in \mathcal{A} : \nu(A) = \int_A f\, d\mu$.
- ν is *absolutely continuous* w.r.t. μ, in short $\nu \ll \mu$, if $\forall A \in \mathcal{A} : \mu(A) = 0 \Rightarrow \nu(A) = 0$.

Obviously $\nu = f\mu$ implies $\nu \ll \mu$. The converse is guaranteed by the *Randon-Nikodym* theorem[10, Satz 11.13].

1.2.2 Theorem *If* $\nu \ll \mu$, *then* ν *has density w.r.t.* μ. *This density is uniquely determined* $\bmod \mu$, *and it is denoted by* $\frac{d\nu}{d\mu}$.

A surprising application is the following: Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $0 \leqslant f \in L^1(\Omega, \mathcal{A}, \mu)$. Then $f\mu$ is a finite measure on $(\Omega, \mathcal{A})$ and $f\mu \ll \mu$. Now consider a sub-σ-algebra $\mathcal{F} \subseteq \mathcal{A}$ and restrict both measures to $\mathcal{F}$. Then $(f\mu)|_{\mathcal{F}} \ll \mu|_{\mathcal{F}}$, and the Radon-Nikodym theorem can be applied to these two measures on $(\Omega, \mathcal{F})$. This results in the density

$$E_\mu[f \mid \mathcal{F}] := \frac{d\left((f\mu)|_{\mathcal{F}}\right)}{d\left(\mu|_{\mathcal{F}}\right)},$$

which is a $\mathcal{F}$ measurable function $\Omega \to \mathbb{R}$ (defined only $\bmod \mu$). This is property (1) in the ergodic theorem. If $A \in \mathcal{F}$, then

$$\int_A E_\mu[f \mid \mathcal{F}]\, d\mu = \int_A E_\mu \mid \mathcal{F}]\, d\mu|_{\mathcal{F}} = \int_A \frac{d\left((f\mu)|_{\mathcal{F}}\right)}{d\left(\mu|_{\mathcal{F}}\right)}\, d\mu|_{\mathcal{F}} = \int_A d(f\mu)|_{\mathcal{F}} = \int_A d(f\mu) = \int_A f\, d\mu\,,$$

which is property (2) in the ergodic theorem. For more on conditional expectations see [10, Kap. 12]. There you read in particular that $E_\mu[f \mid \mathcal{F}]$ is the *orthogonal projection* of f onto the closed linear subspace $L^2(\Omega, \mathcal{F}, \mu)$.

If $\mathcal{F}$ is trivial, i.e. if $\mathcal{F}$ contains only sets of measure 0 or 1, then $E_\mu[f \mid \mathcal{F}]$ is constant because of (1), and the constant value is $\int f\,d\mu$ because of (2). This yields the following corollary of the ergodic theorem:

1.2.3 Corollary *Let $(\Omega, \mathcal{A}, \mu, T)$ be an ergodic mpds and $f \in L^1(\Omega, \mathcal{A}, \mu)$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(\omega)) = \int f\,d\mu \quad \text{for } \mu\text{-a.e. } \omega \in \Omega.$$

1.2.4 Remark A far-reaching generalization of the ergodic theorem will be proved in the next chapter. The starting point for this generalization is that the map $f \mapsto f \circ T$ is a positive linear contraction on $L^1(\Omega, \mathcal{A}, \mu)$:

- $(\alpha f + \beta g) \circ T = \alpha\,(f \circ T) + \beta\,(g \circ T)$
- $f \geqslant 0 \Rightarrow f \circ T \geqslant 0$
- $\|f \circ T\|_1 = \int |f \circ T|\,d\mu = \int |f|\,d\mu = \|f\|_1$ (Here "$\leqslant$" would be sufficient for a contraction.)

Indeed, we will prove very general ergodic theorems for such linear operators on L^1. A special case will be Birkhoff's ergodic theorem, and the more general statements will be used in later chapters.

1.2.1 Exercises

1.2.1 Consider the mpds $(\{0,1\}^{\mathbb{N}}, \mathcal{A}, \mu_p, T)$ from Example 1.1.2 with the left shift T. Determine the limit

$$\lim_{n\to\infty} \frac{1}{n} \#\{k \in \{0, \ldots, n-1\} : \omega_k \omega_{k+1} \omega_{k+2} = 101\}$$

for μ_p-a.e. ω.

1.2.2 Consider the mpds $(\{0,1\}^{\mathbb{N}}, \mathcal{A}, \mu, T)$ from Example 1.1.2 with the left shift T and the measure $\mu = \int_{(0,1)} \mu_p\,dm(p)$. Determine the limit

$$\lim_{n\to\infty} \frac{1}{n} \#\{k \in \{0, \ldots, n-1\} : \omega_k \omega_{k+1} \omega_{k+2} = 101\}$$

for μ-a.e. ω.

1.2.3 Consider $T : \mathbb{T}^1 \to \mathbb{T}^1$, $T(x) = x + a$ with some irrational $a \in \mathbb{R}$ and let $f \in C(\mathbb{T}^1, \mathbb{R})$. Prove that $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(\omega)) = \int_{\mathbb{T}^1} f(x)\,dx$ for <u>all</u> $\omega \in \mathbb{T}^1$.

2 Ergodic theory of positive L^1-contractions

This chapter follows the corresponding parts of the textbooks by Neveu [13] and Krengel [11], mostly [11, §§3.1-3.3]. Throughout the chapter we fix a σ-finite measure space $(\Omega, \mathcal{A}, m)$. We use the short-hand notation L^p for $L^p(\Omega, \mathcal{A}, m)$.

2.1 The Hopf decomposition

2.1.1 Operators and kernels

2.1.1 Definition *a) A map $p : \Omega \times \mathcal{A} \to [0, 1]$ is a substochastic kernel if*

(K1) $\omega \mapsto p(\omega, A)$ is $\mathcal{A}$-measurable for all $A \in \mathcal{A}$, and

(K2) $A \mapsto p(\omega, A)$ is a measure for all $\omega \in \Omega$.

p is a stochastic kernel if $p(\omega, \Omega) = 1$ for all $\omega \in \Omega$.

b) The kernel p is null-preserving, if $p(\omega, A) = 0$ for m-a.e. ω if $A \in \mathcal{A}$ and $m(A) = 0$.

2.1.2 Examples a) $T : \Omega \to \Omega$ is measurable and $m \circ T^{-1} \approx m$. Define $p(\omega, A) = \delta_{T(\omega)}(A)$. If $m(A) = 0$, then $m(T^{-1}(A)) = 0$ and hence $p(\omega, A) = \delta_\omega(T^{-1}(A)) = 0$ for μ-a.e. ω.

b) $\Omega = \mathbb{Z}$, m the counting measure and $p(\omega, .\,) = \frac{1}{2}\delta_{\omega-1} + \frac{1}{2}\delta_{\omega+1}$. This kernel describes the random transitions of a symmetric nearest neighbour random walk on $\mathbb{Z}$. $p(\omega, A)$ is the probability to be in A at time $t + 1$ if the process is in ω at time t.

A null-preserving substochastic kernel p defines the following linear operators:

$$\begin{aligned} P : L^1 \to L^1, \quad & Pf(\omega) = \frac{d}{dm}\left(\int f(\omega')p(\omega', .\,)\, dm(\omega')\right)(\omega) \\ P^* : L^\infty \to L^\infty, \quad & P^*h(\omega) = \int h(\omega')\, p(\omega, d\omega')\ . \end{aligned} \tag{2.1.1}$$

Both operators are positive linear contractions. For P^* this is quite obvious. For P observe first that the integral expression is a finite measure (it is the mixture of the (sub-)probability measures $P(\omega, .\,)$ with the finite measure fm), and that this measure is $\ll m$, because the kernel is null-preserving. Then linearity and positivity of P are obvious, and

$$\|Pf\|_1 = \int_\Omega |Pf|\, dm \leqslant \int_\Omega P|f|\, dm = \int_\Omega |f|(\omega)p(\omega, \Omega)\, dm(\omega) \leqslant \int_\Omega |f|\, dm = \|f\|_1$$

shows that P is a contraction. Note also that for all $A \in \mathcal{A}$

$$\int_A Pf\, dm = \int_A d\left(\int f(\omega')p(\omega', .\,)\, dm(\omega')\right) = \int f(\omega')p(\omega', A)\, dm(\omega') \tag{2.1.2}$$

and that this identity determines P uniquely.

P^* is indeed the dual to P: For $f \in L^1$ and $h \in L^\infty$ we have by Fubini's theorem (version for kernels)

$$\begin{aligned}\int P^*h \cdot f\,dm &= \int \left(\int h(\omega')\,p(\omega, d\omega') \cdot f(\omega) \right) dm(\omega) \\ &= \int h(\omega') \frac{d}{dm} \left(\int f(\omega) p(\omega,\,.\,)\,dm(\omega \right) (\omega')\,dm(\omega') \\ &= \int h \cdot Pf\,dm\end{aligned} \tag{2.1.3}$$

2.1.3 Remark In a later chapter we will see that on most "interesting" probability spaces each linear contraction $P : L^1 \to L^1$ is associated to a null-preserving kernel as above.

2.1.4 Examples a) The kernel $p(\omega, A) = \delta_{T(\omega)}(A)$ from Example 2.1.2a gives rise to operators P and P^* where $P^*h(\omega) = h(T(\omega))$ and $Pf = \frac{d\left((f\,m)\circ T^{-1}\right)}{dm}$. If T is invertible this can also be written as $Pf = f \circ T^{-1} \cdot \frac{d(m\circ T^{-1})}{dm}$.

b) The kernel $p(\omega,\,.\,) = \frac{1}{2}\delta_{\omega-1} + \frac{1}{2}\delta_{\omega+1}$ on $\mathbb{Z}$ from Example 2.1.2b gives rise to operators P and P^* where $P^*h(\omega) = \frac{1}{2}h(\omega-1) + \frac{1}{2}h(\omega+1)$ (conditional expectation of h one time step ahead) and $Pf(\omega) = \frac{1}{2}f(\omega+1) + \frac{1}{2}f(\omega-1)$ (here f is to be interpreted as a probability vector and Pf is the new probability distribution after one step of time). For the determination of $Pf(\omega)$ one has to use the translation invariance of m.

In the next sections it will play no role whether such a dual pair of contractions is defined using a kernel or not. Later we will see that on most spaces all such operators stem from a kernel. Here are two examples where this is not so obvious:

2.1.5 Example Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds and define $P : L^1 \to L^1$ by $Pf := f \circ T$. This is obviously a positive linear contraction, see Remark 1.2.4. Observe that the same composition - but on L^∞ - occured in the previous example. Note that, for invertible T, $P^*f = f \circ T^{-1}$, hence $P^{-1} = P^*$, and in particular $P1 = 1$ and also $P^*1 = 1$ with $1 \in L^1 \cap L^\infty$.

2.1.6 Example Let $(\Omega, \mathcal{A}, \mu)$ be a probability space, $\mathcal{F} \subseteq \mathcal{A}$ a sub-σ-algebra, and define $P : L^1 \to L^1$ by $Pf = E_\mu[f \mid \mathcal{F}]$. Then P is linear, $Pf \geqslant 0$ if $f \geqslant 0$, and $\|Pf\|_1 = \int |E_\mu[f \mid \mathcal{F}]|\,d\mu \leqslant \int E_\mu[|f| \mid \mathcal{F}]\,d\mu = \int |f|\,d\mu = \|f\|_1$, so P is a positive linear L^1-contraction. P^*h is uniquely characterized by duality: for all $f \in L^1$ we have

$$\int P^*h \cdot f\,d\mu = \int h \cdot E_\mu[f \mid \mathcal{F}]\,d\mu = \int E_\mu[h\,f \mid \mathcal{F}]\,d\mu = \int E_\mu[h \mid \mathcal{F}] \cdot f\,d\mu\,,$$

so that also $P^*h = E_\mu[f \mid \mathcal{F}]$.

It follows immediately that for two sub-σ-algebras $\mathcal{F}_1, \mathcal{F}_2$ and their corresponding conditional expectation operators P_1 and P_2 also the composition P_2P_1 is a positive linear contraction on L^1.

2.1.7 Lemma *Let P be a positive linear contraction on $L^1(\Omega, \mathcal{A}, m)$.*

a) If $0 \leqslant f_n \nearrow f \in L^1$ m-a.e., then $0 \leqslant Pf_n \nearrow Pf \in L^1$ m-a.e. and $\|Pf - Pf_n\|_1 \to 0$.

b) If $h_n \searrow 0$ m-a.e., $h_n \in L^\infty$, then $P^*h_n \searrow 0$ m-a.e. Equivalently, if $0 \leqslant h_n \nearrow h \in L^\infty$ m-a.e., then also $P^*h_n \nearrow P^*h$ m-a.e.

Proof:

a) $0 \leqslant Pf_n \nearrow g$ for some g by positivity of P. As $\sup_n Pf_n\,dm \leqslant \sup_n \int f_n\,dm \leqslant \int f\,dm < \infty$, Beppo Levi's theorem implies $\|Pf_n - Pf\|_1 \leqslant \|f_n - f\|_1 \to 0$ and $\|g - Pf\|_1 \leqslant \|g - Pf_n\|_1 + \|Pf_n - Pf\|_1 \to 0$, i.e. $Pf_n \nearrow g = Pf$ m-a.e.

b) $P^*h_n \searrow h \geqslant 0$ m-a.e. for some $h \in L^\infty$ follows from the monotonicity of P^*. Let $A \in \mathcal{A}$ with $m(A) < \infty$. Then $1_A P^*h_n \leqslant 1_A P^*h_1 \leqslant 1_A \|h_1\|_\infty$ and $h_n\, P1_A \leqslant h_1\, P1_A \leqslant \|h_1\|_\infty\, P1_A$ so that we can use dominated convergence twice to get

$$\int_A h\,dm = \lim_{n\to\infty} \int_A P^*h_n\,dm = \lim_{n\to\infty} \int h_n\, P1_A\,dm = 0\ .$$

It follows that $h = 0$ m-a.e. If $0 \leqslant h_n \nearrow h$, consider $0 \leqslant (h - h_n) \searrow 0$.

□

2.1.8 Remark If $h : \Omega \to [0,\infty]$ is measurable, one can define $h \wedge n := \min\{n, h\}$ such that $0 \leqslant h \wedge n \nearrow h$. Then also $0 \leqslant P^*(h \wedge n) \nearrow$, and as we just saw, the limit of this sequence is P^*h if $h \in L^\infty$. Otherwise we define P^*h to be this limit. Then, if $h_k \in L^\infty_+$ and $h_k \nearrow h$, we have

$$\sup_k P^*h_k = \sup_k \sup_n P^*(h_k \wedge n) = \sup_n \sup_k P^*(h_k \wedge n) = \sup_n P^*(h \wedge n) = P^*h\ .$$

2.1.2 The conservative and dissipative part

We continue to work with a σ-finite measure space $(\Omega, \mathcal{A}, m)$ and a positive L^1 contraction P.

2.1.9 Definition *a) A measurable function $h : \Omega \to \mathbb{R}_+$ is called harmonic, if $P^*h = h$.*

b) A measurable function $h : \Omega \to \mathbb{R}_+$ is called superharmonic, if $P^*h \leqslant h$. It is called strictly superharmonic on A *if $P^*h < h$ on A.*

2.1.10 Theorem (Hopf decomposition: 1) *There exists a decomposition $\Omega = C \uplus D$ into measurable sets determined uniquely mod μ by:*
(C1) If h is superharmonic, i.e. $P^*h \leqslant h$, then $P^*h = h$ on C.
(D1) There exists a bounded superharmonic h_0 which is strictly superharmonic on D.
h_0 may be chosen with the additional properties $h_0 = 0$ on C, $h_0 \leqslant 1$, and $P^{*n}h_0 \to 0$ on D.

Proof: As m is σ-finite there exists a probability measure $\mu \approx m$. Let

$$\mathcal{S} := \{A \in \mathcal{A} : \exists g \leqslant 1 \text{ which is superharmonic, strictly on } A\}\ .$$

$\mathcal{S}$ has the obvious property that for any $A \in \mathcal{S}$ and any measurable $A' \subset A$ also $A' \in \mathcal{S}$.

If $A_1, A_2, \dots \in \mathcal{S}$ with strict superharmonics $g_1, g_2, \dots$, then $A := \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{S}$, because $0 \leqslant g := \sum_{n\in\mathbb{N}} 2^{-n} g_n \leqslant 1$ and on each set A_k we have

$$P^*g = \sum_{n\in\mathbb{N}} 2^{-n} P^*g_n < 2^{-k} g_k + \sum_{n\in\mathbb{N}\setminus\{k\}} 2^{-n} g_n < g\ .$$

Therefore the following supremum is attained by some set $D \in \mathcal{S}$ with bounded strict superharmonic g_D:

$$\alpha := \sup\{\mu(A) : A \in \mathcal{S}\} .$$

Let $C := D^c$ and consider any (also unbounded) superharmonic h. If it is not harmonic on C, then there is a measurable $B \subseteq C$ with $\mu(B) > 0$ such that h is strictly superharmonic on B. Choose some $\beta \in \mathbb{R}_+$ for which $B' := B \cap \{P^*h < \beta \leqslant h\}$ has positive measure. Then, because $0 \leqslant P^*1 \leqslant 1$, the function $0 \leqslant g := \min\{\beta^{-1}h, 1\} \leqslant 1$ is superharmonic and strictly superharmonic on B': $P^*g \leqslant \min\{\beta^{-1}P^*h, 1\} \leqslant \min\{\beta^{-1}h, 1\} = g$ with strict inequality on B'. Hence $B' \in \mathcal{S}$ and therefore $B' \subseteq D$ in contradiction to $B' \subseteq B \subseteq C$ and $\mu(B') > 0$. Therefore h is harmonic on C. This proves (C1) and (D1).

Finally, as g_D from above is superharmonic and as P^* is positive, $g_D \geqslant P^*g_D \geqslant P^{*2}g_D \geqslant \ldots$, and the sequence converges to some $\tilde{g} \geqslant 0$. As g_D is strictly superharmonic on D, $\tilde{g} < g_D$ on D. By(C1), $\tilde{g} = g_D$ on C. Finally, $P^*\tilde{g} = \lim_{n\to\infty} P^{*n+1}g_D = \tilde{g}$, and so $h_0 := g_D - \tilde{g}$ has the additional required properties. □

2.1.11 Corollary $P^*1 = 1$ *on* C.

For $f \in L^1$, $h \in L^\infty$ and $n \in \mathbb{N} \cup \{\infty\}$ we use the notation

$$S_n f = \sum_{i=0}^{n-1} P^i f \quad \text{and} \quad S_n^* h = \sum_{i=0}^{n-1} P^{*i} h .$$

2.1.12 Theorem (Hopf decomposition: 2) *The sets C and D of the Hopf decomposition are determined uniquely mod μ by*
(C2) For all $h \in L_+^\infty$: $S_\infty^* h = \infty$ on $C \cap \{S_\infty^* h > 0\}$.
(D2) $\exists h_D \in L_+^\infty$: $\{h_D > 0\} = D$ and $S_\infty^* h_D \leqslant 1$.
h_D can be chosen such that $S_\infty^* h_D = 0$ on C.

Proof: Let h_0 be as in Theorem 2.1.10 and set $h_D := h_0 - P^*h_0$. As h_0 is strictly superharmonic on D, $h_D > 0$ on D, and as $h_0 = 0$ on C, also $h_D = 0$ on C. Finally

$$S_n^* h_D = \sum_{i=0}^{n-1} P^{*i}(h_0 - P^*h_0) = h_0 - P^{*n}h_0 \leqslant h_0$$

and hence $S_\infty^* h_D \leqslant h_0 \leqslant 1$, which proves (D2) and also $S_\infty^* h_D = 0$ on C.

Now let $h \in L_+^\infty$. Consider $\tilde{h} := \min\{1, S_\infty^* h\}$. As $P^*1 \leqslant 1$ and $P^*(S_\infty^* h) = \sum_{i=1}^\infty P^{*i}h \leqslant S_\infty^* h$, we have $P^*\tilde{h} \leqslant \tilde{h} \leqslant 1$ and $P^{*j+1}\tilde{h} \leqslant P^{*j}\tilde{h} \leqslant 1$ for all $j \geqslant 0$. Hence $P^{*j+1}\tilde{h} = P^{*j}\tilde{h}$ on C for all $j \geqslant 0$, i.e. $P^{*j}\tilde{h} = \tilde{h}$ on C for all $j \geqslant 0$. Therefore, on $C \cap \{S_\infty^* h < \infty\}$,

$$\tilde{h} = P^{*j}\tilde{h} \leqslant \sum_{i=j}^{\infty} P^{*i}h \to 0 \text{ as } j \to \infty$$

so that $S_\infty^* h = \tilde{h} = 0$ on this set. This implies $\{S_\infty^* h > 0\} \cap C \cap \{S_\infty^* h < \infty\} = \varnothing \mod m$, which is just (C2). □

2.1.13 Theorem (Hopf decomposition: 3) *The sets C and D of the Hopf decomposition are determined uniquely $\mod \mu$ by*
(C3) For all $f \in L^1_+$: $S_\infty f = \infty$ on $C \cap \{S_\infty f > 0\}$.
(D3) For all $f \in L^1_+$: $S_\infty f < \infty$ on D.

Proof: For h_D from (D2) and $f \in L^1_+$ we have

$$\langle S_\infty f, h_D \rangle = \langle f, S^*_\infty h_D \rangle \leqslant \langle f, 1 \rangle < \infty .$$

As $h_D > 0$ on D this implies (D3).

Now let $f \in L^1_+$. As m is σ-finite, there is $\tilde{h} \in L^1_+$ with $\tilde{h} \leqslant 1$ and $\{\tilde{h} > 0\} = C \cap \{0 < S_\infty f < \infty\}$. Let $h := \frac{\tilde{h}}{1+S_\infty f}$. Then $0 \leqslant h \leqslant 1$, $\{h > 0\} = \{\tilde{h} > 0\}$, and it follows from (C2) that $S^*_\infty h = \infty$ on $C \cap \{S^*_\infty h > 0\} \supseteq C \cap \{h > 0\} = C \cap \{\tilde{h} > 0\} = \{\tilde{h} > 0\}$. Hence, for each $n \in \mathbb{N}$,

$$\langle P^n f, S^*_\infty h \rangle = \langle S_\infty (P^n f), h \rangle \leqslant \langle S_\infty f, h \rangle \leqslant \int_\Omega \tilde{h} \, dm < \infty .$$

As $S^*_\infty h = \infty$ on $\{\tilde{h} > 0\}$, this implies $P^n f = 0$ for all n on $\{\tilde{h} > 0\}$. It follows that $S_\infty f = 0$ on $\{\tilde{h} > 0\} = C \cap \{0 < S_\infty f < \infty\}$, i.e. $m\,(C \cap \{0 < S_\infty f < \infty\}) = 0$. This is (C3). □

2.1.14 Examples a) Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds. Then $P : L^1 \to L^1$, $Pf = f \circ T$ is a L^1-contraction (because $\mu \circ T^{-1} = \mu$). P is conservative, i.e. $C = \Omega \mod \mu$, because $1 \in L^1_+$, $P^n 1 = 1 \circ T^n = 1$ for all n and hence $S_\infty 1 = \infty$ everywhere. (If $\mu \circ T^{-1} \approx \mu$ but if μ is not invariant, then this need no longer hold.)

b) Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $P : L^1 \to L^1$ be a composition of conditional expectations, see Example 2.1.6. Then $P1 = 1$, and as before it follows that P is conservative.

2.1.15 Example Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $T : \Omega \to \Omega$ be measurable with $\mu \circ T^{-1} \approx \mu$. The kernel $p(\omega, A) = \delta_{T(\omega)}(A)$ determines a L^1-contraction P, see Example 2.1.2a.

a) For $s \in \mathbb{N}$ let $D_s := \{S_\infty 1 \leqslant s\}$. Then $D_s \nearrow D$ by (D3). For each $s \in \mathbb{N}$

$$\sum_{n=0}^{\infty} \mu\{\omega : T^n(\omega) \in D_s\} = \sum_{n=0}^{\infty} \int_\Omega 1_{D_s} \circ T^n \, d\mu = \sum_{n=0}^{\infty} \int_\Omega P^{*n} 1_{D_s} \, d\mu = \sum_{n=0}^{\infty} \int_{D_s} P^n 1 \, d\mu$$
$$= \int_{D_s} S_\infty 1 \, d\mu \leqslant s < \infty .$$

By the Borel-Cantelli Lemma, μ-a.e. trajectory visits each D_s only finitely often.

b) Let $U \subseteq C$ be measurable. Then (C2) applied to $h = 1_U$ yields

$$\sum_{n=0}^{\infty} 1_U(T^n \omega) = \sum_{n=0}^{\infty} P^{*n} 1_U(\omega) = \infty$$

for μ-a.e. $\omega \in C \cap \{S^*_\infty 1_U > 0\} = C \cap \bigcup_{n=0}^{\infty} T^{-n}(U)$. It follows that μ-a.e. ω that visits U at all, visits U infinitely often.

2.1.16 Definition *a) For $B \in \mathcal{A}$ we denote $L^p(B) := \{f \in L^p : \{f \neq 0\} \subseteq B\}$.*

b) $B \in \mathcal{A}$ *is* *P-absorbing*, *if* $P(L^1(B)) \subseteq L^1(B)$.

c) $\mathcal{C} := \{B \in \mathcal{A} : \; B \subseteq C$ *and* B *is* P*-absorbing*$\}$.

d) P *is called* *ergodic*, *if* $P^* 1_A = 1_A$ *implies* $A = \varnothing$ *or* $A = \Omega \mod \mu$.

2.1.17 Theorem *$C \in \mathcal{C}$, i.e. the conservative part C of a positive L^1-constraction P is P-absorbing.*

Proof: Let h_D be as in (D2) with $\{h_D > 0\} = D$ and $S^*_\infty h_D = 0$ on C. Then, for each $f \in L^1(C)$,

$$\langle |Pf|, h_D \rangle \leqslant \langle P|f|, h_D \rangle = \langle |f|, P^* h_D \rangle \leqslant \int |f| \, S^*_\infty h_D \, dm = 0$$

so that $Pf = 0$ on D, i.e. $Pf \in L^1(C)$. □

2.1.18 Theorem *a) B is P-absorbing if and only if $P^* 1_{B^c} \leqslant 1_{B^c}$, i.e. 1_{B^c} is superharmonic.*

b) If $C = \Omega$, then $B \in \mathcal{C}$ if and only if $P^* 1_B = 1_B$.

c) $\mathcal{C}|_C$ is a σ-algebra.

Proof: a) Let B be P-absorbing. There is $f \in L^1_+(B)$ such that $\{f > 0\} = B$. Then $\langle f, P^* 1_{B^c} \rangle = \langle Pf, 1_{B^c} \rangle = 0$, because also $Pf \in L^1(B)$. Hence $P^* 1_{B^c} = 0$ on B. As $P^* 1_{B^c} \leqslant P^* 1 \leqslant 1$, this proves $P^* 1_{B^c} \leqslant 1_{B^c}$. On the other hand, if this inequality holds and if g belongs to $L^1_+(B)$, then

$$0 \leqslant \langle Pg, 1_{B^c} \rangle = \langle g, P^* 1_{B^c} \rangle \leqslant \langle g, 1_{B^c} \rangle = 0 \, ,$$

so that $Pg \in L^1(B)$. For general $g \in L^1(B)$ use the decomposition $g = g^+ - g^-$.
b) $P^* 1 = 1$ by Corollary 2.1.11. Hence

$$B \in \mathcal{C} \Leftrightarrow P^* 1_{B^c} \leqslant 1_{B^c} \Leftrightarrow P^* 1_{B^c} = 1_{B^c} \Leftrightarrow P^* 1 - P^* 1_B = 1 - 1_B \Leftrightarrow P^* 1_B = 1_B$$

where we used (C1) of Theorem 2.1.10 for the second equivalence.
c) Suppose first that $\Omega = C$. Then $P^* 1 = 1$ and $B \in \mathcal{C} \Leftrightarrow P^* 1_B = 1_B$. It follows that $\mathcal{C}$ is closed under passing to the complement, and also under finite intersection: If $A, B \in \mathcal{C}$, then $P^* 1_{A \cap B} \leqslant P^* 1_A = 1_A$ and $P^* 1_{A \cap B} \leqslant P^* 1_B = 1_B$. Hence $P^* 1_{A \cap B} \leqslant 1_{A \cap B}$ so that $(A \cap B)^c \in \mathcal{C}$. Hence $\mathcal{C}$ is an algebra, and in view of Lemma 2.1.7b it is even a σ-algebra. (One can use the lemma to see that it is a Dynkin-system or a monotone class.)

If C is a strict subset of Ω, then the previous arguments apply to $P|_{L^1(C)}$. This operator is clearly conservative, and $B \subseteq C$ is P-absorbing if and only if it absorbing under this restriction. □

2.1.19 Corollary *If $C = \Omega$, then P is ergodic if and only if $\mathcal{C}$ is trivial.*

2.1.20 Theorem *If P is invertible and if also P^{-1} is a positive L^1-contraction, then $P^* 1_B = 1_B$ for each $B \in \mathcal{C}$. (So $\mathcal{C}$ is the σ-algebra of invariant sets).*

Proof: As $P^*1 \leqslant 1$ and $P^{*-1}1 \leqslant 1$, we have $P^*1 = 1$. Let $B \in \mathcal{C}$. Then $P^*1_{B^c} \leqslant 1_{B^c}$ so that $P^*1_B = P^*1 - P^*1_{B^c} \geqslant 1 - 1_{B^c} = 1_B$. Let $h := P^*1_B - 1_B$. Then $0 \leqslant h \leqslant 1$. Let $\tilde{h} := P^{*-1}h = 1_B - P^{*-1}1_B$. Then also $0 \leqslant \tilde{h} \leqslant 1$. Now

$$\sum_{k=0}^{n} P^{*k}\tilde{h} = P^{*n}1_B - P^{*-1}1_B \leqslant 1 \text{ for all } n \geqslant 1,$$

so that also $\sum_{k=0}^{\infty} P^{*k}\tilde{h} \leqslant 1$. Therefore, by (D2) of Theorem 2.1.12, $\tilde{h} = 0$ on C. As $B \subseteq C$, this implies $P^{*-1}1_B = 1_B$, i.e. $P^*1_B = 1_B$. □

2.1.3 Exercises

2.1.1 Let $(\Omega, \mathcal{A}, m)$ be a σ-finite measure space and $T : \Omega \to \Omega$ a measurable map that satisfies $m \circ T^{-1} \approx m$. Consider the L^1-contraction P whose dual is given by $P^*h = h \circ T$ (see Example 2.1.4), and denote its dissipative part by D. Prove that

1. $\#\{n \in \mathbb{N} : T^n(\omega) \in D\} = 0$ for m-a.e. $\omega \in C$ and
2. there are $D_k \nearrow D$ such that $\#\{n \in \mathbb{N} : T^n(\omega) \in D_k\} \leqslant k$ for all $k \in \mathbb{N}$ and m-a.e. $\omega \in \Omega$.

2.1.2 Let $(\Omega, \mathcal{A}, \mu)$ be a probability space.

a) Prove that P is conservative if $\int P^*h \, d\mu = \int h \, d\mu$ for each $h \in L^\infty$.

b) Prove that each L^1-contraction given by a conditional expectation is conservative.

c) Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and consider a L^1-contraction which is a composition of two (or more) conditional expectation operators. Prove that it is conservative. (The composition of arbitrary conservative operators need not be conservative, though.)

