IV. The Mean Ergodic Theorem

"Ergodic theory is the study of transformations from the point of view of ... dynamical properties connected with *asymptotic behavior*" (Walters [1975], p. 1). Here, the asymptotic behavior of a transformation φ is described by

$$\text{“}\lim_{n\to\infty}\text{”}\varphi^n$$

where it is our task first to make precise in which sense the "lim" has to be understood and second to prove its existence. Motivated by the original problem "time mean equals space mean" (see III.D.6) we investigate in this lecture the existence of the limit for $n \to \infty$ not of the powers φ^n but of the "Cesàro means"

$$\frac{1}{n}\sum_{i=0}^{n-1} f \circ \varphi^i$$

where f is an "observable" (see physicist's answer in Lecture I) contained in an appropriate function space. With a positive answer to this question - for convergence in L^2-space - ergodic theory was born as an independent mathematical discipline.

IV.1 Theorem (J. von Neumann, 1931):
Let $(X, \Sigma, \mu; \varphi)$ be and MDS and denote by T_φ the induced (unitary) operator on $L^2(X, \Sigma, \mu)$. For any $f \in L^2(\mu)$ the sequence of functions

$$f_n := \frac{1}{n}\sum_{i=0}^{n-1} T_\varphi^i f, \quad n \in \mathbb{N}$$

(norm-)converges to a T_φ-invariant function $\bar{f} \in L^2(\mu)$.

It was soon realized that only a few of the above assumptions are really necessary, while the assertion makes sense in a much more general context. Due to the importance of the concept and the elegance of the results, an axiomatic and purely functional-analytic approach seems to be the most appropriate.

IV.2 Definition:
An FDS $(E; T)$ (resp. a bounded linear operator T) is called *mean ergodic*, if the sequence

$$T_n := \frac{1}{n}\sum_{i=0}^{n-1} T^i, \quad n \in \mathbb{N}$$

converges in $\mathscr{L}(E)$ for the strong operator topology.

As above, the operators T_n will be called the *"Cesàro means"* of the powers T^i. Moreover we call $P := \lim_{n\to\infty} T_n$, if it exists, the *"projection corresponding to T"*. This language is justified by the following elementary properties of mean ergodic operators.

IV.3 Proposition:
(0) $(\mathrm{id} - T)T_n = \frac{1}{n}(\mathrm{id} - T^n)$ for every $n \in \mathbb{N}$

If T is mean ergodic with corresponding projection P, we have

(1) $TP = PT = P = P^2$.

(2) $PE = F := \{f \in E : Tf = f\}$.

(3) $P^{-1}(0) = \overline{(\mathrm{id} - T)E}$.

(4) The adjoints T_n' converge to P' in the weak*-operator topology of $\mathscr{L}(E')$ and $P'E' = F' := \{f' \in E' : T'f' = f'\}$.

(5) $(PE)'$ is (as a topological vector space) isomorphic to $P'E'$.

Proof.

(0) is obvious from the definition of T_n.

(1) Clearly, $(n+1)T_{n+1} - \mathrm{id} = nT_nT = nTT_n$ holds. Dividing by n and letting n tend to infinity we obtain $P = PT = TP$. From this we infer that $T_nP = P$ and thus $P^2 = P$.

(2) $PE \subseteq F$ follows from $TP = P$, and $F \subseteq PE$ from $P = \lim_{n\to\infty} T_n$.

(3) By the relations in (1), $(\mathrm{id} - T)E$ and (by the continuity of P) its closure is contained in $P^{-1}(0)$. Now take $f \in P^{-1}(0)$. Then

$$f = f - Pf = f - PTf = \lim_{n\to\infty} (\mathrm{id} - T_nT)f = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (\mathrm{id} - T^i)f$$
$$= \lim_{n\to\infty} (\mathrm{id} - T)\frac{1}{n} \sum_{i=1}^{n} iT_if \in \overline{(\mathrm{id} - T)E}.$$

(4) By the definition of the weak* operator topology, T'_n converges to P' if $\langle T_nf, f'\rangle = \langle f, T'_nf'\rangle \to \langle f, P'f'\rangle = \langle Pf, f'\rangle$ for $f \in E$ and $f' \in E'$. This follows from the convergence of T_n to P in the strong operator topology. Together with $(PT)' = T'P' = P'$ this implies the remaining property as in (2).

(5) This statement holds for every projection on a Banach space (see B.7, Proposition).

■

Our main result contains a list of surprisingly different, but equivalent characterizations of mean ergodicity at least for operators with bounded powers.

IV.4 Theorem:
If $(E;T)$ is an FDS with $\|T^n\| \leqslant c$ for every $n \in \mathbb{N}$ the following assertions are equivalent:

(a) T is mean ergodic.

(b) T_n converges in the weak operator topology.

(c) $\{T_nf : n \in \mathbb{N}\}$ has a weak accumulation point for all $f \in E$

(d) $\overline{\mathrm{co}}\{T^if : i \in \mathbb{N}_0\}$ contains a T-fixed point for all $f \in E$.

(e) The T-fixed space F separates points of the T'-fixed space F'.

Proof. The implications (a) ⇒ (b) ⇒ (c) are trivial.

(c)⇒ (d): Take $f \in E$ and let g be a weak accumulation point of $\{T_nf : n \in \mathbb{N}\}$, i.e. $g \in \overline{\{T_nf : n > n_0\}}^{\sigma(E,E')}$ for all $n_0 \in \mathbb{N}$. Certainly, g is contained in $\overline{\mathrm{co}}\{T^if : i \in \mathbb{N}_0\}$, and we shall show that g is fixed under T: For any $n_0 \in \mathbb{N}$ we obtain

$$g - Tg = (\mathrm{id} - T)g \in (\mathrm{id} - T)\overline{\{T_nf : n > n_0\}}^{\sigma} \subseteq \overline{\{(\mathrm{id} - T)T_nf : n > n_0\}}^{\sigma}$$
$$= \overline{\{\tfrac{1}{n}(\mathrm{id} - T^n)f : n > n_0\}}^{\sigma} \subseteq \tfrac{1}{n_0}(1 + c)\|f\|U,$$

where U is the closed unit ball in E – we used the fact that $(\mathrm{id} - T)$ is continuous for the weak topology and that U is weakly closed (see B.7 and B.3).

(d) ⇒ (e) : Choose $f', g' \in F'$, $f' \neq g'$, and $f \in E$ with $\langle f, f'\rangle \neq \langle f, g'\rangle$. For all elements $f_0 \in \overline{\mathrm{co}}\{T^if : i \in \mathbb{N}_0\}$ we have $\langle f_0, f'\rangle = \langle f, f'\rangle$ and $\langle f_0, g'\rangle = \langle f, g'\rangle$ Therefore the T-fixed point $f_1 \in \overline{\mathrm{co}}\{T^if : i \in \mathbb{N}\}$ which exists by (d), satisfies

$\langle f_1, f' \rangle = \langle f, f' \rangle \neq \langle f, g' \rangle = \langle f_1, g' \rangle$, i.e. it separates f' and g'.

(e) ⇒ (a): Consider

$$G := F \oplus \overline{(\mathrm{id} - T)E}$$

and assume that $f' \in E'$ vanishes on G. Since it vanishes on $(\mathrm{id} - T)E$ it follows immediately that $f' \in F'$. Since it also vanishes on F, which is supposed to separate F', we conclude that $f' = 0$, hence that $G = E$. But $T_n f$ converges for every $f \in F \oplus (\mathrm{id} - T)E$, and the assertion follows from the equicontinuity of $\{T_n : n \in \mathbb{N}\}$. ∎

The standard method of applying the above theorem consists in concluding mean ergodicity of an operator from the apparently "weakest" condition (IV.4.c) and the weak compactness of certain sets in certain Banach spaces. This settles the convergence problem for the means T_n as long as the operator T is defined on the right Banach space E.

IV.5 Corollary:
Let $(E; T)$ be an FDS where E is a reflexive Banach space, and assume that $\|T^n\| \leqslant c$ for all $n \in \mathbb{N}$. Then T is mean ergodic.

