SOME APPLICATIONS OF JOININGS IN ERGODIC THEORY

Measure-Preserving systems

$\mathbf{X} = (X, \mathcal{B}, \mu, T)$ where X is a compact metric space, $\mathcal{B}$ is its Borel sigma-algebra, μ is a probability measure on $\mathcal{B}$ and $T : X \to X$ is an invertible measurable map preserving μ. T also acts by composition on functions defined on X: $Tf = f \circ T$. In particular T is a unitary operator on $L^2(\mu)$. $\mathbf{X}$ is **ergodic**, means $T(A) = A$ implies $\mu(A) = 0$ or 1, equivalently $Tf = f$ implies f is μ-a.e. constant.

(Universal) example: T is the shift on a stationary stochastic process (SSP) (discrete time) taking values in $\mathbb{R}$ or $\mathbb{C}$.

Bernoulli shifts: T is the shift on a iid process.

Rotations: T is the rotation of the circle by angle $2\pi\alpha$ where α is irrational.

Joinings

A **joining** of $\mathbf{X}$ and $\mathbf{Y} = (Y, \mathcal{C}, \nu, S)$ is a $T \times S$-invariant Borel measure λ on $X \times Y$ with marginals μ and ν. This means

$$\lambda(A \times Y) = \mu(A) \text{ and } \lambda(X \times B) = \nu(B),$$

equivalently

$$\lambda(f \otimes 1_X) = \mu(f) \text{ and } \lambda(1_X \times g) = \nu(g).$$

If $\mathbf{X}$ and $\mathbf{Y}$ are thought of as the shifts on SSP's then a joining is a stationary coupling of the two processes.
More generally a joining of systems $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}$, is an invariant measure on the product space with the appropriate marginals on each factor.
$\mu \times \nu$ is always a joining. If $\mathbf{X} = \mathbf{Y}$, diagonal measure μ_Δ is a joining.

Notation: $\mathcal{J}(\mathbf{X}, \mathbf{Y})$ denotes the set of joinings, $\mathcal{J}e(\mathbf{X}, \mathbf{Y})$ the subset of ergodic joinings.

Compact topology on $J(X, Y)$: $X \times Y$ is a compact metric space and $J(X, Y)$ is a closed subset of the set of complex Borel measures on $X \times Y$, which is compact in the weak-* topology.

Proposition. *$\lambda_n \to \lambda$ iff*

$$\lambda_n(A \times B) \to \lambda(A \times B) \quad \forall \text{ measurable } A, B.$$

This follows easily from the following lemma.

Lemma. *If $|f|, |f'| and |g|$ are all bounded by 1 then for all $\alpha > 0$*

$$|\lambda(f \otimes g) - \lambda(f' \otimes g)| < \alpha + \mu\{|f - f'| < \alpha\}$$

Definition: $\mathbf{X}$ *and* $\mathbf{Y}$ *are disjoint* ($\mathbf{X} \perp \mathbf{Y}$) *if* $\mu \times \nu$ *is the only joining.*

Theorem. $\mathbf{X}$ *ergodic,* $S = \mathrm{id} \Rightarrow \mathbf{X} \perp \mathbf{Y}$.

Proof: $f \to f \otimes 1$ embeds $L^2(\mu)$ isometrically as a subspace of $L^2(\lambda)$. If $g \in L^2(\nu)$ then $1 \otimes h$ is $T \times \mathrm{id}$ -invariant, so $g_0 = P_{L^2(\mu)} g$ is T-invariant and $g_0 = \nu(g)$. Now

$$\lambda(f \otimes g) = \langle f \otimes 1, 1 \otimes g \rangle_{L^2(\lambda)} = \langle f, g_0 \rangle_{L^2(\mu)} = \mu(f)\nu(g).$$

$\square$

An Application

Weak Ergodic Theorem. *If* $\mathbf{X}$ *is ergodic then* $\forall\ A, B$ $\frac{1}{n} \sum \mu(A \cap T^n B) \to \mu(A)\mu(B)$.