2.1.3 Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds and $P : L^1 \to L^1$, $Pf = f \circ T$ as in Example 2.1.5. Determine P^*h.

2.1.4 Give an example of two conservative positive L^1-contractions whose composition is not conservative. *Hint:* The simplest examples are based on composition operators.

2.1.5 Let $\Omega = \mathbb{Z}$ with the counting measure m and consider $Pf(\omega) = \frac{1}{2}f(\omega+1) + \frac{1}{2}f(\omega-1)$ as in Example 2.1.4.

1. Determine the conservative and the dissipative part of P.
2. Do the same for the operator $Pf(\omega) = \alpha f(\omega+1) + (1-\alpha)f(\omega-1)$ when $\alpha \in (0,1)\backslash\{\frac{1}{2}\}$.

2.2 Ergodic theorems

2.2.1 The Chacon-Ornstein theorem

Let $f \in L^1$, $g \in L^1_+$. In this subsection we prove that $\frac{S_n f}{S_n g}$ converges m-a.e. on $\{S_\infty g > 0\}$.

The filling scheme We need a number of notations.

- For $f, g \in L^1_+$ we write $f \stackrel{1}{\to} g$ if $\exists r, s \in L^1_+ : f = r + s$ and $g = r + Ps$.
- $f \stackrel{n}{\to} g$ if $\exists g_1 : f \stackrel{n-1}{\to} g_1$ and $g_1 \stackrel{1}{\to} g$.
- $U : L^1 \to L^1, Uh := Ph^+ - h^-$ is the nonlinear (!) filling operator.
- $M_n f := \max\{S_1 f, \dots, S_n f\}$.

2.2.1 Lemma *For $f, g \in L^1_+$, $h = f - g$ and $n \in \mathbb{N}$ there exists $f_n \in L^1_+$ such that $U^n h = f_n - g$ and $f \stackrel{n}{\to} f_n$.*

Proof: For $n = 0$ take $f_0 = f$. Determine f_{n+1} inductively: let $t_n := f_n - (U^n h)^+$ and observe that $t_n \geqslant 0$ and $(U^n h)^- = (U^n h)^+ - U^n h = f_n - t_n - f_n + g = g - t_n$. Hence $U^{n+1} h = P((U^n h)^+) - (U^n h)^- = P((U^n h)^+) + t_n - g =: f_{n+1} - g$ and $f_n \stackrel{1}{\to} f_{n+1}$ as $f_n = (U^n h)^+ + t_n$.
□

2.2.2 Lemma *For $h \in L^1$ and $n \in \mathbb{N}$ we have $U^n h \geqslant 0$ on $\{M_{n+1} h > 0\}$.*

Proof: By definition, $(Uh)^- \leqslant h^-$ so that inductively $h^- \geqslant (Uh)^- \geqslant (U^2 h)^- \geqslant \dots$. Therefore it is sufficient to show that for m-a.e. $\omega \in \{M_{n+1} h > 0\}$ there exists some $k \in \{0, \dots, n\}$ such that $(U^k h)^+(\omega) > 0$. This is equivalent to showing that $\varphi_\ell(\omega) := \sum_{k=0}^{\ell} (U^k h)^+(\omega) > 0$ for some $\ell \in \{0, \dots, n\}$. We will show that

$$\varphi_\ell \geqslant S_{\ell+1} h \quad \text{for all } \ell. \tag{2.2.1}$$

This implies $\max_{\ell=0,\dots,n} \varphi_\ell(\omega) \geqslant \max_{\ell=1,\dots,n+1} S_\ell h(\omega) = M_{n+1} h(\omega)$ which proves the claim.

We turn to the proof of (2.2.1). For $\ell = 0$ it is the trivial fact that $h^+ \geqslant 0$. Inductively,

$$\begin{aligned}
S_{\ell+2} h &= h + P S_{\ell+1} h \\
&\leqslant h + P\varphi_\ell = h + \sum_{k=0}^{\ell} P((U^k h)^+) = h + \sum_{k=0}^{\ell} (U^{k+1} h + (U^k h)^-) \\
&= h + \sum_{k=0}^{\ell} ((U^{k+1} h)^+ - (U^{k+1} h)^- + (U^k h)^-) \\
&= h + \varphi_{\ell+1} - h^+ + h^- - (U^{\ell+1} h)^- = \varphi_{\ell+1} - (U^{\ell+1} h)^- \leqslant \varphi_{\ell+1}
\end{aligned}$$

□

2.2.3 Lemma *for $f \in L^1$ and $g \in L^1_+$ we have $\lim_{n\to\infty} \frac{P^n f}{S_{n+1} g} = 0$ m-a.e. on $\{S_\infty g > 0\}$.*

Proof: We may consider f^+ and f^- separately and thus assume that $f \geqslant 0$. We can also assume that $\int g \, dm = 1$ so that gm is a probability measure. Fix $\epsilon > 0$ and set $r_n = P^n f - \epsilon S_{n+1} g$ and

$$A_n := \left\{ \frac{P^n f}{S_{n+1} g} > \epsilon \right\} \cap \{g > 0\} = \{r_n > 0\} \cap \{g > 0\} .$$

Then

$$r_n = P r_{n-1} - \epsilon g \leqslant P r_{n-1}^+ - \epsilon g \quad (n \geqslant 1)$$

so that indeed

$$r_n^+ \leqslant Pr_{n-1}^+ - \epsilon 1_{A_n} g \ .$$

Hence

$$\int \epsilon 1_{A_n} g\, dm \leqslant \int Pr_{n-1}^+\, dm - \int r_n^+\, dm \leqslant \int r_{n-1}^+\, dm - \int r_n^+\, dm \ .$$

Summing over $n \geqslant 1$ we obtain

$$\sum_{n=1}^{\infty} (gm)(A_n) \leqslant \epsilon^{-1} \int r_0^+\, dm \leqslant \epsilon^{-1} \int f^+\, dm < \infty \ .$$

By Borel-Cantelli, $\#\left\{n \in \mathbb{N} : \frac{P^n f(\omega)}{S_{n+1} g(\omega} > \epsilon\right\} < \infty$ holds for gm-a.e. ω and hence for m-a.e. $\omega \in \{g > 0\}$. As $\epsilon > 0$ was arbitrary this proves $0 \leqslant \limsup_{n\to\infty} \frac{P^n f}{S_{n+1} g} \leqslant 0$ m-a.e. on $\{g > 0\}$. Applying this argument to $P^j f$ and $P^j g$ (any $j \in \mathbb{N}$) we obtain

$$0 \leqslant \limsup_{n\to\infty} \frac{P^n f}{S_{n+1} g} \leqslant \limsup_{n\to\infty} \frac{P^{n+j} f}{S_{n+1}(P^j g)} \leqslant \limsup_{n\to\infty} \frac{P^{n+j} f}{S_{n+j+1} g} \leqslant 0 \ .$$

Hence $\lim_{n\to\infty} \frac{P^n f}{S_{n+1} g} = 0$ m-a.e. on $\{S_\infty g > 0\}$. □

2.2.4 Lemma *Let $f, g \in L^1_+$, $k \geqslant 1$, $g \xrightarrow{k} g_1$ and $\gamma > 1$. Then*

$$\left\{\limsup_{n\to\infty} S_n(f - g) > 0\right\} \subseteq \left\{\limsup_{n\to\infty} S_n(\gamma f - g_1) > 0\right\} \tag{2.2.2}$$

Proof: We prove the lemma first for $k = 1$. Then this result can be applied repeatedly with $\gamma^{1/k}$ instead of γ to treat general $k \in \mathbb{N}$.

As $g \xrightarrow{1} g_1$, there are $r, s \in L^1_+$ such that $g = r + s$ and $g_1 = r + Ps$. Hence

$$\begin{aligned} S_n(\gamma f - g_1) &= \sum_{i=0}^{n-1} P^i(\gamma f - g + g - r - Ps) = S_n(\gamma f - g) + \sum_{i=0}^{n-1} P^i(s - Ps) \\ &= S_n(\gamma f - g) + s - P^n s \geqslant S_n(f - g) + (\gamma - 1) S_n f - P^{n-1} Ps \ . \end{aligned}$$

By Lemma 2.2.3, $P^{n-1}(Ps)(\omega) < (\gamma - 1) S_n f(\omega)$ if $S_\infty f(\omega) > 0$ and n is suffciently large (depending on ω). Hence $\limsup_{n\to\infty} S_n(\gamma f - g_1) \geqslant \limsup_{n\to\infty} S_n(f - g)$ on $\{S_\infty f > 0\}$ and, restricted to $\{S_\infty f = 0\}$, the inclusion (2.2.2) is trivial, because both sets are empty. □

To proceed we need some more notation:

- For $H \in \mathcal{A}$, $f \in L^1_+$ and $n \in \mathbb{N}$ define

$$\Psi_H^n f := \sup\left\{\int_H g\, dm : \ f \xrightarrow{n} g\right\} \quad \text{and} \quad \Psi_H f := \lim_{n\to\infty} \Psi_H^n f \ .$$

  (Observe that $0 \leqslant \Psi_H^1 f \leqslant \Psi_H^2 f \leqslant \dots$, because $f \xrightarrow{n} g$ implies $f \xrightarrow{n+1} g$.)

- As all g in the definition of Ψ_H^n are nonnegative, we have $\Psi_{H_1}^n \leqslant \Psi_{H_2}^n$ if $H_1 \subseteq H_2$.

- As $f \xrightarrow{1} g$ implies $\int_H g\,dm \leqslant \int r\,dm + \int Ps\,dm \leqslant \int r\,dm + \int s\,dm = \|f\|_1$, it is clear that $\Psi_H f \leqslant \|f\|_1$.
- For $\alpha \geqslant 0$ holds $\Psi_H^n(\alpha f) = \alpha \Psi_H^n(f)$, and the same holds for Ψ_H.
- $E_\infty(h) := \bigcup_{n=1}^\infty \{M_n h > 0\}$.

2.2.5 Lemma *Let $f, g \in L^1_+$, $h := f - g$, $n \in \mathbb{N}$ and $H \in \mathcal{A}$.*
a) $H \subseteq \{M_{n+1}h > 0\}$ implies $\Psi_H^n f \geqslant \int_H g\,dm$.
b) $H \subseteq E_\infty(h)$ implies $\Psi_H f \geqslant \int_H g\,dm$.
c) $H \subseteq \{\limsup_{n\to\infty} S_n h > 0\}$ implies $\Psi_H f \geqslant \Psi_H g$.

Proof: By Lemma 2.2.2, $U^n h \geqslant 0$ on $\{M_{n+1}h > 0\}$, and by Lemma 2.2.1, $U^n h = f_n - g$ for some $f_n \in L^1_+$ with $f \xrightarrow{n} f_n$.
a) If $H \subseteq \{M_{n+1}h > 0\}$, then $0 \leqslant \int_H U^n h\,dm = \int_H (f_n - g)\,dm$ so that $\int_H g\,dm \leqslant \int_H f_n\,dm \leqslant \Psi_H^n f$.
b) Let $H_k := H \cap \{M_{k+1}h > 0\}$. Then $H_k \nearrow H$ and assertion a) implies

$$\Psi_H f = \sup_n \Psi_H^n f \geqslant \sup_n \Psi_{H_n}^n f \geqslant \sup_n \int_{H_n} g\,dm = \int_H g\,dm\ .$$

c) Suppose that $g \xrightarrow{k} g_1$ for some $k \in \mathbb{N}$ and $g_1 \in L^1_+$, and choose $\gamma > 1$ arbitrary. By Lemma 2.2.4,

$$H \subseteq \left\{\limsup_{n\to\infty} S_n(\gamma f - g_1)\right\} \subseteq E_\infty(\gamma f - g_1)\ .$$

By assertion b), $\gamma \Psi_H f = \Psi_H(\gamma f) \geqslant \int_H g_1\,dm$. As $\gamma > 1$ was arbitrary, $\Psi_H f \geqslant \int_H g_1\,dm$. As this holds for all k and all g_1 with $g \xrightarrow{k} g_1$, we conclude that $\Psi_H f \geqslant \Psi_H g$.

□

2.2.6 Theorem (Chacon-Ornstein) *Let P be a positive L^1-contraction, $f \in L^1$ and $g \in L^1_+$. Then*

$$L(f,g) := \lim_{n\to\infty} \frac{S_n f}{S_n g} \in (-\infty, \infty) \quad \textit{exists } m\textit{-a.e. on } \{S_\infty g > 0\}\ .$$

Proof: Again we may assume that $f \geqslant 0$. Observe that $\Psi_H g > 0$ for any $H \subseteq \{S_\infty g > 0\}$ with $m(H) > 0$: in this situation there is some $n \in \mathbb{N}$ with $m(H \cap \{P^n g > 0\}) > 0$, and as $g \xrightarrow{n} P^n g$, it follows that $\Psi_H(g) \geqslant \int_H P^n g\,dm > 0$.

On $\{S_\infty g > 0\}$ we define $\underline{h} := \liminf_{n\to\infty} \frac{S_n f}{S_n g}$ and $\overline{h} := \limsup_{n\to\infty} \frac{S_n f}{S_n g}$. On $\{\overline{h} > \alpha\}$ we have $\limsup_{n\to\infty} \frac{S_n(f-\alpha g)}{S_n g} > 0$. As $S_n g$ increases, this shows that $\{\overline{h} > \alpha\} \subseteq \{\limsup_{n\to\infty} S_n(f - \alpha g) > 0\}$.

Let $H := \{\overline{h} = \infty\} \cap \{S_\infty g > 0\}$. Then, by Lemma 2.2.5, we obtain for all $\alpha > 0$

$$\alpha \Psi_H g = \Psi_H(\alpha g) \leqslant \Psi_H f \leqslant \|f\|_1 < \infty\ ,$$

so that $\Psi_H g = 0$. This implies $m(H) = 0$, i.e. $\overline{h} < \infty$ on $\{S_\infty g > 0\}$.

In order to prove that $\underline{h} = \overline{h}$, let $\alpha < \beta$ in $\mathbb{Q}$ and define $H := \{\underline{h} < \alpha < \beta < \overline{h}\} \cap \{S_\infty g > 0\}$. Then

$$H \subseteq \left\{\limsup_{n\to\infty} S_n(f - \beta g) > 0\right\} \cap \left\{\limsup_{n\to\infty} S_n(\alpha g - f) > 0\right\}$$

and Lemma 2.2.5 implies $\Psi_H f \geqslant \beta \Psi_H g$ and $\alpha \Psi_H g \geqslant \Psi_H f$. As $\alpha < \beta$ this implies $\Psi_H g = 0$ and hence $m(H) = 0$. As this is true for all $\alpha < \beta$ in $\mathbb{Q}$, it follows that $\underline{h} = \overline{h}$ m-a.e. and the theorem is proved. □

The identification of the limit $L(f, g)$ in full generality is a bit technical, see [11, §3.3]. Below we provide it in several special cases that are sufficient for our purposes.

2.2.2 The identification of the limit: Birkhoff's ergodic theorem

Suppose that $(\Omega, \mathcal{A}, \mu)$ is a probability space and P is a positive contraction on $L^1(\Omega, \mathcal{A}, \mu)$. If $P1 = 1$, the Chacon-Ornstein theorem, applied with $g = 1 \in L^1$, simplifies to

$$\bar{f} := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k f \quad \text{exists and is finite } \mu\text{-a.e.} \tag{2.2.3}$$

One has indeed L^1-convergence: For $\epsilon > 0$ choose $s > 0$ such that $\int (|f| - s)^+ \, d\mu < \epsilon$. Then $f = u + v$ where $|u| \leqslant s$ and $\int |v| \, d\mu < \epsilon$. Hence $P^k f = P^k u + P^k v$, and from $P1 = 1$, the positivity and the contraction property of P it follows that $|P^k u| \leqslant s$ and $\int |P^k v| \, d\mu < \epsilon$ for all k. Thus the $P^k f$ are uniformly integrable and hence also the sequence $\left(\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k\right)_{n \geqslant 1}$ is uniformly integrable so that L^1-convergence follows at once, see [10, Lemma 6.17b, Satz 6.18]. Hence $P\bar{f} = \bar{f}$ μ-a.e.

2.2.7 Corollary (Birkhoff's ergodic theorem) *If $(\Omega, \mathcal{A}, \mu, T)$ is a mpds and $f \in L^1$, then*

$$\bar{f} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = E_\mu[f \mid \mathcal{I}(T)] \quad \mu\text{-a.e. and in } L^1. \tag{2.2.4}$$

Proof: $Pf := f \circ T$ is a positive contraction satisfying $P1 = 1$, see Example 2.1.5. Therefore the almost sure convergence follows from (2.2.3).

As $\bar{f} \circ T = P\bar{f} = \bar{f}$ μ-a.e., the limit $\bar{f}$ is T-invariant mod μ and hence $\mathcal{I}(T)$-measurable mod μ. For any $A \in \mathcal{I}(T)$ we have in view of the L^1-convergence and the T-invariance of μ

$$\int_A \bar{f} \, d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_{T^{-k}A} f \circ T^k \, d\mu = \int_A f \, d\mu \, .$$

Hence $\bar{f} = E_\mu[f \mid \mathcal{I}(T)]$. □

2.2.3 The identification of the limit: Compositions of conditional expectations

Recall the notion of a complete probability space and of a completion: If $(\Omega, \mathcal{A}, \mu)$ is a probability space, its completion is the space $(\Omega, \mathcal{A}_\mu, \mu)$ where

$$\mathcal{A}_\mu := \{A \subseteq \Omega : \exists A_0, A_1 \in \mathcal{A} \text{ with } A_0 \subseteq A \subseteq A_1 \text{ and } \mu(A_1 \backslash A_0) = 0\}$$

is the *μ-completion* of $\mathcal{A}$. The obvious extension of μ to $\mathcal{A}_\mu$ is again denoted by μ. The space $(\Omega, \mathcal{A}, \mu)$ is complete, if $\mathcal{A}_\mu = \mathcal{A}$.

2.2.8 Remark Let $(\Omega, \mathcal{A}_1, \mu)$ be a probability space and let $\mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2$ σ-algebras. If $\mathcal{A}_2 \subseteq (\mathcal{A}_0)_\mu$, then the three spaces $L^p(\Omega, \mathcal{A}_i, \mu)$ $(i = 0, 1, 2)$ are naturally identified for any $p \in [1, \infty]$ – we treat them as identical.

2.2.9 Corollary *Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{A}$ be sub-σ-algebras, $P_i := E_\mu[\,.\,|\,\mathcal{F}_i]$ (see Example 2.1.6) and $f \in L^1$. Then*

$$\bar{f} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (P_2 P_1)^k f = E_\mu[f \mid (\mathcal{F}_1)_\mu \cap (\mathcal{F}_2)_\mu] \quad \mu\text{-a.e. and in } L^1. \tag{2.2.5}$$

Proof: Let $P = P_2 P_1$. Then $P1 = 1$ and the convergence as well as $P\bar{f} = \bar{f}$ follow again from (2.2.3).

As each $\frac{1}{n}\sum_{k=0}^{n-1}(P_2P_1)^k f$ is $(\mathcal{F}_2)_\mu$-measurable, so is $\bar{f}$ no matter how it is defined on the null set where there may be no convergence. Hence $P_2\bar{f} = \bar{f} = P\bar{f} = P_2P_1\bar{f}$ μ-a.e. For bounded f (and hence also bounded $\bar{f}$) this implies

$$\langle \bar{f}, \bar{f}\rangle = \langle \bar{f}, P_2P_1\bar{f}\rangle = \langle P_2\bar{f}, P_1\bar{f}\rangle = \langle \bar{f}, P_1\bar{f}\rangle$$

so that $P_1\bar{f} = \bar{f}$ μ-a.e., i.e. $\bar{f}$ is also $(\mathcal{F}_1)_\mu$-measurable. General $f \in L^1$ can be L^1-approximated by bounded ones: $f = u + v$ with bounded u and $\|v\|_1 < \epsilon$. Then $\bar{f} = \bar{u} + \bar{v}$ with a $(\mathcal{F}_1)_\mu \cap (\mathcal{F}_2)_\mu$-measurable $\bar{u}$ and $\|\bar{v}\|_1 \leqslant \|v\|_1$.

Finally, for $A \in (\mathcal{F}_1)_\mu \cap (\mathcal{F}_2)_\mu$ we have

$$\int_A \bar{f}\,d\mu = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\int_A (P_2P_1)^k f\,d\mu = \int_A f\,d\mu\,,$$

because $\int_A (P_2P_1)^k f\,d\mu = \int (P_2P_1)^k 1_A\, f\,d\mu = \int_A f\,d\mu$. □

2.2.4 The identification of the limit: Hurewicz's ergodic theorem

Suppose that $(\Omega, \mathcal{A}, \mu)$ is a probability space and that $T : \Omega \to \Omega$ is an invertible bi-measurable transformation with $\mu \circ T^{-1} \approx \mu$. We consider the stochastic kernel $p(\omega, A) = \delta_{T(\omega)}(A)$ as in Example 2.1.4a. Recall that the associated contraction P satisfies $Pf = f \circ T^{-1} \cdot \frac{d(\mu\circ T^{-1})}{d\mu}$. It is an invertible linear operator, and its inverse is given by $P^{-1}f = f \circ T \cdot \frac{d(\mu\circ T)}{d\mu}$.

2.2.10 Theorem (Hurewicz's ergodic theorem) *For each $f \in L^1$*

$$L(f,1) = \lim_{n\to\infty}\frac{S_n f}{S_n 1} = E_\mu[f \mid \mathcal{C}] \quad \mu\text{-a.e. on } C. \tag{2.2.6}$$

(This contains, as a special case, also Birkhoff's theorem.)

Proof: We begin with some preparatory observations:

- $B \in \mathcal{C}$ implies $1_{T^{-1}B} = 1_B \circ T = P^* 1_B = 1_B$ and hence $T^{-1}B = B \mod \mu$ by Theorem 2.1.20.
- If $f \xrightarrow{1} g$, i.e. if $f = r + s$ and $g = r + Ps$ for some $r, s \in L^1_+$, and if $B \in \mathcal{C}$, then $\int_B g\,d\mu = \int_B r\,d\mu + \int_B Ps\,d\mu = \int_B r\,d\mu + \int P^*1_B \cdot s\,d\mu = \int_B r\,d\mu + \int 1_B \cdot s\,d\mu = \int_B f\,d\mu$.

- Inductively: If $f \xrightarrow{n} g$ and $B \in \mathcal{C}$, then $\int_B f\,du = \int_B g\,d\mu$. Hence: $\Psi_B f = \int_B f\,d\mu$.

Next observe that for all $j, k \in \mathbb{Z}$

$$\frac{d(\mu\circ T^k)}{d\mu}\circ T^j = \frac{d(\mu\circ T^k)}{d(\mu\circ T^{-j})}\circ T^j \cdot \frac{d(\mu\circ T^{-j})}{d\mu}\circ T^j = \frac{d(\mu\circ T^{k+j})}{d\mu}\cdot\frac{d\mu}{d(\mu\circ T^j)} \tag{2.2.7}$$

so that for all $n \in \mathbb{N}$

$$S_n f\circ T = \sum_{k=0}^{n-1}\left(f\circ T^{-k}\cdot\frac{d(\mu\circ T^{-k})}{d\mu}\right)\circ T = f\circ T + S_{n-1}f\cdot\frac{d\mu}{d(\mu\circ T)}\ . \tag{2.2.8}$$

As $S_n 1 \nearrow \infty$ on C by (C3) of Theorem 2.1.13, it follows that

$$L(f,1)\circ T = \lim_{n\to\infty}\frac{S_n f\circ T}{S_n 1\circ T} = \lim_{n\to\infty}\frac{f\circ T + S_{n-1}f\cdot\frac{d\mu}{d(\mu\circ T)}}{1 + S_{n-1}1\cdot\frac{d\mu}{d(\mu\circ T)}} = L(f,1)\ .$$

In particular, $L(f,1)|_C$ is $\mathcal{I}(T)$-measurable mod μ and hence $\mathcal{C}$-measurable mod μ, see Theorem 2.1.20.

In order to prove that $L(f,1) = E_\mu[f \mid \mathcal{C}]$ on C, it suffices to show that for alle rational $\alpha < \beta$ we have $B_{\alpha,\beta} := C \cap \{\alpha < L(f,1) < \beta\} \subseteq \{\alpha \leqslant E_\mu[f \mid \mathcal{C}] \leqslant \beta\}$ mod μ on C.

To this end let $H := B_{\alpha,\beta} \cap \{E_\mu[f \mid \mathcal{C}] < \alpha\}$. Then $H \in \mathcal{C}$ so that $\Psi_H f = \int_H f\,d\mu$ for all $f \in L^1_+$ by the above considerations. Suppose $\mu(H) > 0$. As $H \subseteq B_{\alpha,\beta} \subseteq \{\limsup_{n\to\infty} S_n(f-\alpha) > 0\}$, Lemma 2.2.5 implies

$$\int_H f\,d\mu = \Psi_H f \geqslant \Psi_H \alpha = \int_H \alpha\,d\mu > \int_H E_\mu[f \mid \mathcal{C}]\,d\mu = \int_H f\,d\mu\ ,$$

a contradiction. Hence $\mu(H) = \mu(B_{\alpha,\beta}\backslash\{E_\mu[f \mid \mathcal{C}] < \alpha\}) = 0$. similarly one proves $\mu(B_{\alpha,\beta}\backslash\{E_\mu[f \mid \mathcal{C}] > \beta\}) = 0$. This implies the required inclusion. □

In Chapter 3 we will need the following variant of the Hurewicz theorem:

2.2.11 Theorem (Proposition 7.4.2 in [3]) *For every $f \in L^1(\Omega,\mathcal{A},\mu)$ and μ-a.e. ω,*

$$\lim_{n\to\infty}\frac{\sum_{k=-n}^{n} P^k f}{\sum_{k=-n}^{n} P^k 1} = E_\mu[f \mid \mathcal{I}(T)]\ . \tag{2.2.9}$$

Proof: On the conservative part C the claim follows by applying the Hurewicz theorem separately to P and to P^{-1} and observing that $\mathcal{I}(T) = \mathcal{I}(T^{-1})$ is at the same time the σ-algebra of P-absorbing sets and that of P^{-1}-absorbing sets (mod 0), see Theorem 2.1.20.

On the dissipative part D, both the enumerator and the denominator are a.s. converging series. Equation (2.2.7) implies that

$$P^k g\circ T^j = g\circ T^{-(k-j)}\cdot\frac{d(\mu\circ T^{-(k-j)})}{d\mu}\cdot\frac{d\mu}{d(\mu\circ T^j)} = \frac{P^{k-j}g}{P^{-j}1}$$

both for $g = f$ and $g = 1$. Therefore the limit of the quotient in (2.2.9) is T-invariant on D, and for every $A \in \mathcal{I}(T)|_D$ we have

$$\begin{aligned}\int_A \frac{\sum_{k\in\mathbb{Z}} P^k f}{\sum_{j\in\mathbb{Z}} P^j 1}\,d\mu &= \sum_{k\in\mathbb{Z}}\int_A \frac{P^k f}{\sum_{j\in\mathbb{Z}} P^j 1}\,d\mu = \sum_{k\in\mathbb{Z}}\int_A \frac{f}{\sum_{j\in\mathbb{Z}} P^j 1\circ T^k}\,d\mu \\ &= \sum_{k\in\mathbb{Z}}\int_A \frac{f}{\sum_{j\in\mathbb{Z}} P^j 1}\cdot P^{-k}1\,d\mu = \int_A f\,d\mu\ .\end{aligned}$$

On D the quotient of the series is therefore equal to $E_\mu[f \mid \mathcal{I}(T)]$. □

2.2.5 Exercises

1. Let $\Omega = \mathbb{Z}$ with the counting measure m and consider $Pf(\omega) = \frac{1}{2}f(\omega+1) + \frac{1}{2}f(\omega-1)$ as in Exercise 2.1.5. Prove:
   a) $S_n 1_{\{0\}} \to \infty$ but $\lim_{n\to\infty} n^{-1} S_n 1_{\{0\}} = 0$ pointwise.
   b) $\lim_{n\to\infty} n^{-1} S_n f = 0$ pointwise for each $f \in L^1_+$.
   c) $S^*_n 1_{\{0\}} \to \infty$ but $\lim_{n\to\infty} n^{-1} S^*_n 1_{\{0\}} = 0$ pointwise.
2. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space, let P_1 be a linear L^1-contraction, and let $P_2 := E_\mu[\,.\mid \mathcal{F}]$ for some sub-σ-algebra $\mathcal{F}$. Then the Chacon-Ornstein theorem applies to $P = P_2 P_1$. In Corollary 2.2.9 the limit was identified for the case where also P_1 is a conditional expectation operator. For what other P_1 can you identify the limit?

3 Ergodic decompositions

3.1 Background from measure theory

In a topological context, the continuous image of a closed set need not be closed, and similarly in measure theoretical context the measurable image of a measurable set need not be measurable. However, in topological context, the continuous image of a compact set is always compact (and hence also closed). This section is devoted to a class of measure spaces that play the role of compact spaces in the category of measure spaces. The theory goes back to Rohlin, but we follow here mostly the presentation of de la Rue [4], who also gives further references. One should also note that much more elaborateed versions of this theme - based on descriptive set theory - are available, see e.g. [2].

3.1.1 Lebesgue-Rohlin spaces

In this section a triple $(\Omega, \mathcal{A}, \mu)$ always denotes a probability space. If $\mathcal{C} \subset \mathcal{A}$, then $\sigma(\mathcal{C})$ is the smallest sub-σ-algebra of $\mathcal{A}$ that contains $\mathcal{C}$, and $\sigma(\mathcal{C})_\mu$ denotes its μ-completion.

3.1.1 Definition *a) A probability space $(\Omega, \mathcal{A}, \mu)$ is a pre-Lebesgue-Rohlin space (pLR space) if there is a Hausdorff topology τ on Ω with a countable basis and such that*

$$\sigma(\tau) = \mathcal{A} \quad \text{and} \tag{3.1.1}$$

$$\forall A \in \sigma(\tau): \ \mu(A) = \sup\{\mu(K) : K \subseteq A, K \text{ compact for } \tau\}. \tag{3.1.2}$$

b) It is a Lebesgue-Rohlin space (LR space, also standard probability space or just Lebesgue space), if (3.1.1) is replaced by

$$\tau \subset \mathcal{A} \text{ and } \sigma(\tau)_\mu = \mathcal{A}. \tag{3.1.3}$$

(In particular $\mathcal{A}$ must be complete.)

In both situations each such topology τ is called adapted to μ.

3.1.2 Remark a) In the case of a LR space, (3.1.2) extends to all $A \in \mathcal{A}$.

b) If $(\Omega, \mathcal{A}, \mu)$ is pLR, then $(\Omega, \mathcal{A}_\mu, \mu)$ is LR (where μ denotes also its extension to $\mathcal{A}_\mu$).

3.1.3 Lemma *If each $G \in \tau$ can be written as the union of countably many τ-closed sets, then (3.1.2) can be replaced by*

$$1 = \sup\{\mu(K) : K \subseteq \Omega, K \text{ compact for } \tau\}. \tag{3.1.4}$$

(This holds in particular, if τ is metrizable.)