Proof. Bounded subsets of reflexive Banach spaces are relatively weakly compact (see B.4). Since $\{T_n f : n \in \mathbb{N}\}$ is bounded for every $f \in E$, it has a weak accumulation point. ∎

Besides matrices with bounded powers on $\mathbb{R}^n$ we have the following concrete applications:

Example 1: Let E be a Hilbert space and $T \in \mathscr{L}(E)$ be a contraction. Then T is mean ergodic and the corresponding projection P is orthogonal: By (IV.5) the Cesàro means T_n of T converge to P and the Cesaró means T_n^* of the (Hilbert space) adjoint T^* converge to a projection Q. If $(\cdot|\cdot)$ denotes the scalar product on E, we obtain from $(T_n^* f|g) \to (Qf|g)$ and $(f|T_n g) \to (f|Pg)$ for all $f, g \in E$ that $Q = P^*$. The fixed space $F = PE$ of T and the fixed space $F^* = P^*E$ of T^* are identical: Take $f \in F$. Since $\|T\| = \|T^*\| \leqslant 1$, the relation $(f|f) = (Tf|f) = (f|T^*f)$ implies $(f|f) \leqslant \|f\| \cdot \|T^*f\| \leqslant \|f\|^2 = (f|f)$, hence $T^*f = f$. The other conclusion $F^* \subseteq F$ follows by symmetry. Finally we conclude from $P = P^*P = (P^*P)^* = P^*$ that P is orthogonal.

Example 2: Let $(X, \Sigma, \mu; \varphi)$ be an MDS. The induced operator T_φ on $L^p(X, \sigma, \mu)$ for $1 < p < \infty$ is mean ergodic, and the corresponding projection P is a "conditional expectation" (see B.24):

For $f, g \in L^\infty$ and $T_\varphi f = f$ we obtain $T_\varphi(fg) = T_\varphi \cdot T_\varphi g = f \cdot T_\varphi g$. The same holds for $(T_\varphi)_n$, and therefore $P(fg) = f \cdot Pg$.

Both examples contain the case of the original von Neumann theorem (IV.1).

IV.6 Corollary:
Let $(E; T)$ be an FDS where $E = L^1(X, \Sigma, \mu)$, $\mu(X) < \infty$, and T is a positive contraction such that $T\mathbf{1} \leqslant \mathbf{1}$. Then T is mean ergodic.

Proof. The order interval $[-\mathbf{1}, \mathbf{1}] := \{f \in L^1(\mu) : -\mathbf{1} \leqslant f \leqslant \mathbf{1}\}$ is the unit ball of the dual $L^\infty(\mu)$ of $L^1(\mu)$ and therefore $\sigma(L^\infty, L^1)$-compact. The topology induced by $\sigma(L^1, L^\infty)$ in $[-\mathbf{1}, \mathbf{1}]$ is coarser than that induced by $\sigma(L^\infty, L^1)$ – since $L^\infty(\mu) \subseteq L^1(\mu)$ – but still Hausdorff. Therefore the two topologies coincide (see A.2) and

$[\mathbf{-1},\mathbf{1}]$ is weakly compact. By assumption, T and therefore the Cesàro means T_n map $[\mathbf{-1},\mathbf{1}]$ into itself, hence (IV.4.c) is satisfied for all $f \in L^\infty(\mu)$. As shown in (B.14) the same property follows for all $f \in L^1(\mu)$. ∎

Using deeper functional-analytic tools one can generalize the above corollary still further: Let T be a positive contraction on $L^1(X, \Sigma, \mu)$ and assume that the set $\{T_n u : n \in \mathbb{N}\}$ is relatively compact for some strictly positive function $u \in L^1(\mu)$. By [Schaefer, II.8.8] it follows that $\bigcup_{n\in\mathbb{N}}\{g \in l^1(\mu) : 0 \leqslant g \leqslant T_n u\}$ is also relatively weakly compact. From $0 \leqslant T_n f \leqslant T_n u$ fo $0 \leqslant f \leqslant u$, (B.14) and (IV.4.c) we conclude that T is mean ergodic (see Ito [1965], Yeadon[1980]).

Example 3: Let $(X, \Sigma, \mu; \varphi)$ be an MDS. The induced operator T_φ in $L^2(X, \Sigma, \mu)$ is mean ergodic, and the corresponding projection is a conditional expectation: The first assertion follows from (IV.6) while the second is proved as in Example 2 above.

Example 4: Let $E = L^1([0,1], \mathcal{B}, m)$, m the Lebesgue measure, and $k : [0,1]^2 \to \mathbb{R}_+$ be a measurable function, such that $\int_0^1 k(x,y)\,\mathrm{d}y = 1$ for all $x \in [0,1]$. Then the kernel operator

$$T : E \to E, \quad f \mapsto Tf(x) := \int_0^1 k(x,y) f(y)\,\mathrm{d}y$$

is mean ergodic.

Even though there is still much to say about the functional-analytic properties of mean ergodic operators, we here concentrate on their ergodic properties as defined in Lecture III. A particularly satisfactory result is obtained for MDSs, since the induced operators are automatically mean ergodic on $L^p(\mu)$, $1 \leqslant p < \infty$.

IV.7 Proposition:
Let $(X, \Sigma, \mu; \varphi)$ be an MDS and $E = L^p(X, \Sigma, \mu)$, $1 \leqslant p < \infty$. Then T_φ is mean ergodic and the following properties are equivalent:

(a) φ is ergodic.
(b) The projection corresponding to T_φ has the form $P = \mathbf{1} \otimes \mathbf{1}$, i.e. $Pf = \langle f, \mathbf{1}\rangle \cdot \mathbf{1}$ for all $f \in E$
(c) $\frac{1}{n}\sum_{i=0}^{n-1} \int_X (f \circ \varphi^i) \cdot g\,\mathrm{d}\mu$ converges to $\int_X f\,\mathrm{d}\mu \cdot \int_X g\,\mathrm{d}\mu$ for all $f \in L^p(\mu)$, $g \in L^p(\mu)' = L^q(\mu)$ with $\frac{1}{p} + \frac{1}{q} = 1$.
(d) $\frac{1}{n}\sum_{i=0}^{n-1} \mu(A \cap \varphi^{-1}(B))$ converges to $\mu(A) \cdot \mu(B)$ for all $A, B \in \Sigma$.
(e) $\frac{1}{n}\sum_{i=0}^{n-1} \mu(A \cap \varphi^{-1}(A))$ converges to $\mu(A)^2$ for all $A \in \Sigma$.

Proof.
(a) ⇒ (b): Since φ is ergodic and T_φ is mean ergodic, the fixed spaces of T_φ and T'_φ are one-dimensional (III.4 and IV.4.e). Since P is a projection onto the T_φ-fixed space it must be of the form $f \mapsto Pf = \langle f, f'\rangle\mathbf{1}$ for some $f' \in E'$. But

$$\int_X f\,\mathrm{d}\mu = \langle f, \mathbf{1}\rangle = \langle f, T'_\varphi \mathbf{1}\rangle = \langle f, P'\mathbf{1}\rangle = \langle Pf, \mathbf{1}\rangle = \langle f, f'\rangle \cdot \langle \mathbf{1}, \mathbf{1}\rangle = \langle f, f'\rangle$$

shows that $P = \mathbf{1} \otimes \mathbf{1}$.

(b) ⇒ (c): Condition (c) just says that $\frac{1}{n}\sum_{i=0}^{n-1} T_\varphi^i$ converges toward $\mathbf{1}\otimes\mathbf{1}$ in the weak operator topology for the particular space $L^p(\mu)$ and its dual $L^q(\mu)$.

(c) ⇒ (d): This follows if we take $f = \mathbf{1}_A$ and $g = \mathbf{1}_B$. The implication

(d) ⇒ (e): is trivial.

(e) ⇒ (a): Assume that $\varphi(A) = A \in \Sigma$. Then $\frac{1}{n}\sum_{i=0}^{n-1} \mu(A \cap \varphi^{-i}(A))$ is equal to $\mu(A)$ and converges to $\mu(A)^2$. Therefore $\mu(A)$ must be equal to 0 or 1. ∎

Remark: Further equivalences in (IV.7) are easily obtained by taking in (c) the functions f, g only from total subsets, resp. in (d) or (e) the sets A, B only from a subalgebra generating Σ.