Proof: Let μ_Δ = diagonal measure on $X \times X$, so

$$\mu_\Delta(A \times B) = \mu(A \cap B)\ .$$

What we must show translates as

$$A_n \mu_\Delta \equiv \frac{1}{N} \sum_{n=1}^{N} (\mathrm{id} \times T)^n \mu_\Delta \to \mu \times \mu\ .$$

It suffices to show that any limit point λ of $\{A_n \mu_\Delta\}$ is $\mu \times \mu$. Note that both marginals of $A_n \mu_\Delta$ are μ. $A_n \mu_\Delta$ "approximately invariant" for $\mathrm{id} \times T$ implies that any λ is $\mathrm{id} \times T$-invariant, that is λ is a joining of id_X and T, so $\lambda = \mu \times \mu$. $\square$

Corollary (Mean ergodic theorem). *If* $f \in L^2(\mu)$ *then* $A_n f = \frac{1}{n} \sum_{n=1}^{N} f \circ T^n \to \int f\ d\mu$ *in* L^2 *norm.*

Definition: $\mathbf{X}$ is **weakly mixing** if T, as an operator on $L^2_{\mathbb{C}}(\mu)$, has continuous spectrum (= no eigenfunctions) on the subspace $L^2_0(\mu) = \{c1\}^\perp$.

A standard use of the spectral theorem shows that if T and S are weakly mixing than so are $T \times S$ and T^2. Using the fact that $T \times T^2$ is ergodic an argument similar to that in the proof of the weak ergodic theorem gives

Theorem (Furstenberg). *If* T *is weakly mixing then* $\forall\ A, B, C$

$$\frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^n B \cap T^{2n} C) \to \mu(A)\mu(B)\mu(C)\ .$$

Proof: Let μ_Δ denote the diagonal measure on X^3,

$$\mu_\Delta(A \times B \times C) = \mu(A \cap B \cap C)$$

and let λ be any limit point of the sequence

$$\left\{ A_n \mu_\Delta = \frac{1}{N} \sum_{n=0}^{N-1} (\mathrm{id} \times T \times T^2) \mu_\Delta \right\} ,$$

so λ is $\mathrm{id} \times T \times T^2$ - invariant. Moreover

$$\begin{aligned} \lambda(X \times B \times C) &= \lim_i \frac{1}{N_i} \sum_{n=0}^{N_i-1} \mu(X \cap T^{-n} B \cap T^{-2n} C) \\ &= \lim_i \frac{1}{N_i} \sum_{i=0}^{N_i-1} \mu(B \cap T^{-n} C) \\ &= (\mu \times \mu)(B \times C) , \end{aligned}$$

by the weak ergodic theorem. So, λ is a joining of the systems

$$(\mu, \mathrm{id}_X) \quad \text{and} \quad (\mu \times \mu, T \times T^2) .$$

Since the second one is ergodic we must have $\lambda = \mu \times \mu \times \mu$, as desired. $\square$

Now, letting $a_n = \mu(A \cap T^n B \cap T^{2n} C)$ and $a = \mu(A)\mu(B)\mu(C)$, the above result says that

$$a_n \to a \quad C_1 \quad \text{(Cesaro convergence)}$$

Applying the same result to the weakly mixing system $T \times T$ and sets $A \times A$, $B \times B$, $C \times C$ we find that

$$a_n^2 \to a^2 \quad C_1 ,$$

so

$$(a_n - a)^2 \to 0 \quad C_1 .$$

It follows that $a_n \to a$ in density by which we mean there is a subset $D \subset \mathbb{N}$ of density one such that

$$\lim_{\substack{n \to \infty \\ n \in D}} a_n = a .$$

We have proved Furstenberg's "multiple weak mixing" theorem:

Theorem. *If* **X** *is weakly mixing then*

$$\mu(A \cap T^{-n} B \cap T^{-2n} C) \to \mu(A)\mu(B)\mu(C)$$

in density.

X is called **mixing** if

$$\lim_{n \to \infty} \mu(A \cap T^{-n} B) = \mu(A)\mu(B)$$

(Weak mixing is equivalent to the above convergence in density.)

