Since m is fixed, it suffices to show that

$$\forall \underline{s},\underline{t}\in S^{m+1} \quad \frac{1}{N}\sum_{n=0}^{N-1}\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2 \longrightarrow \mu(\mathcal{C}(\underline{s}))\mu(\mathcal{C}(\underline{t}))$$

Note that

$$\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2 = \mu(T^{-n}\mathcal{C}(\underline{s})\cap\mathcal{C}(\underline{t}))$$

$$\overset{n>m}{=} p(t_0)\cdot P_{t_0t_1}\cdot\ldots\cdot P_{t_{m-1}t_m}\cdot P^{(n-m)}_{t_ms_0}\cdot P_{s_0s_1}\cdot\ldots\cdot P_{s_{m-1}s_m},$$

Hence

$$\frac{1}{N}\sum_{n=0}^{N-1}\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2 = \frac{1}{N}\sum_{n=0}^{m}\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2$$

$$+\,p(t_0)P_{t_0t_1}\cdot\ldots\cdot P_{t_{m-1}t_m}\cdot P_{s_0s_1}\cdot\ldots\cdot P_{s_{m-1}s_m}\cdot\underbrace{\frac{1}{N}\sum_{n=m+1}^{N-1}P^{(n-m)}_{t_ms_0}}_{\to Q_{t_ms_0}}$$

Now $pQ=p \Rightarrow p(s_0)=\sum_{t\in S}p(t)Q_{ts_0}\overset{(ii)}{=}Q_{\cdot s_0}\sum_{t\in S}p(t)=Q_{\cdot s_0}$

$$\therefore\ \frac{1}{N}\sum_{n=0}^{N-1}\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2 \xrightarrow{N\to\infty} p(t_0)\prod_{k=0}^{m-1}P_{t_kt_{k+1}}\cdot p(s_0)\prod_{\ell=0}^{m-1}P_{t_\ell t_{\ell+1}}$$

$$=\mu(\mathcal{C}(\underline{t}))\mu(\mathcal{C}(\underline{s})).$$

Exercise: Show that (ii) $\Leftrightarrow$ (v).

End of lecture 20

(Notions of) Mixing

Recall: $(X,\mathcal{B},\mu,T)$ is ergodic $\Leftrightarrow \forall f,g\in L^2(\mu)\ \ \frac{1}{N}\sum_{n=0}^{N-1}\langle f\circ T^n,g\rangle_2 \longrightarrow \langle f,1\rangle_2\langle 1,g\rangle_2$.

$$\Leftrightarrow \forall A,B\in\mathcal{B}\quad \frac{1}{N}\sum_{n=0}^{N-1}\mu(T^{-n}A\cap B)\longrightarrow\mu(A)\mu(B).$$

Definition ((Strong) mixing)

Let $(X,\mathcal{B},\mu,T)$ a pmps. T (or μ) is called (strong) mixing if

$$\forall A,B\in\mathcal{B}\qquad \mu(T^{-n}A\cap B)\xrightarrow{n\to\infty}\mu(A)\mu(B).$$

Lemma

Let $(X,\mathcal{B},\mu,T)$ a pmps. T is mixing if

$$\forall f,g\in L^2(X,\mu)\quad \langle U_T^n f,g\rangle_2 \longrightarrow \langle f,1\rangle_2\langle 1,g\rangle_2$$

Proof: Skipped; cf. characterization of ergodicity in the exercise sheets.

This lemma, for invertible T, gives an important connection between dynamical systems and unitary representation theory, which is about (strongly continuous) group homomorphisms

$$\pi: G\longrightarrow\mathcal{U}(H).$$

For example, if $\Gamma=SL_2(\mathbb{Z})$, $G=SL_2(\mathbb{R})$, then $\Gamma\backslash G$ admits a G-invariant probability measure and every element $g\in G$ induces an (effectively) mixing transformation, i.e.,

$$\forall\varphi,\psi\in C_c^\infty(\Gamma\backslash G)\quad |\langle\varphi\circ g,\psi\rangle_2-\langle\varphi,1\rangle_2\langle 1,\psi\rangle_2| \ll \xi(g)S(\varphi)S(\psi)$$

Unfortunately, this is basically the whole theory for this notion.