The "automatic" mean ergodicity of T_φ in $L^p(\mu)$, $1 \leqslant p < \infty$ (by Example 2 and 5) is the reason why ergodic MDSs are characterized by the one-dimensional fixed spaces (see III.4). In fact, mean ergodicity is a rather weak property for operators on $L^p(\mu)$, $p \neq \infty$, in the sense that many operators (e.g. all contractions for $p \neq 1$ or all positive contractions satisfying $T\mathbf{1} \leqslant \mathbf{1}$ for $p = 1$) are mean ergodic.

For operators on spaces $C(X)$ the situation is quite different and mean ergodicity of $T \in \mathscr{L}(C(X))$ is a very strong property. The reason is that the sup-norm $\|\cdot\|_\infty$ is much finer than $\|\cdot\|_p$, therefore it is more difficult to identify weakly compact orbits (in order to apply IV.4.c) or the dual fixed space (in order to apply IV.4.e). Even for operators T_φ on $C(X)$ induced by a TDS one has mean ergodicity only if one makes additional assumptions, e.g. (IV.8 below or VIII.2). This non-convergence of the Cesàro means of T_φ accounts for many of the differences and additional complications in the topological counterparts to measure theoretical theorems. A first example is the characterization of minimality by one-dimensional fixed spaces.

IV.8 Proposition:
For a TDS $(X; \varphi)$ the following are equivalent:

(a) T_φ is mean ergodic in $C(X)$ and φ is minimal.
(b) There exists a unique φ-invariant probability measure, and this measure is strictly positive.

Proof. (a) ⇒ (b): From (III.7.i) and (IV.4.e) we conclude that $\dim F = \dim F' = 1$ for the fixed spaces F in $C(X)$, resp. F' in $C(X)'$. Since T_φ is a positive operator, so is P and hence P'. Every element in $C(X)'$ is a difference of positive elements, the same is true for $F' = P'C(X)'$ and therefore F' is the subspace generated by a single probability measure called ν.

Let $0 \leqslant f \in C(X)$ with $\langle f, \nu\rangle = 0$ and define $Y := \bigcap\{[f \circ \varphi^n = 0] : n \in \mathbb{Z}\}$. Then is closed and φ-invariant, and therefore $Y = \varnothing$ or $Y = X$. If $Y = X$, then $f = 0$, whereas if $Y = \varnothing$ implies that for all $x \in X$ one has $f \circ \varphi^n(x) > 0$ for some $n \in \mathbb{Z}$. Since $\langle f, \nu\rangle = \langle F \circ \varphi^n, \nu\rangle$ for all $n \in \mathbb{N}$, this shows that $\langle f, \nu\rangle > 0$.

(b) ⇒ (a): Let $f' \in C(X)'$ be T_φ'-invariant. Since T_φ' is positive, we obtain

$$|f'| = |T_\varphi' f'| \leqslant T_\varphi' |f'|$$

and $\langle \mathbf{1}, |f'|\rangle \leqslant \langle \mathbf{1}, T'_\varphi|f'|\rangle = \langle T_\varphi \mathbf{1}, |f'|\rangle = \langle \mathbf{1}, |f'|\rangle$. Hence $\langle \mathbf{1}, T'_\varphi|f'| - |f'|\rangle = \langle T_\varphi \mathbf{1}|f'|\rangle - \langle \mathbf{1}, |f'|\rangle = 0$, therefore $|f'|$ is T'_φ-invariant, and the dual fixed space F' is a vector lattice. Consequently every element in F' is difference of positive elements and – by assumption – F' is one-dimensional and spanned by the unique φ-invariant probability measure ν. Apply now (IV.4.e) to conclude that T_φ is mean ergodic. Again the corresponding projection is of the form $P = \mathbf{1} \otimes \nu$. Assume now that $Y \subseteq X$ is closed and φ-invariant. There exists $0 < f \in C(X)$ with $f(Y) \subseteq \{0\}$, $T_\varphi f(Y) \subseteq \{0\}$, therefore $(Pf)(Y) \subseteq \{0\}$. Hence $(\int_X f \, \mathrm{d}\nu)\mathbf{1}(Y) \subseteq \{0\}$ and Y must be empty. ∎

Example 5: The rotation φ_a induces a mean ergodic operator T_{φ_a} on $C(\Gamma)$: If $a^{n_0} = 1$ for some $n_0 \in \mathbb{N}$, the operator T_{φ_a} is periodic (i.e. $T_{\varphi_a}^{n_0} = \mathrm{id}$) and therefore mean ergodic (see IV.D.3).

In the other case, every probability measure invariant under φ_a is invariant under φ_{a^n} for all $n \in \mathbb{N}$ and therefore under all rotations. By (D.5) the normalized Lebesgue measure is the unique probability measure having this property, and the assertion follows by (IV.8.b).

The previous example may also be understood without reference to the uniqueness of Haar measure: Let G be a compact group. The mapping

$$G \to \mathscr{L}_s(C(G)) : h \mapsto T_{\varphi_h} \qquad \text{(see II.2.2)}$$

is continuous, hence the orbits – as well as their convex hulls – of any operator T_{φ_h} are relatively (norm)compact in $C(G)$. Then apply (IV.4.c) to obtain the following result.

IV.9 Proposition:
Any rotation operator on $C(G)$, G a compact group, is mean ergodic.

Exercise: The fixed space of T_{φ_g} in $C(G)$, where φ_g is the rotation by g on the compact group G, is one-dimensional if and only if $\{g^k : k \in \mathbb{Z}\}$ is dense in G.

IV.D Discussion

IV.D.0 Proposition:
Assume that $a \in \Gamma$ is not a root of unity. The induced rotation operator T_{φ_a} is mean ergodic on the Banach space $R(\Gamma)$ of all bounded Riemann integrable functions on Γ (with sup-norm), and the (normalized) Riemann integral is the unique rotation invariant normalized positive linear form on $R(\Gamma)$.

Proof. First, we consider characteristic functions χ of "segments" on Γ and show that the Cesàro means

$$T_n\chi := \frac{1}{n}\sum_{i=0}^{n-1} T_{\varphi_a}^i \chi$$

converge in sup-norm $\|\cdot\|_\infty$
For $\varepsilon > 0$ choose $f_\varepsilon, g_\varepsilon \in C(\Gamma)$ such that

$$0 \leqslant f_\varepsilon \leqslant \chi \leqslant g_\varepsilon$$

and

$$\int_\Gamma (g_\varepsilon - f_\varepsilon)\, \mathrm{d}m < \varepsilon, \quad m \text{ Lebesgue measure on } \Gamma.$$

But $T := T_{\varphi_a}$ is mean ergodic (with one-dimensional fixed space) on $C(\Gamma)$, i.e.

$$T_n g_\varepsilon \xrightarrow{\|\cdot\|_\infty} \int_\Gamma g_\varepsilon \, \mathrm{d}m \cdot \mathbf{1}$$

and

$$T_n f_\varepsilon \xrightarrow{\|\cdot\|_\infty} \int_\Gamma f_\varepsilon \, \mathrm{d}m \cdot \mathbf{1}.$$

From $T_n f_\varepsilon \leqslant T_n \chi \leqslant T_n g_\varepsilon$ we conclude that $\|\cdot\|_\infty - \lim_{n\to\infty} T_n\chi$ exists and is equal to $\int_\Gamma \chi \, \mathrm{d}m \cdot \mathbf{1}$. Now, let f be a bounded Riemann integrable function on Γ. Then for every $\varepsilon > 0$ there exist functions g_1, g_2 being linear combinations of segments such that

$$g_1 \leqslant f \leqslant g_2 \quad \text{and} \quad \int_\Gamma (g_2 - g_1) \, \mathrm{d}m < \varepsilon,$$

and an easy calculation shows that

$$\|\cdot\|_\infty - \lim_{n\to\infty} T_n f = \Big(\int_\Gamma f \, \mathrm{d}m \Big) \cdot \mathbf{1}.$$

Finally, since the fixed space of T in $R(\Gamma)$, which is equal to the fixed space under all rotations on Γ, has dimension one, the mean ergodicity implies the one-dimensionality of the dual fixed space. ∎

The preceding result is surprising, has interesting applications (see IV.D.6) and is optimal in a certain sense:

Example 6: The rotation operator T_{φ_a} induced by φ_a, $a \in \Gamma$ not a root of unity, is mean ergodic

neither on (i) $L^\infty(\Gamma, \mathcal{B}, m)$

nor on (ii) $B(\Gamma)$, the space of all bounded Borel measurable

functions on Γ endowed with the sup-norm.

