IX. Mixing

Now we return to the investigation of "mixing properties" of dynamical systems, and the following experiment might serve as an introduction to the subsequent problems and results: two glasses are taken, one filled with red wine, the other with water, and one of the following procedures is performed once a minute.

A. The glasses are interchanged.
B. Nothing is done.
C. Simultaneously, a spoonful of the liquid in the right glass is added to the left glass and vice versa.

Intuitively, the process A is not really mixing because it does not approach any invariant "state", B is not mixing either because it stays in an invariant "state" which is not the equidistribution of water and wine, while C is indeed mixing. However, if in A the glasses are changed very rapidly it will appear to us, as if A were mixing, too.

It is our task to find correct mathematical models of the mixing procedures described above, i.e. we are looking for dynamical systems which are *converging* (in some sense) toward an *"equidistribution"*. The adequate framework will be that of MDSs (compare IV.8 and the remark proceeding it). More precisely, we take an MDS $(X, \Sigma, \mu; \varphi)$. The operator $T := T_\varphi$ induced on $L^p(X, \Sigma, \mu)$, $1 \leqslant p < \infty$, generates a compact semigroup

$$\mathscr{S} := \overline{\{T^n : n \in \mathbb{N}_0\}}$$

in $\mathscr{L}(L^p(\mu))$ for the weak operator topology. Moreover, if we assume $L^p(\mu)$ to be separable, this semigroup is metrizable (see VII.D.4).

The above experiments lead to the following mathematical questions:

convergence: under which conditions and in which sense do the powers T^n converge as $n \to \infty$?

If convergence of T^n holds in any reasonable topology then $P := \lim_{n\to\infty} T^n$ is projection onto the T-fixed space in $L^p(\mu)$. Therefore, the second property describing "mixing" may be expressed as follows.

equidistribution: under which conditions does the T-fixed space contain only the constant functions ?

One answer to these questions – in analogy to the case of the fast version of A – has already been given in Lecture IV, but will be repeated here.

IX.1 Theorem:
An MDS $(X, \Sigma, \mu; \varphi)$ is ergodic if and only if one of the following equivalent properties is satisfied:

(a) $T_n \to \mathbf{1} \otimes \mathbf{1}$ in the weak operator topology.

(b) $\langle T_n f, g\rangle \to \big(\int f \,\mathrm{d}\mu\big)\big(\int g \,\mathrm{d}\mu\big)$ for all $f, g \in L^\infty(X, \Sigma, \mu)$.

(c) $\displaystyle\frac{1}{n}\sum_{i=0}^{n-1} \mu(\varphi^{-i}A \cap B) \to \mu(A)\cdot\mu(B)$ for all $A, B \in \Sigma$.

(d) 1 is simple eigenvalue of T.

Proof. See (III.4) and (IV.7) including the remark. ∎

The really mixing case C is described by the (weak operator) convergence of the powers of T toward the projection $\mathbf{1}\otimes\mathbf{1}$. In analogy to the theorem above we obtain the following result.

IX.2 Theorem:
For an MDS $(X, \Sigma, \mu; \varphi)$ the following are equivalent.

(a) $T^n \to \mathbf{1}\otimes\mathbf{1}$ in the weak operator topology.
(b) $\langle T^n f, g\rangle \to \big(\int f \, \mathrm{d}\mu\big)\big(\int g \, \mathrm{d}\mu\big)$ for all $f, g \in L^\infty(X, \Sigma, \mu)$.
(c) $\mu(\varphi^{-n} A \cap B) \to \mu(A)\cdot\mu(B)$ for all $A, B \in \Sigma$

IX.3 Definition:
An MDS $(X, \Sigma, \mu; \varphi)$, resp. the transformation φ, satisfying one of the equivalent properties of (IX.2) is called *strongly mixing*.

Even if this concept perfectly describes the mixing-procedure C which seems to be the only one of some practical interest, we shall introduce one more concept:

Comparing the equivalences of (IX.1) and (IX.2) one observes that there is lacking a (simple) spectral characterization of strongly mixing. Obviously, the existence of an eigenvalue $\lambda \neq 1$, $|\lambda| = 1$, of T excludes the convergence of the powers T^n. Therefore, we may take this non-existence of non-trivial eigenvalues as the defining property of another type of mixing which possibly might coincide with strong mixing.

