27. Weak and strong mixing

In this section we will expand on the notions of ergodicity and recurrence and show that these are part of a hierarchy of mixing properties of measure preserving systems. Throughout we assume that the underlying measure is a probability.

27.1 Definitions and examples.

In order to introduce the desired notions in proper context, we start by recalling that a m.p.s. $(\mathscr{X},\mathcal{G},\mu,\varphi)$ with $\mu(\mathscr{X})=1$ is

(0) recurrent if

$$\forall B\in\mathcal{G}\colon\ \mu(B)>0\ \Rightarrow\ \exists n\geqslant 1\colon\ \mu(\varphi^{-n}(B)\cap B)>0 \tag{27.1}$$

(1) ergodic if

$$\forall A,B\in\mathcal{G}\colon\quad \frac1n\sum_{k=0}^{n-1}\mu\bigl(\varphi^{-k}(A)\cap B\bigr)\underset{n\to\infty}{\longrightarrow}\mu(A)\mu(B) \tag{27.2}$$

Recurrence follows automatically for finite measure spaces from the Poincaré Recurrence Theorem. For (27.2) we note that the Pointwise/Mean Ergodic Theorem gives

$$\frac1n\sum_{k=0}^{n-1}\mu\bigl(\varphi^{-k}(A)\cap B\bigr)=E\biggl(1_B\frac1n\sum_{k=0}^{n-1}1_A\circ\varphi^k\biggr)\underset{n\to\infty}{\longrightarrow}E(1_B\bar 1_A) \tag{27.3}$$

and ergodicity then forces $\bar 1_A=\int 1_A\mathrm{d}\mu=\mu(A)$ implying (27.2). Conversely, if A is an invariant event then (27.2) with $B:=A$ forces $\mu(A)=\mu(A)^2$; i.e., $\mu(A)\in\{0,1\}$.

We now introduce names for stronger versions of the convergence in (27.2):

Definition 27.1 (Weak and strong mixing) *A m.p.s. $(\mathscr{X},\mathcal{G},\mu,\varphi)$ with $\mu(\mathscr{X})=1$ is*

(2) weakly mixing *if*

$$\forall A,B\in\mathcal{G}\colon\quad \frac1n\sum_{k=0}^{n-1}\Bigl|\mu\bigl(\varphi^{-k}(A)\cap B\bigr)-\mu(A)\mu(B)\Bigr|\underset{n\to\infty}{\longrightarrow}0 \tag{27.4}$$

(3) strongly mixing *(or sometimes simply* mixing*) if*

$$\forall A,B\in\mathcal{G}\colon\quad \mu\bigl(\varphi^{-k}(A)\cap B\bigr)\underset{n\to\infty}{\longrightarrow}\mu(A)\mu(B) \tag{27.5}$$

For practical use it is useful to note that we only need to check the above properties for a suitable generating class of sets:

Lemma 27.2 *Let $\mathcal{S}$ is a semi-algebra such that $\mathcal{G}\subseteq\sigma(\mathcal{S})$. If any of* (27.2), (27.4) *and* (27.5) *holds for all $A,B\in\mathcal{S}$ then it holds for all $A,B\in\mathcal{G}$.*

Proof. Let $\mathcal{S}$ be a semialgebra and let $\mathcal{A}$ be the set of all finite disjoint unions of elements from $\mathcal{S}$. Then $\mathcal{A}$ is an algebra. This shows that if (27.2), (27.4) and (27.5) hold for all $A,B\in\mathcal{S}$, then in holds for all $A,B\in\mathcal{A}$. Next, by the construction of μ on $\sigma(\mathcal{A})$ via the outer measure, for each $A,B\in\sigma(\mathcal{A})$ and each $\epsilon>0$ there exists $A',B'\in\mathcal{A}$ such that $\mu(A\triangle A')<\epsilon$ and $\mu(B\triangle B')<\epsilon$. Then $|\mu(\varphi^{-k}(A)\cap B)-\mu(\varphi^{-k}(A')\cap B')|<2\epsilon$ and

$|\mu(A)\mu(B) - \mu(A')\mu(B')| < 2\epsilon$. Using this we extend the corresponding limit claim to all $A, B \in \sigma(\mathcal{A})$, which contains $\mathcal{G}$. □