Definition (Weak mixing)

Let $(X,\mathcal{B},\mu,T)$ a pmps. T is weak mixing if

$$\forall A,B\in\mathcal{B}\quad \frac{1}{N}\sum_{n=0}^{N-1}|\mu(T^{-1}A\cap B)-\mu(A)\mu(B)|\xrightarrow{N\to\infty}0.$$

Weak mixing is surprisingly interesting/important which is due to the many equivalent characterisations. We need a few more notions:

Defininition (Density)

A subset $J\subseteq\mathbb{N}_0$ has density zero if

$$0=\lim_{N\to\infty}\frac{1}{N}|J\cap[0,N)|$$

Definition

A function $f \in L^2(X,\mu) - \{0\}$ is an eigenfunction of $(X,\mathcal{B},\mu,T)$ (or just of T) if $\exists \lambda \in \mathbb{C}$ $U_T f = \lambda f$. The system $(X,\mathcal{B},\mu,T)$ (or just T) has continuous spectrum if the constant functions are the only eigenfunctions.

Theorem

Let $(X,\mathcal{B},\mu,T)$ a pmps. The following are equivalent.

(i) T is weak mixing.

(ii) $T\times T$ is ergodic w.r.t. $\mu\otimes\mu$.

(iii) $T\times T$ is weak mixing w.r.t. $\mu\otimes\mu$.

(iv) For every ergodic system $(Y,\mathcal{C},\sigma,S)$, the product system

$$(X\times Y, \mathcal{B}\otimes\mathcal{C}, \mu\otimes\sigma, T\times S)$$

is ergodic.

(v) T has continuous spectrum.

(vi) $\forall A,B\in\mathcal{B}\ \exists J_{A,B}\subseteq\mathbb{N}_0$ of density zero such that

$$\lim_{n\in\mathbb{N}_0 - J_{A,B}} \mu(T^{-n}A\cap B) = \mu(A)\mu(B).$$

Corollaries: • Let $(X_i,\mathcal{B}_i,\mu_i,T_i)$ weak mixing. Then

$$(X_1\times\dots\times X_n, \mathcal{B}_1\otimes\dots\otimes\mathcal{B}_n, \mu_1\otimes\dots\otimes\mu_n, T_1\times\dots\times T_n)$$

is weak-mixing.

Proof: We prove it for $n=2$. The claim then follows by induction.

Let T,S weak mixing and R ergodic. Then $S\times R$ is ergodic, hence

$$T\times(S\times R) = (T\times S)\times R$$

is ergodic. Since R was arbitrary, $T\times S$ is weak mixing. $\blacksquare$

• Let $(X,\mathcal{B},\mu,T)$ weak mixing. Let $n\in\mathbb{N}$. Then $(X,\mathcal{B},\mu,T^n)$ is weak mixing.

Proof: Let $A,B\in\mathcal{B}$ and $J_{A,B}\subseteq\mathbb{N}$ s.t.

$$\lim_{m\in\mathbb{N}_0 - J_{A,B}} \mu(T^{-m}A\cap B) = \mu(A)\mu(B)$$

Then

$$\lim_{k\in\mathbb{N}_0 - \frac{1}{n}J_{A,B}} \mu((T^n)^{-k}A\cap B) = \mu(A)\mu(B).$$

Hence it suffices to show that $\frac{1}{n}J_{A,B}\cap\mathbb{N}_0$ has density zero.

Indeed

$$\frac{1}{N}\left|\frac{1}{n}J_{A,B}\cap\mathbb{N}_0\cap[0,N)\right| = \frac{1}{N}\left|J_{A,B}\cap n\mathbb{N}_0\cap[0,nN)\right|$$

$$\leq \frac{n}{nN}\left|J_{A,B}\cap[0,nN)\right| \xrightarrow{N\to\infty} 0.$$

For the proof of the theorem, we need the following Lemma:

Lemma

Let $(a_n)_{n\in\mathbb{N}}\in[0,\infty)^{\mathbb{N}}$ bounded. The following are equivalent.

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

(ii) $\exists J\subseteq\mathbb{N}$ of density zero s.t. $a_n\longrightarrow 0$ as $n\to\infty$ in $\mathbb{N}-J$.

(iii) $\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} a_n^2 = 0$.

Proof: (i) $\Rightarrow$ (ii): Let $J_k = \{n\in\mathbb{N} : a_n > \frac{1}{k}\}$. Then $J_k\subseteq J_{k+1}$ and

$$\frac{1}{N}|J_k\cap[0,N)| = \frac{k}{N}\cdot\frac{1}{k}|J_k\cap[0,N)|$$

$$< \frac{k}{N}\sum_{n\in J_k\cap[0,N)} a_n \xrightarrow{N\to\infty} 0;$$

in particular J_k has density zero. Set $N_0 := 0$ and inductively choose a w.l.o.g increasing sequence $N_k < N_{k+1} < \dots$ s.t.

$$\forall N\geq N_k \quad \frac{1}{N}|J_k\cap[0,N)|\leq\frac{1}{k}.$$

Set $J := \bigcup_{k=1}^{\infty} J_k\cap[N_k,N_{k+1})$.