Proof: (3.1.4) implies for each closed A that

$$\mu(A) = \sup\{\mu(A \cap K) : K \subseteq \Omega, K \text{ compact for } \tau\} = \sup\{\mu(K) : K \subseteq A, K \text{ compact for } \tau\},$$

i.e. (3.1.2). The same follows for open sets, because they are countable increasing unions of of closed subsets.

Denote by $\mathcal{A}_0$ the family of all sets $A \in \mathcal{A}$ such that (3.1.2) holds for A and also for A^c. We just saw that $\tau \subseteq \mathcal{A}_0$. It remains to show that $\mathcal{A}_0$ is a σ-algebra: $\Omega \in \mathcal{A}_0$ because of (3.1.4), and $A \in \mathcal{A}_0 \Leftrightarrow A^c \in \mathcal{A}_0$ by definition. So let $A_1, A_2, \dots \in \mathcal{A}_0$. We must show that $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_0$. To this end let $\epsilon > 0$ and choose compact sets $K_n \subseteq A_n$ and $L_n \subseteq A_n^c$ such that $\mu(A_n \backslash K_n) < 2^{-n}\epsilon$ and $\mu(A_n^c \backslash L_n) < 2^{-n}\epsilon$. Fix $N \in \mathcal{N}$ so large that $\mu\left(A \backslash \bigcup_{n=1}^{N} A_n\right) < \epsilon$. Then $\bigcup_{n=1}^{N} K_n$ and $\bigcap_{n=1}^{\infty} L_n$ are compact subsets of A and A^c, respectively, and

$$\mu\left(A \backslash \bigcup_{n=1}^{N} K_n\right) \leqslant \epsilon + \sum_{n=1}^{N} \mu\left(A_n \backslash K_n\right) < 2\epsilon$$

and

$$\mu\left(A^c \backslash \bigcap_{n=1}^{\infty} L_n\right) = \mu\left(\bigcup_{n=1}^{\infty} (A^c \cap L_n^c)\right) \leqslant \sum_{n=1}^{\infty} \mu\left(A^c \backslash L_n\right) \leqslant \sum_{n=1}^{\infty} \mu\left(A_n^c \backslash L_n\right) < \epsilon\,.$$

□

3.1.4 Example Let $\mathbb{X} := \{0,1\}^{\mathbb{N}}$. Equipped with its product topology $\tau_{\mathbb{X}}$, $\mathbb{X}$ is compact metrizable, in particular Hausdorff. The collection $\mathcal{Z}$ of cylinder sets

$$[a_1, \dots, a_n] := \{x \in \mathbb{X} : x_i = a_i \ (i = 1, \dots, n)\} \quad (n \in \mathbb{N}, a_i \in \{0,1\})$$

forms a countable basis of $\tau_{\mathbb{X}}$ and $\sigma(\tau_{\mathbb{X}}) = \sigma(\mathcal{Z})$. Each cylinder is at the same time open and closed. If μ is a probability on $\sigma(\mathcal{Z})$, then (3.1.4) holds trivially because $\mathbb{X}$ is compact. So $(\mathbb{X}, \sigma(\mathcal{Z}), \mu)$ is a pLR space and hence $(\mathbb{X}, \sigma(\mathcal{Z})_\mu, \mu)$ is LR.

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and $\mathcal{B} = (B_n)_{n \in \mathbb{N}}$ a countable family of sets in $\mathcal{A}$ that separates points. Such a $\mathcal{B}$ defines an *injection*

$$\phi_{\mathcal{B}} : \Omega \to \mathbb{X}, \quad \omega \mapsto (1_{B_n}(\omega))_{n \in \mathbb{N}}\,. \tag{3.1.5}$$

Observe that

$$\phi_{\mathcal{B}}^{-1}([a_1, \dots, a_n]) = \bigcap_{k=1}^{n} B_k^{a_k}$$

where $B_k^1 := B_k$ and $B_k^0 := B_k^c$. Here and in the sequel $B^c := \Omega \backslash B$. In particular, $\phi_{\mathcal{B}}$ is $\sigma(\mathcal{B})$-$\sigma(\tau_{\mathbb{X}})$-measurable. Denote

$$\mu_{\mathcal{B}} := \mu \circ \phi_{\mathcal{B}}^{-1}, \quad \mathcal{A}_{\mathcal{B}} := \sigma(\tau_{\mathbb{X}})_{\mu_{\mathcal{B}}}.$$

Then $(\mathbb{X}, \sigma(\tau_{\mathbb{X}}), \mu_{\mathcal{B}})$ is pLR and $(\mathbb{X}, \mathcal{A}_{\mathcal{B}}, \mu_{\mathcal{B}})$ is LR.

3.1.2 Polish spaces provide (p)LR spaces

3.1.5 Definition *A topological space (Ω, τ) is* polish, *if τ can be generated by a complete separable metric d. (So it is in particular a Hausdorff space with a countable basis.)*

3.1.6 Examples a) $\mathbb{R}^d$ equipped with its usual topology is a polish space.

b) $(\mathbb{X}, \tau_{\mathbb{X}})$ is polish.

c) A separable Hilbert space is a polish space.

d) $C([0,1], \mathbb{R})$ equipped with the topology of uniform convergence and $C(\mathbb{R}_+, \mathbb{R})$ equipped with the topology of uniform convergence on compact subsets are polish spaces.

3.1.7 Lemma *Let Ω be polish and let $\mathcal{B}$ be a countable basis for its topology τ. Then $\phi_{\mathcal{B}}(\Omega) \in \sigma(\tau_{\mathbb{X}})$.*

Proof: We characterize the points belonging to $\phi_{\mathcal{B}}(\Omega)$ in a measurable way: Let $y = (y_n)_{n\in\mathbb{N}} \in \mathbb{X}$. Suppose that $y = \phi_{\mathcal{B}}(\omega)$ for some $\omega \in \Omega$. As $\mathcal{B}$ is a basis of τ, each ball around ω contains some set $B_n \ni \omega$. Therefore,

(1) $\exists n \in \mathbb{N} : y_n = 1$ and $\mathrm{diam}(B_n) \leqslant 1$

(2) $\forall n \in \mathbb{N} : y_n = 1 \Rightarrow \left[\exists m \in \mathbb{N} : y_m = 1,\ \overline{B_m} \subset B_n \text{ and } \mathrm{diam}(B_m) \leqslant \mathrm{diam}(B_n)/2\right]$.

(3) $\forall n \in \mathbb{N} : \bigcap_{k=1}^n B_k^{y_k} \neq \varnothing$.

We claim that, vice versa, if y satisfies (1) - (3), then $y = \phi_{\mathcal{B}}(\omega)$ for some $\omega \in \Omega$. Indeed: for each $n \in \mathbb{N}$ there is $\omega_n \in \bigcap_{k=1}^n B_k^{y_k}$ because of (3). Because of (1) and (2), $(\omega_n)_{n\in\mathbb{N}}$ is a Cauchy sequence that has a limit ω since the metric d is complete. We show that $\omega \in \bigcap_{k=1}^\infty B_k^{y_k}$, i.e. $y = \phi_{\mathcal{B}}(\omega)$. Suppose this is not the case. Then there is $k \in \mathbb{N}$ such that $\omega \notin B_k^{y_k}$. If $y_k = 0$, then $\omega \in B_k$ and hence $\omega_n \in B_k$ for all large n, contradicting $\omega_n \in \bigcap_{k=1}^n B_k^{y_k}$. If $y_k = 1$, then $\omega \in B_k^c$, and because of (2) there is $m \in \mathbb{N}$ such that $\omega \in \Omega\backslash\overline{B_m}$. As before this leads to a contradiction, because $\omega_n \in \Omega\backslash\overline{B_m}$ for all large n.

It remains to note that (1) - (3) specify a Borel measurable set of points in $\mathbb{X}$. □

3.1.8 Lemma *Let (Ω, τ) be polish with Borel-σ-algebra $\mathcal{A}$, let μ be a probability measure on $\mathcal{A}$, and let $(F_p)_{p\in\mathbb{N}}$ be a family of τ-closed subsets of Ω. Denote by τ' the topology on Ω generated by τ and the F_p. Then τ' is adapted to μ, in particular $(\Omega, \mathcal{A}_\mu, \mu)$ is a LR space.*

Proof: $\tau \subset \tau' \subset \mathcal{A}$, so $\mathcal{A} = \sigma(\tau) \subseteq \sigma(\tau') \subseteq \mathcal{A}$. Hence (3.1.1) holds for τ'. Let $\mathcal{B}$ be a countable basis for τ and let $\mathcal{B}'$ be the countable family of all finite intersections of sets from $\mathcal{B}$ and sets F_p. Then $\mathcal{B}'$ is a countable basis for τ' and each element of $\mathcal{B}'$ can be written as a countable union of τ-closed sets - and hence of τ'-closed sets. As each set in τ' is the countable union of sets in $\mathcal{B}'$, this shows that each set in τ' can be written as a countable union of τ'-closed sets. Hence, by Lemma 3.1.3, it suffices to verify (3.1.4) instead of (3.1.2).

Suppose now w.l.o.g. that all F_p^c belong to the basis $\mathcal{B} = \{B_1, B_2, \dots\}$ of τ, and observe that $\phi_{\mathcal{B}}^{-1}(\tau_{\mathbb{X}})$ is a topology on Ω. Observe also that $B_n^a = \phi_{\mathcal{B}}^{-1}\{y \in \mathbb{X} : y_n = a\} \in \phi_{\mathcal{B}}^{-1}(\tau_{\mathbb{X}})$ for all $B_n \in \mathcal{B}$ and $a \in \{0,1\}$, in particular $\mathcal{B} \subset \phi_{\mathcal{B}}^{-1}(\tau_{\mathbb{X}})$ and $F_p \in \phi_{\mathcal{B}}^{-1}(\tau_{\mathbb{X}})$ for all p. Hence $\tau' \subseteq \phi_{\mathcal{B}}^{-1}(\tau_{\mathbb{X}})$. Therefore, if $K \subseteq \phi_{\mathcal{B}}(\Omega)$ is $\tau_{\mathbb{X}}$-compact, then $\phi_{\mathcal{B}}^{-1}(K)$ is τ'-compact: To see this, suppose $\phi_{\mathcal{B}}^{-1}(K) \subseteq \bigcup_{i\in I} O_i$ and $O_i \in \tau'$. Then $O_i = \phi_{\mathcal{B}}^{-1}(U_i)$ for some $U_i \in \tau_{\mathbb{X}}$ and hence $\phi_{\mathcal{B}}^{-1}(K) \subseteq \phi_{\mathcal{B}}^{-1}\left(\bigcup_{i\in I} U_i\right)$. As $\phi_{\mathcal{B}}$ is injective and as $K \subseteq \phi_{\mathcal{B}}(\Omega)$, this implies $K \subseteq \bigcup_{i\in I} U_i$, and as K is $\tau_{\mathbb{X}}$-compact, there is a finite subset J of I such that $K \subseteq \bigcup_{i\in J} U_i$. Hence $\phi_{\mathcal{B}}^{-1}(K) \subseteq \bigcup_{i\in J} O_i$.

As $(\mathbb{X}, \sigma(\tau_{\mathbb{X}}), \mu_{\mathcal{B}})$ is a pLR space and as $\phi_{\mathcal{B}}(\Omega) \in \sigma(\tau_{\mathbb{X}})$ by Lemma 3.1.7, we finally have (3.1.4):

$$\begin{aligned} 1 = \mu(\Omega) = \mu_{\mathcal{B}}(\phi_{\mathcal{B}}(\Omega)) &= \sup\{\mu_{\mathcal{B}}(K) : K \subseteq \phi_{\mathcal{B}}(\Omega), K\ \tau_{\mathbb{X}}\text{-compact}\} \\ &= \sup\{\mu(\phi_{\mathcal{B}}^{-1}K) : K \subseteq \phi_{\mathcal{B}}(\Omega), K\ \tau_{\mathbb{X}}\text{-compact}\} \\ &\leqslant \sup\{\mu(K') : K' \subseteq \Omega, K'\ \tau'\text{-compact}\}\ . \end{aligned}$$

□

3.1.9 Theorem *If (Ω, τ) is polish and if $(\Omega, \mathcal{A}, \mu)$ satisfies $\sigma(\tau)_\mu = \mathcal{A}$, then it is a LR space.*

Proof: Apply the previous lemma with no sets F_p added. □

3.1.10 Examples (Continuation of Examples 3.1.6)

a) $\mathbb{R}^d$ and, more generally, any separable Hilbert space equipped with a Borel probability measure and the completed Borel-σ-algebra is a LR space.

b) $C([0,1],\mathbb{R})$ and $C(\mathbb{R}_+,\mathbb{R})$ equipped with the Wiener measure are LR spaces. (Observe that one has to make sure that the Borel-σ-algebra of C as a polish space coincides with the trace on C of the product σ-algebra on $\mathbb{R}^{[0,1]}$ and $\mathbb{R}^{\mathbb{R}_+}$, respectively. See for example [1, Satz 38.6 and Korollar 40.4].)

3.1.3 The measurable image property in LR spaces

3.1.11 Lemma *Let $(\Omega, \mathcal{A}, \mu)$ be a LR space and $B_n \in \mathcal{A}$ $(n \in \mathbb{N})$. Then there is a topology τ' on Ω which is adapted to μ and such that $B_n \in \tau'$ for all $n \in \mathbb{N}$.*

Proof: Let τ be a μ-adapted topology on Ω. Because of (3.1.2), there are K_σ-sets $S_n \subseteq B_n$ (countable unions of τ-compact sets) such that $\mu(B_n \backslash S_n) = 0$ for all $n \in \mathbb{N}$. Denote by $\mathcal{K}$ the countable family of compact sets used to construct the countably many S_n and denote by τ' the topology generated by τ, $\mathcal{K}$ and the B_n.

As τ has a countable basis, so does τ'. As $\tau \subset \tau'$ and τ separates points and satisfies (3.1.3), so does τ'. It remains to prove (3.1.2) for τ'. Let

$$N = \bigcup_{n \in \mathbb{N}} (B_n \backslash S_n) \quad \text{and} \quad \Omega_0 = N^c .$$

Then $\mu(N) = 0$ so that $\mu(\Omega_0) = 1$, and $B_n \cap \Omega_0 = S_n \cap \Omega_0$ for all n, because

$$B_n \cap \Omega_0 = B_n \cap \bigcap_{k \in \mathbb{N}} \left(B_k^c \cup S_k\right) \subseteq B_n \cap S_n = S_n .$$

Now let $A \in \mathcal{A}$ and $\epsilon \in (0, \mu(A))$. There exists a τ-compact set $K \subset A \cap \Omega_0$ such that $\mu(K) \geqslant \mu(A) - \epsilon$. With its trace topology $\tau_{|K}$, K is a compact space with a countable basis that separates points. Hence it is metrizable, and so it is polish. As $K \subseteq \Omega_0$, all intersections $K \cap (B_n \backslash S_n)$ are empty, and so the trace topology $\tau'_{|K}$ is generated by $\tau_{|K}$ and the countable family $\mathcal{K}_{|K} := \{K' \cap K : K' \in \mathcal{K}\}$. Therefore, by Lemma 3.1.8 there exists a $\tau'_{|K}$-compact set $K' \subseteq K$ such that $\mu(K') \geqslant \mu(K) - \epsilon \geqslant \mu(A) - 2\epsilon$. To conclude observe that each $\tau'_{|K}$-compact set K' is also τ'-compact. □

3.1.12 Theorem *Let $(\Omega, \mathcal{A}, \mu)$ and $(\Omega', \mathcal{A}', \mu')$ be complete probability spaces and $h : \Omega \to \Omega'$ measurable such that $\mu \circ h^{-1} = \mu'$. Let also $\mathcal{B}' = (B'_n)_{n \in \mathbb{N}}$ be a family of $\mathcal{A}'$-measurable sets that separates points in Ω'. Then, if $(\Omega, \mathcal{A}, \mu)$ is LR, also $(\Omega', \mathcal{A}', \mu')$ is LR, $\mathcal{A}' = \sigma(\mathcal{B}')_{\mu'}$ and $h(\Omega_0) \in \mathcal{A}'$ for each $\Omega_0 \in \mathcal{A}$ with $\mu(\Omega_0) = 1$.*

Proof: W.l.o.g. one can assume that $\mathcal{B}'$ contains with each B'_n also $(B'_n)^c$. As $\mathcal{B}'$ separates points, $\phi_{\mathcal{B}'} : \Omega' \to \mathbb{X}$ is injective and measurable, see also (3.1.5). Let $\tau' := \phi_{\mathcal{B}'}^{-1}(\tau_{\mathbb{X}})$. The family $\mathcal{B}'' := \{B'_{i_1} \cap \dots \cap B'_{i_n} : n \in \mathbb{N}, B'_{i_1}, \dots, B'_{i_n} \in \mathcal{B}'\}$ is a countable basis for τ' and $\sigma(\tau') = \sigma(\mathcal{B}')$.

Let $B_n = h^{-1}(B'_n)$ for all n. By Lemma 3.1.11 there is a μ-adapted topology τ on Ω such that $B_n \in \tau$ for all n. Then $h^{-1}(\mathcal{B}'') \subseteq \tau$ so that $h : \Omega \to \Omega'$ is τ-τ' continuous. We are going to show that τ' is adapted to μ'.

Let $A' \in \mathcal{A}'$ and denote $A := h^{-1}(A')$. Then

$$\mu'(A') = \mu(h^{-1}A') = \mu(A) = \sup\{\mu(K) : K \subseteq A, K\ \tau\text{-compact}\} .$$

Let $\epsilon > 0$ and choose a τ-compact $K \subseteq A$ such that $\mu(K) \geqslant \mu'(A') - \epsilon$. Let $K' := h(K)$. Then $K' \subseteq A'$ is τ'-compact and

$$\mu'(K') = \mu(h^{-1}(K')) \geqslant \mu(K) \geqslant \mu'(A') - \epsilon ,$$

so that

$$\mu'(A') = \sup\left\{\mu'(K') : K' \subseteq A', K'\ \tau'\text{-compact}\right\} .$$

As each τ'-compact set belongs to $\sigma(\tau')$, it follows that $\sigma(\mathcal{B}')_{\mu'} = \sigma(\tau')_{\mu'} = \mathcal{A}'$. This shows that τ' is adapted to μ'.

Finally, because τ is adapted to μ, there is an increasing sequence of τ-compact sets $K_n \subseteq \Omega_0$ such that $\mu(K_n) \nearrow 1$. As h is τ-τ'-continuous, the sets $h(K_n)$ are τ'-compact. Then $S := \bigcup_{n\in\mathbb{N}} h(K_n)$ is $\sigma(\tau')$-measurable, $S \subseteq h(\Omega_0) \subseteq \Omega'$ and $\mu'(S) = \sup_n \mu'(h(K_n)) = \sup_n \mu(h^{-1}(h(K_n))) \geqslant \sup_n \mu(K_n) = 1$. This proves $h(\Omega_0) \in \mathcal{A}'$. □

3.1.13 Definition *A basis of a LR space $(\Omega, \mathcal{A}, \mu)$ is a countable family of $\mathcal{A}$-measurable sets that separates the points of Ω. (One can always assume that the basis is closed under finite intersections.)*

3.1.14 Remark Each LR space has a basis: It has a Hausdorff topology with a countable basis.

3.1.15 Theorem *If $\mathcal{B}$ is a basis of the LR space $(\Omega, \mathcal{A}, \mu)$, then $\sigma(\mathcal{B})_\mu = \mathcal{A}$.*

Proof: Apply Theorem 3.1.12 to $\Omega' = \Omega$ and $h = \mathrm{id}_\Omega$. □

3.1.16 Theorem *Let $(\Omega, \mathcal{A}, \mu)$ and $(\Omega', \mathcal{A}', \mu')$ be LR spaces. If $h : \Omega \to \Omega'$ is injective and measurable and satisfies $\mu \circ h^{-1} = \mu'$, then $h(A) \in \mathcal{A}'$ for each $A \in \mathcal{A}$.*

Proof: Let $\mathcal{B}' = (B_n)_{n\in\mathbb{N}}$ be a basis of $(\Omega', \mathcal{A}', \mu')$ and set $B_n := h^{-1}(B'_n)$ for all n. As the B'_n separate points in Ω' and as h is injective, the family $\mathcal{B} := (B_n)_{n\in\mathbb{N}}$ separates points in Ω. Then, by Theorem 3.1.12, $h(\Omega) \in \mathcal{A}'$. Therefore, for all $n \in \mathbb{N}$, $h(B_n) = B'_n \cap h(\Omega) \in \mathcal{A}'$. As a formula,

$$\Omega \in \mathcal{H} := \{A \in \mathcal{A} : h(A) \in \mathcal{A}'\} \quad \text{and} \quad \mathcal{B} \subset \mathcal{H} .$$

As h is injective, the family $\mathcal{H}$ is a σ-algebra. Hence $\sigma(\mathcal{B}) \subset \mathcal{H}$ so that finally $\mathcal{A} = \sigma(\mathcal{B})_\mu \subseteq \mathcal{H}_\mu$. Hence, for each $A \in \mathcal{A}$ there are $H_0, H_1 \in \mathcal{H}$ with $H_0 \subseteq A \subseteq H_1$ and $\mu(H_1 \backslash H_0) = 0$ so that $h(H_0) \subseteq h(A) \subseteq h(H_1)$ with $h(H_0), h(H_1) \in \mathcal{A}'$ and

$$\mu'(h(H_1)\backslash h(H_0)) = \mu\left(h^{-1}(h(H_1))\backslash h^{-1}(h(H_0))\right) = \mu(H_1\backslash H_0) = 0 .$$

As $\mathcal{A}'$ is μ-complete, this shows that $h(A) \in \mathcal{A}'$. □

3.1.4 The classification of Lebesgue-Rohlin spaces

3.1.17 Definition *Two probability spaces* $(\Omega, \mathcal{A}, \mu)$ *and* $(\Omega', \mathcal{A}', \mu')$ *are* isomorphic mod 0 *if there are* $\Omega_0 \in \mathcal{A}$ *and* $\Omega_0' \in \mathcal{A}'$ *with* $\mu(\Omega_0) = \mu'(\Omega_0') = 1$ *and a bi-measurable bijection* $h : \Omega_0 \to \Omega_0'$ *such that* $\mu \circ h^{-1} = \mu'$.

3.1.18 Theorem *Let* $\mathcal{B}$ *be a basis of the LR space* $(\Omega, \mathcal{A}, \mu)$ *and define* $\phi_{\mathcal{B}} : \Omega \to \mathbb{X}$ *as in (3.1.5). Denote again* $\mu_{\mathcal{B}} = \mu \circ \phi_{\mathcal{B}}^{-1}$ *and* $\mathcal{A}_{\mathcal{B}} = \sigma(\tau_{\mathbb{X}})_{\mu_{\mathcal{B}}}$. *Then* $(\Omega, \mathcal{A}, \mu)$ *is isomorphic mod 0 to* $(\mathbb{X}, \mathcal{A}_{\mathcal{B}}, \mu_{\mathcal{B}})$.

Proof: By Theorem 3.1.16, $\phi_{\mathcal{B}} : \Omega \to \phi_{\mathcal{B}}(\Omega)$ is a bi-measurable bijection and $\phi_{\mathcal{B}}(\Omega) \in \mathcal{A}_{\mathcal{B}}$. By definition, $\mu_{\mathcal{B}} = \mu \circ \phi_{\mathcal{B}}^{-1}$, and finally $\mu_{\mathcal{B}}(\phi_{\mathcal{B}}(\Omega)) = \mu(\phi_{\mathcal{B}}^{-1}(\phi_{\mathcal{B}}(\Omega))) = \mu(\Omega) = 1$. □

In the remaining part of this chapter we classify LR spaces $(\Omega, \mathcal{A}, \mu)$ up to isomorphism mod 0. To this end we associate to such a space a decreasing sequence $(m_n)_{n\in\mathbb{N}}$ of reals in $[0, 1]$ as follows: μ has at most countably many atoms $\omega_1, \omega_2, \ldots$, i.e. points with $\mu(\{\omega_n\}) > 0$. (This may be a finite or even empty sequence.) They can be enumerated in such a way that $\mu(\{\omega_1\}) \geqslant \mu(\{\omega_2\}) \geqslant \ldots$. For $n \geqslant 1$ let $m_n := \mu(\{\omega_n\})$ if there are at least n atoms and otherwise let $m_n = 0$. The m_n are called the invariants of $(\Omega, \mathcal{A}, \mu)$. Obviously, if two spaces are isomorphic mod 0, then they have the same invariants. The converse is part of the following theorem.

3.1.19 Theorem *A LR space* $(\Omega, \mathcal{A}, \mu)$ *with invariants* $(m_n)_{n\in\mathbb{N}}$ *is isomorphic mod 0 to the LR space* $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mu})$ *where*

- $\tilde{\Omega} = [0, m_0] \cup \bigcup_{n\in\mathbb{N}}\{1 + \frac{1}{n}\}$ *with* $m_0 = 1 - \sum_n m_n$,
- $\tilde{\mu}_{|[0,m_0]}$ *is the Lebesgue measure and* $\tilde{\mu}(\{1 + \frac{1}{n}\}) = m_n$,
- $\tilde{\mathcal{A}}$ *is the* $\tilde{\mu}$*-completion of the Borel* σ*-algebra on* $\tilde{\Omega}$.

Proof: We restrict to the case where there are no atoms so that $m_0 = 1$.

I) **Reduction to** $\Omega = \mathbb{X}$**:** Because of Theorem 3.1.18 it suffices to consider $\Omega = \mathbb{X} = \{0, 1\}^{\mathbb{N}}$ with $\mathcal{A} = \sigma(\tau_{\mathbb{X}})_{\mu}$.

II) **Reduction to** $\Omega = [0, 1]$**:** Define $h : \mathbb{X} \to [0, 1]$ by $h(\omega) = \sum_{n=1}^{\infty} \omega_n 2^{-n}$ (binary expansion). h is continuous, hence $\sigma(\tau_{\mathbb{X}})$-$\mathcal{F}$-measurable, where $\mathcal{F} := \sigma(\tau_{[0,1]})$ denotes the Borel σ-algebra on $[0, 1]$. Its restriction to the set $\Omega_0 := \mathbb{X} \backslash S$ is injective where $S := \{\omega \in \mathbb{X} : \omega_n = 0 \text{ for at most finitely many } n\}$. Obviously, $h(\Omega_0) = [0, 1)$. As S is countable, it is measurable and $\mu(S) = 0$ because μ has no atoms. Let $\mu' := \mu \circ h^{-1}$ and denote $\mathcal{A}' = \mathcal{F}_{\mu'}$. Then h is $\mathcal{A}$-$\mathcal{A}'$-measurable because $\mathcal{A} = \sigma(\tau_{\mathbb{X}})_{\mu}$.

   As the countable family of all subintervals of $[0, 1]$ with rational endpoints separates points, we can apply Theorems 3.1.12 and 3.1.16 and conclude that $([0, 1], \mathcal{A}', \mu')$ is a LR space and that $h : \Omega_0 \to [0, 1)$ is a bi-measurable bijection.

III) **Final step:** Now $\Omega = [0, 1]$ and $\mathcal{A} = \sigma(\tau_{[0,1]})_{\mu}$. Define $h : [0, 1] \to [0, 1]$, $h(x) := \mu([0, x])$ to be the distribution function of μ. It is well known (from basic probability theory) that h is monotone and continuous (μ has no atoms!). Let $I_1, I_2, \ldots$ be the at most countable many maximal non-trivial closed intervals on which h is constant and denote $\Omega_0 = \bigcup_n I_n$. Then $\mu(\Omega_0) = 1$, $h_{|\Omega_0} : \Omega_0 \to h(\Omega_0)$ is bijective, and $[0, 1] \backslash h(\Omega_0)$ is at most countable. Finally, $\mu \circ h^{-1}([0, a]) = \mu\{x \in [0, 1] : h(x) = \mu([0, x]) \leqslant a\} = a = \tilde{\mu}([0, a])$.

□

3.1.20 Remark For the LR space $[0,1]$ equipped with Lebesgue measure it is not so hard to prove that conditional probabilities with respect to a sub-σ-algebra can be represented by regular conditional probability distributions. (This fact is proved in most textbooks on probability theory, see also [10, Satz 13.10].) In view of Theorem 3.1.19 the same is true for all LR spaces. In chapter 2 we will provide an independent proof of a far reaching generalization of this fact.

3.1.5 Measure separability of Lebesgue-Rohlin spaces

3.1.21 Theorem *Let* $(\Omega, \mathcal{A}, \mu)$ *be a LR space. Then:*
a) $(\Omega, \mathcal{A}, \mu)$ *is separable, i.e. there is a countable sub-algebra* $\mathcal{B}$ *of* $\mathcal{A}$ *such that*

$$\forall A \in \mathcal{A}\, \forall \epsilon > 0\, \exists A' \in \mathcal{B} : \; \mu(A \triangle A') < \epsilon.$$

b) $L^2(\Omega, \mathcal{A}, \mu)$ *is separable.*

Proof: Recall that a LR space has a countable family $\mathcal{B}_0$ of (open) sets such that $\sigma(\mathcal{B}_0)_\mu = \mathcal{A}$. Denote by $\mathcal{B}$ the algebra generated by $\mathcal{B}_0$ and observe that also $\mathcal{B}$ is countable. It is a corollary to the Carathéodory construction that $\mathcal{B}$ is dense in $\mathcal{A}$, see e.g. [10, Satz 2.29]. Denote by G the space of all rational linear combinations of indicator functions of sets from $\mathcal{B}$ and note that G is a countable dense subset of L^1. Now let $f \in L^2$. For $n \in \mathbb{N}$ define

$$f_n(\omega) = \max\{-n, \min\{n, f(\omega)\}\}\,.$$

Then $\|f - f_n\|_2 \to 0$ as $n \to \infty$, and there are $g_n \in G$ such that $\|f_n - g_n\|_1 \leqslant n^{-3}$, so that $\|f_n - g_n\|_2 \leqslant \sqrt{n\, \|f_n - g_n\|_1} \leqslant n^{-1}$. Hence $\|f - g_n\|_2 \to 0$ as $n \to \infty$. □

3.1.22 Remark Let $(\Omega, \mathcal{A}, \mu)$ be a probability space. Define $d : \mathcal{A} \times \mathcal{A} \to [0,1]$, $d(A,B) = \mu(A \triangle B)$. It is easily seen that d is a pseudo-metric on $\mathcal{A}$. The equivalence relation $A \sim B :\Leftrightarrow d(A,B) = 0$ defines equivalence classes $[A]$ such that $d([A],[B]) := d(A,B)$ is a well defined metric on the space $\hat{\mathcal{A}} := \{[A] : A \in \mathcal{A}\}$. One can easily prove:

(1) $(\hat{\mathcal{A}}, d)$ is complete.

(2) If $\mathcal{F}$ is a sub-σ-algebra of $\mathcal{A}$, then $\hat{\mathcal{F}}$ is closed in $(\hat{\mathcal{A}}, d)$.

(3) If $(\Omega, \mathcal{A}, \mu)$ is a LR space, then $(\hat{\mathcal{A}}, d)$ is separable. (It is compact if and only if μ is purely atomic.)

(4) If $h : \Omega \to \Omega'$ is measurable and $\mu' = \mu \circ h^{-1}$, then $\phi : (\hat{\mathcal{A}}', d') \to (\hat{\mathcal{A}}, d)$, $[A'] \mapsto [h^{-1}(A')]$, defines an isometric embedding.

Suppose now that $(\Omega, \mathcal{A}, \mu)$ and $(\Omega', \mathcal{A}', \mu')$ are LR spaces. (For the following assertions one may need some additional assumptions on the isometry, but I see a chance that they can be proved without.)