Proof. (i) The rotation φ_a is ergodic on Γ, hence the fixed space of $T := T_{\varphi_a}$ in $L^1(m)$ and a fortiori in $L^\infty(m)$ has dimension one. We show that the dual fixed space F' is at least two-dimensional: Consider $A := \{a^n : n \in \mathbb{Z}\}$ and $I := \{\check{f} \in L^\infty(m) :$ there is $f \in \check{f}$ vanishing on some neighbourhood (depending on f) of $A\}$. Then I is $\neq \{0\}$, T-invariant and generates a closed (lattice or algebra) ideal J in $L^\infty(m)$. From the definition follows that $TJ \subseteq J$ and $\mathbf{1} \notin J$. Consequently, there exists $\nu \in (L^\infty(m))'$ such that $\langle \mathbf{1}, \nu\rangle = 1$, but ν vanishes on J. The same is true for $T'\nu$ and $T_n'\nu$ for all $n \in \mathbb{N}$. By the weak* compactness of the dual unit ball the sequence $\{T_n'\nu\}_{n\in\mathbb{N}}$ has a weak* accumulation point ν_0. As in (IV.4), c $\Rightarrow$ d we show that $\nu_0 \in F'$. Since $\langle \mathbf{1}, \nu_0\rangle = 1$ and $\langle f, \nu_0\rangle = 0$ for $f \in J$, we conclude $0 \neq \nu_0 \neq m$.

(ii) Take a 0-1-sequence $(c_i)_{i\in\mathbb{N}_0}$ which is not Cesàro summable, i.e.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} c_i$$

does not exist. The characteristic function χ of the set

$$\{a^n : c_n = 1\}$$

is a Borel function for which

$$T_n\chi(a)$$

does not converge, hence the functions $T_n\chi$ do not converge in $B(\Gamma)$. ∎

IV.D.1. "Mean ergodic" vs. "ergodic":
The beginner should carefully distinguish these concepts. "Ergodicity" is a mixing property of an MDS $(X, \Sigma, \mu; \varphi)$ (or a statement on the fixed space of T_φ in $L^p(X, \Sigma, \mu)$), while "mean ergodicity" is a convergence property of the Cesàro means of a linear operator on a Banach space. More systematically we agree on the following terminology: "Ergodicity" of a linear operator $T \in \mathscr{L}(E)$, E Banach space, refers to the convergence of the Cesàro means T_n with respect to the uniform, strong or weak operator topology and such operators will be called "uniformly ergodic", "strongly ergodic", resp. "weakly ergodic". For $\{T^n : n \in \mathbb{N}\}$ bounded, it follows from Theorem (IV.4) that weakly ergodic and strongly ergodic operators coincide. Therefore and in order to avoid confusion with "strongly ergodic" transformations (see IX.D.4) we choose a common and different name for such operators and called them "mean ergodic". Here, the prefix "mean" refers to the convergence in the L^2-mean in von Neumann's original ergodic theorem (IV.1). "Uniform ergodicity" is a concept much stronger than "mean ergodicity" and will be discussed in Appendix W in detail.

IV.D.2. Mean ergodic semigroups:
Strictly speaking it is not the operator T which is mean ergodic but the semigroup $\{T^n : n \in \mathbb{N}_0\}$ of all powers of T. More precisely, in the bounded case, mean ergodicity of T is equivalent by (IV.4.d) to the following property of the semigroup $\{T^m : n \in \mathbb{N}_0\}$: the closed convex hull

$$\overline{\text{co}}\{T^n : n \in \mathbb{N}_0\}$$

of $\{T^n : n \in \mathbb{N}_0\}$ in $\mathscr{L}_s(E)$, which is still a semigroup, contains a zero element, i.e. contains P such that

$$SP = PS = P$$

for all $S \in \overline{\text{co}}\{T^n : n \in \mathbb{N}_0\}$ (Remark: $PT = TP = P$ is sufficient!). This point of view is well suited for generalizations which shall, be carried out in Appendix Y. As an application of this method we show that every root of a mean ergodic operator is mean ergodic, too.

Theorem: Let E be a Banach space and $S \in \mathscr{L}(E)$ be a mean ergodic operator with bounded powers. Then every root of S is mean ergodic.

Proof. Assume that $S := T^k$ is mean ergodic with corresponding projection P_S. Define $P := \big(\frac{1}{k}\sum_{j=0}^{k-1} T^j\big)P_S$ and observe that $P \in \overline{\text{co}}\{T^i : i \in \mathbb{N}_0\}$ and $TP := \big(\frac{1}{k}\sum_{j=0}^{k-1} T^{j+1}\big)P_S = P$, $(T^k P_S = P_S)$. Therefore, T is mean ergodic (see IV.4.d) and P is the projection corresponding to T. ∎

On the contrary, it is possible that no power of a mean ergodic operator is mean ergodic.

Example: Let $S : (x_n)_{n\in\mathbb{N}_0} \mapsto (x_{n+1})_{n\in\mathbb{N}_0}$ be the (left)shift on $\ell^\infty(\mathbb{N}_0)$ and take a 0-1-sequence $(a_n)_{n\in\mathbb{N}_0}$ which is not Cesàro summable.

For $k > 1$ we define elements $x_k \in \ell^\infty(\mathbb{N}_0)$:

$$x_k := (x_{k,n})_{n\in\mathbb{N}_0} \text{ by } \begin{cases} x_{k,n} := a_{\frac{n}{k}} & \text{for } n = ki \ (i \in \mathbb{N}_0) \\ x_{k,n} := -a_{\frac{n-1}{k}} & \text{for } n = ki+1 \ (i \in \mathbb{N}_0) \\ x_{k,n} := 0 & \text{otherwise.} \end{cases}$$

Consider the closed S-invariant subspace E generated by $\{S^i x_k : i \in \mathbb{N}_0,\ k > 1\}$ in $\ell^\infty(\mathbb{N}_0)$ and the restriction $T := S|_E$. By construction we obtain $\|T_n x_k\| \leqslant \frac{2}{n}$ for all $k > 1$. Consequently, T is mean ergodic with corresponding projection $P = 0$. On the other hand the sequence $\left(\frac{1}{m}\sum_{i=0}^{m-1} x_{k,ki}\right)_{m\in\mathbb{N}} = \left(\frac{1}{m}\sum_{i=0}^{m-1} a_i\right)_{m\in\mathbb{N}}$ is not convergent for $k > 1$, i.e. the Cesaró means $T_m^k(x_k)$ of the powers T^{ik}, $i \in \mathbb{N}$, applied to x_k, do not converge. Therefore, no power T^k $(k > 1)$ is mean ergodic.

References: Sine [1976].

IV.D.3. Examples:

(i) A linear operator T on the Banach space $E = \mathbb{C}$ is mean ergodic if and only if $\|T\| \leqslant 1$. Express this fact in a less cumbersome way!

(ii) The following operators $T \in \mathscr{L}(E)$, E a Banach space, are mean ergodic with corresponding projection P:

(a) T periodic with $T^{n_0} = \mathrm{id}$, $n_0 \in \mathbb{N}$, implies $P = \frac{1}{n_0}\sum_{i=0}^{n_0-1} T^i$.

(b) T with spectral radius $r(T) < 1$ (e.g. $\|T\| < 1$) implies $P = 0$.

(c) T has bounded powers and maps bounded sets into relatively compact sets.

(d) $T(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$ on ℓ^p, $1 < p < \infty$.

(e) $T(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots)$ on ℓ^p, $1 \leqslant p < \infty$.

(f) $Tf(x) = \int_0^x f(y)\,\mathrm{d}y$ for $f \in C([0,1])$.