Yet another simple fact is the invariance under isomorphisms:

Lemma 27.3 *The properties (1), (2) and (3) are isomorphism invariants: If $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ and $(\mathscr{X}', \mathcal{G}', \mu', \varphi')$ are m.p.s. $\Phi\colon \mathscr{X} \to \mathscr{X}'$ is a bi-measurable bijection such that $\mu' = \mu \circ \Phi^{-1}$ and $\Phi \circ \varphi = \varphi' \circ \Phi$, then if any of (1-3) above hold for one m.p.s. then it holds for the other m.p.s. as well.*

Proof. This is immediate from the assumptions. □

We proceed by examples of strongly mixing m.p.s.:

Lemma 27.4 *Left-shifts of i.i.d. sequences are strongly mixing.*

Proof. Let $\{X_k\}_{k\geqslant 0}$ be i.i.d. random variables defined on a suitable product space and let φ be the left shift. Since $\bigcup_{k\geqslant 1} \sigma(X_0, X_1, \dots, X_k)$ is an algebra, it suffices to prove (27.5) for $A, B \in \sigma(X_0, X_1, \dots, X_k)$. But this follows by noting that $\varphi^{-n}(A) \in \sigma(X_n, X_{n+1}, \dots)$ and so $\varphi^{-n}(A)$ and B are independent once $n > k$. □

Corollary 27.5 *The "continued fractions" m.p.s. (Example 20.10) and "Baker's transform" m.p.s. (Example 20.12) are strongly mixing.*

Proof. Since both of these systems are isomorphic to i.i.d. sequences with the left shift, this follows from Lemmas 27.3 and 27.4. □

Another example of a strongly mixing system is an irreducible, aperiodic countable-state Markov chain which we will discuss in the upcoming lectures. In both cases we can infer strong mixing by showing a yet stronger property. We start with a definition:

Definition 27.6 (Tail triviality) *A m.p.s. $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ with $\mu(\mathscr{X}) = 1$ is said to be*

(4) tail trivial *if μ is trivial on the tail σ-field*

$$\mathcal{T} := \bigcap_{n\geqslant 1} \{\varphi^{-n}(A)\colon A \in \mathcal{G}\} \tag{27.6}$$

in the sense

$$\forall B \in \mathcal{T}\colon \;\; \mu(B) \in \{0, 1\} \tag{27.7}$$

We say that φ is exact *if $\mathcal{T} = \{\emptyset, \mathscr{X}\}$. (This implies tail triviality.)*

Lemma 27.7 *Let $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ be a m.p.s. with $\mu(\mathscr{X}) = 1$. Then φ is tail-trivial if and only if*

$$\forall B \in \mathcal{G}\colon \quad \sup_{A\in\mathcal{G}} \Bigl| \mu\bigl(\varphi^{-n}(A) \cap B\bigr) - \mu(A)\mu(B) \Bigr| \underset{n\to\infty}{\longrightarrow} 0 \tag{27.8}$$

Proof. Suppose $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ is tail trivial. Denote $\mathscr{T}_n := \{\varphi^{-n}(A)\colon A \in \mathcal{G}\}$ and note that $\mathscr{T}_n \downarrow \mathscr{T}$. Then

$$\begin{aligned}\Bigl|\mu\bigl(\theta^{-n}(A) \cap B\bigr) - \mu(A)\mu(B)\Bigr| &= \Bigl|E\Bigl(1_{\varphi^{-n}(A)}\bigl(1_B - \mu(B)\bigr)\Bigr)\Bigr| \\ &= \Bigl|E\Bigl(1_{\varphi^{-n}(A)}\bigl(E(1_B|\mathscr{T}_n) - \mu(B)\bigr)\Bigr)\Bigr| \leqslant E\bigl|E(1_B|\mathscr{T}_n) - \mu(B)\bigr|. \end{aligned} \tag{27.9}$$

The Backward Lévy Theorem implies $\mathbb{E}(1_B|\mathscr{T}_n) \to \mathbb{E}(1_B|\mathscr{T})$ (both a.s. and in L^1) and by tail-triviality $\mathbb{E}(1_B|\mathscr{T}) = \mu(B)$ a.s. Hence, the right-hand side converges to zero uniformly in A, proving (27.8).