- Each isometric embedding $\phi : (\hat{\mathcal{A}}', d') \to (\hat{\mathcal{A}}, d)$ with $\phi([\varnothing]) = [\varnothing]$ is of the form in (4).
- If ϕ is even an isometry, then the two LR spaces are isomorphic mod 0 with an isomorphism that determines ϕ as in (4).

3.1.23 Corollary *Let $(\Omega, \mathcal{A}, \mu)$ be a LR space and let $(\mathcal{F}_i)_{i \in I}$ be an arbitrary family of μ-complete sub-σ-algebras of $\mathcal{A}$. Then there is an at most countable subset $J \subseteq I$ such that $\bigcap_{i \in I} \mathcal{F}_i = \bigcap_{j \in J} \mathcal{F}_j$.*

Proof: As the $\mathcal{F}_i$ are all μ-complete, it suffices to prove the corresponding statements for the closed subsets $\hat{\mathcal{F}}_i \subseteq \hat{\mathcal{A}}$. But that is a simple consequence of the separability of the metric space $(\hat{\mathcal{A}}, d)$, because $\hat{\mathcal{A}} \backslash \bigcap_{i \in I} \hat{\mathcal{F}}_i = \bigcup_{i \in I} (\hat{\mathcal{A}} \backslash \hat{\mathcal{F}}_i)$ is open and hence the union of at most countably many sets from a given countable basis of the topology.

□

For later use we also note the following lemma:

3.1.24 Lemma *Let $(\mathcal{F}_k)_{k \in \mathbb{N}}$ be a sequence of sub-σ-algebras of the LR space $(\Omega, \mathcal{A}, \mu)$ satisfying $\bigcap_{k=1}^{n} (\mathcal{F}_k)_\mu = \left(\bigcap_{k=1}^{n} \mathcal{F}_k\right)_\mu$ for all $n \in \mathbb{N}$. Then $\bigcap_{k=1}^{\infty} (\mathcal{F}_k)_\mu = \left(\bigcap_{k=1}^{\infty} \mathcal{F}_k\right)_\mu$.*

Proof: The inclusion $\supseteq$ is trivial. For the other direction let $f \in L^1(\Omega, \mathcal{A}, \mu)$. From a two-fold application of the decreasing martingale theorem we have μ-a.e.

$$
\begin{aligned}
E_\mu\left[f \mid \bigcap_{k=1}^{\infty} (\mathcal{F}_k)_\mu\right] &= \lim_{n \to \infty} E_\mu\left[f \mid \bigcap_{k=1}^{n} (\mathcal{F}_k)_\mu\right] = \lim_{n \to \infty} E_\mu\left[f \mid \left(\bigcap_{k=1}^{n} \mathcal{F}_k\right)_\mu\right] \\
&= \lim_{n \to \infty} E_\mu\left[f \mid \bigcap_{k=1}^{n} \mathcal{F}_k\right] = E_\mu\left[f \mid \bigcap_{k=1}^{\infty} \mathcal{F}_k\right] = E_\mu\left[f \mid \left(\bigcap_{k=1}^{\infty} \mathcal{F}_k\right)_\mu\right] .
\end{aligned}
$$

As this holds for any $f \in L^1$, the sub-σ-algebras $\bigcap_{k=1}^{\infty} (\mathcal{F}_k)_\mu$ and $\left(\bigcap_{k=1}^{\infty} \mathcal{F}_k\right)_\mu$ coincide mod μ, so that $\bigcap_{k=1}^{\infty} (\mathcal{F}_k)_\mu \subseteq \left(\bigcap_{k=1}^{\infty} (\mathcal{F}_k)_\mu\right)_\mu = \left(\bigcap_{k=1}^{\infty} \mathcal{F}_k\right)_\mu$. □

3.1.6 Sub-σ-algebras and factors of Lebesgue-Rohlin spaces

In general it is not true that a sub-σ-algebra of a countably generated σ-algebra is itself countably generated. In LR spaces, however, this holds at least modulo null sets:

3.1.25 Definition *A sub-σ-algebra $\mathcal{F}$ of $(\Omega, \mathcal{A}, \mu)$ is countably generated mod 0 if there is a countable family $\mathcal{B} \subset \mathcal{F}$ such that $\mathcal{F} \subseteq \sigma(\mathcal{B})_\mu$ (equivalently: $\mathcal{F}_\mu = \sigma(\mathcal{B})_\mu$).*

3.1.26 Theorem *Each sub-σ-algebra $\mathcal{F}$ of a LR space $(\Omega, \mathcal{A}, \mu)$ is countably generated mod 0.*

Proof: Let F be the space $L^2(\Omega, \mathcal{F}, \mu) = L^2(\Omega, \mathcal{F}_\mu, \mu)$ considered as a closed subspace of $L^2(\Omega, \mathcal{A}, \mu)$. Denote by π the orthogonal projection on F. Let $(\phi_n)_{n \in \mathbb{N}}$ be a countable dense subset of $L^2(\Omega, \mathcal{A}, \mu)$ (see Theorem 3.1.21), and choose $\mathcal{F}$-measurable representatives ψ_n of the $\pi(\phi_n)$. Denote by $\mathcal{B}$ the countable family of sets $\{\psi_n < \alpha\} \in \mathcal{F}$ ($n \in \mathbb{N}$, $\alpha \in \mathbb{Q}$). Obviously each ψ_n is $\sigma(\mathcal{B})$-measurable. Now, for each $A \in \mathcal{F}$ there is a sequence $(\phi_{n_i})_{i \in \mathbb{N}}$ such that $\|1_A - \phi_{n_i}\|_2 \to 0$ as $i \to \infty$. It follows that $\|\psi_{n_i} - 1_A\|_2 = \|\pi(\phi_{n_i} - 1_A)\|_2 \leqslant \|\phi_{n_i} - 1_A\|_2 \to 0$ as $i \to \infty$. Therefore $A \in \sigma(\mathcal{B})_\mu$, □

3.1.27 Corollary *Each sub-σ-algebra of a pLR space $(\Omega, \mathcal{A}, \mu)$ is countably generated mod 0.*

Proof: Let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{A}$ and consider the LR-space $(\Omega, \mathcal{A}_\mu, \mu)$ (see Remark 3.1.2b). Then $\mathcal{F} \subset \mathcal{A}_\mu$ and so $\mathcal{F}$ is countably generated mod 0 in $(\Omega, \mathcal{A}_\mu, \mu)$ which implies at once that it is countably generated mod 0 in $(\Omega, \mathcal{A}, \mu)$. □

Consider now two LR spaces $(\Omega_i, \mathcal{A}_i, \mu_i)$ $(i = 1, 2)$ and a measurable map $h : \Omega_1 \to \Omega_2$ with $\mu_2 = \mu_1 \circ h^{-1}$. Then $h^{-1}(\mathcal{A}_2)$ is a sub-σ-algebra of $\mathcal{A}_1$. Conversely, the next theorem will show that (up to completion) each sub-σ-algebra arises from a factor system.

3.1.28 Theorem *Let $(\Omega, \mathcal{A}, \mu)$ be a LR space, $\mathcal{F} \subseteq \mathcal{A}$ a sub-σ-algebra. Then there is a $\mathcal{F}$-$\sigma(\tau_{\mathbb{X}})$-measurable map $\phi : \Omega \to \mathbb{X}$ such that $\mathcal{F}_\mu = \sigma(\phi^{-1}(\tau_{\mathbb{X}}))_\mu$.*

Proof: Let $\mathcal{B} \subset \mathcal{F}$ be a countable family such that $\sigma(\mathcal{B})_\mu = \mathcal{F}_\mu$ (Theorem 3.1.26). Define $\phi_{\mathcal{B}} : \Omega \to \mathbb{X}$ as in (3.1.5). Then

$$\mathcal{F}_\mu = \sigma(\mathcal{B})_\mu = \sigma(\phi_{\mathcal{B}}^{-1}(\tau_{\mathbb{X}}))_\mu \, .$$

□

3.1.29 Remark Suppose now $T : \Omega \to \Omega$ leaves the measure class of μ invariant, i.e. $\mu \circ T^{-1} \approx \mu$, and consider $\mathcal{F} := \mathcal{I}(T)$. Let $\phi : \Omega \to \mathbb{X}$ be as in Theorem 3.1.28, and denote $\Gamma_x := \phi^{-1}(\{x\})$ for $x \in \mathbb{X}$. Then

i) Ω is the uncountable disjoint union of the sets Γ_x, $x \in \mathbb{X}$.

ii) $\forall x \in \mathbb{X} : \; \Gamma_x \in \mathcal{I}(T)$.

iii) $\forall A \in \mathcal{I}(T) \; \exists A' \in \phi^{-1}(\sigma(\tau_{\mathbb{X}})) : \; \mu(A \triangle A') = 0$.

The goal of the next two subsections is to define "restrictions" μ_x of μ to the sets Γ_x, i.e. conditional measures, and to show that the $(\Gamma_x, \mathcal{A}_{|\Gamma_x}, \mu_x, T_{|\Gamma_x})$ are non-singular ergodic dynamical systems.

3.1.7 From positive L^1-contractions to kernels on Lebesgue-Rohlin spaces

In this subsection we prove that each positive linear L^1-contraction on a (p)LR space stems from a substochastic kernel as in (2.1.1). It suffices to treat only LR spaces because of the following simple fact: if $(\Omega, \mathcal{A}, \mu)$ is a probability space and if $\mathcal{F} \subset \mathcal{A}$ is a σ-algebra with $\mathcal{A} \subseteq \mathcal{F}_\mu$, then $L^1(\Omega, \mathcal{A}, \mu)$ is trivially identified with $L^1(\Omega, \mathcal{F}, \mu|_{\mathcal{F}})$.

3.1.30 Theorem *Let $(\Omega, \mathcal{A}, \mu)$ be a LR space with basis $\mathcal{B}$. Let P be a positive linear contraction on $L^1(\Omega, \mathcal{A}, \mu) = L^1(\Omega, \sigma(\mathcal{B}), \mu)$. Then there is a substochastic kernel p on $(\Omega, \sigma(\mathcal{B}))$ that determines P by equation (2.1.2), i.e. that satisfies*

$$\int_A Pf \, d\mu = \int_\Omega f(\omega) p(\omega, A) \, d\mu(\omega) \quad \text{for all } A \in \sigma(\mathcal{B}) \text{ and all } f \in L^1 \tag{3.1.6}$$

and

$$P^* 1_A = p(\,.\,, A) \quad \text{for all } A \in \sigma(\mathcal{B}). \tag{3.1.7}$$

If $P^*(L^\infty(\Omega, \mathcal{A}, \mu)) \subseteq L^\infty(\Omega, \mathcal{F}, \mu)$ for some sub-σ-algebra $\mathcal{F}$, then $p(\,.\,, A)$ can be chosen $\mathcal{F}$-measurable for each $A \in \sigma(\mathcal{B})$.

Proof: Assume first that $\Omega = \mathbb{X}$ and that $\mathcal{A}$ is the μ-completion of $\sigma(\tau_{\mathbb{X}}) = \sigma(\mathcal{Z})$.

- For each $A \in \mathcal{Z}$ fix a $\mathcal{F}$-measurable representative $p(\,.\, , A) : \mathbb{X} \to \mathbb{R}$ of $P^*(1_A) \in L^\infty(\mathbb{X}, \mathcal{F}, \mu)$.
- As P^* is linear and positive, we can choose these representatives such that $p(x, \varnothing) = 0$ and $0 \leqslant p(x, A) \leqslant 1$ for all $x \in \mathbb{X}$ and $A \in \mathcal{Z}$.
- As $\mathcal{Z}$ is countable, there is $N \in \mathcal{F}$ with $\mu(N) = 0$ such that for all $n \in \mathbb{N}$, for all $x \in \mathbb{X}\backslash N$ and for all pairwise disjoint $A_1, \ldots, A_n \in \mathcal{Z}$ holds

$$p(x, A_1 \cup \cdots \cup A_n) = P^*(1_{A_1} + \ldots 1_{A_n})(x) = \sum_{i=1}^{n} P^*(1_{A_i})(x) = \sum_{i=1}^{n} p(x, A_i) \; .$$

- For $x \in N$ redefine $p(x, A) = \mu(A)$.

Then, for each $x \in \mathbb{X}$, $p(x, \,.\,) : \mathcal{Z} \to [0, 1]$ is a finitely additive and subadditive set function. (For the subadditivity see e.g. [10, Lemma A.6].) The $p(x, \,.\,)$ are even σ-subadditive on $\mathcal{Z}$: suppose $A \subset \bigcup_{k=1}^{\infty} A_k$ for some $A, A_k \in \mathcal{Z}$. As A is compact and all A_k are open, there is $n \in \mathbb{N}$ such that $A \subseteq \bigcup_{k=1}^{n} A_k$. Hence

$$p(x, A) \leqslant \sum_{k=1}^{n} p(x, A_k) \leqslant \sum_{k=1}^{\infty} p(x, A_k) \; .$$

As $\mathcal{Z}$ is a semi-ring, all $p(x, \,.\,)$ can be extended uniquely to finite measures on $(\mathbb{X}, \sigma(\mathcal{Z}))$, see e.g. [10, Beispiel 1.3 and Satz 2.26]. By definition of p we have

$$\int_A P 1_B \, d\mu = \int_B P^* 1_A \, d\mu = \int 1_B(x) \, p(x, A) \, d\mu(x)$$

for all $A \in \mathcal{Z}$ and $B \in \sigma(\mathcal{Z})$.

- For fixed $B \in \sigma(\mathcal{Z})$, both sides of this equation describe finite measures in A. Hence the identity extends to all $A \in \sigma(\mathcal{Z})$.
- For fixed $A \in \sigma(\mathcal{Z})$, both sides of this equation are linear, positive and σ-additive in 1_B. Hence equation (3.1.6) follows along standard lines (see e.g. the proof of [10, Satz 5.21]).

It remains to transfer the result from LR spaces with $\Omega = \mathbb{X}$ to arbitrary LR spaces $(\Omega, \mathcal{A}, \mu)$. By Theorem 3.1.18, $(\Omega, \mathcal{A}, \mu)$ is isomorphic mod 0 to $(\mathbb{X}, \mathcal{A}_{\mathcal{B}}, \mu_{\mathcal{B}})$ via a bi-measurable bijection $\phi_{\mathcal{B}} : \Omega \to \phi_{\mathcal{B}}(\Omega) \subseteq \mathbb{X}$. (Recall that $\mu_{\mathcal{B}} = \mu \circ \phi_{\mathcal{B}}^{-1}$ and $\mathcal{A}_{\mathcal{B}} = \sigma(\tau_{\mathbb{X}})_{\mu_{\mathcal{B}}}$.) Then $Uf := f \circ \phi_{\mathcal{B}}$ is a positive isometric isomorphism from $L^1(\mathbb{X}, \mathcal{A}_{\mathcal{B}}, \mu_{\mathcal{B}})$ to $L^1(\Omega, \mathcal{A}, \mu)$, and $\tilde{P} := U^{-1} P U$ is a positive linear contraction from $L^1(\mathbb{X}, \mathcal{A}_{\mathcal{B}}, \mu_{\mathcal{B}})$ to $L^1(\mathbb{X}, \phi_{\mathcal{B}}(\mathcal{F}), \mu_{\mathcal{B}})$. $\phi_{\mathcal{B}}(\mathcal{F})$ is a sub-σ-algebra of $\mathcal{A}_{\mathcal{B}}$ by Theorem 3.1.16. Therefore the first part of the proof yields a $\phi_{\mathcal{B}}(\mathcal{F})$-measurable substochastic kernel $\tilde{p}$ for $\tilde{P}$. Define $p(\omega, A) := \tilde{p}(\phi_{\mathcal{B}}(\omega), \phi_{\mathcal{B}}(A))$. As $\phi_{\mathcal{B}}(\sigma(\mathcal{B})) = \sigma(\tau_{\mathbb{X}})$, p is a substochastic kernel on $(\Omega, \sigma(\mathcal{B}), \mu)$, measurable w.r.t. $\phi_{\mathcal{B}}^{-1}(\phi_{\mathcal{B}}(\mathcal{F})) = \mathcal{F}$, and for all $\sigma(\mathcal{B})$-measurable $f \in L^1$ and $A \in \sigma(\mathcal{B})$

$$\begin{aligned}
\int_A Pf \, d\mu &= \int_{\phi_{\mathcal{B}}^{-1}(\phi_{\mathcal{B}}(A))} (\tilde{P}(U^{-1} f)) \circ \phi_{\mathcal{B}} \, d\mu = \int_{\phi_{\mathcal{B}}(A)} \tilde{P}(U^{-1} f) \, d\mu_{\mathcal{B}} \\
&= \int_{\mathbb{X}} (U^{-1} f)(x) \tilde{p}(x, \phi_{\mathcal{B}}(A)) \, d\mu_{\mathcal{B}}(x) = \int_{\Omega} f(\omega) \tilde{p}(\phi_{\mathcal{B}}(\omega), \phi_{\mathcal{B}}(A)) \, d\mu(\omega) \\
&= \int_{\Omega} f(\omega) p(\omega, A) \, d\mu(\omega) \; .
\end{aligned}$$

□

3.1.8 Regular conditional probabilities on pre-Lebesgue-Rohlin spaces

Let $(\Omega, \mathcal{A}, \mu)$ be a pLR space. Fix a countable basis $\mathcal{B}$ of the Hausdorff topology τ on Ω. Then $\mathcal{A} = \sigma(\tau) = \sigma(\mathcal{B})$. Observe that $\mathcal{B}$ is also a basis of the LR space $(\Omega, \mathcal{A}_\mu, \mu)$.

3.1.31 Definition *Let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{A}$. Any stochastic kernel p such that $p(\,.\,, A)$ is a $\mathcal{F}$-measurable representative of $E_\mu[1_A \mid \mathcal{F}]$ for all $A \in \mathcal{A}$ is called a regular conditional probability w.r.t. $\mathcal{F}$ and μ.*

3.1.32 Remark A $\mathcal{F}$-measurable stochastic kernel p is a regular conditional probability w.r.t. $\mathcal{F}$ and μ if and only if

$$\forall A \in \mathcal{A}\ \forall F \in \mathcal{F} :\ \mu(A \cap F) = \int_F p(\omega, A)\, d\mu(\omega)\ ,$$

because this is equivalent to $\forall \in \mathcal{A} :\ E_\mu[1_A \mid \mathcal{F}] = \frac{d(1_A\mu)|_\mathcal{F}}{d\mu|_\mathcal{F}} = p(\,.\,, A)$.

3.1.33 Corollary *Let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{A}$. There is a regular conditional probability $\mu_\mathcal{F}$ w.r.t. $\mathcal{F}$ and μ. For each $A \in \mathcal{A}$, $\mu^\mathcal{F}(\,.\,, A)$ is a representative of the conditional expectation $E_\mu[1_A \mid \mathcal{F}]$.*

Proof: Apply Theorem 3.1.30 to the positive L^1-contraction $P : f \mapsto E_\mu[f \mid \mathcal{F}]$ (Example 2.1.6) on the LR space $(\Omega, \mathcal{A}_\mu, \mu)$. Recall that also $E_\mu[1_A \mid \mathcal{F}] = P^* 1_A$. As $\mathcal{A} = \sigma(\mathcal{B})$, equation (3.1.7) implies for all $A \in \mathcal{A}$

$$E_\mu[1_A \mid \mathcal{F}] = P^* 1_A = p(\,.\,, A)\ .$$

□

Notation: We write $\mu^\mathcal{F}_\omega(A)$ instead of $\mu^\mathcal{F}(\omega, A)$, if convenient. We also denote $\mathcal{A}^1 := \{A \in \mathcal{A} : \mu(A) = 1\}$.

3.1.34 Lemma *If $\mu^\mathcal{F}$ and $\nu^\mathcal{F}$ are regular conditional probabilities w.r.t. $\mathcal{F}$ and μ, then there is $\Omega_0 \in \mathcal{A}^1$ such that*

i) $\mu^\mathcal{F}_\omega(A) = \nu^\mathcal{F}_\omega(A)$ for all $A \in \mathcal{A} = \sigma(\mathcal{B})$ and all $\omega \in \Omega_0$, and

ii) For all $\omega \in \Omega_0$ and $\mu^\mathcal{F}_\omega$-a.e. ω' the probabilities $\mu^\mathcal{F}_\omega$ and $\mu^\mathcal{F}_{\omega'}$ coincide on $\mathcal{A}$.

iii) If $\mathcal{F}_0 \subseteq \mathcal{F}$ is a countably generated sub-σ-algebra with $\mathcal{F} \subseteq (\mathcal{F}_0)_\mu$, Ω_0 can be chosen such that $\forall \omega \in \Omega_0\ \forall F \in \mathcal{F}_0\ \forall A \in \mathcal{A} :\ \mu^\mathcal{F}_\omega(F \cap A) = \delta_\omega(F) \cdot \mu^\mathcal{F}_\omega(A)$.

Proof: i) W.l.o.g. let $\mathcal{B} = \{B_1, B_2, \dots\}$ be closed under finite intersections. Then, for each B_i and each $F \in \mathcal{F}$,

$$\int_F \mu^\mathcal{F}(\omega, B_i)\, d\mu(\omega) = \int_F E_\mu[f \mid \mathcal{F}](\omega)\, d\mu(\omega) = \int_F \nu^\mathcal{F}(\omega, B_i)\, d\mu(\omega)\ .$$

As the $\mu^\mathcal{F}(\,.\,, B_i)$ and $\nu^\mathcal{F}(\,.\,, B_i)$ are $\mathcal{F}$-measurable, it follows that there is a set $\Omega_0 \in \mathcal{A}^1$ such that $\mu^\mathcal{F}(\omega, B_i) = \nu^\mathcal{F}(\omega, B_i)$ for the countably many $B_i \in \mathcal{B}$ and all $\omega \in \Omega_0$, so that $\mu^\mathcal{F}(\omega, .\,)|_{\sigma(\mathcal{B})} = \nu^\mathcal{F}(\omega, .\,)|_{\sigma(\mathcal{B})}$ for all $\omega \in \Omega_0$.

ii) For μ-a.e. $\omega \in \Omega$ and each $B \in \mathcal{B}$,

$$\int (\mu^\mathcal{F}(\,.\,, B))^2\, d\mu^\mathcal{F}_\omega = E_\mu[(\mu^\mathcal{F}(\,.\,, B))^2 \mid \mathcal{F}](\omega) = (\mu^\mathcal{F}(\omega, B))^2 = (E_\mu[\mu^\mathcal{F}(\,.\,, B) \mid \mathcal{F}](\omega))^2$$
$$= \left(\int \mu^\mathcal{F}(\,.\,, B)\, d\mu^\mathcal{F}_\omega\right)^2$$

The Cauchy-Schwarz (or Jensen)-inequality implies that for these ω the function $\omega' \mapsto \mu^{\mathcal{F}}_{\omega'}(B)$ is $\mu^{\mathcal{F}}_{\omega}$-a.e. constant and equal to $\mu^{\mathcal{F}}_{\omega}(B)$. Hence, for μ-a.e. ω and $\mu^{\mathcal{F}}_{\omega}$-a.e. ω', the probabilities $\mu^{\mathcal{F}}_{\omega}$ and $\mu^{\mathcal{F}}_{\omega'}$ coincide on the countable $\cap$-stable family $\mathcal{B}$ and hence on $\mathcal{A} = \sigma(\mathcal{B})$.

iii) Let $F \in \mathcal{F}$. Then for μ-a.e. $\omega \in \Omega$ and each $B \in \mathcal{B}$

$$\mu^{\mathcal{F}}_{\omega}(F \cap B) = E_\mu[1_F 1_B \mid \mathcal{F}](\omega) = 1_F(\omega)\, E_\mu[1_B \mid \mathcal{F}](\omega) = \delta_\omega(F) \cdot \mu^{\mathcal{F}}_{\omega}(B)\ . \tag{3.1.8}$$

As, for fixed F, both sides of this identity define a finite measure on $\mathcal{A} = \sigma(\mathcal{B})$, the claim follows for any $B \in \mathcal{A}$ and all F from a countable generator of $\mathcal{F}_0$. But for fixed $B \in \mathcal{A}$, both sides of the identity define finite measures in F, so the identity extends to $F \in \mathcal{F}_0$. □

3.1.35 Remark For $F = \Omega$, item iii) of this lemma implies in particular

$$\forall \omega \in \Omega_0\ \forall F \in \mathcal{F}_0 :\ \mu^{\mathcal{F}}_{\omega}(F) = \delta_\omega(F)\ . \tag{3.1.9}$$

On a technical level, the problem treated in the next section is to extend this to all $F \in \mathcal{F}$, when $\mathcal{F}$ is a suitable σ-algebra of invariant sets.

3.1.36 Remark Recall the setting from Remark 3.1.29: $T : \Omega \to \Omega$ is such that $\mu \circ T^{-1} \approx \mu$, and $\phi : \Omega \to \mathbb{X} = \{0,1\}^{\mathbb{N}}$ is such that $\mathcal{I}(T)_\mu = \sigma(\phi^{-1}(\tau_{\mathbb{X}}))_\mu$. Recall also the notation $\Gamma_x = \phi^{-1}(\{x\})$. Then

i) $\Omega = \bigcup_{x \in \mathbb{X}} \Gamma_x$.

ii) $\forall x \in \mathbb{X} :\ \Gamma_x \in \mathcal{I}(T)$.

Equation (3.1.9), applied to $\mathcal{F}_0 = \phi^{-1}(\sigma(\tau_{\mathbb{X}}))$, implies that for μ-a.e. $\omega \in \Omega$ and each $x \in \mathbb{X}$

$$\mu^{\mathcal{F}}_{\omega}(\Gamma_x) = \delta_\omega(\phi^{-1}(\{x\})) = \begin{cases} 1 & \text{if } \phi(\omega) = x \\ 0 & \text{otherwise}\ . \end{cases}$$

Therefore the conditional measures $\mu^{\mathcal{F}}_{\omega}$ can be localized on the sets Γ_x.

3.1.9 Exercises

3.1.1 a) Let $(\Omega_1, \mathcal{A}_1, \mu_1)$ and $(\Omega_2, \mathcal{A}_2, \mu_2)$ be pLR spaces. Prove that also $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \otimes \mu_2)$ is a pLR space.

b) Let $(\Omega_n, \mathcal{A}_n, \mu_n)$ $(n \in \mathbb{N})$ be pLR spaces. Prove that also $(\bigtimes_{n\in\mathbb{N}} \Omega_n, \bigotimes_{n\in\mathbb{N}} \mathcal{A}_n, \bigotimes_{n\in\mathbb{N}} \mu_n)$ is a pLR space.

3.1.2 Recall the shift example with the mixed Bernoulli measure from Remark 1.1.14. Define explicitly a map $\phi : \{0,1\}^{\mathbb{Z}} \to \mathbb{X}$ such that $\mathcal{I}(T)_\mu = \sigma(\phi_{\mathcal{B}}^{-1}(\tau_{\mathbb{X}}))_\mu$, see Remark 3.1.29. What is $\mu \circ \phi^{-1}$? Describe the sets Γ_x. Do the same with $[0,1]$ instead of $\mathbb{X}$. (This may be more meaningful than for $\mathbb{X}$, in particular the measure $\mu \circ \phi^{-1}$ may have a simple and clear interpretation if ϕ is chosen properly.)

3.1.3 Verify i) - iii) of Remark 3.1.29.

3.1.4 Let $(\Omega, \mathcal{A}, \mu)$ be a LR space and $\mathcal{F} \subset \mathcal{A}$ a μ-complete sub-σ-algebra. Suppose that the space $(\Omega, \mathcal{F}, \mu_{|\mathcal{F}})$ has no atoms. Prove that $(\Omega, \mathcal{A}, \mu)$ is isomorphic modulo 0 to $([0,1]^2, \sigma(\tau)_{\lambda^2}, \lambda^2)$, where $\sigma(\tau)$ is the Borel-σ-algebra on $[0,1]^2$ and λ^2 the two-dimensional Lebesgue measure, with an isomorphism that maps $\mathcal{F}$-measurable sets to sets of the form $B \times [0,1]$.

3.1.5 Let $(\Omega, \mathcal{A}, \mu)$ be a LR space and let $\mathcal{G}$ be a countable family of $\mathcal{A}$-measurable sets. Prove the following statements, which are essentially reformulations of a number of results of this section:
There is an (uncountable) partition $\Gamma = \{\Gamma_x : x \in \mathbb{X}_0 \subseteq \mathbb{X}\}$ of Ω such that

a) $\sigma(\mathcal{G}) = \{A \in \mathcal{A} : \forall x \in \mathbb{X}_0 : \Gamma_x \subseteq A \text{ or } \Gamma_x \cap A = \emptyset\}$.

b) There are probabiliy measures μ_x on $(\Gamma_x, \mathcal{A}|_{\Gamma_x})$ such that all $(\Gamma_x, \mathcal{A}|_{\Gamma_x}, \mu_x)$ are Lebesgue spaces.

c) $x \mapsto \mu_x(A \cap \Gamma_x)$ is $\sigma(\tau_{\mathbb{X}})$-measurable for each $A \in \mathcal{A}$.

d) There is a probability measure ν on $(\mathbb{X}, \sigma(\tau_{\mathbb{X}}))$ such that $\mu(A) = \int_{\mathbb{X}} \mu_x(A \cap \Gamma_x)\, d\nu(x)$ for all $A \in \mathcal{A}$.

3.1.6 Verify the claims made in Remark 3.1.22.

3.2 Raugi's ergodic decomposition theorem

In this section we prove an ergodic decomposition theorem along the lines of the purely measure theoretic/probabilistic approach of Raugi [15]. But observe also the paper by Greschonig and Schmidt [7].

3.2.1 Ergodically decomposing sub-σ-algebras

Let $(\Omega, \mathcal{A}, \mu)$ be a pre-Lebesgue-Rohlin space.

3.2.1 Definition *Let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{A}$ and let $\mu^{\mathcal{F}}$ be a regular conditional probability w.r.t. $\mathcal{F}$ and μ.*

a) $\mathcal{F}$ decomposes μ ergodically, if there is a set $\Omega_1 \in \mathcal{A}^1$ such that

$$\forall F \in \mathcal{F}\ \forall \omega \in \Omega_1 :\ \mu^{\mathcal{F}}_\omega(F) \in \{0,1\}\ .$$

b) $\mathcal{F}$ is conditionally countably generated mod 0, if there are a countable family $\mathcal{E} \subset \mathcal{F}$ and a set $\Omega_2 \in \mathcal{A}^1$ such that

$$\forall F \in \mathcal{F}\ \exists E \in \sigma(\mathcal{E})\ \forall \omega \in \Omega_2 :\ \mu^{\mathcal{F}}_\omega(F \triangle E) = 0\ .$$

Both definitions are independent of the particular choice of the regular conditional probability $\mu^{\mathcal{F}}$ - only the sets Ω_1 and Ω_2 depend on it, see Lemma 3.1.34i.