IX.4 Definition:
An MDS $(X, \Sigma, \mu; \varphi)$, resp. the transformation φ, is called *weakly mixing* if 1 is a simple and the unique eigenvalue of T in $L^p(X, \Sigma, \mu)$.

The results of Lecture VII applied to the compact semigroup

$$\mathscr{S} := \overline{\{T^n : n \in \mathbb{N}\}}^{\sigma}$$

will clarify the structural significance of this definition:
Let P be the projection corresponding to the mean ergodic operator T, i.e. $\{P\}$ is the minimal ideal of $\overline{\text{co}}\mathscr{S}$, and denote by $Q \in \mathscr{S}$ the projection generating the minimal ideal

$$\mathscr{K} = Q\mathscr{S}$$

of $\mathscr{S}$. The fact that 1 is a simple eigenvalue of T corresponds to the fact that $P = \mathbf{1}\otimes\mathbf{1}$, see (IV.7), hence

$$\mathbf{1}\otimes\mathbf{1} \in \overline{\text{co}}\mathscr{S}.$$

In (VII.5) we proved that Q is a projection onto the subspace spanned by all unimodular eigenvectors, hence

$$QE = PE = \langle \mathbf{1} \rangle.$$

From $Q \in \mathscr{S}$ it follows as in (IV.7) that

$$Q = P = \mathbf{1}\otimes\mathbf{1},$$

or equivalently

$$\{\mathbf{1}\otimes\mathbf{1}\} = \mathscr{K}$$

is the minimal ideal in $\mathscr{S}$. Briefly, weakly mixing systems are those for which the mean ergodic projection is already contained in $\mathscr{S}$ and is of the form $\mathbf{1} \otimes \mathbf{1}$. The following theorem shows in which way weak mixing lies between ergodicity (IX.1) and strong mixing (IX.2).

IX.5 Theorem:
Let $(X, \Sigma, \mu; \varphi)$ be an MDS. If $E := L^p(X, \Sigma, \mu)$, $1 \leqslant p < \infty$ is separable, the following assertions are equivalent:

(a) $T^{n_i} \to \mathbf{1} \otimes \mathbf{1}$ for the weak operator topology and for some subsequence $\{n_i\} \subseteq \mathbb{N}$.

(a′) $T^{n_i} \to \mathbf{1} \otimes \mathbf{1}$ for the weak operator topology and for some subsequence $\{n_i\} \subseteq \mathbb{N}$ having density 1.

(a″) $\frac{1}{n} \sum_{i=0}^{n-1} |\langle T^i f, g\rangle - \langle f, \mathbf{1}\rangle \cdot \langle \mathbf{1}, g\rangle| \to 0$ for all $f \in E$, $g \in E'$.

(b) $\langle T^{n_i} f, g\rangle \to \left(\int f \, \mathrm{d}\mu\right) \cdot \left(\int g \, \mathrm{d}\mu\right)$ for all $f, g \in L^\infty(X, \Sigma, \mu)$ and for some subsequence $\{n_i\} \subseteq \mathbb{N}$.

(c) $\mu(\varphi^{-n_i} A \cap B) \to \mu(A) \cdot \mu(B)$ for all $A, B \in \Sigma$ and for some subsequence $\{n_i\} \subseteq \mathbb{N}$.

(d) φ is weakly mixing.

(e) $\varphi \otimes \varphi$ is ergodic.

(f) $\varphi \otimes \varphi$ is weakly mixing.

IX.6 Remarks:

1. A subsequence $\{n_i\} \subseteq \mathbb{N}$ has density 1 if
$$\lim_{k \to \infty} \frac{1}{k} \left|\{n_i\} \cap \{1, 2, \ldots, k\}\right| = 1 \qquad \text{(see App.E.1)}.$$
2. The definition $\varphi \otimes \varphi : (x, y) \mapsto (\varphi(x), \varphi(y)$ makes $(X \times X, \Sigma \otimes \Sigma, \mu \otimes \mu; \varphi \otimes \varphi)$ an MDS.
3. (a) and (a′) are formally weaker than (IX.2.a), while (a″) (called "strong Cesàro convergence") is formally stronger than (IX.1.a).
4. "Primed" versions of (b) and (c) analogous to (a) are easily deduced.
5. Further equivalences are easily obtained by taking in (b) the functions f, g only from a subset of $L^\infty(\mu)$ which is total in $L^1(\mu)$, resp. in (c) the sets A, B only from a subalgebra generating Σ.