To prove the converse, pick $B \in \mathscr{T}$. Then for each $n \geqslant 1$ there exists $A_n \in \mathcal{G}$ such that $B = \varphi^{-n}(A_n)$. Using uniformity in A of the limit (27.8), this along with $\mu(A_n) = \mu(B)$ implies

$$\mu(B) - \mu(B)^2 = \mu\bigl(\varphi^{-n}(A_n) \cap B\bigr) - \mu(A_n)\mu(B) \underset{n\to\infty}{\longrightarrow} 0 \tag{27.10}$$

i.e., $\mu(B)^2 = \mu(B)$. Hence μ is trivial on $\mathscr{T}$. □

27.2 Separating the mixing concepts.

It is easy to check that (4) ⇒ (3) ⇒ (2) ⇒ (1) ⇒ (0). Our next goal is to demonstrate that each concept is different from the others. The following example shows that tail triviality is stronger than strong mixing:

Lemma 27.8 *Let $\{X_n\}_{n\in\mathbb{Z}}$ be Bernoulli(1/2) and let $Z_n := \sum_{m\geqslant 0} 2^{-m+1}X_{n-m}$. Then the left shift of $\{Z_n\}_{n\in\mathbb{Z}}$ is strongly mixing but not tail trivial, due to*

$$\mathscr{T} := \bigcap_{n\geqslant 0} \sigma(Z_n, Z_{n+1}, \dots) = \sigma(X_n\colon n \in \mathbb{Z}) = \sigma(Z_n\colon n \in \mathbb{Z}) \tag{27.11}$$

Proof. Since $\{X_k\}_{k\leqslant n}$ is determined by the binary expansions of Z_n, for each $n \in \mathbb{Z}$ we have $\sigma(Z_n) \supseteq \sigma(X_k\colon k \leqslant n)$ and so

$$\forall n \geqslant 0\colon \quad \sigma(Z_n, Z_{n+1}, \dots) \supseteq \sigma(X_n\colon n \in \mathbb{Z}) = \sigma(Z_n\colon n \in \mathbb{Z}) \tag{27.12}$$

Hence also $\mathscr{T} = \sigma(Z_n\colon n \in \mathbb{Z})$.

To see that the m.p.s. is strongly mixing, note that $\{Z_k \geqslant \ell 2^{-n}\}$, where $\ell, n \in \mathbb{N}$ is independent of $\{X_j\colon j \leqslant -n\}$ and thus of $\{Z_j\colon j \leqslant -n\}$. As Z_n only depends on $\{X_k\colon k \leqslant n\}$, for any choices of $m \in \mathbb{N}$, $\ell_{-m}, \dots, \ell_m \in \mathbb{N}$ and $\ell'_{-m}, \dots, \ell'_m \in \mathbb{N}$, the events

$$A := \bigcap_{i=-m}^{m} \{Z_i \geqslant \ell_i 2^{-n}\} \quad \text{and} \quad B := \bigcap_{i=-m}^{m} \{Z_i \geqslant \ell'_i 2^{-n}\} \tag{27.13}$$

obey

$$\forall k > n + 2m\colon\ \mu\bigl(\varphi^{-k}(A) \cap B\bigr) = \mu(A)\mu(B) \tag{27.14}$$

implying (27.5). Since the above events form a semialgebra that generates $\mathcal{G}$, Lemma 27.2 shows that the left shift of $\{Z_n\}_{n\in\mathbb{Z}}$ is strongly mixing. □

Next we give an example of an ergodic, but non-mixing transformation:

Lemma 27.9 *Irrational rotations of the unit circle are ergodic but not strongly mixing.*

Proof. Ergodicity follows from the observation that all invariant functions are constant, as established in the proof of Weyl's Equidistribution Theorem (Theorem 21.5). To invalidate strong mixing, let $A = B := [0, 1/4]$. As ergodicity forces $\{x + \alpha n \bmod 1 \colon n \in \mathbb{N}\}$ to be dense in $[0,1]$ for at least one $x \in [0,1)$, there exists $n_k \to \infty$ such that $n_k\alpha \bmod 1 \to 1/2$. But then $\varphi^{-n_k}(A)$ is an interval eventually contained in $[1/3, 2/3]$ and so $\varphi^{-n_k}(A) \cap B = \emptyset$ for k sufficiently large. As $\mu(A) = \mu(B) = 1/4$, the transformation is not strongly mixing. (An enhanced version of this argument shows that weak mixing fails as well.) □