3.2.2 Remark a) If $\mathcal{F}$ is conditionally countably generated mod 0, then it is countably generated mod 0, because $\mu(F \triangle E) = \int \mu^{\mathcal{F}}_\omega(F \triangle E)\, d\mu(\omega)$.

b) In view of (3.1.9) there is always a sub-σ-algebra $\mathcal{F}_0 \subseteq \mathcal{F} \subseteq (\mathcal{F}_0)_\mu$ which decomposes μ ergodically. This does not imply that $\mathcal{F}$ decomposes μ ergodically, as the following example shows: Let $\Omega = [0,1]^2$, let μ be the two-dimensional Lebesgue measure on Ω and $\mathcal{A}$ the σ-algebra of Lebesgue-measurable sets. Consider the sub-σ-algebra $\mathcal{F}$ consisting of all sets $A \times [0,1] \cup N_1 \times [0,1/2) \cup N_2 \times [1/2,1]$ where $A \subset [0,1]$ is Lebesgue-measurable and $N_1, N_2 \subset [0,1]$ are at most countable. Denote also by $\mathcal{F}_0 \subset \mathcal{F}$ the sub-σ-algebra of all sets of the form $B \times [0,1]$ where B is a Borel-subset of $[0,1]$. Then $\mathcal{F} \subset (\mathcal{F}_0)_\mu$, $\mathcal{F}_0$ is countably generated and the measures $A \mapsto \lambda(A_x)$ (λ the one-dimensional Lebesgue measure) serve as regular conditional probabilities both for $\mu_x^{\mathcal{F}_0}$ and for $\mu_x^{\mathcal{F}}$. Obviously, $\mathcal{F}_0$ decomposes μ ergodically. But for each $x \in [0,1]$ each set $F \in \mathcal{F}_0$ can be modified on its x-fibre into a set A from $\mathcal{F}_0$ such that $\mu_x^{\mathcal{F}}(A) = \frac{1}{2}$. So $\mathcal{F}$ does not decompose μ ergodically.

We will see that such local modifications are often impossible when $\mathcal{F}$ is a σ-algebra of invariant sets.

3.2.3 Theorem *Let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{A}$ and let $\mu^{\mathcal{F}}$ be a regular conditional probability w.r.t. $\mathcal{F}$ and μ. The following are equivalent:*

(i) $\mathcal{F}$ decomposes μ ergodically.

(ii) For any countably generated sub-σ-algebra $\tilde{\mathcal{F}} \subseteq \mathcal{F}$ with $\mathcal{F} \subseteq \tilde{\mathcal{F}}_\mu$ holds: There is $\Omega_2 \in \mathcal{A}^1$ such that

$$\forall F \in \mathcal{F}\ \exists E \in \tilde{\mathcal{F}}\ \forall \omega \in \Omega_2 :\ \mu_\omega^{\mathcal{F}}(F \triangle E) = 0\ .$$

(iii) $\mathcal{F}$ is conditionally countably generated mod 0.

(iv) There is a set $\Omega_3 \in \mathcal{A}^1$ such that for all $\omega \in \Omega_3$ holds: the kernel $\mu^{\mathcal{F}}$ is a regular conditional probabilitiy for the probability measure $\mu_\omega^{\mathcal{F}}$, i.e.

$$\forall \omega \in \Omega_3\ \forall A \in \mathcal{A} :\ E_{\mu_\omega^{\mathcal{F}}}[1_A \mid \mathcal{F}](\omega') = \mu_{\omega'}^{\mathcal{F}}(A) \quad \text{for } \mu_\omega^{\mathcal{F}}\text{-a.e. } \omega', \text{ equivalently}$$

$$\forall \omega \in \Omega_3\ \forall A \in \mathcal{A}\ \forall F \in \mathcal{F} :\ \mu_\omega^{\mathcal{F}}(A \cap F) = \int_F \mu_{\omega'}^{\mathcal{F}}(A)\, d\mu_\omega^{\mathcal{F}}(\omega')\ .$$

As both sides of the last identity define a probability measure in A, it suffices to check either of these conditions for all A from a countable $\cap$-stable generator of $\mathcal{A}$.

(The equivalence between (i) and (iii) is from [15, Proposition 3.3].)

Proof: $(i) \Rightarrow (ii)$: Let $\tilde{\mathcal{F}} \subseteq \mathcal{F}$ be a countably generated sub-σ-algebra with $\mathcal{F} \subseteq \tilde{\mathcal{F}}_\mu$. By Corollary 3.1.33, there is a regular conditional probability $\mu^{\tilde{\mathcal{F}}}$. As $\tilde{\mathcal{F}} \subseteq \mathcal{F} \subseteq \tilde{\mathcal{F}}_\mu$, this serves also as regular conditional probability for $\mathcal{F}$and we denote it simply by $\mu^{\mathcal{F}}$.

Let Ω_1 be as in the definition of ergodic decomposition by $\mathcal{F}$. Then, for each $F \in \mathcal{F}$, we can decompose Ω disjointly as $\Omega = F^0 \cup F^1 \cup N^F$ where $F^0 := \{\omega \in \Omega : \mu_\omega^{\mathcal{F}}(F) = 0\}$, $F^1 := \{\omega \in \Omega : \mu_\omega^{\mathcal{F}}(F) = 1\}$, and $N^F \subseteq \Omega \backslash \Omega_1$. As $\mu^{\mathcal{F}}$ could be chosen $\tilde{\mathcal{F}}$-measurable, we have $F^0, F^1 \in \tilde{\mathcal{F}}$. In particular, for all $\omega \in \Omega_1$ and all $F \in \mathcal{F}$, $\mu_\omega^{\mathcal{F}}(F \cap F^0) \in \{0,1\}$ and also $\mu_\omega^{\mathcal{F}}(F \cap F^1) \in \{0,1\}$.

Let $\Omega_2 = \{\omega \in \Omega_1 : \mu_\omega^{\mathcal{F}}(\Omega_1) = 1\}$. Then $\mu(\Omega_2) = 1$. For $\omega \in \Omega_2$ and each $F \in \mathcal{F}$ we have $\mu_\omega^{\mathcal{F}}(N^F) \leqslant \mu_\omega^{\mathcal{F}}(\Omega \backslash \Omega_1) = 0$. Therefore, in view of Lemma 3.1.34iii

$$\mu_\omega^{\mathcal{F}}(F \backslash F^1) = \mu_\omega^{\mathcal{F}}(F \cap F^0) = \delta_\omega(F^0) \cdot \mu_\omega^{\mathcal{F}}(F) = 0$$

and

$$\mu_\omega^{\mathcal{F}}(F^1 \backslash F) = \mu_\omega^{\mathcal{F}}(F^c \cap F^1) = \delta_\omega(F^1) \cdot \mu_\omega^{\mathcal{F}}(F^c) = 0 ,$$

so that $\mu_\omega^{\mathcal{F}}(F \triangle F^1) = 0$ for all $\omega \in \Omega_2$ and all $F \in \mathcal{F}$ with $F^1 \in \tilde{\mathcal{F}}$.
$(ii) \Rightarrow (iii)$: This follows from Corollary 3.1.27 which asserts the existence of at least one countably generated sub-σ-algebra as in (ii).
$(iii) \Rightarrow (iv)$: Let $\tilde{\mathcal{F}} = \sigma(\mathcal{E})$, $\mathcal{E}$ from Definition 3.2.1b. By Lemma 3.1.34ii and iii, for all $\omega \in \Omega_0$, all $F \in \tilde{\mathcal{F}}$ and all $A \in \mathcal{A}$,

$$\int_F \mu_{\omega'}^{\mathcal{F}}(A)\, d\mu_\omega^{\mathcal{F}}(\omega') = \mu_\omega^{\mathcal{F}}(A) \cdot \mu_\omega^{\mathcal{F}}(F) = \mu_\omega^{\mathcal{F}}(A) \cdot \delta_\omega(F) = \mu_\omega^{\mathcal{F}}(A \cap F)$$

As $\mathcal{F}$ is conditionally countably generated mod 0, this extends to all $F \in \mathcal{F}$.
$(iv) \Rightarrow (i)$: Let $\omega \in \Omega_0 \cap \Omega_3$ with Ω_0 from Lemma 3.1.34. Then $\mu_{\omega'}^{\mathcal{F}} = \mu_\omega^{\mathcal{F}}$ for $\mu_\omega^{\mathcal{F}}$-a.e. ω', and for each $F \in \mathcal{F}$ and $\mu_\omega^{\mathcal{F}}$-a.e. ω' holds

$$\mu_\omega^{\mathcal{F}}(F) = \mu_{\omega'}^{\mathcal{F}}(F) = E_{\mu_\omega^{\mathcal{F}}}[1_F \mid \mathcal{F}](\omega') = 1_F(\omega') .$$

Hence $\mu_\omega^{\mathcal{F}}(F) \in \{0, 1\}$. □

3.2.4 Corollary *Let $\mathcal{G} \subseteq \mathcal{F} \subseteq \mathcal{A}$ be sub-σ-algebras. If $\mathcal{F} \subseteq \mathcal{G}_\mu$ and if $\mathcal{F}$ decomposes μ ergodically, then also $\mathcal{G}$ decomposes μ ergodically.*

Proof: By Corollary 3.1.27 there is a countably generated sub-σ-algebra $\tilde{\mathcal{G}} \subseteq \mathcal{G}$ such that $\mathcal{G} \subseteq \tilde{\mathcal{G}}_\mu$, hence $\tilde{\mathcal{G}} \subseteq \mathcal{F}$ and $\mathcal{F} \subseteq \mathcal{G}_\mu = \tilde{\mathcal{G}}_\mu$. As $\mathcal{F}$ decomposes μ ergodically, Theorem 3.2.3(ii) implies that there is $\Omega_2 \in \mathcal{A}^1$ such that

$$\forall F \in \mathcal{F}\ \exists E \in \tilde{\mathcal{G}}\ \forall \omega \in \Omega_2 :\ \mu_\omega^{\mathcal{F}}(F \triangle E) = 0 .$$

This holds a fortiori for all $F \in \mathcal{G}$, so $\mathcal{G}$ is conditionally countably generated mod 0, and another application of Theorem 3.2.3 shows that $\mathcal{G}$ decomposes μ ergodically.

□

We finish this subsection by the following observation: for two sub-σ-algebras $\mathcal{F}$ and $\mathcal{G}$ of $\mathcal{A}$ with regular conditional probabilities $\mu^{\mathcal{F}}$ and $\mu^{\mathcal{G}}$ define the stochastic kernel $\mu^{\mathcal{G}} \ltimes \mu^{\mathcal{F}}$ by

$$\mu^{\mathcal{G}} \ltimes \mu^{\mathcal{F}}(\omega, A) := \int_\Omega \mu_{\omega'}^{\mathcal{F}}(A)\, d\mu_\omega^{\mathcal{G}}(\omega') . \tag{3.2.1}$$

In this situation we have the following lemma:

3.2.5 Lemma *a) If $\mathcal{F}$ decomposes μ ergodically and if*

$$(\mu^{\mathcal{G}} \ltimes \mu^{\mathcal{F}})_\omega = \mu_\omega^{\mathcal{G}} \quad \text{for } \mu\text{-a.e. } \omega, \tag{3.2.2}$$

then $\mu^{\mathcal{F}}$ is a regular conditional probability w.r.t. $\mathcal{F}$ and $\mu_\omega^{\mathcal{G}}$ for μ-a.e. ω.
b) (3.2.2) holds in particular, if $\mathcal{G} \subseteq \mathcal{F}$.

Proof: a) As $\mathcal{F}$ decomposes μ ergodically, there is, by Theorem 3.2.3(iv), a set $\Omega_3 \in \mathcal{A}^1$ such that for all $\tilde{\omega} \in \Omega_3$

$$\forall A \in \mathcal{A}\, \forall F \in \mathcal{F}: \; \mu^{\mathcal{F}}_{\tilde{\omega}}(A \cap F) = \int_F \mu^{\mathcal{F}}_{\omega'}(A)\, d\mu^{\mathcal{F}}_{\tilde{\omega}}(\omega') \; .$$

As $\mu^{\mathcal{F}}$ is a regular conditional probability w.r.t. $\mathcal{F}$ and μ, there is $\Omega_4 \in \mathcal{A}^1$ such that this identity holds for all $\omega \in \Omega_4$ and $\mu^{\mathcal{F}}_{\omega}$-a.e. $\tilde{\omega}$. Hence, integrating w.r.t. $\mu^{\mathcal{F}}_{\omega}$ we have for all $\omega \in \Omega_4$ that also satisfy $(\mu^{\mathcal{G}} \ltimes \mu^{\mathcal{F}})_\omega = \mu^{\mathcal{G}}_\omega$ and for all $A \in \mathcal{A}$ and $F \in \mathcal{F}$

$$\begin{aligned} \mu^{\mathcal{G}}_\omega(A \cap F) = (\mu^{\mathcal{G}} \ltimes \mu^{\mathcal{F}})_\omega(A \cap F) &= \int_\Omega \mu^{\mathcal{F}}_{\tilde{\omega}}(A \cap F)\, d\mu^{\mathcal{G}}_\omega(\tilde{\omega}) \\ &= \int_\Omega \left(\int_F \mu^{\mathcal{F}}_{\omega'}(A)\, d\mu^{\mathcal{F}}_{\tilde{\omega}}(\omega') \right) d\mu^{\mathcal{G}}_\omega(\tilde{\omega}) = \int_F \mu^{\mathcal{F}}_{\omega'}(A)\, d(\mu^{\mathcal{G}} \ltimes \mu^{\mathcal{F}})_\omega(\omega') \\ &= \int_F \mu^{\mathcal{F}}_{\omega'}(A)\, d\mu^{\mathcal{G}}_\omega(\omega') \; . \end{aligned}$$

Hence, for μ-a.e. ω, $\mu^{\mathcal{F}}$ is a regular conditional probability w.r.t. $\mathcal{F}$ and $\mu^{\mathcal{G}}_\omega$.
b) Let $A \in \mathcal{A}$. We have the μ-a.e. identities

$$\begin{aligned} (\mu^{\mathcal{G}} \ltimes \mu^{\mathcal{F}})_\omega(A) &= \int_\Omega \mu^{\mathcal{F}}_{\omega'}(A)\, d\mu^{\mathcal{G}}_\omega(\omega') = E_\mu[\mu^{\mathcal{F}}_{\cdot}(A) \mid \mathcal{G}](\omega) = E_\mu\left[E_\mu[1_A \mid \mathcal{F}] \mid \mathcal{G}\right](\omega) \\ &= E_\mu\left[1_A \mid \mathcal{G}\right](\omega) = \mu^{\mathcal{G}}_\omega(A) \; . \end{aligned}$$

Hence for μ-a.e. ω the measures $(\mu^{\mathcal{G}} \otimes \mu^{\mathcal{F}})_\omega$ and $\mu^{\mathcal{G}}_\omega$ coincide on a countable generator of $\mathcal{A}$ and therefore on all of $\mathcal{A}$.

□

3.2.2 The ergodic decomposition for a single non-singular automorphism

3.2.6 Definition *Let $(\Omega, \mathcal{A}, m)$ be a σ-finite measure space. A bijective and bi-measurable map T : $\Omega \to \Omega$ satisfying $m \circ T^{-1} \approx m$ is called a non-singular automorphism.*

3.2.7 Theorem *Let T be a non-singular automorphism of the pLR space $(\Omega, \mathcal{A}, \mu)$. Recall that $\mathcal{I}(T) = \{A \in \mathcal{A} : T^{-1}(A) = A\}$.*

a) Let $\mathcal{F} \subseteq \mathcal{I}(T)$ be a sub-σ-algebra with a regular conditional probability $\mu^{\mathcal{F}}$. Then $\mu^{\mathcal{F}}_\omega \circ T^{-1} \approx \mu^{\mathcal{F}}_\omega$ for μ-a.e. ω.

b) $\mathcal{I}(T)$ decomposes μ ergodically.

Proof (following [15, Proposition 5.1]):
a) Let ψ be a version of the Radon-Nikodym derivative $\frac{d(\mu \circ T^{-1})}{d\mu}$. We will prove that for μ-a.e. ω it is also a version of $\frac{d(\mu^{\mathcal{F}}_\omega \circ T^{-1})}{d\mu^{\mathcal{F}}_\omega}$: let $A \in \mathcal{A}$ and $F \in \mathcal{F}$. Then

$$\begin{aligned} \int_F \mu^{\mathcal{F}}_\omega \circ T^{-1}(A)\, d\mu(\omega) &= \int_F E_\mu\left[1_{T^{-1}(A)} \mid \mathcal{F}\right] d\mu = \mu(F \cap T^{-1}(A)) = \mu(T^{-1}(F \cap A)) \\ &= \int_{F \cap A} \psi\, d\mu = \int_F E_\mu\left[1_A\, \psi \mid \mathcal{F}\right] d\mu = \int_F \left(\int_A \psi\, d\mu^{\mathcal{F}}_\omega \right) d\mu(\omega) \; . \end{aligned}$$

Hence, for each given $A \in \mathcal{A}$, $\mu_\omega^\mathcal{F} \circ T^{-1}(A) = \int_A \psi \, d\mu_\omega^\mathcal{F}$ for μ-a.e. ω. Let $\mathcal{E}$ be a countable $\cap$-stable generator for $\mathcal{A}$. Then there is $\Omega' \in \mathcal{A}^1$ such that the measures $\mu_\omega^\mathcal{F} \circ T^{-1}$ and $\psi\mu_\omega^\mathcal{F}$ coincide on $\mathcal{E}$ - and hence on $\mathcal{A} = \sigma(\mathcal{E})$ – for all $\omega \in \Omega'$.
b) Recall from Example 2.1.4a that $Pf := f \circ T^{-1} \cdot \psi$ defines a positive L^1-contraction. Let again $\mathcal{B}$ be a countable $\cap$-stable generator for $\mathcal{A}$. From Theorem 2.2.11 (a variant of Hurewicz's theorem) we know that there is $\Omega'' \in \mathcal{A}^1$ such that

$$\forall \omega' \in \Omega'' \; \forall A \in \mathcal{E} \; : \lim_{n\to\infty} \frac{\sum_{k=-n}^{n}(P^k 1_A)(\omega')}{\sum_{k=-n}^{n}(P^k 1)(\omega)} = E_\mu[1_A \mid \mathcal{I}](\omega') = \mu_{\omega'}^\mathcal{I}(A) \,,$$

where the regular conditional probability $\mu^\mathcal{I}$ exists in view of Corollary 3.1.33. Changing Ω'' (if necessary), we can replace the "$\forall \omega' \in \Omega''$" by "$\forall \omega \in \Omega''$ and $\mu_\omega^\mathcal{I}$-a.e. ω'", because $\mu^\mathcal{I}$ is a regular conditional probability w.r.t. $\mathcal{I}$ and μ.

The same theorem, applied to the non-singular measures $\mu_\omega^\mathcal{I}$ ($\omega \in \Omega'$) from part a), tells us that for each $\omega \in \Omega'$ and $\mu_\omega^\mathcal{I}$-a.e. ω'

$$\lim_{n\to\infty} \frac{\sum_{k=-n}^{n}(P^k 1_A)(\omega')}{\sum_{k=-n}^{n}(P^k 1)(\omega')} = E_{\mu_\omega^\mathcal{I}}[1_A \mid \mathcal{I}](\omega') \quad \text{for all } A \in \mathcal{E}.$$

Hence, for all $\omega \in \Omega' \cap \Omega''$ and all $\mathcal{A} \in \mathcal{E}$, we have

$$\mu_\omega^\mathcal{I}(A) = E_{\mu_\omega^\mathcal{I}} \left[1_A \mid \mathcal{I}\right](\omega') \quad \text{for } \mu_\omega^\mathcal{I}\text{-a.e. } \omega', \tag{3.2.3}$$

and this extends to all $A \in \sigma(\mathcal{E}) = \mathcal{A}$, because the family of sets A satisfying (3.2.3) forms a Dynkin-System. Recall from Lemma 3.1.34ii that there is $\Omega_0 \in \mathcal{A}^1$ such that $\mu_\omega^\mathcal{I} = \mu_{\omega'}^\mathcal{I}$ for all $\omega \in \Omega_0$ and $\mu_\omega^\mathcal{I}$-a.e. ω'. Hence, for $\omega \in \Omega_0 \cap \Omega' \cap \Omega''$ and all $A \in \mathcal{A}$,

$$\mu_{\omega'}^\mathcal{I}(A) = E_{\mu_\omega^\mathcal{I}} \left[1_A \mid \mathcal{I}\right](\omega') \quad \text{for } \mu_\omega^\mathcal{I}\text{-a.e. } \omega',$$

But his means that $\mu^\mathcal{I}$ is a regular contitional probability for the probability measure $\mu_\omega^\mathcal{I}$, and Theorem 3.2.3 implies that $\mathcal{I}(T)$ decomposes μ ergodically. □

3.2.3 The ergodic decomposition for countable families of non-singular automorphisms

In this section $(\Omega, \mathcal{A}, \mu)$ is a pLR space.

3.2.8 Theorem *[15, Corollary 5.2] Let $(T_i : i \in I)$ be an at most countable family of non-singular automorphisms of the pLR space $(\Omega, \mathcal{A}, \mu)$. Then*
a) the sub-σ-algebra $\mathcal{I} := \bigcap_{i\in I} \mathcal{I}(T_i)$ decomposes μ ergodically, and
b) for μ-a.e. $\omega \in \Omega$ holds: $\forall i \in I : \; \mu_\omega^\mathcal{I} \circ T_i^{-1} \approx \mu_\omega^\mathcal{I}$.

This theorem is essentially a corollary to the next theorem whose proof is postponed to the next subsection.

3.2.9 Theorem *[15, Theorem 3.4] Let $(\mathcal{F}_i : i \in I)$ be an at most countable nonempty family of sub-σ-algebras that decompose μ ergodically. If*

$$\left(\bigcap_{i\in J} \mathcal{F}_i\right)_\mu = \bigcap_{i\in J} (\mathcal{F}_i)_\mu \quad \text{for all finite subsets } J \subseteq I, \tag{3.2.4}$$

then also $\bigcap_{i\in I} \mathcal{F}_i$ decomposes μ ergodically.

3.2.10 Remark Raugi [15] formulates condition (3.2.4) not in terms of the μ-completions but in terms of *relative μ-completions* $\overline{\mathcal{F}} := \mathcal{F}_\mu \cap \mathcal{A}$. His condition reads

$$\overline{\bigcap_{i\in J}\mathcal{F}_i} = \bigcap_{i\in J}\overline{\mathcal{F}_i} \quad \text{for all finite subsets } J \subseteq I. \tag{3.2.5}$$

Obviously (3.2.5) follows from (3.2.4) by intersecting both sides of (3.2.4) with $\mathcal{A}$. The converse is an immediate consequence of the following two assertions:

a) $\left(\overline{\mathcal{F}}\right)_\mu = \mathcal{F}_\mu$, which follows from $\mathcal{F} \subseteq \overline{\mathcal{F}} \subseteq \mathcal{F}_\mu$.

b) $\left(\bigcap_{i\in J}\overline{\mathcal{F}_i}\right)_\mu = \bigcap_{i\in J}(\mathcal{F}_i)_\mu$. This can be seen as follows: $\bigcap_{i\in J}\overline{\mathcal{F}_i} = \bigcap_{i\in J}(\mathcal{F}_i)_\mu \cap \mathcal{A} \subseteq (\mathcal{F}_j)_\mu$ for each $j \in J$. On the other hand, if $A \in \bigcap_{i\in J}(\mathcal{F}_i)_\mu$, then there are $A_i \in \mathcal{F}_i \subseteq \mathcal{A}$ such that $\mu(A\triangle A_i) = 0$ $(i \in J)$. In particular $A_i \in \mathcal{A}$ and $\mu(A_i\triangle A_j) = 0$ for all $i, j \in J$. Hence $A_j \in \bigcap_{i\in J}\overline{\mathcal{F}_i}$ for each $j \in J$, so that $A \in \left(\bigcap_{i\in J}\overline{\mathcal{F}_i}\right)_\mu$.

One may notice that the above arguments remain valid for infinite sets $J \subseteq I$.

3.2.11 Remark Raugi claims both theorems for arbitrary (also uncountable) families of automorphisms or σ-algebras, respectively. Below we give an example which shows that Theorem 3.2.9 does not extend to uncountable families – at least not without additional assumptions. I do not know whether Theorem3.2.8 extends to uncountable families. It would do in cases where the relative completions $\overline{\mathcal{I}(T_i)}$ decompose μ ergodically, but I have no indication that this is generally true.

3.2.12 Example Let $(\Omega, \mathcal{A}, \mu)$ be $[0, \frac{1}{2}) \times \{1, 2\}$ equipped with its product Borel-σ-algebra and the probability μ = Lebesgue $\times$ counting measure. Denote by $\mathcal{N}$ the family of all uncountable Borel subsets of $[0, \frac{1}{2})$ with zero Lebesgue measure. For $N \in \mathcal{N}$ let

$$\begin{aligned}\mathcal{F}_N := \{&N_0 \times \{1,2\} \cup M_1 \times \{1\} \cup M_2 \times \{2\} \cup A_1 \times \{1\} \cup A_2 \times \{2\} :\\ &N_0, M_1, M_2 \subseteq N, A_1, A_2 \subseteq \Omega\backslash N, N_0, A_1, A_2 \text{ Borel}, M_1, M_2 \text{ at most countable}\}.\end{aligned}$$

Then $\mathcal{F}_N$ is a σ-algebra and $\overline{\mathcal{F}_N} = \mathcal{A}$ decomposes μ ergodically for all $N \in \mathcal{N}$, because $\mu^{\mathcal{A}}_\omega = \delta_\omega$. If $N_1, \dots, N_k \in \mathcal{N}$, then

$$\mathcal{F}_{N_1} \cap \dots \cap \mathcal{F}_{N_k} = \mathcal{F}_{N_1\cup\dots\cup N_k}.$$

Hence

$$\overline{\mathcal{F}_{N_1} \cap \dots \cap \mathcal{F}_{N_k}} = \overline{\mathcal{F}_{N_1\cup\dots\cup N_k}} = \mathcal{A} = \overline{\mathcal{F}_{N_1}} \cap \dots \cap \overline{\mathcal{F}_{N_k}}.$$

Consider now $\mathcal{F} := \bigcap_{N\in\mathcal{N}}\mathcal{F}_N$. We will show that $\mathcal{F}$ does not decompose μ ergodically: Observe first that

$$\mathcal{F} = \bigcap_{N\in\mathcal{N}}\mathcal{F}_N = \{A \times \{1,2\} \cup M_1 \times \{1\} \cup M_2 \times \{2\} : A \text{ Borel}, M_1, M_2 \text{ at most countable}\}.$$

(This follows from the fact that a Borel set $B \subseteq [0, \frac{1}{2})$ that has at most countable intersection with any Borel Lebesgue-null set is itself at most countable. If B is a nul set, this is obvious. Otherwise one uses the fact that each Borel set of positive Lebesgue measure contains an uncountable Borel null set.) We have $\left(\pi_1^{-1}(\text{Borel sets})\right)_\mu \subseteq \mathcal{F} \subseteq \left(\pi_1^{-1}(\text{Borel sets})\right)_\mu$, hence the family $\mu^{\mathcal{F}}_{(x,a)} := \frac{1}{2}\left(\delta_{(x,1)} + \delta_{(x,2)}\right)$ serves as regular conditional probability for $\mathcal{F}$. (Let A be a Borel subset of Ω and $F = F_0 \times \{1, 2\}$. Then $\int_F \mu^{\mathcal{F}}_{(x,a)}(A)\, d\mu(x, a) = \frac{1}{2}\int_{F_0}(1_A(x, 1) + 1_A(x, 2))\, d(\mu \circ \pi_1^{-1})(x) = \int(1_{F_0\cap A_1}(x) + 1_{F_0\cap A_2}(x))\, dx = \mu(F \cap A)$.) But for each $\omega = (x, a) \in \Omega$ the set $\Omega\backslash\{(x, a)\}$ belongs to $\mathcal{F}$ and $\mu^{\mathcal{F}}_{(x,a)}(\Omega\backslash\{(x, a)\}) = \frac{1}{2} \notin \{0, 1\}$.

Proof that Theorem 3.2.8 follows from Theorem 3.2.9: Observing that I is at most countable, assertion b) is a direct consequence of Theorem 3.2.7a. We turn to assertion a). By Theorem 3.2.7b each $\mathcal{I}(T_i)$ decomposes μ ergodically. So it suffices to prove that $\bigcap_{i\in J}\mathcal{I}(T_i)_\mu \subseteq (\bigcap_{i\in J}\mathcal{I}(T_i))_\mu$ for all finite subsets $J \subseteq I$. For this we will use the general identity among sets $A\backslash(A\triangle B) = A \cap B$.

Let $A \in \bigcap_{i\in J}\mathcal{I}(T_i)_\mu$. There are $A_i \in \mathcal{I}(T_i)$ $(i \in J)$ with $\mu(A\triangle A_i) = 0$. Hence

$$\mu(A\triangle T_i^{-1}A) \leqslant \mu(A\triangle A_i) + \mu(T_i^{-1}(A_i\triangle A)) = 0 \quad \text{for all } i \in J.$$

Denote by G the countable algebraic subgroup of all non-singular automorphisms on $(\Omega, \mathcal{A}, \mu)$ which is generated by the $T_i (i \in J)$ and set $N := \bigcup_{g\in G}\bigcup_{i\in J} g^{-1}(A\triangle T_i^{-1}A)$. As $\mu \circ g^{-1} \approx \mu$ for all $g \in G$, we have $\mu(N) = 0$. As $G = \{g \circ T_i : g \in G\}$, we have $T_i^{-1}(N) = N$ for all $i \in J$, so that

$$T_i^{-1}(A\backslash N) = T_i^{-1}(A)\backslash N = \left(T_i^{-1}A\backslash(A\triangle T_i^{-1}A)\right)\backslash N = \left(A \cap T_i^{-1}A\right)\backslash N$$

and

$$A\backslash N = \left(A\backslash(A\triangle T_i^{-1}A)\right)\backslash N = \left(A \cap T_i^{-1}A\right)\backslash N\ .$$

Hence $A\backslash N \in \bigcap_{i\in J}\mathcal{I}(T_i)$, and as $A\backslash N \subseteq A \subseteq A \cup N$ it follows that $A \in (\bigcap_{i\in J}\mathcal{I}(T_i))_\mu$. □

3.2.4 Proof of Theorem 3.2.9 (Raugi's main theorem)

The case of two sub-σ-algebras

For $i \in I := \{1, 2\}$ and $A \in \mathcal{A}$ let $(P_i 1_A)(\omega) := \mu_\omega^{\mathcal{F}_i}(A)$. Then $P_i 1_A$ is a version of the conditional expectation $E_\mu[1_A \mid \mathcal{F}_i]$. From Corollary 2.2.9 we know that for each $A \in \mathcal{A}$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}(P_2P_1)^k 1_A = E_\mu[1_A \mid (\mathcal{F}_1)_\mu \cap (\mathcal{F}_2)_\mu] \quad \mu\text{-a.e.} \tag{3.2.6}$$

By the assumption of the theorem we have $(\mathcal{F}_1)_\mu \cap (\mathcal{F}_2)_\mu = (\mathcal{F}_1 \cap \mathcal{F}_2)_\mu$. Let $\mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2$ and denote $\mu^{\mathcal{F}}$ the corresponding regular conditional probability w.r.t. μ. Let $\mathcal{E}$ be a countable $\cap$-stable generator for $\mathcal{A}$. Then (3.2.6) implies that there is $\Omega' \in \mathcal{A}^1$ such that

$$\forall\omega' \in \Omega'\ \forall A \in \mathcal{E} :\ \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}(P_2P_1)^k 1_A(\omega') = \mu_{\omega'}^{\mathcal{F}}(A)\ . \tag{3.2.7}$$

Changing Ω' (if necessary), we can replace the "$\forall\omega' \in \Omega'$" by "$\forall\omega \in \Omega'$ and $\mu_\omega^{\mathcal{F}}$-a.e. ω'", because $\mu^{\mathcal{F}}$ is a regular conditional probability w.r.t. $\mathcal{F}$ and μ.