Claim

J has density zero and $a_n\longrightarrow 0$ as $n\to\infty$ in $\mathbb{N}-J$.

Proof of claim:

Density: Let $N\in\mathbb{N}$ and choose $K\in\mathbb{N}$ s.t. $N\in[N_K,N_{K+1})$. Then

$$\frac{1}{N}|J\cap[0,N)| = \frac{1}{N}\sum_{k=1}^{K-1}|J_k\cap[N_k,N_{k+1})| + \frac{1}{N}|J_K\cap[N_K,N)| \leq \frac{1}{N}|J_K\cap[0,N)|$$

Hence $\forall N \in [N_k, N_{k+1})$ $\frac{1}{N}|J \cap [0,N)| \leq \frac{1}{K}$.

Convergence of a_n to 0: If $n \geq N_{k+1}$ and $n \notin J$, then $a_n \leq \frac{1}{k+1}$. Indeed, let $\ell \geq k+1$ such that $n \in [N_\ell, N_{\ell+1})$. Then $n \notin J \Rightarrow n \notin J_\ell \Rightarrow a_n \leq \frac{1}{\ell} \leq \frac{1}{k+1}$.

(ii)$\Rightarrow$(i): Fix $R>0$ s.t. $\forall n \in \mathbb{N}$ $a_n < R$ (which exists by assumption). We know that $a_n \longrightarrow 0$ as $n \longrightarrow \infty$ in $\mathbb{N}-J$. Hence, given $k \in \mathbb{N}$, let $N_k \in \mathbb{N}$ s.t.

- $\forall n \in \mathbb{N}-J$ $n \geq N_k \Rightarrow a_n < \frac{1}{k}$
- $\forall n \in \mathbb{N}$ $N \geq N_k \Rightarrow \frac{1}{N}|J \cap [0,N)| < \frac{1}{k}$.

Suppose $N \geq kN_k$, then

$$\frac{1}{N}\sum_{n=0}^{N-1} a_n = \frac{1}{N}\left(\sum_{n=0}^{N_k-1} a_n + \sum_{n \in J \cap [N_k,N)} a_n + \sum_{n \in [N_k,N)-J} a_n\right)$$

$$< \frac{1}{N}\left(N_k R + |J \cap [N_k,N)|\cdot R + N \cdot \frac{1}{k}\right) \leq \frac{2R+1}{k} \xrightarrow{k \to \infty} 0.$$

(iii)$\Leftrightarrow$(ii): $\exists J \subseteq \mathbb{N}$ of density 0 s.t. $a_n^k \longrightarrow 0$ as $n \longrightarrow \infty$ in $\mathbb{N}-J$ if and only if $a_n \longrightarrow 0$ as $n \longrightarrow \infty$ in $\mathbb{N}-J$. $\square$

Proof of the theorem

(i)$\Leftrightarrow$(vi): this is an immediate application of the lemma.

(vi)$\Rightarrow$(iii): Using the characterization of weak mixing given by

$$T \text{ is weak mixing} \Leftrightarrow \forall f,g \in L^2(\mu)\ \frac{1}{N}\sum_{n=0}^{N-1}|\langle f \circ T^n, g\rangle_2 - \langle f,1\rangle_2\langle 1,g\rangle_2| \longrightarrow 0$$

and after redoing the exercise that it suffices to prove this for dense set of functions, e.g., linear combinations of simple functions, we note that by definition of the product σ-algebra, it suffices to show that

$\forall A_1,A_2,B_1,B_2 \in \mathcal{B}\ \exists J_{A_1\times A_2, B_1 \times B_2} \subseteq \mathbb{N}$ of density 0 s.t.

$$\mu \otimes \mu((T\times T)^{-n}(A_1 \times A_2) \cap B_1 \times B_2) \longrightarrow \mu(A_1)\mu(A_2)\mu(B_1)\mu(B_2)$$

as $n \longrightarrow \infty$ in $\mathbb{N} - J_{A_1\times A_2, B_1\times B_2}$.