It is not at all easy to demonstrate an example of a weakly mixing transformation that is not strongly mixing. An early example was that of Chacon in 1968. That being said, weakly mixing transformations are in a sense typical so we will at least give their characterization by a number of properties. We begin a lemma that demonstrates better the difference between convergence in Cezaro sense and ordinary convergence:

Lemma 27.10 (Koopman and von Neumann, 1932) *Let $f \colon \mathbb{N} \to \mathbb{R}$ be non-negative and bounded. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(k) = 0 \tag{27.15}$$

is equivlaent to

$$\exists E \subseteq \mathbb{N}\colon \quad \limsup_{n\to\infty} \frac{|E \cap \{1,\dots,n\}|}{n} = 0 \quad \wedge \quad \lim_{\substack{n\to\infty \\ n\notin E}} f(n) = 0 \tag{27.16}$$

Proof. To prove (27.16) ⇒ (27.15) we note that, if E is a set as in (27.16), then for any $\epsilon > 0$ there exists $n_0 \geqslant 1$ such that $\forall n \geqslant n_0 \colon n \in E \Rightarrow f(n) < \epsilon$. With M such that $f \leqslant M$ we then have

$$\frac{1}{n} \sum_{k=1}^{n} f(k) \leqslant M \frac{|E \cap \{1,\dots,n\}|}{n} + M\frac{n_0}{n} + \epsilon \tag{27.17}$$

As $n \to \infty$, the right-hand side converges to ϵ. Since ϵ was arbitrary, we thus get (27.15).

For (27.15) ⇒ (27.16) we have to work harder. For $m \geqslant 1$ set $E_m := \{n \colon f(n) > 1/m\}$ and note that $m \mapsto E_m$ is non-decreasing. Using that $1_{E_m}(k) \leqslant mf(k)$ (in conjunction with $f \to 0$ in Cezaro sense) implies that E_m has zero density. This allows us to recursively define a sequence $i_1 < i_2 < \dots$ such that

$$\forall n \geqslant i_{m-1} \colon \quad \frac{1}{n} \sum_{k=0}^{n-1} 1_{E_m}(k) < \frac{1}{m} \tag{27.18}$$

Now define

$$E := \bigcup_{m\geqslant 1} \big(E \cap (i_{m-1}, \dots, i_m]\big) \tag{27.19}$$

and let us check that E has the desired properties. First we claim that E has zero density. Indeed, if $i_m < n < i_{m+1}$ then $E \cap \{1, \dots, n\} \subseteq E_{m+1}$ and so

$$\frac{1}{n}\sum_{k=0}^{n-1} 1_E(k) \leqslant \frac{1}{n}\sum_{k=0}^{n-1} 1_{E_{m+1}}(k) < \frac{1}{m+1} \tag{27.20}$$

As $i_m \to \infty$ as $m \to \infty$, the left-hand side tends to zero as $n \to \infty$.

Next let us check that $f \to 0$ in $\mathbb{N} \smallsetminus E$. Again consider $i_m < n < i_{m+1}$ and note that $n \notin E$ implies $n \notin E_{m+1}$ giving us $f(n) \leqslant \frac{1}{m+1}$. Hence $\sup\{f(k) \colon k \geqslant i_m,\, k \notin E\} \leqslant \frac{1}{m+1}$. As this holds for all $m \geqslant 1$, we get $f \to 0$ in $\mathbb{N} \smallsetminus E$. □

We will refer to the above limit using the following concept:

Definition 27.11 (Limit in density) *Let $\{a_n\}_{n\geqslant 0}$ be an $\mathbb{R}$-valued sequence. We say that the sequence has a limit in density if*

$$\exists a \in \mathbb{R}\colon \quad \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} |a_n - a| = 0 \tag{27.21}$$

We will write this as $\text{D-lim}_{n\to\infty} f(n) = a$. *(The limit is necessarily unique.)*