Proof. The general considerations above imply that (d) is equivalent to $\mathbf{1} \otimes \mathbf{1} \in \mathscr{S} = \overline{\{T^n : n \in \mathbb{N}\}}$. But by (VII.D.4), $\mathscr{S}$ is metrizable for the weak operator topology, hence there even exists a subsequence in $\{T^n : n \in \mathbb{N}\}$ converging to $\mathbf{1} \otimes \mathbf{1}$, which shows the equivalence of (a) and (d).

(a) $\Rightarrow$ (a′): We recall again that $\mathscr{S}$ is a commutative compact semigroup containing $\mathbf{1} \otimes \mathbf{1}$ as a zero, i.e. $R \cdot (\mathbf{1} \otimes \mathbf{1}) = \mathbf{1} \otimes \mathbf{1}$ for all $R \in \mathscr{S}$. Define the operator
$$\tilde{T} : C(\mathscr{S}) \to C(\mathscr{S})$$

induced by the rotation by T on $\mathscr{S}$, i.e.

$$\tilde{T}\tilde{f}(R) = \tilde{f}(TR) \quad \text{for } R \in \mathscr{S},\ \tilde{f} \in C(\mathscr{S}).$$

First, we show that this operator is mean ergodic with projection $\tilde{P}$ defined as

$$\tilde{P}\tilde{f}(R) = \tilde{f}(\mathbf{1} \otimes \mathbf{1}) \quad \text{for } R \in \mathscr{S},\ \tilde{f} \in C(\mathscr{S}).$$

Since multiplication by T is (uniformly) continuous on $\mathscr{S}$, the mapping from $\mathscr{S}$ into $\mathscr{L}(C(\mathscr{S}))$ which associates to every $R \in \mathscr{S}$ its rotation operator $\tilde{R}$ is well defined. Consider a sequence $(S_k)_{k\in\mathbb{N}}$ in $\mathscr{S}$ converging to S. Then $\tilde{S}_k\tilde{f}(R) = \tilde{f}(S_kR)$ converges to $\tilde{f}(SR) = \tilde{S}\tilde{f}(R)$ for all $R \in \mathscr{S}$, $\tilde{f} \in C(\mathscr{S})$. But the pointwise convergence and the boundedness of $\tilde{S}_k\tilde{f}$ imply weak convergence (see App.B.18), hence $\tilde{S}_k \to \tilde{S}$ in $\mathscr{L}_w(C(\mathscr{S}))$, and the mapping $S \mapsto \tilde{S}$ is continuous from $\mathscr{S}$ into $\mathscr{L}_w(C(\mathscr{S}))$. Therefore, from $T^{n_i} \to \mathbf{1} \otimes \mathbf{1}$ we obtain $\tilde{T}^{n_i} \to \widetilde{\mathbf{1} \otimes \mathbf{1}} = \tilde{P} \in \mathscr{L}_w(C(\mathscr{S}))$. Applying (IV.4.d) we conclude that the Cesàro means of $\tilde{T}^n$ converge strongly to $\tilde{P}$. Take now $f \in E$, $g \in E'$ and define a continuous function $\tilde{f} \in C(\mathscr{S})$ by

$$\tilde{f}(R) := |\langle Rf, h\rangle - \langle f, \mathbf{1}\rangle \cdot \langle \mathbf{1}, g\rangle|.$$

Obviously, we have $\tilde{P}\tilde{f}(T) = \tilde{f}(\mathbf{1} \otimes \mathbf{1}) = 0$. Therefore

$$0 = \lim_{n\to\infty} \tilde{T}_n\tilde{f}(T) \lim_{n\to\infty} = \frac{1}{n}\sum_{i=0}^{n-1} \big|\langle T^i f, g\rangle - \langle f, \mathbf{1}\rangle \cdot \langle \mathbf{1}, g\rangle\big|.$$