Let $J_1 = J_{A_1,B_1}$, $J_2 = J_{A_2,B_2}$ as in (vi). Then $J = J_1 \cup J_2$ has density zero and

$$\mu \otimes \mu((T\times T)^{-n}(A_1\times A_2) \cap B_1 \times B_2) = \mu(T^{-n}A_1 \cap B_1)\mu(T^{-n}A_2 \cap B_2)$$

implies the desired convergence as $n \longrightarrow \infty$ in $\mathbb{N}-J$.

(iii)$\Rightarrow$(i): Suppose $T\times T$ is weak mixing. Let $A, B \in \mathcal{B}$, $A' = A\times X$, $B' = B \times X$. Then

$$\mu\otimes\mu((T\times T)^{-n}A' \cap B') = \mu\otimes\mu(T^{-n}A \times X \cap B \times X) = \mu(T^{-n}A \cap B)$$

and

$$\mu\otimes\mu(A') = \mu(A), \quad \mu\otimes\mu(B') = \mu(B).$$

Hence the claim.

(i)$\Rightarrow$(iv): Let $A_1, B_1 \in \mathcal{B}$, $A_2, B_2 \in \mathcal{C}$. Then

$$\frac{1}{N}\sum_{n=0}^{N-1} \mu\otimes\sigma((T\times S)^{-n}(A_1 \times A_2) \cap B_1 \times B_2)$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \mu(T^{-n}A_1 \cap B_1)\sigma(S^{-n}A_2 \cap B_2)$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \mu(A_1)\mu(B_1)\sigma(S^{-n}A_2 \cap B_2)$$

$$+ \frac{1}{N}\sum_{n=0}^{N-1}(\mu(T^{-n}A_1 \cap B_1) - \mu(A_1)\mu(B_1))\sigma(S^{-n}A_2 \cap B_2)$$

$$\leq \frac{1}{N}\sum_{n=0}^{N-1}|\mu(T^{-n}A_1 \cap B_1) - \mu(A_1)\mu(B_1)| \xrightarrow{N\to\infty} 0$$

$$\xrightarrow{N \to \infty} \mu\otimes\sigma(A_1 \times A_2)\mu\otimes\sigma(B_1 \times B_2).$$

(iv)$\Rightarrow$(ii): weak mixing $\Rightarrow$ ergodic.

(ii)$\Rightarrow$(vi): Let $A, B \in \mathcal{B}$. We'll show that

$$\frac{1}{N}\sum_{n=0}^{N-1}|\mu(T^{-n}A \cap B) - \mu(A)\mu(B)|^2 \xrightarrow{N\to\infty} 0,$$

which by the lemma implies (vi). Note

$$|\mu(T^{-n}A \cap B) - \mu(A)\mu(B)|^2 = \mu\otimes\mu((T\times T)^{-n}(A \times A) \cap B \times B)$$
$$- 2\mu(A)\mu(B)\mu(T^{-n}A \cap B)$$
$$+ \mu(A)^2\mu(B)^2$$

Since $(X, \mathcal{B}, \mu, T)$ is a factor of $(X\times X, \mathcal{B}\otimes\mathcal{B}, \mu\otimes\mu, T\times T)$, it is ergodic by the assumptions of (ii). Hence

$$\frac{1}{N}\sum_{n=0}^{N-1}|\mu(T^{-n}A \cap B) - \mu(A)\mu(B)|^2 \xrightarrow{N\to\infty} 0.$$

since ergodicity implies

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu\otimes\mu\left((T\times T)^{-n}(A\times A)\cap B\times B\right)\xrightarrow{N\to\infty}\mu\otimes\mu(A\times A)\mu\otimes\mu(B\times B)=\mu(A)^2\mu(B)^2$$

and

$$\mu(A)\mu(B)\cdot\frac{1}{N}\sum_{n=0}^{N-1}\mu(T^{-n}A\cap B)\xrightarrow{N\to\infty}\mu(A)^2\mu(B)^2.$$

End of lecture 21

(ii) $\Rightarrow$ (v): Suppose $f\in L^2(\mu)$ is an eigenfunction of T with eigenvalue $\lambda\in\mathbb{C}$. Since U_T is an isometry, we have

$$|\lambda|\,\|f\|_2=\|\lambda f\|_2=\|U_Tf\|_2=\|f\|_2,$$

i.e., $\lambda\in\mathbb{S}^1$. Define $\varphi\in L^2(\mu\otimes\mu)$ by

$$\varphi(x_1,x_2)=f(x_1)\overline{f(x_2)}.$$

Then

$$U_{T\times T}\varphi(x_1,x_2)=U_Tf(x_1)\overline{U_Tf(x_2)}=\varphi(x_1,x_2),$$

i.e., φ is an eigenfunction for eigenvalue 1 / invariant. Since $T\times T$ is ergodic by assumption, φ is constant and thus so is f (by Fubini).