Next we will show that weak mixing is preserved under taking products. Here we adopt the convention that if φ and ψ are two m.p.t.'s, then $\varphi \times \psi$ denotes the unique transformation induced thereby on the Cartesian product of the corresponding measure spaces. With this understanding we give:

Theorem 27.12 (Characterizations of weak mixing) *Let φ be a m.p.t. on $(\mathscr{X}, \mathcal{G}, \mu)$ with $\mu(\mathscr{X}) = 1$. Then the following are equivalent:*

(1) *φ is weakly mixing*
(2) *For all $A, B \in \mathcal{F}$,*

$$\underset{n\to\infty}{\text{D-lim}}\; \mu(\varphi^{-n}(A) \cap B) = \mu(A)\mu(B) \tag{27.22}$$

(3) *$\varphi \times \varphi$ is weakly mixing*
(4) *$\varphi \times \psi$ is ergodic for any ergodic ψ*
(5) *$\varphi \times \varphi$ is ergodic*

Proof. For (1) ⇔ (2), we invoke Lemma 27.10.

For (2) ⇒ (3) we use that $\text{D-lim}_{n\to\infty} f(n) = a$ and $\text{D-lim}_{n\to\infty} g(n) = b$ implies that $\text{D-lim}_{n\to\infty} f(n)g(n) = ab$ by the product law for ordinary limits and the fact that the union of two zero density sets is a zero density set. Then the implication follows from Lemma 27.10. (One can do the argument directly for Cezaro averages.)

For (3) ⇒ (4), let φ be a m.p.t. on $(\mathscr{X}_1, \mathcal{G}_1, \mu_1)$ and ψ an m.p.t. on $(\mathscr{X}_2, \mathcal{G}_2, \mu_2)$. We will first prove that (3) ⇒ (2). Indeed, assuming $\varphi \times \varphi$ is weakly mixing, for each $A, B \in \mathcal{G}_1$

we have

$$\begin{aligned}\operatorname*{D-lim}_{n\to\infty}\ &\mu_1(\varphi^{-n}(A)\cap B)\\ &=\operatorname*{D-lim}_{n\to\infty}\ \mu_1\otimes\mu_1\big((\varphi\times\varphi)^{-n}(A\times\mathscr{X}_1)\cap(B\times\mathscr{X}_1)\big)\\ &=\mu_1\otimes\mu_1(A\times\mathscr{X}_1)\mu_1\otimes\mu_1(B\times\mathscr{X}_1)=\mu_1(A)\mu_1(B)\end{aligned}\tag{27.23}$$

proving that φ is weakly mixing.

Next we will show that (3)$\wedge$(1) $\Rightarrow$ (4). For that assume ψ to be ergodic and pick $A_1, B_1 \in \mathcal{G}_1$ and $A_2, B_2 \in \mathcal{G}_2$. Our goal is to show that

$$\frac{1}{n}\sum_{k=0}^{n-1}\Big(\mu_1(\varphi^{-k}(A_1)\cap B_1)\mu_2(\psi^{-k}(A_2)\cap B_2)-\mu_1(A_1)\mu_1(B_1)\mu_2(A_2)\mu_2(B_2)\Big)\tag{27.24}$$

tends to zero as $n\to\infty$. By adding and subtracting appropriate terms, the absolute value of this quantity is bounded by the sum of two terms:

$$\mathrm{I}_n := \mu_1(A_1)\mu_1(B_1)\left|\frac{1}{n}\sum_{k=0}^{n-1}\Big(\mu_2(\psi^{-k}(A_2)\cap B_2)-\mu_2(A_2)\mu_2(B_2)\Big)\right|\tag{27.25}$$

which tends to zero as $n\to\infty$ because ψ is ergodic, and

$$\mathrm{II}_n := \frac{1}{n}\sum_{k=0}^{n-1}\Big|\mu_1(\varphi^{-k}(A_1)\cap B_1)-\mu_1(A_1)\mu_1(B_1)\Big|\mu_2(\psi^{-k}(A_2)\cap B_2)\tag{27.26}$$

which tends to zero because $\mu_2(\psi^{-k}(A_2)\cap B_2)\leqslant 1$ and φ is weakly mixing.