(iii) The following operators are not mean ergodic:

(a) $Tf(x) = x \cdot f(x)$ on $C([0,1])$: $F = \{0\}$ but $\|T_n\| = 1$ for all $n \in \mathbb{N}$.

(b) $Tf(x) = f(x^2)$ on $C([0,1])$: $F = \langle \mathbf{1} \rangle$ but Dirac measures δ_0, δ_1 are contained in F'

(c) $T(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$ on ℓ^1: $F = \{0\}$ but $\|T_n(x_k)\| = \|(x_k)\|$ for $0 \leqslant (x_k) \in \ell^1$.

(d) $T(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots)$ on ℓ^∞: $0-1$-sequence which is not Cesàro summable.

IV.D.4. Convex combinations of mean ergodic operators:

Examples of "new" mean ergodic operators can be obtained by convex combinations of mean ergodic operators. Our first lemma is due to Kakutani (see Sakai [1977], 1.6.6)

Lemma 1: Let E be a Banach space. Then the identity operator id is an extreme point of the closed unit ball in $\mathscr{L}(E)$

Proof. Take $T \in \mathscr{L}(E)$ such that $\|\mathrm{id} + T\| \leqslant 1$ and $\|\mathrm{id} - T\| \leqslant 1$. Then the same is true for the adjoints: $\|\mathrm{id}' + T'\| \leqslant 1$ and $\|\mathrm{id}' - T'\| \leqslant 1$. For $f' \in E'$ define $f_1' := (\mathrm{id}' + T')f'$. resp. $f_2' := (\mathrm{id}' - T')f'$, and conclude $f' = \frac{1}{2}(f_1' + f_2')$ and $\|f_1'\|, \|f_2'\| \leqslant \|f'\|$. A soon as f' is an extreme point of the unit ball in E' we obtain $f' = f_1' = f_2'$ and hence $T'f' = 0$. But by the Krein-Milman theorem this is sufficient to yield $T' = 0$, and hence $T = 0$. Now assume that $\mathrm{id} = \frac{1}{2}(R + S)$

for contractions $R, S \in \mathscr{L}(E)$, and define $T := \mathrm{id} - R$. This implies $\mathrm{id} - T = R$ and $\mathrm{id} + T = 2\mathrm{id} - R = S$. By the above considerations it follows that $T = 0$, i.e. $\mathrm{id} = R = S$. ∎

Lemma 2: Let R, S be two commuting operators with bounded powers on a Banach space E, and consider

$$T := \alpha R + (1-\alpha)S$$

for $0 < \alpha < 1$. Then the fixed spaces $F(T)$, $F(R)$ and $F(S)$ of T, R and S are related by

$$F(T) = F(R) \cap F(S).$$

Proof. Only the inclusion $F(T) \subseteq F(R) \cap F(S)$ is not obvious. Endow E with an equivalent norm $\|x\|_1 := \sup\{\|R^n S^m x\| : n, m \in \mathbb{N}_0\}$, $x \in E$ and observe that R and S are contractive for the corresponding operator norm. From the definition of T we obtain

$$\mathrm{id}_{F(T)} = T|_{F(T)} = \alpha R|_{F(T)} + (1-\alpha)S|_{F(T)}$$

and $R|_{F(T)}, S|_{F(T)} \in \mathscr{L}(F(T))$, since R and S commute. Lemma 1 implies $R_{F(T)} = S|_{F(T)} = \mathrm{id}_{F(T)}$, i.e. $F(T) \subseteq F(R) \cap F(S)$. ∎

Now we can prove the main result.

Theorem:
Let E be a Banach space and R, S two commuting operators on E with $\|R^n\|, \|S^n\| \leqslant c$ for all $n \in \mathbb{N}$. If R and S are mean ergodic, so is every convex combination

$$T := \alpha R + (1-\alpha)S, \quad 0 \leqslant \alpha \leqslant 1.$$

Proof. Let $0 < \alpha < 1$. By Lemma 2 we have $F(T) = F(R) \cap F(S)$ and $F(T') = F(R') \cap F(S')$, and by (IV.4.e) it suffices to show that $F(R) \cap F(S)$ separates $F(R') \cap F(S')$: For $f' \neq g'$ both contained in $F(R') \cap F(S')$ there is $f \in F(R)$ with $\langle f, f' \rangle \neq \langle f, g' \rangle$. Since $SF(R) \subseteq F(R)$ we have $P_S f \in F(R) \cap F(S)$ where P_S denotes the projection corresponding to S. Consequently

$$\langle P_S f, f' \rangle = \langle f, P_S' f' \rangle = \langle f, P_{S'} f' \rangle = \langle f, f' \rangle \neq \langle f, g' \rangle = \langle P_S f, g' \rangle.$$

∎

The following corollaries are immediate consequences.

Corollary 1:
For T, R and S as above denote by P_R, resp. P_S the corresponding projections. Then the projection P_T corresponding to T is obtained as

$$P_T = P_R P_S = P_S P_R = \lim_{n\to\infty} (R_n S_n).$$

Corollary 2:
Let $\{R_i : 1 \leqslant i \leqslant m\}$ be a family of commuting mean ergodic operators with bounded powers. Then every convex combination $T := \sum_{i=1}^m \alpha_i R_i$ is mean ergodic.

IV.D.5. Mean ergodic operators with unbounded powers:
A careful examination of the proof of (IV.4) shows that the assumption

$$\|T^n\| \leqslant c \quad \text{for all } n \in \mathbb{N}_0,$$

may be replaced by the weaker requirements

$$\lim_{n\to\infty} \frac{1}{n}\|T^n\| = 0 \quad \text{and} \quad \|T_n\| \leqslant c \quad \text{for all } n \in \mathbb{N}.$$

The following example (Sato [1977]) demonstrates that such situations may occur. We define two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$.

$$a_1 = 1, \quad a_n = 2 \cdot 4^{n-2} \quad \text{for } n \geqslant 2$$

and $$b_n = \sum_{i=1}^{n} a_i = \tfrac{1}{3}(2 \cdot 4^{n-1} + 1) \quad \text{for } n \in \mathbb{N}.$$

Endow $$X := \{(n,i) : n \in \mathbb{N}, 1 \leqslant i \leqslant b_n\}$$

with the power set as σ-algebra Σ, and consider the measure μ defined by

$$\nu(\{(n,i)\}) := \begin{cases} 2^{1-n} & \text{if } 1 \leqslant i \leqslant a_n \\ \nu(\{(n-1, i-a_n)\}) & \text{if } a_n < i \leqslant b_n. \end{cases}$$

Observing that $\sum_{i=1}^{b_n} \nu(\{(n,i)\}) = 2^{n-1}$ we obtain a probability measure μ on Σ by

$$\mu(\{(n,i)\}) := 2 \cdot 4^{-n} \cdot \nu(\{(n,i)\}).$$

The measurable (not measure-preserving!) transformation

$$\varphi : (n,i) \mapsto \begin{cases} (n, i+1) & \text{for } 1 \leqslant i < b_n \\ (n+1, 1) & \text{for } i = b_n \end{cases}$$

on X induces the desired operator $T := T_\varphi$ on $L^1(X, \Sigma, \mu)$.

First, it is not difficult to see that $\|T^k\| = 2^n$ for $k = b_n, b_n+1, \ldots, b_{n+1}-1$. This shows that $\sup\{\|T^k\| : k \in \mathbb{N}\} = \infty$ and $\lim_{k\to\infty} \frac{1}{k}\|T^k\| = 0$.

Second, for $b_n + 1 \leqslant k \leqslant b_{n+1}$ we estimate the norm of the Cesaró means

$$\|T_k\| \leqslant \frac{1}{b_n \cdot \nu(\{(m+1,1)\})} \sum_{i=1}^{b_{n+1}} \nu(\{(n+1,i)\}) = \frac{2^n}{\frac{1}{3}(2 \cdot 4^{n-1}+1) \cdot 2^{-n}} \leqslant 6.$$

Finally, T is mean ergodic: With the above remark this follows from (IV.4.c) as in (IV.6).