Because of Lemma 3.2.5, there is a set $\Omega'' \in \mathcal{A}^1$ such that for all $\omega \in \Omega''$ holds: $\mu^{\mathcal{F}_i}$ is a regular conditional probability w.r.t. $\mathcal{F}_i$ and $\mu_\omega^{\mathcal{F}}$ $(i = 1, 2)$ so that we have for $\omega \in \Omega''$

$$(P_i 1_A)(\omega') = \mu_{\omega'}^{\mathcal{F}_i}(A) = E_{\mu_\omega^{\mathcal{F}}}[1_A \mid \mathcal{F}_i](\omega') \quad \text{for } \mu_\omega^{\mathcal{F}}\text{-a.e. } \omega'.$$

Hence, another application of Corollary 2.2.9 tells us that, for any $\omega \in \Omega' \cap \Omega''$ and for any $A \in \mathcal{E}$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}(P_2P_1)^k 1_A(\omega') = E_{\mu_\omega^{\mathcal{F}}}[1_A \mid \tilde{\mathcal{F}}_1 \cap \tilde{\mathcal{F}}_2](\omega') \quad \text{for } \mu_\omega^{\mathcal{F}}\text{-a.e. } \omega', \tag{3.2.8}$$

where $\tilde{\mathcal{F}}_i = (\mathcal{F}_i)_{\mu_\omega^{\mathcal{F}}}$. Therefore Lemma 3.1.34ii implies that therer is $\Omega''' \in \mathcal{A}^1$ such that for all $\omega \in \Omega' \cap \Omega'' \cap \Omega'''$, all $A \in \mathcal{E}$ and $\mu_\omega^{\mathcal{F}}$-a.e. ω'

$$\mu_\omega^{\mathcal{F}}(A) = \mu_{\omega'}^{\mathcal{F}}(A) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (P_2 P_1)^k 1_A(\omega') = E_{\mu_\omega^{\mathcal{F}}}[1_A \mid \tilde{\mathcal{F}}_1 \cap \tilde{\mathcal{F}}_2](\omega') \,. \tag{3.2.9}$$

Hence, for $\omega \in \Omega' \cap \Omega'' \cap \Omega'''$, $F \in \mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2 \subseteq \tilde{\mathcal{F}}_1 \cap \tilde{\mathcal{F}}_2$ and $A \in \mathcal{E}$,

$$\mu_\omega^{\mathcal{F}}(A \cap F) = \int_F 1_A \, d\mu_\omega^{\mathcal{F}} = \int_F E_{\mu_\omega^{\mathcal{F}}}[1_A \mid \tilde{\mathcal{F}}_1 \cap \tilde{\mathcal{F}}_2](\omega') \, d\mu_\omega^{\mathcal{F}}(\omega') = \int_F \mu_\omega^{\mathcal{F}}(A) \, d\mu_\omega^{\mathcal{F}}(\omega') \,,$$

and as for fixed F both sides of this identity define finite measures in A, the identity extends to all $A \in \sigma(\mathcal{E}) = \mathcal{A}$ and in particular to $A = F$. Therefore $\mu_\omega^{\mathcal{F}}(F) = \mu_\omega^{\mathcal{F}}(F)^2 \in \{0,1\}$ for all $\omega \in \Omega' \cap \Omega'' \cap \Omega''' \in \mathcal{A}^1$, i.e. $\mathcal{F}$ decomposes μ ergodically.

The case of a sequence of sub-σ-algebras

Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a sequence of sub-σ algebras that are ergodic w.r.t. μ and satisfy condition (3.2.4). For $n \in \mathbb{N}$ let $\mathcal{G}_n = \bigcap_{i=1}^n \mathcal{F}_i$. Then, for any $n \geqslant 2$ we have $\mathcal{G}_n = \mathcal{G}_{n-1} \cap \mathcal{F}_n$ and

$$(\mathcal{G}_{n-1})_\mu \cap (\mathcal{F}_n)_\mu = \bigcap_{i=1}^{n-1} (\mathcal{F}_i)_\mu \cap (\mathcal{F}_n)_\mu = \left(\bigcap_{i=1}^{n} \mathcal{F}_i \right)_\mu = (\mathcal{G}_{n-1} \cap \mathcal{F}_n)_\mu \,.$$

Hence, from the case of two sub-σ-algebras it follows inductively that all $\mathcal{G}_n$ decompose μ ergodically.

Let $\mathcal{G} = \bigcap_{n\in\mathbb{N}} \mathcal{G}_n$ and consider again a countable $\cap$-stable generator $\mathcal{E}$ of $\mathcal{A}$. By the decreasing martingale theorem [10, Satz C.4], there is $\Omega' \in \mathcal{A}^1$ such that

$$\forall \omega' \in \Omega' \; \forall A \in \mathcal{E} : \; \lim_{n\to\infty} \mu_{\omega'}^{\mathcal{G}_n}(A) = \lim_{n\to\infty} E_\mu[1_A \mid \mathcal{G}_n](\omega') = E_\mu[1_A \mid \mathcal{G}](\omega') = \mu_{\omega'}^{\mathcal{G}}(A) \,.$$

Observing Lemma 3.1.34ii and changing Ω' (if necessary) this implies

$$\forall \omega \in \Omega' \; \forall A \in \mathcal{E} : \; \lim_{n\to\infty} \mu_{\omega'}^{\mathcal{G}_n}(A) = \mu_{\omega'}^{\mathcal{G}}(A) = \mu_\omega^{\mathcal{G}}(A) \quad \text{for } \mu_\omega^{\mathcal{G}}\text{-a.e. } \omega'.$$

The dominated convergence theorem implies that for all $G \in \mathcal{G}$, $A \in \mathcal{E}$, and $\omega \in \Omega'$,

$$\lim_{n\to\infty} \int_G \mu_{\omega'}^{\mathcal{G}_n}(A) \, d\mu_\omega^{\mathcal{G}}(\omega') = \int_G \lim_{n\to\infty} \mu_{\omega'}^{\mathcal{G}_n}(A) \, d\mu_\omega^{\mathcal{G}}(\omega') = \int_G \mu_\omega^{\mathcal{G}}(A) \, d\mu_\omega^{\mathcal{G}}(\omega') = \mu_\omega^{\mathcal{G}}(A) \cdot \mu_\omega^{\mathcal{G}}(G) \,.$$

Because of Lemma 3.2.5 applied with $\mathcal{G} \subseteq \mathcal{F} := \mathcal{G}_n$, there is $\Omega'' \in \mathcal{A}^1$ such that $\mu^{\mathcal{G}_n}$ is a regular conditional probability w.r.t. $\mathcal{G}$ and $\mu_\omega^{\mathcal{G}_n}$ for all $\omega \in \Omega''$, i.e.

$$\forall A \in \mathcal{A} \; \forall G \in \mathcal{G} : \; \int_G \mu_{\omega'}^{\mathcal{G}_n}(A) \, d\mu_\omega^{\mathcal{G}}(\omega') = \mu_\omega^{\mathcal{G}}(A \cap G) \,.$$

Hence

$$\forall \omega \in \Omega' \cap \Omega'' \; \forall G \in \mathcal{G} \; \forall A \in \mathcal{E} : \; \mu_\omega^{\mathcal{G}}(A \cap G) = \mu_\omega^{\mathcal{G}}(A) \cdot \mu_\omega^{\mathcal{G}}(G) \,.$$

Again, for fixed G and ω, both sides of the identity are finite measures as functions of A, which coincide on the $\cap$-stable generator $\mathcal{E}$ of $\mathcal{A}$, so that the identity extends to all $A \in \mathcal{A}$, in particular to $A = G$. Therefore, $\mu_\omega^{\mathcal{G}}(G) = \mu_\omega^{\mathcal{G}}(G)^2 \in \{0,1\}$, i.e. $\mathcal{G}$ decomposes μ ergodically.

Remarks on the case of an uncountable family of sub-σ-algebras

Let $(\mathcal{F}_i : i \in I)$ be a nonempty family of sub-σ-algebras decomposing μ ergodically and satisfying

$$\left(\bigcap_{i\in J} \mathcal{F}_i\right)_\mu = \bigcap_{i\in J} (\mathcal{F}_i)_\mu \quad \text{for all finite subsets } J \subseteq I, \tag{3.2.10}$$

By Corollary 3.1.23, there is a countable subset $J \subset I$ such that

$$\bigcap_{i\in I} (\mathcal{F}_i)_\mu = \bigcap_{i\in J} (\mathcal{F}_i)_\mu \tag{3.2.11}$$

and by Lemma 3.1.24, the identity (3.2.10) carries over to this countable set J, so that

$$\bigcap_{i\in I} (\mathcal{F}_i)_\mu = \left(\bigcap_{i\in J} \mathcal{F}_i\right)_\mu , \tag{3.2.12}$$

but there is no way to prove that $\bigcap_{i\in J} \mathcal{F}_i \subseteq \left(\bigcap_{i\in I} \mathcal{F}_i\right)_\mu$, which would be necessary for the further proof. Indeed, Example 3.2.12 shows that this is not true in general: In that case $\left(\bigcap_{i\in I} \mathcal{F}_i\right)_\mu \subseteq \left(\pi_1^{-1}(\text{Borel sets})\right)_\mu$ but $\left(\bigcap_{i\in J} \mathcal{F}_i\right)_\mu$ for countable J is again of the form $(\mathcal{F}_k)_\mu = \mathcal{A}_\mu$ for some $k \in I$.

However, it the $\mathcal{F}_i$ are σ-algebras of invariant sets, such counter-examples may not exist.

3.2.5 The corresponding Hilbert space decomposition

Generalizing the idea of Remark 3.1.36, there is an injective map $\phi : \Omega \to \mathbb{X}$ such that $\phi^{-1}(\sigma(\tau_{\mathbb{X}})) \subseteq \mathcal{I} \subseteq (\phi^{-1}(\sigma(\tau_{\mathbb{X}})))_\mu$ and, writing $\Gamma_x := \phi^{-1}(\{x\})$, we have

i) $\Omega = \bigcup_{x\in\mathbb{X}} \Gamma_x$ (disjointly),

ii) $\forall x \in \mathbb{X} :\ \Gamma_x \in \phi^{-1}(\sigma(\tau_{\mathbb{X}})) \subseteq \mathcal{I}$, and

iii) for μ-a.e. $\omega \in \Omega$ and each $x \in \mathbb{X}$

$$\mu_\omega^{\mathcal{I}}(\Gamma_x) = \delta_\omega(\phi^{-1}(\{x\})) = \begin{cases} 1 & \text{if } \phi(\omega) = x \\ 0 & \text{otherwise .} \end{cases}$$

Therefore the conditional measures $\mu_\omega^{\mathcal{I}}$ can be localized on the sets Γ_x. In particular, there is a set $\Omega_0 \in \mathcal{A}^1$ such that for all $\omega, \omega' \in \Omega_0$ and all $i \in I$ holds: If $x = \phi(\omega) = \phi(\omega')$, then $\mu_x := \mu_{\phi(\omega)}$ is uniquely defined and $T_i|_{\Gamma_x}$ is an ergodic non-singular automorphism of the probability space $(\Gamma_x, \mathcal{A}|_{\Gamma_x}, \mu_x)$. Furthermore, $\mu(A) = \int_{\mathbb{X}} \mu_x(A \cap \Gamma_x)\, d\nu(x)$ for each $A \in \mathcal{A}$ where $\nu = \mu \circ \phi^{-1}$. In the rest of this subsection we transfer this decomposition to the corresponding L^2 spaces.

Let $\mathcal{H} = \int_{\mathbb{X}}^{\oplus} L^2(\Gamma_x, \mathcal{A}|_{\Gamma_x}, \mu_x)\, d\nu(x)$. That means $\mathcal{H}$ is the linear space of all families $(f_x)_{x\in\mathbb{X}}$ with the following properties:

(1) $\forall x \in \mathbb{X} :\ f_x$ is $\mathcal{A}|_{\Gamma_x}$-measurable,

(2) $x \mapsto \|f_x\|_2 := \left(\int_{\Gamma_x} f_x \overline{f_x}\, d\mu_x\right)^{1/2}$ is $\sigma(\tau_{\mathbb{X}})_\nu$-measurable,

(3) $\|(f_x)_{x\in\mathbb{X}}\|^2 := \int_{\mathbb{X}} \|f_x\|_2^2\, d\nu(x) < \infty$.

In view of the polarization identity, property (3) implies for any $(f_x)_{x\in\mathbb{X}}, (\tilde{f}_x)_{x\in\mathbb{X}} \in \mathcal{H}$:

(4) $x \mapsto \langle f_x, \tilde{f}_x\rangle := \int_{\Gamma_x} f_x \overline{\tilde{f}_x}\, d\mu_x$ is $\sigma(\tau_{\mathbb{X}})_\nu$-measurable.

This allows to define a scalar product

$$\langle (f_x)_{x\in\mathbb{X}}, (\tilde{f}_x)_{x\in\mathbb{X}}\rangle := \int_{\mathbb{X}} \langle f_x, \tilde{f}_x\rangle\, d\nu(x)$$

on $\mathcal{H}$. (At this point we do not yet claim that $\mathcal{H}$ is complete, i.e. that $\mathcal{H}$ is a Hilbert space.)

3.2.13 Proposition *$G : L^2(\Omega, \mathcal{A}, \mu) \to \mathcal{H}$, $G(f) := (f|_{\Gamma_x})_{x\in\mathbb{X}}$, is unitary.*

Proof: The family $G(f)$ clearly satisfies (1). As $\omega \mapsto \|f|_{\Gamma_{\phi(\omega)}}\|^2 = \int_{\Gamma_{\phi(\omega)}} f\overline{f}\, d\mu_{\phi(\omega)} = \int_\Omega f\overline{f}\, d\mu_\omega^{\mathcal{I}}$ is $\mathcal{I}$-measurable and as $\mathcal{I} = \phi^{-1}(\sigma(\tau_{\mathbb{X}}))_\mu$, the map $x \mapsto \|f|_{\Gamma_x}\|^2$ is $\sigma(\tau_{\mathbb{X}})_\nu$-measurable, i.e. $G(f)$ satisfies (2). Finally,

$$\|G(f)\|^2 = \int_{\mathbb{X}} \|f|_{\Gamma_x}\|_2^2\, d\nu(x) = \int_{\mathbb{X}} \left(\int_{\Gamma_x} f\overline{f}\, d\mu_x\right) d\nu(x) = \int_\Omega \left(\int_\Omega f\overline{f}\, d\mu_\omega^{\mathcal{I}}\right) d\mu(\omega) = \|f\|_2^2\,,$$

so that also (3) is satisfied and G is well defined and isometric. Obviously, G is linear. Therefore $V := G(L^2(\Omega, \mathcal{A}, \mu))$ is a linear subspace of $\mathcal{H}$. As $L^2(\Omega, \mathcal{A}, \mu)$ is a complete space, so is V. We must show that $V^\perp = \{0\}$.

So let $(f_x)_{x\in\mathbb{X}} \in V^\perp$. As $(\Omega, \mathcal{A}, \mu)$ is a pLR space, there is a $\cap$-stable countable generator $\mathcal{E}$ of $\mathcal{A}$. For $A \in \mathcal{E}$ define $g_A : \mathbb{X} \to \mathbb{R}$,

$$g_A(x) = \int_{A\cap\Gamma_x} f_x\, d\mu_x = \langle f_x, 1_A|_{\Gamma_x}\rangle\,,$$

By property (4), all g_A are measurable. Observing that $\phi(\omega) = \phi(\omega')$ for $\mu^{\mathcal{I}}_{\phi(\omega')}$-a.e. ω, we have

$$\begin{aligned}
0 &= \langle G\left(1_A \cdot (g_A \circ \phi)\right), (f_x)_{x\in\mathbb{X}}\rangle \\
&= \int_{\mathbb{X}} \left(\int_{A\cap\Gamma_x} g_A(\phi(\omega))\, \overline{f_x(\omega)}\, d\mu_x(\omega)\right) d\nu(x) \\
&= \int_\Omega \left(\int_A g_A(\phi(\omega))\, \overline{f_{\phi(\omega')}(\omega)}\, d\mu^{\mathcal{I}}_{\phi(\omega')}(\omega)\right) d\mu(\omega') \\
&= \int_\Omega g_A(\phi(\omega')) \left(\int_A \overline{f_{\phi(\omega')}(\omega)}\, d\mu^{\mathcal{I}}_{\phi(\omega')}(\omega)\right) d\mu(\omega') \\
&= \int_{\mathbb{X}} g_A(x) \left(\int_{A\cap\Gamma_x} \overline{f_x(\omega)}\, d\mu_x(\omega)\right) d\nu(x) \\
&= \int_{\mathbb{X}} g_A(x)\, \overline{g_A(x)}\, d\nu(x)\,.
\end{aligned}$$

It follows that $g_A(x) = 0$ for ν-almost all $x \in \mathbb{X}$. Hence there is a Borel set $Y \subseteq \mathbb{X}$ with $\nu(Y) = 1$ and such that $0 = g_A(x) = \int_{A\cap\Gamma_x} f_x\, d\mu_x$ for all $A \in \mathcal{E}$ and all $x \in Y$. As for fixed x, this expression defines a finite measure as a function of A, and as the family $\mathcal{B}$ is a $\cap$-stable generator of $\sigma(\tau_{\mathbb{X}})$, it follows that $\int_{A\cap\Gamma_x} f_x\, d\mu_x = 0$ for all $A \in \sigma(\tau_{\mathbb{X}})$ and all $x \in Y$ so that $f_x = 0$ μ_x-almost everywhere for all $x \in Y$. But this means that $(f_x)_{x\in\mathbb{X}} = 0$ in $\mathcal{H}$. □

4 The structure of (ergodic) mpds

This chapter deals exclusively with measure preserving dynamical systems $(\Omega, \mathcal{A}, \mu, T)$ where T need not be invertible. We give short glimpses of the following spects:

- System constructions like factors, extensions and products,
- Weak mixing and its various characterizations
- Kolmogorov-Sinai entropy.

4.1 Factors, extensions and products

Let $(\Omega_i, \mathcal{A}_i, \mu_i, T_i)$ $(i = 1, 2)$ be mpds.

4.1.1 Products

- The *product* of these two systems is the mpds $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \otimes \mu_2, T_1 \times T_2)$. $T_1 \times T_2$ obviously preserves the measure of product sets, and as the collection of these sets is a $\cap$-stable generator for the product σ-algebra, the product measure is preserved.
- If both probability spaces are LR, one may pass to the completed σ-algebra $(\mathcal{A}_1 \otimes \mathcal{A}_2)_{\mu_1 \otimes \mu_2}$ to obtain a LR space as product space.
- To simplify the notation we will often just talk about $T_1 \times T_2$ and take the (completed) product probability space for granted.

4.1.2 Factors and extensions

- If there are $\Omega_1' \subseteq \Omega_1$ and $\Omega_2' \subseteq \Omega_2$ of full measure and a measurable map $\phi : \Omega_1' \to \Omega_2'$ such that $\mu_2 = \mu_1 \circ \phi^{-1}$ and $\phi \circ T_1|_{\Omega_1'} = T_2 \circ \phi$, then $(\Omega_2, \mathcal{A}_2, \mu_2, T_2)$ is a *factor* of $(\Omega_1, \mathcal{A}_1, \mu_1, T_1)$. We often just say that T_2 is a factor of T_1 or that T_1 is an *extension* of T_2.
- Dealing with LR spaces we know that $\phi(\Omega_1') \in \mathcal{A}_2$ (Theorem 3.1.16), so that we can choose $\Omega_2' = \phi(\Omega_1')$.
- Both, T_1 and T_2 are factors of $T_1 \times T_2$.

Sometimes, if a mpds is not invertible, it is desirable to have a smallest invertible extension. An example of this situation is a two-sided Bernoulli shift which is clearly an extension of the corresponding one-sided Bernoulli shift (Example 1.1.2).

4.1.3 The natural extension

4.1.1 Theorem *Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds on a pLR or LR space $(\Omega, \mathcal{A}, \mu)$. Up to isomorphism mod 0 there is a unique invertible extension $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mu}, \hat{T})$ of $(\Omega, \mathcal{A}, \mu, T)$ by a pLR or LR space with the property that each invertible extension $(\Omega', \mathcal{A}', \mu', T')$ of $(\Omega, \mathcal{A}, \mu, T)$ is also an extension of $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mu}, \hat{T})$. More precisely: Each factor map from $(\Omega', \mathcal{A}', \mu', T')$ to $(\Omega, \mathcal{A}, \mu, T)$ factorizes over $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mu}, \hat{T})$. $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mu}, \hat{T})$ is called the* natural extension *of $(\Omega, \mathcal{A}, \mu, T)$. It is explicitly constructed in the proof.*

Attention: *For the present proof I need an additional assumption on the existence of pre-images that is discussed below!*

Proof: In view of Theorem 3.1.12, $T^n(\Omega) \in \mathcal{A}_\mu$ for all $n \in \mathbb{N}$. As $\mu(T^n(\Omega)) = \mu(T^{-n}(T^n(\Omega))) \geqslant \mu(\Omega) = 1$, we have $\Omega_0 := \bigcap_{n\in\mathbb{N}} T^n(\Omega) \in \mathcal{A}_\mu^1$. By definition, $T(\Omega_0) \subseteq \Omega_0$, but I think it may happen that $T(\Omega_0) \neq \Omega_0$. Of course one can repeat this construction with Ω_0 instead of Ω etc., but I do not see that countably many repetitions lead to a surjective map. So I will make the additional assumption that $T(\Omega_0) = \Omega_0$. As the theorem only makes assertions about isomorphisms mod 0, we can then assume w.l.o.g. that $T(\Omega) = \Omega$.

Consider $\hat{\Omega} := \{\hat{\omega} = (\omega_0, \omega_1, \dots) \in \Omega^{\mathbb{N}} : T(\omega_{i+1}) = \omega_i \ \forall i \in \mathbb{N}\}$ equipped with the trace of the product σ-algebra $\mathcal{A}^{\otimes\mathbb{N}}$. Define

$$\hat{T} : \hat{\Omega} \to \hat{\Omega}, \quad (\omega_0, \omega_1, \omega_2, \dots) \mapsto (T(\omega_0), \omega_0, \omega_1, \dots) \ .$$

$\hat{T}$ is obviously invertible and bi-measurable. Denote

$$\pi_i : \hat{\Omega} \to \Omega,\ \hat{\omega} \mapsto \omega_i \ ,$$
$$\mathcal{F}_n := \sigma(\pi_0, \dots, \pi_n) \ .$$

Let $\mathcal{C}_n$ be the family of all sets

$$[A_0, \dots, A_n] := \{\hat{\omega} \in \hat{\Omega} : \omega_i \in A_i \ (i = 0, \dots, n)\} = \left\{\hat{\omega} \in \hat{\Omega} :\ \omega_n \in \bigcap_{i=0}^n T^{-i}(A_{n-i})\right\} = \pi_n^{-1}\left(\bigcap_{i=0}^n T^{-i}(A_{n-i})\right)$$

with $A_i \in \mathcal{A}$. It is for the last identity that we need $T(\Omega) = \Omega$, because otherwise we can extend a point $\omega_n \in \bigcap_{i=0}^n T^{-i}(A_{n-i})$ to a potential starting segment $(\omega_0, \dots, \omega_n)$ of a point $\hat{\omega} \in \hat{\Omega}$, but we cannot guarantee that the segment can indeed be extended to such a point.

By definition, $\mathcal{C}_n \subseteq \sigma(\pi_n)$, and as $\pi_n^{-1}(A) = [\Omega, \dots, \Omega, A]$ for each $A \in \mathcal{A}$, we have indeed that $\mathcal{C}_n = \sigma(\pi_n)$. So we can define a probability measure $\hat{\mu}_n$ on $\mathcal{C}_n$ by

$$\hat{\mu}_n([A_0, \dots, A_n]) = \hat{\mu}_n\left(\pi_n^{-1}\left(\bigcap_{i=0}^n T^{-i}(A_{n-i})\right)\right) := \mu\left(\bigcap_{i=0}^n T^{-i}(A_{n-i})\right) \ .$$

Obviously $\mathcal{C}_n \subseteq \mathcal{C}_{n+1}$, and we will show that $(\hat{\mu}_n)_{n\in\mathbb{N}}$ is a projective family of probability measures: If $[A_0, \dots, A_n] = [B_0, \dots, B_m]$ and $m < n$, we extend the sequence $B_0, \dots, B_m$ to a sequence $B_0, \dots, B_m, \Omega, \dots, \Omega$ of length $n + 1$ and note that

$$\bigcap_{i=0}^m T^{-i}(B_{m-i}) = \bigcap_{i=0}^{n-m-1} T^{-i}(\Omega) \cap \bigcap_{i=n-m}^n T^{-i}(B_{n-i})$$

to see that $\hat{\mu}_n([A_0,\dots,A_n]) = \hat{\mu}_m([B_0,\dots,B_m])$.

Now suppose first that $\Omega = [0,1]$ and that $\mathcal{A}$ is the Borel-σ-algebra on $\mathbb{R}$. Then Kolmogorov's extension theorem (a minor variant of [10, Satz 10.3]) guarantees the existence of a probability $\hat{\mu}$ on $\hat{\Omega}$ such that $\hat{\mu}([A_0,\dots,A_n]) = \hat{\mu}_n([A_0,\dots,A_n])$ for all n and all sets from $\mathcal{C}_n$. In particular, $\hat{\mu}\circ\pi_0^{-1} = \hat{\mu}_0\circ\pi_0^{-1} = \mu$ and

$$\begin{aligned}
\hat{\mu}\circ\hat{T}^{-1}([A_0,\dots,A_n]) &= \hat{\mu}_{n-1}([A_1\cap T^{-1}A_0, A_2, A_3,\dots,A_n]) \\
&= \mu\left(T^{-(n-1)}(A_1\cap T^{-1}(A_0))\cap\bigcap_{i=0}^{n-2}T^{-i}(A_{n-i})\right) \\
&= \mu\left(\bigcap_{i=0}^{n}T^{-i}(A_{n-i})\right) = \hat{\mu}_n([A_0,\dots,A_n]) \\
&= \hat{\mu}([A_0,\dots,A_n])
\end{aligned}$$

for all sets in $\bigcup_n \mathcal{C}_n$. Hence $\hat{\mu}\circ\hat{T}^{-1} = \hat{\mu}$.

If $(\Omega,\mathcal{A},\mu)$ is a non-atomic pLR or LP space, then it is isomorphic mod 0 to $[0,1]$ equipped with Lebesgue measure λ. Denote the isomorphism by $h:\Omega_0\to[0,1]$. Then $\tau := h\circ T\circ h^{-1}$ is defined Lebesgue-a.e. on $[0,1]$ and $\lambda\circ\tau^{-1} = \lambda\circ h\circ T^{-1}\circ h^{-1} = \mu\circ T^{-1}\circ h^{-1} = \mu\circ h^{-1} = \lambda$. As $\tau(h(\omega_{i+1})) = h(T(\omega_{i+1})) = h(\omega_i)$, h extends to $\hat{h}:\widehat{\Omega_0}\to\widehat{[0,1]}$, $\hat{\omega}\mapsto(h(\omega_n))_{n\in\mathbb{N}}$ and

$$\hat{\tau}(\hat{h}(\hat{\omega})) = (\tau(h(\omega_0)), h(\omega_0), h(\omega_1),\dots) = (h(T(\omega_0)), h(\omega_0), h(\omega_1),\dots) = \hat{h}(\hat{T}(\hat{\omega}))\,.$$

Let $\hat{\mu} = \hat{\lambda}\circ\hat{h}$. Then $\hat{h}$ is an isomorphism mod 0 and

$$\hat{\mu}\circ\hat{T}^{-1} = \hat{\lambda}\circ(\hat{T}\circ\hat{h}^{-1})^{-1} = \hat{\lambda}\circ(\hat{h}^{-1}\circ\hat{\tau})^{-1} = \hat{\lambda}\circ\hat{\tau}^{-1}\circ\hat{h} = \hat{\lambda}\circ\hat{h} = \hat{\mu}\,.$$

It remains to show that each invertible extension of $(\Omega,\mathcal{A},\mu,T)$ factorizes over $(\hat{\Omega},\hat{\mathcal{A}},\hat{\mu},\hat{T})$. So let

$$\phi:(\Omega',\mathcal{A}',\mu',T')\to(\Omega,\mathcal{A},\mu,T)$$

be such an invertible extension. Define $\hat{\phi}:\Omega'\to\hat{\Omega}$, $\hat{\phi}(\omega') = (\phi(\omega'), \phi(T'^{-1}(\omega')), \phi(T'^{-2}(\omega')),\dots)$. Then

$$\begin{aligned}
(\mu'\circ\hat{\phi}^{-1})([A_0,\dots,A_n]) &= (\mu'\circ\hat{\phi}^{-1})\left(\pi_n^{-1}\left(\bigcap_{i=0}^{n}T^{-i}(A_{n-i})\right)\right) \\
&= (\mu'\circ T'^{-n})\left(\phi^{-1}\left(\bigcap_{i=0}^{n}T^{-i}(A_{n-i})\right)\right) \\
&= \mu\left(\bigcap_{i=0}^{n}T^{-i}(A_{n-i})\right) \\
&= \mu([A_0,\dots,A_n])
\end{aligned}$$

so that $\mu'\circ\hat{\phi}^{-1} = \mu$, and obviously $\pi_0\circ\hat{\phi} = \phi$. □

4.1.2 Remark In Example 1.1.2 we introduced the left shifts $T_{\mathbb{N}}$ on $\{0,1\}^{\mathbb{N}}$ and $T_{\mathbb{Z}}$ on $\{0,1\}^{\mathbb{Z}}$, equipped with invariant the p-Bernoulli measures $\mu_{p,\mathbb{N}}$ and $\mu_{p,\mathbb{Z}}$, respectively. One can show that $(\{0,1\}^{\mathbb{Z}},\mu_{p,\mathbb{Z}},T_{\mathbb{Z}})$ is isomorphic mod 0 to the natural extension of $(\{0,1\}^{\mathbb{N}},\mu_{p,\mathbb{N}},T_{\mathbb{N}})$, both equipped with their product σ-algebras.

4.1.4 Exercises

1. Prove Remark 4.1.2.
2. Prove that a factor of an ergodic mpds is ergodic.
3. Prove that a mpds is ergodic if and only if its natural extension is.

4.2 Weak mixing

4.2.1 Weak mixing and its various characterizations

Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds and denote $L_0^2 := \{f \in L^2(\Omega, \mathcal{A}, \mu) : \langle f, 1 \rangle = 0\}$ and $Uf := f \circ T$. The following lemma is a simple consequence of the L^1-convergence in Birkhoff's ergodic theorem:

4.2.1 Lemma $(\Omega, \mathcal{A}, \mu, T)$ *is ergodic if and only if*

$$\forall A, B \in \mathcal{A} : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(A \cap T^{-k}(B)) = \mu(A) \cdot \mu(B) .$$

An obviously stronger property than ergodicity is the following one, called *mixing*:

$$\forall A, B \in \mathcal{A} : \lim_{n\to\infty} \mu(A \cap T^{-n}(B)) = \mu(A) \cdot \mu(B) .$$

More important than this is the property defined next - situated between ergodicity and mixing:

4.2.2 Definition *If* $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(A \cap T^{-k}(B)) - \mu(A)\mu(B)| = 0$ *for all* $A, B \in \mathcal{A}$*, the system* $(\Omega, \mathcal{A}, \mu, T)$ *is called weakly mixing.*

The following theorem provides various characterization of weak mixing. Proofs of its different parts can be found in the text-books [5, 14, 17]. I will follow mostly [5].