The proof of (4) $\Rightarrow$ (5) is easy: Pick $\mathscr{X}_2 := \{1\}$ and let μ_2 be the unique probability measure on $\mathscr{X}_2$ and let $\psi(1) := 1$. Then ψ is ergodic and since $\varphi\times\psi$ is isomorphic to φ, so is φ. But then (4) implies that $\varphi\times\varphi$ is ergodic as well.

The final step is the proof of (5) $\Rightarrow$ (2). Suppose that φ is an m.p.t. on $(\mathscr{X},\mathcal{G},\mu)$ with $\mu(\mathscr{X})=1$ and assume that $\varphi\times\varphi$ is ergodic. The same argument as above shows that also φ is ergodic. Then for all $A,B\in\mathcal{G}$ we have

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1}&\Big|\mu(\varphi^{-k}(A)\cap B)-\mu(A)\mu(B)\Big|^2\\ &=\frac{1}{n}\sum_{k=0}^{n-1}\mu\otimes\mu\big((\varphi\times\varphi)^{-k}(A\times A)\cap(B\times B)\big)\\ &\quad-2\mu(A)\mu(B)\Big(\frac{1}{n}\sum_{k=0}^{n-1}\mu\big(\varphi^{-n}(A)\cap B\big)\Big)+\mu(A)^2\mu(B)^2\end{aligned}\tag{27.27}$$

Thanks the ergodicity of φ and $\varphi\times\varphi$, each term appearing on the right-hand side converges to a constant times $\mu(A)^2\mu(B)^2$ and the total thus converges to zero. Then (2) follows again by invoking Lemma 27.10 for the sum on the left. (Alternatively, we could infer (1) by way of Jensen's inequality.) □

Here is a demonstration of how these conditions can be put to immediate use:

Corollary 27.13 *Irrational rotations of the unit circle are ergodic but not weakly mixing.*

Proof. Let $\mathscr{X} := [0,1)$ be endowed with Borel sets and the Lebesgue measure. Let $\varphi\colon [0,1) \to [0,1)$ be defined by $\varphi(x) := x + \alpha \bmod 1$ where $\alpha \in \mathbb{R} \smallsetminus \mathbb{Q}$. Weyl's Equidistribution Theorem implies that φ is ergodic. To show that φ is not weakly mixing, let $f\colon [0,1] \to \mathbb{R}$ be bounded and measurable. Then $g(x,y) := f(x-y)$ is $\varphi\times\varphi$-invariant for any f but g is constant a.e. only if f is constant a.e. Hence $\varphi\times\varphi$ is not ergodic. Theorem 27.12 implies that φ is not weakly mixing. □

27.3 Further properties of weak mixing.

We now give a few more interesting facts about weakly mixing transformations. We start with the following observation:

Lemma 27.14 *Suppose $(\mathscr{X},\mathcal{G},\mu)$ with μ a probability and $\mathcal{G}$ countably generated. Then an m.p.t. φ on $(\mathscr{X},\mathcal{G},\mu)$ is weakly mixing iff there exists $E \subseteq \mathbb{N}$ of zero density such that*

$$\forall A,B \in \mathcal{G}\colon\ \lim_{\substack{n\to\infty\\ n\notin E}} \mu\bigl(\varphi^{-n}(A)\cap B\bigr) = \mu(A)\mu(B), \tag{27.28}$$

(In short, the zero-density set E can be chosen uniformly for all A,B.)

Lemma 27.15 *If φ is weakly mixing then so is φ^n for all $n \geqslant 1$. The same is true also for $\sqrt[n]{\varphi}$ which is a notation for any m.p.t. ψ such that $\psi^n = \varphi$.*

Lemma 27.16 *An m.p.t. on $(\mathscr{X},\mathcal{G},\mu)$ with $\mu(\mathscr{X}) = 1$ is weakly mixing if and only if*

$$\forall f,g \in L^2\colon\ \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}\Bigl|\langle T^n f, g\rangle - \langle f,1\rangle\langle 1,g\rangle\Bigr| = 0 \tag{27.29}$$

where $Tf := f\circ\varphi$ and $\langle\cdot,\cdot\rangle$ is the inner product in L^2.