IV.D.6. Equidistribution mod 1 (Kronecker, 1884; Weyl, 1916):
Mean ergodicity of an operator T with respect to the supremum norm in some function space is a strong and useful property. For example, if $T = T_\varphi$ for some $\varphi : X \to X$ and if $\chi = \mathbf{1}_A$ is the characteristic function of a subset $A \subseteq X$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} T^i \chi(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi(\varphi^i(x)), \quad x \in X$$

is the "mean frequency" of visits of $\varphi^n(x) \in A$. Therefore, if χ is contained in some function space on which T is mean ergodic (for $\|\cdot\|_\infty$), then this mean frequency exists (uniformly in $x \in X$). Moreover, if the corresponding projection P is one-dimensional, hence of the form $P = \mu \otimes \mathbf{1}$, the mean frequency of visits in A is equal to $\mu(A)$ for every $x \in X$.

This observations may be applied to the "irrational rotation" φ_a on Γ and to the Banach space $R(P)$ of all bounded Riemann integrable functions on Γ (see IV.D.0). Thus we obtain the following classical result on the equidistribution of sequences mod 1.

Theorem (Weyl, 1916):
Let $\xi \in [0,1]\backslash\mathbb{Q}$. The sequence $(\xi_n)_{n\in\mathbb{N}} := n\xi \mod 1$ is (uniformly) equidistributed in $[0,1]$, i.e. for every interval $[\alpha,\beta] \subseteq [0,1]$ holds

$$\lim_{n\to\infty} \frac{N(\alpha,\beta,n)}{n} = \beta - \alpha,$$

where $N(\alpha,\beta,n)$ denotes the number of elements $\xi_i \in [\alpha,\beta]$ for $1 \leqslant i \leqslant n$.

This theorem H. Weyl [1916] is the first example of number-theoretical consequences of ergodic theory. A first introduction into this circle: of ideas can be found in Jacobs [1972] or Hlawka [1979], while Furstenberg [1981] presents more and deeper results.

IV.D.7. Irreducible operators on L^p-spaces:
The equivalent statements of Proposition (IV.7) express essentially mean ergodicity and some "irreducibility" of the operator T_φ corresponding to the transformation φ. Using more operator theory, further generalizations should be possible (see also III.D.11). Here we shall generalize (IV.7) to FDSs $(E;T)$, where $E = L^p(X,\Sigma,\mu)$, $\mu(X) = 1$, $1 \leqslant p < \infty$, and $T \in \mathscr{L}(E)$ is positive satisfying $T\mathbf{1} = \mathbf{1}$ and $T'\mathbf{1} = \mathbf{1}$.

First, an operator-theoretical property naturally corresponding to "ergodicity" of a bi-measure-preserving transformation has to be defined.

Definition:
Let $(E;T)$ be an FDS as explained above. A set $A \in \Sigma$ is called *T-invariant* if $T\mathbf{1}_A(x) = 0$ for almost all $x \in X\backslash A$. The positive operator T is called irreducible if every T-invariant set has measure 0 or 1.

Remarks:
1. It is obvious that for an operator T_φ induced by an MDS $(X,\Sigma,\mu;\varphi)$ irreducibility of T_φ is equivalent to ergodicity of φ
2. If E is finite-dimensional, i.e. $X = \{x_1,\ldots,x_n\}$, and T is reducible, i.e. not irreducible, then there exists a non-trivial T-invariant subset A of X. After a permutation of the points in X we may assume $A = \{x_1,\ldots,x_k\}$ for $1 \leqslant k < n$. Then $T\mathbf{1}_A(x) = 0$ for all $x \in X\backslash A$ means that the matrix associated with T has the form

$$\begin{array}{c} \qquad\quad k \\ \left(\begin{array}{cc|c} \star & \star & \star \\ \star & \star & \star \\ \hline \multicolumn{2}{c|}{0} & \star \end{array}\right) k \end{array}$$

Proposition: Let $(E;T)$ be an FDS formed by $E = L^p(X,\Sigma,\mu)$, $\mu(X) = 1$, $1 \leqslant p < \infty$, and a positive operator T satisfying $T\mathbf{1} = \mathbf{1}$ and $T'\mathbf{1} = \mathbf{1}$. Then T is mean ergodic and the following statements are equivalent:

(a) T is irreducible.
(a′) The fixed space F of T is one-dimensional, i.e. $F = \langle \mathbf{1} \rangle$.
(b) The corresponding mean ergodic projection has the form $P = \mathbf{1} \otimes \mathbf{1}$.
(c) $\langle T_n f, g \rangle$ converges to $\int_X f \, \mathrm{d}\mu \cdot \int_X g \, \mathrm{d}\mu$ for every $f \in L^p(\mu)$, $g \in L^1(\mu)$.
(d) $\langle T_n \mathbf{1}_A, \mathbf{1}_B \rangle$ converges to $\mu(A) \cdot \mu(B)$ for every $A, B \in \Sigma$.
(e) $\langle T_n \mathbf{1}_A, \mathbf{1}_A \rangle$ converges to $\mu(A)^2$ for every $A \in \Sigma$.

Proof. Observe first that the assumptions $T\mathbf{1} = \mathbf{1}$ and $T'\mathbf{1} = \mathbf{1}$ imply that T naturally induces contractions on $L^1(\mu)$, resp. $L^\infty(\mu)$. From the Riesz convexity theorem (e.g. Schaefer [1974], V.8.2) it follows that $\|T\| \leqslant 1$. Consequently, T is mean ergodic by (IV.5) or (IV.6)

(a)⇒ (a′): Assume that the T-fixed space F contains a function f which is not constant. By adding an appropriate multiple of $\mathbf{1}$ we may obtain that f assumes positive and negative values. Its absolute value satisfies

$$|f| = |Tf| \leqslant T|f| \quad \text{and} \quad \int_X |f| \, \mathrm{d}\mu = \int_X T|f| \, \mathrm{d}\mu,$$

hence $|f| \in F$ and also $0 < f^+ := \frac{1}{2}(|f| + f) \in F$ and $0 < f^- := \frac{1}{2}(|f| - f) \in F$. Analogously we conclude that for every $n \in \mathbb{N}$ the function

$$f_n^+ := \inf(n \cdot f^+, \mathbf{1}) = \frac{1}{2}(n \cdot f^+ + \mathbf{1} - |n \cdot f^+ - \mathbf{1}|)$$

is contained in F. From the positivity of T we obtain

$$\mathbf{1}_A = \sup\{f_n^+ : n \in \mathbb{N}\} \in F$$

where $A := [f^+ > 0]$. Obviously, A is a non-trivial T-invariant set.

The implications (a′) ⇒ (b) ⇒ (c) ⇒ (d) ⇒ (e) follow as in the proof of (IV.7).

(e) ⇒ (a): If A is T-invariant the hypothesis $T\mathbf{1} = \mathbf{1}$ implies $T\mathbf{1}_A \leqslant \mathbf{1}_A$ and the hypothesis $T'\mathbf{1} = \mathbf{1}$ implies that $T\mathbf{1}_A = \mathbf{1}_A$. Therefore,

$$\langle T_n \mathbf{1}_A, \mathbf{1}_A \rangle = \langle T\mathbf{1}_A, \mathbf{1}_A \rangle = \langle \mathbf{1}_A, \mathbf{1}_A \rangle = \mu(A)$$

and the condition (e) implies $\mu(A) \in \{0,1\}$. ∎

IV.D.8. Ergodicity of the Markov shift:
As an application of (IV.7) we show that the ergodicity of the Markov shift $(\hat{X}, \hat{\Sigma}, \hat{\mu}; \tau)$ (see II.6) with transition matrix $T = (a_{ij})$ and strictly positive invariant distribution $\mu = (p_0, \ldots, p_{k-1})^\top$ can be characterized by an elementary property of the $k \times k$- matrix T.

Proposition: The following are equivalent:

(a) The transition matrix T is irreducible.
(b) The Markov shift $(\hat{X}, \hat{\Sigma}, \hat{\mu}; \tau)$ is ergodic.