4.2.3 Theorem *The following properties are equivalent:*

(i) $(\Omega, \mathcal{A}, \mu, T)$ *is weakly mixing.*

(ii) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (\mu(A \cap T^{-k}(B)) - \mu(A)\mu(B))^2 = 0$ *for all* $A, B \in \mathcal{A}$.

(iii) Given $A, B \in \mathcal{A}$*, there is a subset* $L \subset \mathbb{N}$ *of asymptotic density* 0 *such that* $\lim_{n\to\infty, n\notin L} \mu(A \cap T^{-n}(B)) = \mu(A)\mu(B)$*, i.e.* $\lim_{n\to\infty, n\notin L} |\mu(A \cap T^{-n}(B)) - \mu(A)\mu(B)| = 0$

(iv) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle U^k f, g \rangle - \langle f, 1 \rangle \cdot \langle 1, g \rangle| = 0$ *for all* $f, g \in L^2$.

(v) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle U^k f, f \rangle| = 0$ *for all* $f \in L_0^2$.

(vi) If $Uf = \lambda f$*, then* $\lambda = 1$ *and* $f = const.$

(vii) $T \times T$ *is weakly mixing.*

(viii) $T \times S$ *is ergodic on* $\Omega \times Y$ *for each ergodic mpds* $(Y, \mathcal{B}, \nu, S)$.

(ix) $T \times T$ *is ergodic.*

(x) The system is ergodic and has no rational or irrational rotation as a factor, except for the identity.

4.2.4 Remark It suffices to check property (i) or (iii) for all A, B from a $\cap$-stable generator of the σ-algebra $\mathcal{A}$. The proof proceeds via the two-fold application of a Dynkin system argument. (The same holds for the property characterizing ergodicity in Lemma 4.2.1.)

For the proof of Theorem 4.2.3 we need the following lemma about sequences of real numbers:

4.2.5 Lemma *Let $(a_n)_{n\in\mathbb{N}}$ be a bounded sequence of non-negative real numbers. The following are equivalent:*

(i) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k = 0$.

(ii) There is a subset $L \subset \mathbb{N}$ of asymptotic density 0 such that $\lim_{n\to\infty, n\notin L} a_n = 0$.

The elementary *proof* can be found e.g. in [17]. The implication (ii) $\Rightarrow$ (i) is more or less trivial, and for the reverse implication one has to note that non-negativity and boundedness of a sequence satisfying (i) imply that for each $\epsilon > 0$ the fraction of the number of elements $a_n > \epsilon$ tends to zero.

Proof of Theorem 4.2.3: We will prove the following equivalences separately:

1. (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (vi) $\Rightarrow$ (ix) $\Rightarrow$ (ii)
2. (iii) $\Rightarrow$ (vii) $\Rightarrow$ (ix)
3. (i) $\Rightarrow$ (viii) $\Rightarrow$ (ix)
4. (vi) $\Leftrightarrow$ (x)

Most of these implications are trivial or elementary to prove. (vi) $\Rightarrow$ (ix), however, uses a simple consequence of the spectral theory of compact self-adjoint operators.

1. The *equivalence of (i), (iii) and (ii)* follows immediately from Lemma 4.2.5.
(iv) $\Rightarrow$ (i) is obvious.
(i) $\Rightarrow$ (iv): For $f \in L^2$ denote by R_f the set of all $g \in L^2$ that satisfy (iv), and for $g \in L^2$ denote by S_g the set of all $f \in L^2$ that satisfy (iv). All R_f and S_g are closed linear subspaces of L^2. Because of (i), each R_{1_A}, $A \in \mathcal{A}$, contains the space of all measurable indicator functions 1_B. Hence $R_{1_A} = L^2$. This means that for each $g \in L^2$ the space S_g contains all measurable indicator functions 1_A, so $S_g = L^2$.
(iv) $\Rightarrow$ (v) is trivial.
(v) $\Rightarrow$ (vi): Let $Uf = \lambda f$, $\|f\|_2 > 0$. Then $\|f\|_2 = \|Uf\|_2 = |\lambda|\,\|f\|_2$, so that $|\lambda| = 1$. Similarly, $\int f\,d\mu = \int Uf\,d\mu = \lambda \int f\,d\mu$, whence $\lambda = 1$ or $f \in L_0^2$. In both cases, $f_0 := f - \langle f, 1\rangle$ is an eigenfunction for the eigenvalue λ. Now (v) implies that $\|f_0\|_2 = 0$, i.e. $f = \langle f, 1\rangle = \text{const} \notin L_0^2$, so that finally $\lambda = 1$.
(vi) $\Rightarrow$ (ix): All proofs of this implication that are known to me use some variant of spectral theory for Hilbert space operators. The traditional proof uses knwoledge of the spectral measure for unitary operators, see e.g. [14]. Here I present a proof from [5, pp.62] that uses the fact that kernel operators are non-trivial self-adjoint compact operators on L^2 and thus have at least one eigenvalue λ with a finite-dimensional eigenspace V_λ, see [6] for a precise and short account.

The proof proceeds by contradiction: Suppose that $T \times T$ is not ergodic. Then there is a non-constant function $f \in L^2_{\mathbb{C}}(\Omega \times \Omega, \mathcal{A} \otimes \mathcal{A}, \mu \otimes \mu)$ that is almost invariant under $T \times T$. Let

$$f_1(x,y) := f(x,y) + \overline{f(y,x)}\,, \quad f_2(x,y) := i\left(f(x,y) - \overline{f(y,x)}\right).$$

Both functions are almost invariant under $T \times T$ and have the additional symmetry $f_i(x,y) = \overline{f_i(y,x)}$. As $f_1 - if_2 = 2f$ and as f is non-constant, at least one of f_1 and f_2 is non-constant. Hence we can assume w.l.o.g. that $f(x,y) = \overline{f(y,x)}$. We may further assume (by subtracting $\int f\,d\mu^2$) that $\int f\,d\mu^2 = 0$. Define the linear operator F on $L^2_{\mathbb{C}}(\Omega, \mathcal{A}, \mu)$ by

$$(F(g))(x) := \int_\Omega f(x,y)g(y)\,d\mu(y)\ .$$

It is a non-trivial self-adjoint compact operator [6, Theorem 2.0.1] which, as such, has at least one eigenvalue λ with a finite-dimensional space V_λ of eigenfunctions [6, Theorem 1.0.2]. We prove that $U_T(V_\lambda) \subseteq V_\lambda$: assume that $F(g) = \lambda g$. Then

$$\begin{aligned} F(U_T g)(x) &= \int_\Omega f(x,y)(U_T g)(y)\,d\mu(y) \\ &= \int_\Omega f(Tx,Ty)g(Ty)\,d\mu(y) \quad (\text{ as } f \text{ is } T\times T\text{-invariant}) \\ &= \int_\Omega f(Tx,y)g(y)\,d\mu(y) \quad (\text{ as } \mu\circ T^{-1} = \mu) \\ &= (F(g))(Tx) = U_T(\lambda g)(x) = \lambda(U_T g)(x)\ . \end{aligned}$$

It follows that $U_T|_{V_\lambda}$ is a non-trivial linear map of finite-dimensional vector space and therefore has a non-trivial eigenfunction $g \in L^2_{\mathbb{C}}(\Omega, \mathcal{A}, \mu)$. By assumption (vi) g is constant, $g = \int_\Omega g\,d\mu$. But then $\lambda \int_\Omega g\,d\mu = \int_\Omega F(g)\,d\mu = \int_\Omega g\,d\mu \cdot \int_{\Omega^2} f\,d\mu^2 = 0$ so that $g = \int_\Omega g\,d\mu = 0$, a contradiction.
(ix) ⇒ (ii): If $A, B \in \mathcal{A}$, then

$$\begin{aligned} &\frac{1}{n}\sum_{k=0}^{n-1}(\mu(A\cap T^{-k}(B)) - \mu(A)\mu(B))^2 \\ &= \frac{1}{n}\sum_{k=0}^{n-1}\mu(A\cap T^{-k}(B))^2 - 2\mu(A)\mu(B)\frac{1}{n}\sum_{k=0}^{n-1}\mu(A\cap T^{-k}(B)) + (\mu(A)\mu(B))^2 \\ &= \frac{1}{n}\sum_{k=0}^{n-1}\mu\otimes\mu((A\times A)\cap(T\times T)^{-k}(B\times B)) \\ &\quad - 2\mu(A)\mu(B)\frac{1}{n}\sum_{k=0}^{n-1}\mu\otimes\mu((A\times\Omega)\cap(T\times T)^{-k}(B\times\Omega)) + (\mu(A)\mu(B))^2 \end{aligned}$$

Since $T\times T$ is ergodic by assumption, Lemma 4.2.1 implies that, as $n\to\infty$ this tends to

$$\begin{aligned} &\mu\otimes\mu(A\times A)\cdot\mu\otimes\mu(B\times B) - 2\mu(A)\mu(B)\cdot\mu\otimes\mu(A\times\Omega)\cdot\mu\otimes\mu(B\times\Omega) + (\mu(A)\mu(B))^2 \\ &= (\mu(A)\mu(B))^2 - 2(\mu(A)\mu(B))^2 + (\mu(A)\mu(B))^2 = 0\ . \end{aligned}$$

2. *(iii) ⇒ (vii)*: Let $A_1, A_2, B_1, B_2 \in \mathcal{A}$. By (iii) there are sets $L_1, L_2 \subset \mathbb{N}$ of asymptotic density zero such that

$$\lim_{n\to\infty, n\notin L_i}\mu(A_i\cap T^{-n}(B_i)) = \mu(A_i)\mu(B_i) \quad \text{for } i = 1,2\ .$$

Let $L = L_1 \cup L_2$. then L has asymptotic density zero, too, and

$$\begin{aligned} &\lim_{n\to\infty, n\notin L}\mu\otimes\mu((A_1\times A_2)\cap(T\times T)^{-n}(B_1\times B_2)) \\ &= \lim_{n\to\infty, n\notin L}\big(\mu(A_1\cap T^{-n}(B_1))\cdot\mu(A_2\cap T^{-n}(B_2))\big) \\ &= \mu(A_1)\mu(B_1)\cdot\mu(A_2)\mu(B_2) = \mu\otimes\mu(A_1\times B_1)\cdot\mu\otimes\mu(A_2\times B_2)\ . \end{aligned}$$

In view of Remark 4.2.4 this proves property (iii) for $T \times T$ which, as already proved in part 1, is equivalent to weak mixing.
(vii) ⇒ (ix) is trivial in view of Lemma 4.2.1.

3. *(i) ⇒ (viii)*: Let $(Y, \mathcal{B}, \nu, S)$ be an ergodic mpds. In order to prove ergodicity of $T \times S$ the following is sufficient in view of Remark 4.2.4: for $A, B \in \mathcal{A}$ and $C, D \in \mathcal{B}$,

$$\begin{aligned}
&\frac{1}{n}\sum_{k=0}^{n-1} \mu \otimes \nu \left((A \times C) \cap (T \times S)^{-k}(B \times D)\right) \\
&= \frac{1}{n}\sum_{k=0}^{n-1} \mu(A \cap T^{-k}(B)) \cdot \nu(C \cap S^{-k}(D)) \\
&= \frac{1}{n}\sum_{k=0}^{n-1} \left(\mu(A \cap T^{-k}(B)) - \mu(A)\mu(B)\right) \cdot \nu(C \cap S^{-k}(D)) + \mu(A)\mu(B)\frac{1}{n}\sum_{k=0}^{n-1} \nu(C \cap S^{-k}(D)) \\
&\to \mu(A)\mu(B) \cdot \nu(C)\nu(D) = \mu \otimes \nu(A \times C) \cdot \mu \otimes \nu(B \times D) \quad \text{as } n \to \infty ,
\end{aligned}$$

because the first sum is bounded by $\frac{1}{n}\sum_{k=0}^{n-1} \left|\mu(A \cap T^{-k}(B)) - \mu(A)\mu(B)\right|$ which tends to zero as T is weakly mixing.
(viii) ⇒ (ix): Let $(Y, \mathcal{B}, \nu, S)$ be the trivial one-point mpds which is of course ergodic. Then the systems $(\Omega, \mathcal{A}, \mu, T)$ and $(\Omega \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu, T \times S)$ are trivially isomorphic, so that $(\Omega, \mathcal{A}, \mu, T)$ is ergodic by (viii). Hence $T \times T$ is ergodic by (viii).

4. *(vi) ⇒ (x)*: The ergodicity of T follows at once by applying (vi) to $f = 1_A$, $A \in \mathcal{A}$. Suppose now that $R_\alpha : \mathbb{S}^1 \to \mathbb{S}^1$ is a factor of T, $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$. Then there is a measurable $f : \Omega \to \mathbb{S}^1$ such that $f \circ T = R_\alpha \circ f$, that means $Uf(\omega) = f(T(\omega)) = e^{i\alpha} f(\omega)$, so that $e^{i\alpha} = 1$, and the factor is the identity system. As a factor of an ergodic system it is ergodic, so it is a trivial one point system.
(x) ⇒ (vi): Suppose $Uf = \lambda f$ for some $f \in L^2$ with $\|f\|_2 > 0$. Then $|\lambda| = 1$ as in (v) ⇒ (vi), and $|f| \circ T = |Uf| = |f|$. As T is ergodic, Birkhoff's ergodic theorem implies $|f| = \int |f| \, d\mu > 0$. Normalizing f we can assume that $|f| = 1$, i.e. $f : \Omega \to \mathbb{S}^1$. Let $\lambda = e^{i\alpha}$. Then $f \circ T(\omega) = \lambda f(\omega) = R_\alpha \circ f(\omega)$ for μ-a.e. ω, i.e. $R_\alpha : \mathbb{S}^1 \to \mathbb{S}^1$ is a factor of T. It follows from (x) that $\lambda = e^{i\alpha} = 1$. Hence $f \circ T = Uf = f$ and, as before for $|f|$, the ergodicity of T implies that $f = \int f \, d\mu$ a.e., i.e. f is constant in L^2. □

4.2.2 The Kronecker factor

The equivalence of (vi) and (x) in Theorem 4.2.3 shows the close connection between eigenfunctions and rotation factors for mpds. Here we extend this point of view. Our general assumption is that $(\Omega, \mathcal{A}, \mu)$ is a LR-space.

Let $(\Omega, \mathcal{A}, \mu, T)$ be an ergodic mpds. Then:

- If $U_T f = \lambda f$, then $|\lambda| = 1$ as seen in the proof of (v) ⇒ (vi) of Theorem 4.2.3. Hence $U_T|f| = |U_T f| = |f|$ so that $|f|$ is constant, and normalizing it we can assume that $|f| = 1$.

- If $U_T f_i = \lambda_i f_i$ $(i = 1, 2)$, then $U_T(f_1 f_2) = U_T f_1 \cdot U_T f_2 = \lambda_1\lambda_2 \cdot f_1 f_2$, so that the eigenvalues of U_T form a multiplicative subgroup of $\{|z| = 1\}$, and the eigenfunctions form a multiplicative subgroup of L^2. Furthermore, $\lambda_1\overline{\lambda_2}\langle f_1, f_2\rangle = \langle U_T f_1, U_T f_2\rangle = \langle f_1, f_2\rangle$, i.e. eigenfunctions to distinct eigenvalues are orthogonal. Denote by E_T the subspace of L^2 generated by all eigenfunctions of U_T. It has an orthonormal basis of eigenfunctions.

- If $U_T f = \lambda f$, then $\sigma(f)$ is a T-invariant σ-algebra: for each measurable $A \subseteq \mathbb{C}$ we have $T^{-1}\{f \in A\} = \{U_T f \in A\} = \{f \in \lambda^{-1}A\}$. The same holds for $\mathcal{K}_T := \sigma(E_T)$.

- As $(\Omega, \mathcal{A}, \mu)$ is a LR-space, there is a countable $\bmod \mu$-generator $\mathcal{B}$ of $\mathcal{K}_T$. Then also $\tilde{\mathcal{B}} := \{T^{-k}B : B \in \mathcal{B}, k \in \mathbb{N}\}$ is a countable $\bmod \mu$-generator of $\mathcal{K}_T$ and $T^{-1}(\tilde{\mathcal{B}}) \subseteq \tilde{\mathcal{B}}$. Let $\tilde{B} = \{B_0, B_1, B_2, \dots\}$ and consider the asociated coding map $\tilde{\phi} : \Omega \to \mathbb{X}$. As each B_n is the $\tilde{\phi}$-preimage of a cylinder set in $\mathbb{X}$, we have $\tilde{\mathcal{B}} \subseteq \tilde{\phi}^{-1}(\sigma(\tau_{\mathbb{X}}))) \subseteq \sigma(\tilde{\mathcal{B}})$ and hence $\sigma(\tilde{\mathcal{B}}) = \tilde{\phi}^{-1}(\sigma(\tau_{\mathbb{X}}))$ and $\mathcal{K}_T = \tilde{\phi}^{-1}(\sigma(\tau_{\mathbb{X}})) \bmod \mu$. This means that $\tilde{\phi}^{-1}$ describes a 1-1 correspondence between equivalence classes of measurable sets, and standard arguments yield that for every $f \in E = L^2_{\mathbb{C}}(\Omega, \mathcal{K}_T, \mu)$ there is a unique $\tilde{f} \in L^2_{\mathbb{C}}(\mathbb{X}, \sigma(\tau_{\mathbb{X}}), \tilde{\mu})$, $\tilde{\mu} := \mu \circ \tilde{\phi}^{-1}$, such that $\tilde{f} \circ \tilde{\phi} = f$ μ-a.e.

- As $T^{-1}(\tilde{\mathcal{B}}) \subseteq \tilde{\mathcal{B}}$, there is a map $\kappa : \mathbb{N} \to \mathbb{N}$ such that $T^{-1}(B_n) = B_{\kappa(n)}$ for all $n \in \mathbb{N}$. define $S : \mathbb{X} \to \mathbb{X}$, $(S(x))_n := x_{\kappa(n)}$. Then S is Borel-measurable,

$$(S(\tilde{\phi}(\omega)))_n = (\tilde{\phi}(\omega))_{\kappa(n)} = 1_{B_{\kappa(n)}}(\omega) = 1_{B_n}(T(\omega)) = \tilde{\phi}(T(\omega)) ,$$

  and S preserves the measure $\tilde{\mu}$.

- If $U_T f = \lambda f$ for $f \in L^2_{\mathbb{C}}(\Omega, \mathcal{A}, \mu)$, then

$$U_S \tilde{f} \circ \tilde{\phi} = f \circ S \circ \tilde{\phi} = \tilde{f} \circ \tilde{\phi} \circ T = U_T f = \lambda f = \lambda \tilde{f} \circ \tilde{\phi}$$

  so that $\tilde{\mu}\{U_S \tilde{f} \neq \lambda \tilde{f}\} = \mu\{U_S \tilde{f} \circ \tilde{\phi} \neq \lambda \tilde{f} \circ \tilde{\phi}\} = 0$, i.e. $\tilde{f}$ is a $L^2_{\mathbb{C}}(\mathbb{X}, \sigma(\tau_{\mathbb{X}}), \tilde{\mu})$-eigenfunction of U_S.

  On the other hand, if $U_S \tilde{f} = \lambda \tilde{f}$ for some $\tilde{f} \in L^2_{\mathbb{C}}(\mathbb{X}, \sigma(\tau_{\mathbb{X}}), \tilde{\mu})$, then $U_T(\tilde{f} \circ \tilde{\phi}) = \tilde{f} \circ \tilde{\phi} \circ T = U_S \tilde{f} \circ \tilde{\phi} = \lambda(\tilde{f} \circ \tilde{\phi})$.

- Hence $\tilde{\phi}^{-1}(\mathcal{K}_S) = \mathcal{K}_T = \tilde{\phi}^{-1}(\sigma(\tau_{\mathbb{X}})) \bmod \mu$ so that $\mathcal{K}_S = \sigma(\tau_{\mathbb{X}}) \bmod \tilde{\mu}$. The factor system $(\mathbb{X}, \sigma(\tau_{\mathbb{X}}), \tilde{\mu}, S))$ is called the *Kronecker factor* of $(\Omega, \mathcal{A}, \mu, T)$. One can show that it is isomorphic (as a mpds) to a rotation on a compact group equipped with Haar measure.

- Restricting a system to $\mathcal{K}_T$ corresponds to the restriction of the linear operator U_T to the subspace E_T generated by its *geometric* eigenfunctions. In the latter case a better picture of the dynamics associated with the discrete spectrum of the operator is obtained if one restricts the operator to the space generated by its *algebraic* eigenfunctions. In finite-dimensional spaces this is the complete picture described by the Jordan normal form. In infinite-dimensional spaces the situation is more complicated because one may encounter “infinite Jordan blocks”. On the dynamical systems side this situation is described by an invariant sub-σ-algebra containing $\mathcal{K}_T$, which gives rise to another factor called the *maximal distal factor*. In a sense that can be made precise one can say that each mpds is weakly mixing relative to its maximal distal factor.

4.2.3 Exercises

1. Prove that a factor of a weakly mixing mpds is weakly mixing.

2. Prove that a mpds is weakly mixing if and only if its natural extension is.

3. Give an example of an ergodic mpds $(\Omega, \mathcal{A}, \mu, T)$ for which $T \times T$ is not ergodic (and hence T is not weakly mixing).

4. Let $(\mathbb{T}^d, \mathcal{A}, \mu)$ be as in Example 1.1.1, i.e. μ is Lebesgue measure on $\mathbb{T}^d$. Give at least two different proofs based on Theorem 4.2.3 for each of the following assertions:
   a) Let $T(x) = x + a$ for some $a \in \mathbb{R}^d$. Then T is not weakly mixing.
   b) Let $T(x) = Ax$ for some $d \times d$ integer matrix A with $\det(A) \neq 0$. Then T is weakly mixing if and only if it is ergodic.

   Hint: A look at Theorem 1.1.11 may help to understand the situation.

5. Prove that the shift dynamical system $(\{0,1\}^{\mathbb{Z}}, \mathcal{A}, \mu_p, T)$ from Example 1.1.2a is weakly mixing.

6. Prove Remark 4.2.4.

7. [5, See also Exercise 2.7.2] Show that if a mpds $(\Omega, \mathcal{A}, \mu, T)$ has the property that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \sup\left\{ |\mu(A \cap T^{-k}(B)) - \mu(A)\mu(B)| : A, B \in \mathcal{A} \right\} = 0 ,$$

   then the system is trivial in the sense that $\mu(A) \in \{0,1\}$ for all $A \in \mathcal{A}$.

4.3 Entropy

For this section I refer to one of the many textbooks that treat the bascis of entropy theory for mpds, e.g. [14, 17, 16] and to the monograph [12] that is totally devoted to entropy as a concept from probability and information theory. Basics on entropy in a purely probabilistic, non-dynamical context can be found in [9].

In this course I will follow the very compact presentation of the main material in [16, Chapter 4] and refer for some advanced classical results to [12].

4.3.1 Information content and entropy

Motivational material, definitions and notations from [16, Section 4.1]. In particular: The word partition always means a finite or countable decomposition of Ω into measurable sets. For all our results sets of measure zero can be neglected in any respect.

4.3.2 The entropy of a partition

More notation and basic results from [16, Section 4.2], in particular:

$$H_\mu(\alpha \mid \mathcal{F}) = \int_\Omega I_\mu(\alpha \mid \mathcal{F})\, d\mu = -\sum_{A\in\alpha} \int_\Omega \mu(A \mid \mathcal{F}) \log \mu(A \mid \mathcal{F})\, d\mu = \sum_{A\in\alpha} \int_\Omega \varphi(\mu(A \mid \mathcal{F}))\, d\mu$$

where $\varphi(t) := -t \log(t)$.

4.3.1 Theorem *Suppose that α an d β are partitions with $H_\mu(\alpha), H_\mu(\beta) < \infty$ and $\mathcal{F}$ is a sub-σ-algebra of $\mathcal{A}$. Then*
1. $I_\mu(\alpha \vee \beta \mid \mathcal{F}) = I_\mu(\alpha \mid \mathcal{F}) + I_\mu(\beta \mid \mathcal{F} \vee \alpha)$.
2. $H_\mu(\alpha \vee \beta \mid \mathcal{F}) = H_\mu(\alpha \mid \mathcal{F}) + H_\mu(\beta \mid \mathcal{F} \vee \alpha)$.
3. $H_\mu(\alpha \vee \beta) = H_\mu(\alpha) + H_\mu(\beta \mid \alpha)$.

4.3.2 Proposition *Let α, β be partitions with finite entropy and $\mathcal{F}, \mathcal{G}$ be sub-σ-algebras of $\mathcal{A}$. Then*

1. *If $\mathcal{F} \subseteq \mathcal{G}$, then $H_\mu(\alpha \mid \mathcal{F}) \geqslant H_\mu(\alpha \mid \mathcal{G})$.*
2. *If $\alpha \leqslant \beta$, then $H_\mu(\alpha \mid \mathcal{F}) \leqslant H_\mu(\beta \mid \mathcal{F})$.*
3. *$H_\mu(\alpha \vee \beta \mid \mathcal{F}) \leqslant H_\mu(\alpha \mid \mathcal{F}) + H_\mu(\beta \mid \mathcal{F})$.*
4. *$\alpha \subseteq \mathcal{F} \mod \mu$ if and only if $H_\mu(\alpha \mid \mathcal{F}) = 0$.*

This is slightly more general than Proposition 4.2 in [16], but see [8, Theorem 3.1.8].

4.3.3 Theorem *Let α be a partition with $H_\mu(\alpha) < \infty$ and let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a family of sub-σ-algebras. Then*

1. *The family $(I_\mu(\alpha \mid \mathcal{F}_n))_{n\in\mathbb{N}}$ is uniformly integrable.*
2. *if $\mathcal{F}_1 \subseteq \mathcal{F}_2 \cdots \nearrow \mathcal{F}_\infty$ or $\mathcal{F}_1 \supseteq \mathcal{F}_2 \cdots \searrow \mathcal{F}_\infty$, then $\lim_{n\to\infty} I_\mu(\alpha \mid \mathcal{F}_n) = I_\mu(\alpha \mid \mathcal{F}_\infty)$ μ-a.e. and in L^1_μ.*

In both cases $\lim_{n\to\infty} H_\mu(\alpha \mid \mathcal{F}_n) = H_\mu(\alpha \mid \mathcal{F}_\infty)$. (If the σ-algebras are increasing, the family $(I_\mu(\alpha \mid \mathcal{F}_n))_{n\in\mathbb{N}}$ is even dominated by a L^1_μ-function, see [16, Lemma 4.2] or [8, Theorem 3.1.10].)

Proof: Applying the increasing or the decreasing martingale theorem to each $\mu(A \mid \mathcal{F}_n)$ separately we see at once that $I_\mu(\alpha \mid \mathcal{F}_n) = \sum_{A\in\alpha} 1_A \cdot \mu(A \mid \mathcal{F}_n) \to I_\mu(\alpha \mid \mathcal{F}_\infty)$ μ-a.e. For the L^1_μ-convergence (and hence also the convergence of the conditional entropies) it suffices to show that the family $(I_\mu(\alpha \mid \mathcal{F}_n))_{n\in\mathbb{N}}$ is uniformly integrable:
For $t > 1$ we have with $B_{n,t} := \{I_\mu(\alpha \mid \mathcal{F}_n) > t\}$:

$$\begin{aligned}
\int_{B_{n,t}} I_\mu(\alpha \mid \mathcal{F}_n)\, d\mu &= \sum_{A\in\alpha} \int_{B_{n,t}} -1_A \log \mu(A \mid \mathcal{F}_n)\, d\mu \\
&= \sum_{A\in\alpha} \int_{B_{n,t}} -\mu(A \mid \mathcal{F}_n) \log \mu(A \mid \mathcal{F}_n)\, d\mu \\
&\leqslant \sum_{A\in\alpha} \min\left\{\varphi(e^{-t}), \int_\Omega \varphi(\mu(A \mid \mathcal{F}_n))\, d\mu\right\} \\
&\leqslant \sum_{A\in\alpha} \min\left\{\varphi(e^{-t}), \varphi(\mu(A))\right\} \\
&\to 0 \quad \text{as } t \to \infty
\end{aligned}$$

as $\sum_{A\in\alpha} \varphi(\mu(A)) = H_\mu(\alpha) < \infty$. □

4.3.3 The entropy of a mpds (*metric entropy* or *Kolmogorov-Sinai entropy*)

Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds. For a partition α and for integers $k \leqslant \ell$ let $\alpha_k^\ell := T^{-k}\alpha \vee \cdots \vee T^{-(\ell-1)}\alpha$. (When T is not invertible, we require $k \geqslant 0$.) Observe that $H_\mu(T^{-1}\alpha \mid T^{-1}\mathcal{F}) = H_\mu(\alpha \mid \mathcal{F})$ (see e.g. [8, Lemma 3.2.3]).

4.3.4 Definition *a) $h_\mu(T, \alpha) := \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^n)$.*
b) $h_\mu(T) := \sup\{h_\mu(T, \alpha) : H_\mu(\alpha) < \infty\}$.
(The limit in a) exists because $(H_\mu(\alpha_0^n))_{n\in\mathbb{N}}$ is a subadditive sequence: $H_\mu(\alpha_0^{n+m}) = H_\mu(\alpha_0^m \vee T^{-m}\alpha_0^n) \leqslant H_\mu(\alpha_0^m) + H_\mu(T^{-m}\alpha_0^n) = H_\mu(\alpha_0^m) + H_\mu(\alpha_0^n)$.

$h_\mu(T)$ is called the *metric entropy* or *Kolmogorov-Sinai entropy (KS-entropy)* of the mpds $(\Omega, \mathcal{A}, \mu, T)$. It is obviously an isomorphism invariant and can only decrease when passing to a factor system.

4.3.5 Theorem *Let α be a partition with $H_\mu(\alpha) < \infty$ and let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{A}$. Then*
(a) $H_\mu(\alpha_0^n \mid T^{-n}\mathcal{F}) - H_\mu(\alpha \mid \mathcal{F}) = \sum_{k=1}^n H_\mu(\alpha \mid \alpha_1^k \vee T^{-k}\mathcal{F})$
(b) $H_\mu(\alpha_0^n) - H_\mu(\alpha) = \sum_{k=1}^n H_\mu(\alpha \mid \alpha_1^k)$
(c) $h_\mu(T,\alpha) = \lim_{n\to\infty} H_\mu(\alpha \mid \alpha_1^n) = H_\mu(\alpha \mid \alpha_1^\infty)$ *where* $\alpha_1^\infty := \sigma\left(\bigcup_{n=1}^\infty T^{-n}\alpha\right)$.