For our last result, recall that an invertible ergodic transformation corresponds to a unitary operator that may have many eigenvalues but each of these has to be simple. For weakly mixing transformation, we in turn get:

Theorem 27.17 (Spectral characterization of weak mixing) *Let φ be an measurably-invertible m.p.t. on $(\mathscr{X},\mathcal{G},\mu)$ with $\mu(\mathscr{X}) = 1$. Let $T\colon L^2 \to L^2$ be the unitary operator defined by $Tf := f\circ\varphi$. Then φ is weakly mixing if and only if 1 is the only eigenvalue of T and it is simple.*

Proof. Let us first assume that φ is weakly mixing and let $\lambda \in \mathbb{C}$ be an eigenvalue with eigenvector $f \in L^2$; i.e., $Tf = \lambda f$. Since T is unitary, we have $|\lambda| = 1$. Let $g \in L^2$. Then $\langle T^n f, g\rangle = \lambda^n\langle f,g\rangle$ and so $\text{D-lim}_{n\to\infty}\langle T^n f, g\rangle$ exists if and only if $\lambda = 1$.

For the converse we invoke the Spectral Theorem for unitary operators which says that, for each $f,g \in L^2$ there exists complex-valued (finite) measure $\nu_{f,g}$ concentrated on $\{\lambda \in \mathbb{C}\colon |\lambda| = 1\}$ such that

$$\forall n \geqslant 0\colon\ \langle T^n f, g\rangle = \int t^n \nu_{f,g}(\mathrm{d}t) \tag{27.30}$$

Then

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1}\bigl|\langle T^k f,g\rangle\bigr|^2 &= \frac{1}{n}\sum_{k=0}^{n-1}\int (t\bar{s})^k \nu_{f,g}(\mathrm{d}t)\bar{\nu}_{f,g}(\mathrm{d}s)\\ &= \int\biggl(\frac{1}{n}\sum_{k=0}^{n-1}(t\bar{s})^k\biggr)\nu_{f,g}(\mathrm{d}t)\bar{\nu}_{f,g}(\mathrm{d}s)\\ &= \int \frac{1}{n}\frac{1-(t\bar{s})^n}{1-t\bar{s}}\nu_{f,g}(\mathrm{d}t)\bar{\nu}_{f,g}(\mathrm{d}s)\end{aligned}\tag{27.31}$$

where the integrand is to be interpreted as 1 when $t\bar{s} = 1$. Since $|t\bar{s}| = 1$, the integrand converges to $1_{\{1\}}(t\bar{s})$. Since it is also bounded, the Bounded Convergence Theorem implies

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\bigl|\langle T^k f,g\rangle\bigr|^2 &= \int 1_{\{1\}}(t\bar{s})\nu_{f,g}(\mathrm{d}t)\bar{\nu}_{f,g}(\mathrm{d}\bar{s})\\ &= \int \bar{\nu}_{f,g}(\{t\})\nu_{f,g}(\mathrm{d}t) = \sum_t \bigl|\nu_{f,g}(\{t\})\bigr|^2\end{aligned}\tag{27.32}$$

where we used Tonelli's theorem.

If 1 is the only eigenvalue of T, then $\nu_{f,g}(\{t\}) = 0$ for all $t \neq 1$. It follows that if $f \in \mathrm{Ker}(1-T)^\perp$, then the right-hand side of (27.32) vanishes. Since 1 is also assumed to be simple, $f - \langle f,1\rangle 1 \in \mathrm{Ker}(1-T)^\perp$ and so we get (27.29). Specializing to $f := 1_A$ and $g := 1_B$ we get that φ is weakly mixing. □

Notice that this characterization allows us to give another proof of Corollary 27.13. Indeed, as is directly checked, the function $e_n(x) := \mathrm{e}^{2\pi i n x}$ is an eigenfunction of the rotation by α with eigenvalue $\mathrm{e}^{2\pi i n\alpha}$. This equals one for all $n \in \mathbb{Z}$ if and only if $\alpha \in \mathbb{Z}$.

Weakly mixing transformations are useful in the studies of multiple recurrence and other "fancy" topics in ergodic theory. We refer the reader to texts specializing on this subject.