(a″) ⇒ (a): Since $\mathscr{S}$ is metrizable and compact for the topology induced from $\mathscr{L}_w(E)$, there exist countably many $f_k \in E$, $g_l \in E'$ such that the seminorms

$$p_{k,l}(R) := |\langle Rf_k, g_l\rangle|$$

define the topology on $\mathscr{S}$. By the assumption (a″) and by (App.E.2) for every pair (k, l) we obtain a subsequence

$$\{n_i\}^{k,l} \subseteq \mathbb{N}$$

with density 1, such that

$$\langle T^{n_i} f, g\rangle \to \langle f_k, \mathbf{1}\rangle \cdot \langle \mathbf{1}, g_l\rangle.$$

By (App.E.3) we can find a new subsequence, still having density 1, such that the concergence is valid simultaneously for all f_k and g_l. As usual, we apply (App.B.15) to obtain weak operator convergence.

(a′) ⇒ (a) is clear.

The equivalences (a) ⇔ (b) ⇔ (c) follow if we observe that the topologies we are considering in (b) and (c) are Hausdorff and weaker than the weak operator topology for which $\mathscr{S}$ is compact. Therefore, these topologies coincide on $\mathscr{S}$.

(c) ⇒ (f): Take $A, A', B, B' \in \Sigma$. For a suitable but fixed subsequence $(n_i) \subseteq \mathbb{N}$ $\mu(\varphi^{-n_i}A \cap B)$, resp. $\mu(\varphi^{-n_i}A' \cap B')$ converges to $\mu(A) \cdot \mu(B)$, resp. $\mu(A') \cdot \mu(B')$, as $n_i \to \infty$. This implies that

$$(\mu \otimes \mu)(((\varphi \otimes \varphi)^{-n_i} A \times A') \cap (B \times B')) = \mu(\varphi^{-n_i}A \cap B) \cdot \mu(\varphi^{-n_i}A' \cap B')$$

converges to $\mu(A) \cdot \mu(B) \cdot \mu(A') \cdot \mu(B') = (\mu \otimes \mu)(A \times A') \cdot (\mu \otimes \mu)(B \otimes B')$. Since the same assertion holds for disjoint unions of sets of the form $A \times A'$ we obtain the desired convergence for all sets in a dense subalgebra of $\Sigma \otimes \Sigma$. Using an argument as in the above proof of (a) ⇔ (b) ⇔ (c) we conclude that the MDS

$(X \times X, \Sigma \otimes \Sigma, \mu \otimes \mu; \varphi \otimes \varphi)$ satisfies a convergence property as (c), hence it is weakly mixing.

(f) $\Rightarrow$ (e) is clear.

(e) $\Rightarrow$ (d): Assume that $T_\varphi f = \lambda f$, $|\lambda| = 1$, for $0 \neq f \in L^1(\mu)$. Then we have $T_\varphi \bar{f} = \bar{\lambda}\bar{f}$ and, for the function $f \otimes \bar{f} : (x, y) \mapsto f(x) \cdot \bar{f}(y)$, $(x, y) \in X \times X$, we obtain $T_{\varphi\otimes\varphi}(f \otimes \bar{f}) = \lambda f \otimes \bar{\lambda}\bar{f} = |\lambda|^2(f \otimes \bar{f}) = f \otimes \bar{f}$. But 1 is a simple eigenvalue of $T_{\varphi\otimes\varphi}$ with eigenvector $\mathbf{1}_X \otimes \mathbf{1}_X$. Therefore we conclude $f = c\mathbf{1}_X$ and $\lambda = 1$ i.e. φ is weakly mixing. ∎

IX.7 Example: While it is easy to find MDSs which are ergodic but not weakly mixing (e.g. the rotation φ_a, $a^n \neq 1$ for all $n \in \mathbb{N}$, on the circle Γ has all powers of a as eigenvalues of T_{φ_a}), it remained open for a long time whether weak mixing implies strong mixing. That this is not the case will be shown in the next lecture.

The Bernoulli shift $B(p_0, \ldots, p_{k-1})$is strongly mixing as can be seen in proving (IX.2.c) for the rectangles, analogously to (III.5.ii).