Proof: For each $k \in \mathbb{N}$,

$$\begin{aligned} H_\mu(\alpha_0^k \mid T^{-k}\mathcal{F}) - H_\mu(\alpha_0^{k-1} \mid T^{-(k-1)}\mathcal{F}) &= H_\mu(\alpha \vee \alpha_1^k \mid T^{-k}\mathcal{F}) - H_\mu(\alpha_1^k \mid T^{-k}\mathcal{F}) \\ &= H_\mu(\alpha \mid \alpha_1^k \vee T^{-k}\mathcal{F}) \,. \end{aligned}$$

Summing over k we obtain (a), and (b) is just the special case $\mathcal{F} = \{\emptyset, \Omega\}$. Dividing (b) by n and passing to the limit $n \to \infty$, we obtain

$$h_\mu(T,\alpha) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n H_\mu(\alpha \mid \alpha_1^k) = \lim_{n\to\infty} H_\mu(\alpha \mid \alpha_1^n) = H_\mu(\alpha \mid \alpha_1^\infty)$$

where the second identity is due to the fact that the sequnce is decreasing and the third one follows from Theorem 4.3.3. □

4.3.6 Theorem (Shannon - McMillan - Breiman) *Let $(\Omega, \mathcal{A}, \mu, T)$ be an ergodic mpds and α a partition with $H_\mu(\alpha) < \infty$. Then*

$$\lim_{n\to\inf} \frac{1}{n} I_\mu(\alpha_0^n) = h_\mu(T,\alpha) \quad \mu\text{-a.e. and in } L^1_\mu \,.$$

Proof: (See [16, Theorem 4.3] or [8, Theorem 3.2.7b].) The main steps are:

- the purely computational decomposition

  $$I_\mu(\alpha_0^n) = \sum_{k=0}^{n-1} I_\mu(\alpha \mid \alpha_1^{n-k}) \circ T^k = R_n + S_n \quad \text{with}$$

  with

  $$R_n = \sum_{k=0}^{n-1} I_\mu(\alpha \mid \alpha_1^\infty) \circ T^k \quad \text{and} \quad S_n = \sum_{k=0}^{n-1} \Big(I_\mu(\alpha \mid \alpha_1^{n-k}) - I_\mu(\alpha \mid \alpha_1^\infty) \Big) \circ T^k \,,$$

- $\lim_{n\to\infty} \frac{1}{n} R_n = \int I_\mu(\alpha \mid \alpha_1^\infty)\, d\mu = H_\mu(\alpha \mid \alpha_1^\infty) = h_\mu(T,\alpha)$ by Birkhoff's ergodic theorem,
- $\lim_{n\to\infty} \frac{1}{n} S_n = 0$ μ-a.e. and in L^1_μ. The L^1_μ-convergence is easy, because

  $$\int |S_n|\, d\mu \leqslant \frac{1}{n}\sum_{k=0}^{n-1} \int \Big| I_\mu(\alpha \mid \alpha_1^{n-k}) - I_\mu(\alpha \mid \alpha_1^\infty) \Big|\, d\mu \to 0$$

  by Theorem 4.3.5. From this one passes to a.e. convergence as follows:
  Let $F_N := \sup_{n \geqslant N} |I_\mu(\alpha \mid \alpha_1^n) - I_\mu(\alpha \mid \alpha_1^\infty)|$. Then $(F_N)_{N\in\mathbb{N}}$ is uniformly integrable and $F_N \to 0$ μ-a.e. by Theorem 4.3.3, hence $\lim_{N\to\infty} \int F_N\, d\mu = 0$. On the other hand,

  $$\begin{aligned} \limsup_{n\to\infty} \left| \frac{1}{n} S_n \right| &\leqslant \limsup_{n\to\infty} \left(\frac{1}{n}\sum_{k=0}^{n-N} F_N \circ T^k + \frac{1}{n}\sum_{k=n-N+1}^{n-1} F_1 \circ T^k \right) \\ &\leqslant \int F_N\, d\mu + \limsup_{n\to\infty} \frac{1}{n}\sum_{k=n-N+1}^{n-1} F_1 \circ T^k \end{aligned}$$

where, for each fixed N, the last limit is 0 a.e. because $F_1 \in L^1_\mu$.

□

4.3.7 Lemma *For partitions α, β with finite entropy we have $h_\mu(T, \alpha \vee \beta) \leqslant h_\mu(T, \alpha) + h_\mu(T, \beta)$ and*

$$h_\mu(T, \alpha) \leqslant h_\mu(T, \alpha \vee \beta) \leqslant h_\mu(T, \beta) + H_\mu(\alpha \mid \beta) .$$

Proof: In view of Theorem 4.3.1 and Proposition 4.3.2 we have

$$\begin{aligned} H_\mu(\alpha_0^n) &\leqslant H_\mu((\alpha \vee \beta)_0^n) = H_\mu(\alpha_0^n \vee \beta_0^n) = H_\mu(\beta_0^n) + H_\mu(\alpha_0^n \mid \beta_0^n) \\ &\leqslant H_\mu(\beta_0^n) + \sum_{k=0}^{n-1} H_\mu(T^{-k}\alpha \mid \beta_0^n) \leqslant H_\mu(\beta_0^n) + \sum_{k=0}^{n-1} H_\mu(T^{-k}\alpha \mid T^{-k}\beta) \\ &= H_\mu(\beta_0^n) + nH_\mu(\alpha \mid \beta) . \end{aligned}$$

Dividing by n and taking the limit yields the lemma. □

4.3.8 Theorem $h_\mu(T) = \sup\{h_\mu(T, \beta) :\ \beta$ *a finite partition*$\}$.

Proof: Only "$\leqslant$" must be proved: Let $\alpha = \{A_1, A_2, \dots\}$ be any partition with $H_\mu(\alpha) < \infty$. Let $\beta_n := \{A_1, \dots, A_n, \bigcup_{k>n} A_k\}$. It is not hard to prove that $H_\mu(\alpha \mid \beta_n) \to 0$. Then the previous lemma shows that $h_\mu(T, \alpha) \leqslant \sup_n h_\mu(T, \beta_n) \leqslant \sup\{h_\mu(T, \beta) :\ \beta$ a finite partition$\}$. □

4.3.9 Definition *Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds, α a finite or countable partition.*
(a) α is a strong generator *for T, if $\alpha_0^\infty = \mathcal{A} \mod \mu$.*
(b) If T is invertible, then α is a generator *for T, if $\alpha_{-\infty}^\infty = \mathcal{A} \mod \mu$.*

4.3.10 Remark If $(\Omega, \mathcal{A}, \mu)$ is a LR-space and if α separates points under T, i.e. if for all $x \neq y$ in Ω there is $n \in \mathbb{N}$ (or $\mathbb{Z}$) such that $T^n x$ and $T^n y$ are in diffeerent elements of α, then the countable collection $\bigcup_n T^{-n}(\alpha)$ separates points in Ω so that α is a (strong) generator, see Theorem 3.1.15.

4.3.11 Theorem (Sinai's generator theorem) *Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds. If α is a strong generator for T (or, in the invertible case, a generator), then $h_\mu(T) = h_\mu(T, \alpha)$.*

Proof: Fix a finite partition β. It suffices to show that $h_\mu(T, \beta) \leqslant h_\mu(T, \alpha)$. Observe that $(\alpha_0^N)_0^n = \alpha_0^{N+n-1}$ and, in the invertible case, $(\alpha_{-N}^N)_0^n = \alpha_{-N}^{n+N-1}$. Hence

$$h_\mu(T, \alpha_{-N}^N) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_{-N}^{N+n-1}) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{2N+n-1}) = h_\mu(T, \alpha)$$

in the invertible case, and similarly, $h_\mu(T, \alpha_0^N) = h_\mu(T, \alpha)$ in the general case. Therefore,

$$h_\mu(T, \beta) \leqslant h_\mu(T, \alpha_{-N}^N) + H_\mu(\beta \mid \alpha_{-N}^N) \to h_\mu(T, \alpha) + H_\mu(\beta \mid \alpha_{-\infty}^\infty) = h_\mu(T, \alpha) .$$

In the general case one argues similarly with α_0^N instead of α_{-N}^N. □

4.3.12 Remark Using ideas of this proof, it is an easy exercise to show that $h_\mu(T^N, \alpha) = |N| \cdot h_\mu(T, \alpha)$ for all $N \in \mathbb{N}$, and if T is invertible, even for all $N \in \mathbb{Z}$:

$$\alpha_0^N \vee T^{-N}\alpha_0^N \vee \cdots \vee T^{-(n-1)N}\alpha_0^N = \alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(nN-1)}\alpha \,.$$

Taking $H_\mu(\,.\,)$ on both sides, dividing by n and taking the limit $n \to \infty$ yields $h_\mu(T^N, \alpha_0^N) = N\, h_\mu(T, \alpha)$, and the proof is easily finished.

4.3.4 Examples

The first two examples are directly taken from [16, Section 4.4].

4.3.13 Example Bernoulli shifts

4.3.14 Example Markov measures on shift spaces

For the next two examples the following Lemma is useful.

4.3.15 Lemma *Let $(\Omega, \mathcal{A}, \mu, T)$ be a mpds. If α is a* strong *generator for T and if $T_A := T|_A$ is bijective, then there are measurable functions $\varphi_A : \Omega \to [0,1]$ such that $\varphi_A \circ T$ is a version of $\mu(A \mid \alpha_1^\infty)$ $(A \in \alpha)$. The φ_A have the following properties:*

1. *$\varphi_A \circ T > 0$ on A.*
2. *Let $J_\mu T : \Omega \to \mathbb{R}$, $J_\mu T := \sum_{A\in\alpha} \frac{1}{\varphi_A \circ T} 1_A$. Then $J_\mu T = \frac{d(\mu\circ T_A)}{d\mu}$ is the "Jacobian" of T_A w.r.t. μ, in the following sense:*

$$\forall A \in \alpha\ \forall B \in \mathcal{A}:\ \mu(T(A \cap B)) = \int_{A\cap B} J_\mu T \, d\mu \,.$$

3. *$h_\mu(T) = \int \log J_\mu T \, d\mu$.*

Proof: 1. For each $A \in \alpha$ there is a version p_A of the conditional expectation $\mu(A \mid \alpha_1^\infty)$. As $\alpha_1^\infty = T^{-1}\alpha_0^\infty = T^{-1}\mathcal{A}$ (mod μ), p_A can be chosen $T^{-1}\mathcal{A}$-measurable so that there are measurable φ_A such that $\mu(A \mid \alpha_1^\infty) = p_A = \varphi_A \circ T$. As $\mu(A \cap \{\varphi_A \circ T = 0\}) = \int_{\{\varphi_A\circ T=0\}} \mu(A \mid T^{-1}\mathcal{A})\, d\mu = \int_{\{p_A=0\}} p_A \, d\mu = 0$, φ_A can be chosen such that $\varphi_A \circ T_A > 0$ on A. Next, let $B \in \mathcal{A}$ and $A \in \alpha$. Then

$$\begin{aligned}
\mu(T(A \cap B)) &= \int_{\{\varphi_A>0\}} \frac{1_{T(A\cap B)}}{\varphi_A} \cdot \varphi_A \, d\mu = \int_{\{\varphi_A\circ T>0\}} \frac{1_{T(A\cap B)}}{\varphi_A} \circ T \cdot p_A \, d\mu \\
&= \int_{A\cap\{\varphi_A\circ T>0\}} \frac{1_{T(A\cap B)}}{\varphi_A} \circ T \, d\mu = \int_A \frac{1_{A\cap B}}{\varphi_A \circ T} \, d\mu = \int_{A\cap B} \log J_\mu T \, d\mu
\end{aligned}$$

which proves the second assertion. Finally,

$$H_\mu(\alpha \mid \alpha_1^\infty) = -\sum_{A\in\alpha} \int_A \log \mu(A \mid \alpha_1^\infty)\, d\mu = \sum_{A\in\alpha} \int 1_A \log \frac{1}{\varphi_A \circ T} \, d\mu = \int \log J_\mu T \, d\mu \,,$$

so that, by Sinai's generator theorem, $h_\mu(T) = h_\mu(T, \alpha) = H_\mu(\alpha \mid \alpha_1^\infty) = \int \log J_\mu T \, d\mu$. □

4.3.16 Example Irrational rotations $T_a : \mathbb{T}^d \to \mathbb{T}^d$. Let α be the partition into 2^d squares obtained by cutting each coordinate direction in two semi-circles. It is not hard ot see that α is a strong generator if T_a is ergodic w.r.t. Lebesgue measure μ, i.e. if $\forall n \in \mathbb{Z}^d \backslash \{0\} : \langle n, a \rangle \notin \mathbb{Z}$ (see Theorem 1.1.11). As $J_\mu T_a = 1$, we conclude that $h_\mu(T_a) = 0$.

4.3.17 Example Piecewise expanding maps $T : [0,1] \to [0,1]$. These are maps for which there is a partition α of $[0,1]$ into intervals such that the restrictions $T|_A$, $A \in \alpha$ are C^1 and uniformly expanding. Then, for two points $x < y$ in $[0,1]$, $|T^n x - T^n y| = \int_x^y |(T^n)'(t)|\,dt \to \infty$ if $T^n x$ and T^n are always in the same monotonicity interval of T. But this contradicts to $|T^n x - T^n y| \leqslant 1$ for all n. Hence α is a strong generator. Suppose now that $\mu \ll m$ where m denotes Lebesgue measure and denote $h := \frac{d\mu}{dm}$. Then $J_\mu T = |T'|\,\frac{h\circ T}{h}$ so that

$$h_\mu(T) = \int \log J_\mu T\,d\mu = \int (\log h \circ T - \log h)\,d\mu + \int \log |T'|\,d\mu = \int \log |T'|\,d\mu\,.$$

This is obvious, if $\log h$ is μ-integrable. Sometimes it is only possible to prove that the positive or the negative part of $\frac{\log h \circ T}{h}$ is μ-integrable. Also in this case $\int (\log h \circ T - \log h)\,d\mu = 0$, see [8, Lemma 4.1.13].

4.3.18 Example (Continued fraction transformation) A special casse of the previous setting is the *Gauss* map $T : (0,1] \to (0,1]$, $T(x) = \frac{1}{x} \bmod 1$. The associated partition is given by the intervals $A_k = (\frac{1}{k+1}, \frac{1}{k}]$, and $\inf |(T^2)'| > 1$. An elementary calculation shows that $d\mu(x) = \frac{1}{\log 2}\frac{dx}{1+x}$ is an invariant measure. Hence

$$h_\mu(T) = h_\mu(T,\alpha) = \frac{1}{\log 2}\int_0^1 \log(x^{-2})\frac{1}{1+x}\,dx = \frac{-2}{\log 2}\int_0^1 \frac{\log x}{1+x}\,dx = \frac{\pi^2}{6\log 2}\,,$$

as the last integral evaluates to $\frac{-\pi^2}{12}$. Observe that

$$\frac{1}{k(x) + Tx} = \frac{1}{k(x) + \frac{1}{k(Tx)+T^2x}} = \frac{1}{k(x) + \frac{1}{k(Tx)} + \frac{1}{k(T^2x)+T^3x}} = \dots$$

converges to the continued fraction expansuion of x, where $x \in J_{k(x)}$. Hence $h_\mu(T)$ gives the average information content per digit in the continued fraction expansion of a Lebesgue typical point x.

4.3.19 Example Linear torus automorphisms $T : \mathbb{T}^d \to \mathbb{T}^d$, $Tx = A\cdot x$, $|\det(A) = 1|$, no eigenvalue of modulus 1. Then [17], with respect to Lebesgue measure μ,

$$h_\mu(T) = \sum_{|\lambda_i|>1} \log |\lambda_i|$$

where the sum extends over all eigenvalues of A of modulus bigger than 1, counted with multiplicities.

4.3.5 K(olmogorov)-systems and the Pinsker algebra

In this section we study only invertible mpds $(\Omega, \mathcal{A}, \mu, T)$ on a LR-space. The material goes back to Kolmogorov, Pinsker, Rohlin and Sinai and is taken from [12, Section 4.3].

Notation: If α is a finite or countable partition, then $\alpha_{-\infty}^{\infty} := \sigma\left(\bigcup_{n\in\mathbb{Z}} T^{-n}\alpha\right)$. Recall that α is a generator, if $\alpha_{-\infty}^{\infty} = \mathcal{A} \mod \mu$. We also denote $\mathrm{Tail}(T,\alpha) := \bigcap_{n\in\mathbb{N}} \alpha_n^{\infty}$. Observe that

$$T\left(\mathrm{Tail}(T,\alpha)\right) = \mathrm{Tail}(T,\alpha) = T^{-1}\left(\mathrm{Tail}(T,\alpha)\right)$$

4.3.20 Theorem ([12, Theorem 2.52]) *The collection*

$$\mathcal{P}(T) := \{A \in \mathcal{A} : h_\mu(T, \{A, \Omega\backslash A\}) = 0\}$$

of measurable sets is a T-invariant σ-algebra, the Pinsker-algebra *of the system. For each finite or countable measurable partition α with $H_\mu(\alpha) < \infty$ holds: $h_\mu(T,\alpha) = 0$ if and only if $\alpha \subseteq \mathcal{P}(T)$.*

Proof: For $A \in \mathcal{A}$ denote by $\tilde{A}$ the partition $\{A, \Omega \backslash A\}$.

- Let $h_\mu(T,\alpha) = 0$. Then $h_\mu(T,\tilde{A}) = 0$ for all $A \in \alpha$, as $\tilde{A} \leqslant \alpha$. Hence $\alpha \subseteq \mathcal{P}(T)$.
- Let $\alpha \subseteq \mathcal{P}(T)$. Then $\alpha = \bigvee_{A \in \alpha} \tilde{A}$. Hence $h_\mu(T,\alpha) \leqslant \sum_{A\in\alpha} h_\mu(T,\tilde{A}) = 0$. Here we used Lemma 4.3.7 in the case of a finite partition α, and for countable partitions one proceeds as in the proof of Theorem 4.3.8.
- $\mathcal{P}(T)$ is T-invariant, because $h_\mu(T,\widetilde{T^{-1}A}) = h_\mu(T,T^{-1}\tilde{A}) = h_\mu(T,\tilde{A})$.
- By definition, $\Omega \in \mathcal{P}(T)$ and $A \in \mathcal{P}(T) \Leftrightarrow \Omega\backslash A \in \mathcal{P}(T)$. So let $A_1,\dots,A_n \in \mathcal{P}(T)$. Then $A_1 \cup \dots A_n \in \tilde{A}_1 \vee \dots \vee \tilde{A}_n$ so that $h_\mu(T,\widetilde{\bigcup_{k=1}^n A_k}) \leqslant \sum_{k=1}^n h_\mu(T,\tilde{A}_k) = 0$. Hence $A_1 \cup \dots \cup A_n \in \mathcal{P}(T)$. Finally, if $A_1, A_2, \dots \in \mathcal{P}(T)$, then

$$h_\mu\left(T,\widetilde{\bigcup_{k=1}^\infty A_k}\right) \leqslant h_\mu\left(T,\widetilde{\bigcup_{k=1}^n A_k} \vee \widetilde{\bigcup_{k=1}^\infty A_k}\right) \leqslant h_\mu\left(T,\widetilde{\bigcup_{k=1}^n A_k}\right) + H_\mu\left(\widetilde{\bigcup_{k=1}^\infty A_k} \mid \widetilde{\bigcup_{k=1}^n A_k}\right)$$

for each n in n. But the first summand is 0 as $\bigcup_{k=1}^n A_k \in \mathcal{P}(T)$, and the second one tends to 0 for $n \to \infty$ as in the proof of Theorem 4.3.8.

□

4.3.21 Theorem *Let α be a partition with $H_\mu(\alpha) < \infty$. Then*
(a) $\mathrm{Tail}(T,\alpha) \subseteq \mathcal{P}(T)$.
(b) If α is a generator, then $\mathrm{Tail}(T,\alpha) = \mathcal{P}(T)$ *mod* μ.

Proof: (a) Let $B \in \mathrm{Tail}(T,\alpha)$ and denote $\beta = \{B, \Omega\backslash B\}$. We have to show that $h_\mu(T,\beta) = 0$. As $\mathrm{Tail}(T,\alpha)$ is T-invariant, we have $\beta_{-\infty}^\infty \subseteq \mathrm{Tail}(T,\alpha) \subseteq \alpha_1^\infty$, in particular

- $H_\mu(\beta \mid \alpha_1^\infty) = 0$,
- $(\alpha \vee \beta)_1^\infty = \alpha_1^\infty \vee \beta_1^\infty = \alpha_1^\infty$, and
- $\alpha_1^k \vee T^{-k}\beta_{-n}^0 \subseteq \alpha_1^\infty \vee \beta_{-\infty}^\infty = \alpha_1^\infty$.

Hence

$$\begin{aligned} h_\mu(T,\alpha\vee\beta) &= H_\mu(\alpha\vee\beta \mid (\alpha\vee\beta)_1^\infty) = H_\mu(\alpha\vee\beta \mid \alpha_1^\infty) \\ &\leqslant H_\mu(\alpha \mid \alpha_1^\infty) + H_\mu(\beta \mid \alpha_1^\infty) = h_\mu(T,\alpha)\,. \end{aligned}$$

On the other hand, by Theorem 4.3.5(a),

$$\begin{aligned} H_\mu((\alpha\vee\beta)_0^n) &= H_\mu(\beta_0^n) + H_\mu(\alpha_0^n \mid \beta_0^n) = H_\mu(\beta_0^n) + H_\mu(\alpha_0^n \mid T^{-n}\beta_{-n}^0) \\ &= H_\mu(\beta_0^n) + H_\mu(\alpha \mid \beta_{-n}^0) + \sum_{k=0}^{n-1} H_\mu(\alpha \mid \alpha_1^k \vee T^{-k}\beta_{-n}^0) \\ &\geqslant H_\mu(\beta_0^n) + H_\mu(\alpha \mid \beta_{-n}^0) + \sum_{k=0}^{n-1} H_\mu(\alpha \mid \alpha_1^\infty)\,. \end{aligned}$$

Dividing by n and passing to the limit this yields, together with the first estimate,

$$h_\mu(T,\alpha) \geqslant h_\mu(T,\alpha \vee \beta) \geqslant h_\mu(T,\beta) + h_\mu(T,\alpha)$$

As $h_\mu(T,\alpha) < \infty$, we conclude that $h_\mu(T,\beta) = 0$.

Claim A: If π is a $\mathcal{P}(T)$ measurable partition, then $\pi_0^\infty = \pi_1^\infty = \mathrm{Tail}(T,\pi)$.
Proof: The inclusions '$\supseteq$" are trivial, and

$$0 = n \cdot h_\mu(T,\pi) = h_\mu(T^n,\pi) = H_\mu\left(\beta \mid \sigma\left(\bigcup_{k=1}^{\infty} T^{-kn}\pi\right)\right) \geqslant H_\mu(\pi \mid \pi_n^\infty)$$

for all $n \in \mathbb{N}$ so that π is π_n^∞-measurable mod μ for all n and hence $\pi \subseteq \bigcap_{n\in\mathbb{N}} \pi_n^\infty = \mathrm{Tail}(T,\pi)$ mod μ.[1] The claim follows now from the T-invariance of $\mathrm{Tail}(T,\pi)$.

Claim B: $H_\mu(\beta \mid \beta_1^\infty \vee \mathcal{F}) = H_\mu(\beta \mid \beta_1^\infty)$ for any sub-σ-algebra $\mathcal{F}$ of $\mathcal{P}(T)$.
Proof: If β,π are partitions with finite entropy and if $\pi \subseteq \mathcal{P}(T)$, then

$$\begin{aligned}
H_\mu(\beta \mid \beta_1^\infty \vee \pi_0^\infty) &= H_\mu(\beta \mid \beta_1^\infty \vee \pi_1^\infty) \quad \text{(by Claim A)}\\
&= H_\mu(\beta \mid \beta_1^\infty \vee \pi_1^\infty) + H_\mu(\pi \mid \pi_1^\infty) \quad (\text{as } \pi \in \mathcal{P}(T))\\
&\geqslant H_\mu(\beta \mid \beta_1^\infty \vee \pi_1^\infty) + H_\mu(\pi \mid \beta \vee \beta_1^\infty \vee \pi_1^\infty)\\
&= H_\mu(\beta \vee \pi \mid \beta_1^\infty \vee \pi_1^\infty) = H_\mu(\beta \vee \pi \mid (\beta \vee \pi)_1^\infty)\\
&= h_\mu(T,\beta \vee \pi)\\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu((\beta \vee \pi)_0^n)\\
&\geqslant \lim_{n\to\infty} \frac{1}{n} H_\mu(\beta_0^n) = h_\mu(T,\beta)\\
&= H_\mu(\beta \mid \beta_1^\infty)\\
&\geqslant H_\mu(\beta \mid \beta_1^\infty \vee \pi_0^\infty)\ .
\end{aligned}$$

Let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{P}(T)$. As $(\Omega,\mathcal{A},\mu)$ is a LR-space, $\mathcal{F}$ is countably generated mod μ. Hence there is an increasing sequence of finite partitions π_n such that $\sigma(\pi_1) \subseteq \sigma(\pi_2) \subseteq \cdots \nearrow \mathcal{F}$ mod μ. Then also $\sigma((\pi_1)_0^\infty) \subseteq \sigma((\pi_2)_0^\infty) \subseteq \cdots \nearrow \mathcal{F}$ mod μ, and the preceding estimate implies

$$H_\mu(\beta \mid \beta_1^\infty \vee \mathcal{F}) = H_\mu(\beta \mid \beta_1^\infty)\ .$$

(b) Now let α be a generator. As $\mathrm{Tail}(T,\alpha) \subseteq \mathcal{P}$ mod μ by part (a), we conclude that

$$H_\mu(\beta \mid \beta_1^\infty \vee \mathrm{Tail}(T,\alpha)) = H_\mu(\beta \mid \beta_1^\infty) = H_\mu(\beta \mid \beta_1^\infty \vee \mathcal{P}(T))\ .$$

If we apply this identity to T^{2^n} instead of T and if we observe that $\mathrm{Tail}(T^{2^n},\alpha) \subseteq \mathrm{Tail}(T,\alpha)$ and $\mathcal{P}(T^{2^n}) = \mathcal{P}(T)$, then

$$H_\mu\left(\beta \mid \bigvee_{k=1}^{\infty} T^{-k2^n}\beta \vee \mathrm{Tail}(T,\alpha)\right) \leqslant H_\mu\left(\beta \mid \bigvee_{k=1}^{\infty} T^{-k2^n}\beta \vee \mathcal{P}(T)\right)\ .$$

[1]To see this, let $A \in \pi$ so that there are $A_n \in \pi_n^\infty$ such that $\mu(A \triangle A_n) = 0$ for all n. Replacing A_n by $\bigcap_{k=n}^\infty A_k$ we can assume that $A_1 \supseteq A_2 \supseteq \dots$. Let $A_\infty := \bigcap_{n=1}^\infty A_n$. Then $A_n \in \mathrm{Tail}(T,\alpha)$ and $\mu(A \triangle A_\infty) = 0$.

Let $\mathcal{B} := \bigcap_{n=1}^{\infty} \left(\bigvee_{k=1}^{\infty} T^{-k2^n}\beta\right)$. The σ-algebras on both sides of this inequality are decreasing to $\mathcal{B} \vee \mathrm{Tail}(T,\alpha)$ and $\mathcal{B} \vee \mathcal{P}(T)$, respectively, so that Theorem 4.3.3 yields in the limit $n \to \infty$

$$H_\mu\left(\beta \mid \mathcal{B} \vee \mathrm{Tail}(T,\alpha)\right) \leqslant H_\mu\left(\beta \mid \mathcal{B} \vee \mathcal{P}(T)\right) \leqslant H_\mu\left(\beta \mid \mathcal{P}(T)\right) \ .$$

Suppose now that $\beta \subseteq \alpha_{-p}^{\infty}$ for some $p \in \mathbb{N}$. Then $\bigvee_{k=1}^{\infty} T^{-k2^n}\beta \subseteq \alpha_{-p+2^n}^{\infty}$ for all n and hence

$$\mathcal{B} = \bigcap_{n=1}^{\infty} \left(\bigvee_{k=1}^{\infty} T^{-k2^n}\beta\right) \subseteq \bigcap_{n=1}^{\infty} \alpha_{-p+2^n}^{\infty} = \mathrm{Tail}(T,\alpha) \ .$$

Therefore, in this case

$$H_\mu\left(\beta \mid \mathrm{Tail}(T,\alpha)\right) \leqslant H_\mu\left(\beta \mid \mathcal{P}(T)\right) \ . \tag{4.3.1}$$

Now let $B \in \mathcal{P}(T)$. As α is a generator, there are sets $B_p \in \alpha_{-p}^{\infty}$ such that $\mu(B \triangle B_p) \to 0$ as $p \to \infty$. Let $\beta_p := \{B_p, \Omega \backslash B_p\}$ and $\beta = \{B, \Omega \backslash B\}$. Then (4.3.1) holds for all β_p and carries over to β in the limit $p \to \infty$. Hence $H_\mu(\beta \mid \mathrm{Tail}(T,\alpha)) \leqslant H_\mu(\beta \mid \mathcal{P}(T)) = 0$, and we conclude that $B \in \mathrm{Tail}(T,\alpha)$ mod μ. □

4.3.22 Definition *An invertible mpds* $(\Omega, \mathcal{A}, \mu, T)$ *is a K-system, if there is a generator* α *with* $H_\mu(\alpha) < \infty$ *and* $\mathrm{Tail}(T,\alpha) = \{\emptyset, \Omega\}$ *mod* μ.

4.3.23 Theorem *An invertible mpds is a K-system if and only if its Pinsker algebra is trivial.*

The "only if" direction follows immediately from the previous theorem. The "if" direction requires additionally the construction of a generating partition.

Bibliography

[1] H. Bauer. *Wahrscheinlichkeitstheorie*. de Gruyter, 1991.

[2] V. I. Bogachev. *Measure Theory, Volume II*, volume I. Springer Berlin Heidelberg, 2007.

[3] J.-P. Conze and A. Raugi. On the ergodic decomposition for a cocycle. *Colloquium Mathematicum*, 117(1):121–156, 2009.

[4] T. de la Rue. Espaces de Lebesgue. *Séminaire de Probabilités XXVII*, 27:15–21, 1993.

[5] M. Einsiedler and T. Ward. *Ergodic Theory (with a view towards number theory)*. Springer, 2010.

[6] P. Garrett. Compact operators, Hilbert-Schmidt operators. http://www.math.umn.edu/∼garrett/, March 2012.

[7] G. Greschonig and K. Schmidt. Ergodic decomposition of quasiinvariant probability measures. *Colloq. Math*, 84:495–514, 2000.

[8] G. Keller. *Equilibrium States in Ergodic Theory*. Cambridge University Press, 1998.

[9] G. Keller. Entropie und Große Abweichungen. Vorlesungsskript, 2013/14.

[10] G. Keller. Wahrscheinlichkeitstheorie. Vorlesungsskript, 2014.

[11] U. Krengel. *Ergodic Theorems*. de Gruyter, 1985.

[12] N. F. Martin and J. W. England. *Mathematical Theory of Entropy*. Addison-Wesley, 1981.

[13] J. Neveu. *Mathematical Foundations of the Calculus of Probability*. Holden-Day, 1965.

[14] K. Petersen. *Ergodic theory*. Cambridge University Press, 1983.

[15] A. Raugi. A probabilistic ergodic decomposition result. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 45(4):932–942, Nov. 2009.

[16] O. Sarig. Lecture Notes on Ergodic Theory. http://www.wisdom.weizmann.ac.il/ sarigo/506/ErgodicNotes.pdf, 2009.

[17] P. Walters. *An Introduction to Ergodic Theory*. Springer, 1982.