Proof. As remarked (IV.7) ergodicity of τ is equivalent to the fact that the induced operator $\widehat{T}f := f \circ \tau$, $f \in L^1(\widehat{X}, \widehat{\Sigma}, \widehat{\mu})$, satisfies

$$\langle \widehat{T}_n \mathbf{1}_A, \mathbf{1}_B \rangle \to \widehat{\mu}(A) \cdot \widehat{\mu}(B)$$

for all $A, B \in \widehat{\Sigma}$, which are of the form

$$\begin{aligned} & A = [x_{-l} = a_{-l}, \ldots\ldots, x_l = a_l] \\ \text{and} \quad & B = [x_{-m} = b_{-l}, \ldots\ldots, x_l = b_m] \end{aligned}$$

with $a_j, b_j \in \{0, \ldots, k-1\}$.
For $n \in \mathbb{N}$ so large that $n' := n - (m + l + 1) \geqslant 0$, we obtain

$$\begin{aligned} \widehat{\mu}(\tau^n A \cap B) &= \widehat{\mu}[x_{-m} = b_{-m}, \ldots, x_m = b_m, x_{n-l} = a_{-l}, \ldots, x_{n+l} = a_l] \\ &= \sum_{c_1=0}^{k-1} \cdots \sum_{c_{n'}=0}^{k-1} \widehat{\mu}[x_{-m} = b_{-m}, \ldots, x_m = b_m, x_{m+1} = c_1, \ldots, x_{m+n'} = c_{n'}, \\ &\qquad\qquad x_{n-l} = a_{-l}, \ldots, x_{n+l} = a_l] \\ &= \sum_{c_1=0}^{k-1} \cdots \sum_{c_{n'}=0}^{k-1} \Big(p_{b_{-m}} \prod_{i=-m}^{m-1} t_{b_i b_{i+1}}\Big)\Big(t_{b_m c_1} \prod_{i=1}^{n'-1} t_{c_i c_{i+1}} t_{c_{n'} a_{-l}}\Big) \prod_{i=-l}^{l-1} t_{a_i a_{i+1}} \\ &= \widehat{\mu}(B)(T^{n-m-l})_{b_m a_{-l}} \cdot (p_{a_{-l}})^{-1} \widehat{\mu}(A). \end{aligned}$$

Thus $\lim_{n\to\infty} \langle \widehat{T}_n \mathbf{1}_A, \mathbf{1}_B \rangle = \widehat{\mu}(B) \cdot (\lim_{n\to\infty} T_n)_{b_m a_{-l}} \cdot (p_{a_{-l}})^{-1} \widehat{\mu}(A) = \widehat{\mu}(A) \cdot \widehat{\mu}(B)$, iff $(\lim_{n\to\infty} T_n)_{ij} = (\mathbf{1} \otimes \mu)_{ij} = p_j > 0$ for every $i, j \in \{0, \ldots, k-1\}$. By the assertion (b) in (IV.D.7, Proposition) the last condition is equivalent to the irreducibility of T. ∎

IV.D.9. A dynamical system which is minimal but not ergodic:
As announced in (III.D.10) we present a minimal TDS $(X; \varphi)$ such that the MDS $(X, \mathcal{B}, \mu; \varphi)$ is not ergodic for a suitable φ-invariant probability measure $\mu \in M(X)$.

Choose numbers $k_i \in \mathbb{N}$, $i \in \mathbb{N}_0$, such that

$$\begin{aligned} & (*) \qquad k_{i-1} \text{ divides } k_i \text{ for all } i \in \mathbb{N} \\ \text{and} \quad & (**) \qquad \sum_{i=1}^{\infty} \frac{k_{i-1}}{k_i} \leqslant \frac{1}{12}. \end{aligned}$$

For example we may take $k_i = 10^{(3^i)}$.
For $i \in \mathbb{N}$ define $Z_i := \{z \in \mathbb{Z} : |z - n \cdot k_i| \leqslant k_{i-1} \text{ for some } n \in \mathbb{Z}\}$ and observe that $\mathbb{Z} = \bigcup_{i\in\mathbb{N}} Z_i$, since k_i tends to infinity. Therefore

$$i(z) := \min\{j \in \mathbb{N} : z \in \mathbb{Z}_j\}$$

is well-defined for $z \in \mathbb{Z}$. Now take

$$a := (a_z)_{z\in\mathbb{Z}} \quad \text{with} \quad a_z := \begin{cases} 0 & \text{if } i(z) \text{ is even} \\ 1 & \text{if } i(z) \text{ is odd,} \end{cases}$$

and consider the shift

$$\tau : (x_z)_{z\in\mathbb{Z}} \mapsto (x_{z+1})_{z\in\mathbb{Z}}$$

on $\{0, 1\}^{\mathbb{Z}}$.

Proposition: With the above definitions and $X := \overline{\{\tau^s a : s \in \mathbb{Z}\}} \subseteq \{0,1\}^{\mathbb{Z}}$ the TDS $(X; \tau|_X)$ is minimal, and there exists a probability measure $\mu \in M(X)$ such that the MDS $(X, \mathcal{B}, \mu; \tau|_X)$ is not ergodic.

Proof. Clearly, X is T-invariant and $(X; \tau|_X)$ is a TDS. The (product) topology on $\{0,1\}^{\mathbb{Z}}$ – and on X – is induced by the metric

$$d((x_z),(y_z)) := \inf\Big\{\frac{1}{t+1} : x_z = y_z \text{for all } |z| < t\Big\}$$

The assertion is proved in several steps.

(i) Take $i \in \mathbb{N}$. By definition of the sets Z_j, $j = 1, \ldots, i$ the number $i(z)$ only depends on $z \mod k_i$ for $i(z) \leqslant i$, i.e. the finite sequence of 0's and 1's

$$a_{-i}, a_{-i+1}, \ldots, a_0, \ldots, a_{i-1}, a_i$$

reappears in $(a_z)_{z\in\mathbb{Z}}$ with constant period. Using the above metric d, the lemma in (III.D.5) shows that X is minimal

(ii) We prove that the induced operator $T := T_{\tau|_X}$, on $C(X)$ is not mean ergodic by showing that for the function $\in C(X)$ defined by

$$f((x_z)_{z\in\mathbb{Z}}) := x_1$$

the sequence $(T_n f(a))_{n\in\mathbb{N}}$ does not converge:

$$T_n f(a) = \frac{1}{n}\sum_{z=0}^{n-1} f(\tau^z a) = \frac{1}{n}\sum_{z=1}^{n} a_z,$$

and $\sum_{z=1}^n a_z$ is the number of those z $(1 \leqslant z \leqslant n)$ for which $i(z)$ is odd. Consider $n = k_i$ and observe that the set $\{1, \ldots, k_i\} \cap Z_j$ has exactly $\frac{k_i}{k_j}(2k_{j-1}+1)$ elements for $j = 1, \ldots, i$. Now

$$\sum_{j=1}^{i} \frac{k_i}{k_j}(2k_{j-1}+1) \leqslant \sum_{j=1}^{i} \frac{3k_{j-1}k_i}{k_j} \leqslant 3k_i \cdot \frac{1}{12} = \frac{k_i}{4} \qquad \text{(use } (**)),$$

i.e. $\{1, \ldots, k\} \cap \bigcup_{j=1}^{i} Z_j$ contains at most $\frac{k_i}{4}$ numbers. However $\{1, \ldots, k_i\} \subseteq Z_{i+1}$, hence

$$\Big|\{1, \ldots, k_i\} \cap (Z_{i+1} \backslash \bigcup_{j=1}^{i} Z_j)\Big| \geqslant \frac{3}{4}k_i,$$

and for all numbers in that intersection we have $i(z) = i + 1$. In conclusion, one obtains

$$|T_{k_{i+1}} f(a) - T_{k_i} f(a)| \geqslant \frac{1}{2}.$$

(iii) Using (IV.8) and (App.S), Theorem 1, we conclude from (ii) taht there exist at least two different τ-invariant probability measures $\mu_1, \mu_2 \in C(X)'$. For $\mu := \frac{1}{2}(\mu_1 + \mu_2)$ the MDS $(X, \mathcal{B}, \mu; \tau|_X)$ is not ergodic by (App.S).

∎

Remark: For examples on the 2-torus see Parry [1980], and on non-metrizable subsets of the Stone-Čech compactification of $\mathbb{N}$ see Rudin [1958] and Gait-Koo [1972].