(v) $\Rightarrow$ (ii) is more difficult (i.e., deferred)

Excursion to compact operators

Definition

A Banach space is a complete normed vector space over $\mathbb{C}$.

Definition (Compact operator)

Let V,W Banach spaces and $A:V\longrightarrow W$ linear. A is a compact operator if $\overline{A(B_1^V(0))}\subseteq W$ is compact.

Example Let $V=C^1([0,1])$ with $\|f\|:=\|f\|_\infty+\|f'\|_\infty$. Let $W=C^0([0,1])$ with norm $|||f|||:=\|f\|_\infty$. Then the "identity" map

$$i:(C^1([0,1]),\|\cdot\|)\hookrightarrow(C^0([0,1]),|||\cdot|||)$$

is compact by Arzela-Ascoli. To this end note that

$$i(B_1^V(0))\subseteq\{f\in W: |||f|||\le 1,\ f\text{ is }1\text{-Lipschitz}\}.$$

Definition

Let V a Banach space. V is a Hilbert space if the norm is induced by an inner product.

Definition

Let V a Hilbert space. A linear operator $A:V\to V$ is self-adjoint if $\forall v_1,v_2\in V$ $\langle v_1,Av_2\rangle=\langle Av_1,v_2\rangle$.

Theorem (without proof – spectral theorem for self-adj compact operators)

Let V be a separable Hilbert space and let $A:V\to V$ self-adjoint compact. Then there exists an orthonormal basis consisting of eigenvectors of A, i.e., A is (in particular) completely diagonalizable. The eigenvalues are real and for every non-zero eigenvalue λ of A, the eigenspace for λ is finite dimensional. Moreover, the eigenvalues are bounded and if $\lambda_1>\lambda_2>\cdots$, then $\lambda_n\xrightarrow{n\to\infty}0$.

Note: $\sigma:\ell^2(\mathbb{Z})\longrightarrow\ell^2(\mathbb{Z})$, $(x_n)_{n\in\mathbb{N}}\longmapsto(x_{n+1})_{n\in\mathbb{N}}$ has no eigenvectors.

Proposition (Hilbert-Schmidt)

Let $(X,\mathcal{B},\mu)$ and $(Y,\mathcal{C},\sigma)$ σ-finite. Let $k\in L^2(\mu\otimes\sigma)$. Then the Hilbert-Schmidt integral operator $K:L^2(\mu)\longrightarrow L^2(\sigma)$ defined by

$$K(f)(y)=\int_X f(x)k(x,y)\,d\mu(x)\qquad\sigma\text{-a.s.}$$

is compact. It is self-adjoint if and only if $k(x_1,x_2)=\overline{k(x_2,x_1)}$ $\mu\otimes\mu$-a.s.

Proof that (v) $\Rightarrow$ (ii):

Suppose $T\times T$ is not ergodic. Then there exists $B\subseteq X\times X$ s.t. $\mu\otimes\mu(B)\in(0,1)$ and $(T\times T)^{-1}B=B$. Define

$$k_1(x,y)=\frac{1}{2}\left(\mathbb{1}_B(x,y)+\mathbb{1}_B(y,x)\right)$$

$$k_2(x,y)=\frac{1}{2i}\left(\mathbb{1}_B(x,y)-\mathbb{1}_B(y,x)\right).$$

Then $k_1(x,y)=\overline{k_1(y,x)}$ and $k_2(x,y)=\overline{k_2(y,x)}$. Moreover

$$\mathbb{1}_B=k_1+ik_2$$

implies that at least one of the k_i is non-constant (since $\mathbb{1}_B$ is non-constant).

Let k be a non-constant element in $\{k_i-\int k_i\,d\mu\otimes\mu : i=1,2\}$

Note that $U_{T\times T}k=k$, $k\neq 0$, $\int k=0$. Let $A:L^2(\mu)\longrightarrow L^2(\mu)$,

$$f\longmapsto\int_X k(x,y)f(x)\,d\mu(x).$$

Then

$$A(U_Tf)(y)=\int_X k(x,y)f(Tx)\,d\mu(x)=\int_X k(Tx,Ty)f(Tx)\,d\mu(x)$$

$$=\int_X k(x,Ty)f(x)\,d\mu(x)=U_T(Af)(y),$$

i.e., A and U_T commute.