References: Ando [1968], Gait-Koo [1972], Jacobs [1960], Parry [1980], Raimi [1964], Rudin [1958].

IV.D.10. Uniquely ergodic systems and the Jewett-Krieger theorem: For an MDS $(X, \Sigma, \mu; \varphi)$ and for $f \in L^p(X, \Sigma, \mu)$, the means

$$\frac{1}{n} \sum_{i=0}^{n-1} T_\varphi^i f$$

converge with respect to the L^p-norm for $1 \leqslant p < \infty$. Concerning the convergence for L^∞-norm (i.e. sup-norm) we don't have yet a definite answer, but know that in general the sup-norm is too strong to yield mean ergodicity of T_φ on $L^\infty(\mu)$. This was shown in example 6 in Lecture IV for any ergodic rotation φ_a on the unit circle Γ. On the other hand, in this same example there exist T_φ-invariant norm-closed subalgebras $\mathscr{A}$ of $L^\infty(X, \Sigma, \mu)$ which are dense in $L^1(X, \Sigma, \mu)$ and on which T_φ becomes mean ergodic (e.g. take $\mathscr{A} = C(\Gamma)$ or even $R(\Gamma)$, see (IV.D.0)). Such a subalgebra $\mathscr{A}$ is isomorphic to a space $C(Y)$ for some compact space Y and the algebra isomorphism on $C(Y)$ corresponding to T_φ is of the form T_ψ for some homeomorphism $\psi : Y \to Y$ (use the Gelfand-Neumark theorem (C.9) and (II.D.5)). The TDS $(Y; \psi)$ is minimal, since T_ψ is mean ergodic with one-dimensional fixed space, and therefore it possesses a unique ψ-invariant, strictly positive probability measure ν (see IV.8). Such systems will be called *uniquely ergodic*, since they determine a unique ergodic MDS. On the other hand it follows from the denseness of $\mathscr{A}$ in $L^1(\Gamma, \mathcal{B}, \mu)$ that the MDS $(\Gamma, \mathcal{B}, m; \varphi_a)$ is isomorphic to $(Y, \mathcal{B}, \nu; \psi)$ (use VI.2), a fact that will be expressed by saying that the original ergodic MDS is isomorphic to some MDS that is uniquely determined by a uniquely ergodic TDS. In fact, $(\Gamma, \mathcal{B}, m; \varphi_a)$ is uniquely ergodic since $\mathscr{A}$ can be chosen to be $C(\Gamma)$, but this choice is by no means unique and $\mathscr{A} = L^\infty(\Gamma, \mathcal{B}, m)$ would not work. Therefore we pose the following interesting question! Is every ergodic MDS isomorphic to an MDS determined by a uniquely ergodic TDS? As we have explained above, this question is equivalent to the following:

Problem: Let $(X, \Sigma, \mu; \varphi)$ be an ergodic MDS. Does there always exist a T_φ-invariant closed subalgebra $\mathscr{A}$ of $L^\infty(X, \Sigma, \mu)$

(i) T_φ is mean ergodic on $\mathscr{A}$, and
(ii) $\mathscr{A}$ is dense in $L^1(X, \Sigma, \mu)$?

The subsequent answer to this problem shows that the rotation $(\Gamma, \mathcal{B}, m; \varphi_a)$ is quite typical: Isomorphic uniquely ergodic systems always exist, but the algebra $L^\infty(\mu)$ is (almost) always too large for that purpose.

Lemma: For an ergodic MDS $(X, \Sigma, \mu; \varphi)$ the following assertions are equivalent:
(a) φ is mean ergodic on $L^\infty(X, \Sigma, \mu)$.
(b) $L^\infty(X, \Sigma, \mu)$ is finite dimensional.

Proof. In view of the representation theorem in (VI.D.6) it suffices to consider operators

$$T_\psi : C(Y) \to C(Y)$$

induced by a homeomorphism on an extremally disconnected space Y. By assumption (a), T_ψ is mean ergodic with one-dimensional fixed space and strictly positive invariant linear form ν. Prom (IV.8) it follows that ψ has to be minimal, and hence $\{\psi^k(y) : k \in \mathbb{Z}\}$ is dense in Y for every $y \in Y$. The lemma in (VI.D.6) implies that $\{\psi^k(y) : k \in \mathbb{Z}\}$ and hence $\{y\}$ is not a null set for the measure corresponding to ν. Therefore, $\{y\}$ must be open and the compact space Y is discrete.

■

Having seen that T_ψ is not mean ergodic on all of $L^\infty(\mu)$ one might try to find smaller subspaces on which mean ergodicity is guaranteed.
On the other hand

$$F(T) \oplus \overline{(\mathrm{id} - T_\varphi)L^\infty}$$

is the largest subspace of $L^\infty(\mu)$ on which T_φ is mean ergodic (use **??**). Unfortunately, this subspace is "never" a subalgebra. More precisely:

IV.D.11 Proposition:
For any ergodic MDS $(X, \Sigma, \mu; \varphi)$ the following assertions are equivalent:
(a) T_φ is mean ergodic on $L^\infty(\mu)$.
(b) $L^\infty(\mu)$ is finite dimensional.
(c) $\langle \mathbf{1} \rangle \oplus \overline{(\mathrm{id} - T_\varphi)L^\infty}$ is a subalgebra of $L^\infty(\mu)$.

Proof. It suffices to show that (c) implies (a). To that purpose we assume that the Banach algebra $L^\infty(\mu)$ is represented as $C(Y)$, Y compact, and the algebra isomorphism corresponding to T_φ is of the form $T_\psi : C(Y) \to C(Y)$ for some homeomorphism $\psi : Y \to Y$ and $\psi \neq \mathrm{id}$. Denote by $\mathrm{Fix}(\psi)$ the fixed point set of ψ. Then every function $f \in \overline{(\mathrm{id} - T_\psi)C(Y)}$ vanishes on $\mathrm{Fix}(\psi)$. Take $0 \neq g \in (\mathrm{id} - T_\psi)C(Y)$. Its square g^2 is contained in the subspace on which the means of T_ψ^i converge and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} T_\psi^i g^2 = \Big(\int_Y g^2 \,\mathrm{d}\nu\Big)\mathbf{1}_Y$$

for the strictly positive ψ-invariant measure ν. Therefore $\mathrm{Fix}(\psi)$ must be empty. It is now a simple application of Urysohn's lemma to show that $(\mathrm{id} - T_\psi)C(Y)$ separates the points in Y. By the Stone-Weierstrass theorem we obtain that $\langle \mathbf{1} \rangle \oplus (\mathrm{id} - T_\psi)C(Y)$ is dense in $C(Y)$ and therefore that T_ψ is mean ergodic on $L^\infty(\mu)$. ■

After these rather negative results it becomes clear that our task consists in finding "large" subalgebras contained in $\langle \mathbf{1} \rangle \oplus \overline{(\mathrm{id} - T_\varphi)L^\infty(\mu)}$. This has been achieved by Jewett [1970] (in the weak mixing case) and Krieger [1972]. Theirs as well as all other available proofs rest on extremely ingenious combinatorial techniques and we regret not being able to present a functional-analytic proof of this beautiful theorem.

Theorem (Jewett-Krieger, 1970):
Let $(X, \Sigma, \mu; \varphi)$ be an ergodic MDS. There exists a T_φ-invariant closed subalgebra $\mathscr{A}$ of $L^\infty(X, \Sigma, \mu)$, dense in $L^1(X, \Sigma, \mu)$, on which T_φ is mean ergodic.

Applying an argument similar to that used in the proof of (IV.D.0) the algebra of the above theorem can be enlarged and the corresponding structure spaces become totally disconnected. In conclusion we state the following answer to the original question.

Corollary:
Every separable ergodic $(X, \Sigma, \mu; \varphi)$ is isomorphic to an MDS determined by a uniquely ergodic TDS on a totally disconnected compact metric space.

References: Bellow-Furstenberg [1979], Denker [1973], Hansel [1974], Hansel-Raoult [1973], Jewett [1970], Krieger [1972], Petersen [1983].