X is called **3-fold mixing** (along multiples) if

$$\lim_{n\to\infty} \mu(A \cap T^n B \cap T^{2n} C) = \mu(A)\mu(B)\mu(C) \ .$$

Rohlin's question (1949): Does mixing imply 3-fold mixing?

This question can be approached via joinings as follows. Let μ_Δ denote the diagonal measure on X^3:

$$\int_{X^3} f d\mu_\Delta = \int_X f(x,x,x) d\mu(x)$$
$$\mu_\Delta(A \times B \times C) = \mu(A \cap B \cap C) \ .$$

By a compactness argument, to show **X** is 3-fold mixing is to show that any limit point λ of the sequence

$$\left\{(\mathrm{id} \times T^n \times T^{2n})\mu_\Delta\right\} \subset \mathcal{J}(\mathbf{X},\mathbf{X},\mathbf{X})$$

must be μ^3. Now mixing implies that such a λ is 2-fold independent, that is

$$\lambda(A \times B \times C) = \mu(A)\mu(B)\mu(C) \ ,$$

whenever one of A, B, C is X. So, to show that a mixing **X** is 3-fold mixing it's enough to show that any 2-fold independent $\lambda \in \mathcal{J}(\mathbf{X},\mathbf{X},\mathbf{X})$ must be μ^3. Host and Ryzhikov have proved some deep results in this direction. More about this in Talk #2.

Here is a basic result which guarantees the existence of ergodic joinings.

Theorem. *Suppose* **X** *and* **Y** *are ergodic. Then* $\mathcal{J}(\mathbf{X},\mathbf{Y})$ *is a non-empty compact convex subset of the space of probability measures on* $X \times Y$*, with respect to the weak-* topology. Ext* $(\mathcal{J}(\mathbf{X},\mathbf{Y})) = \mathcal{J}e(\mathbf{X},\mathbf{Y})$*. In particular* $\mathcal{J}e(\mathbf{X},\mathbf{Y})$ *is never empty, by the Krein-Milman Theorem.*

Graphic joinings

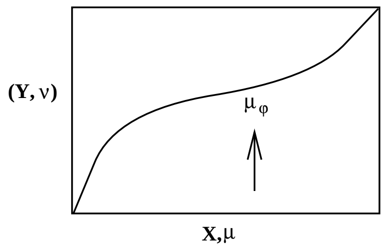
Definition: A **homomorphism** $\varphi : \mathbf{X} \to \mathbf{Y}$ is a measurable map $\varphi : X \to Y$ such that $\varphi(\mu) = \nu$ and $S \circ \varphi = \varphi \circ T$ μ-a.e. If such a φ exists say $\mathbf{Y}$ is a **factor** of $\mathbf{X}$. Say φ is an **isomorphism** if φ^{-1} exists a.e. and is measurable.

Remark: If $\lambda \in \mathcal{J}(\mathbf{X}, \mathbf{Y})$ and

$$\mathbf{Z} = (X \times Y, \mathcal{B} \otimes \mathcal{C}, \lambda, T \times S)$$

then the projections of Z onto X and Y are homomorphisms. So, a joining is "the same as" an embedding of $\mathbf{X}$ and $\mathbf{Y}$ as factors in a common system $\mathbf{Z}$.

If $\varphi : \mathbf{X} \to \mathbf{Y}$ is a homomorphism define a measure μ_φ on $X \times Y$ as the image of μ under $x \mapsto \big(x, \varphi(x)\big)$. Then μ_φ is a joining of $\mathbf{X}$ and $\mathbf{Y}$ which we refer to as a graphic joining.



lift μ up to graph of φ.

Clearly we have

$$X \times C \subset \mathcal{B} \times Y \quad \mu_\varphi\text{-a.e.},$$

with equality if φ is an isomorphism.

Theorem (von Neumann). *Suppose $\lambda \in \mathcal{J}(\mathbf{X}, \mathbf{Y})$ and $X \times \mathcal{C} \subset \mathcal{B} \times Y$ λ a.e. ($X \times \mathcal{C} = \mathcal{B} \times Y$ λ a.e.). Then $\lambda = \mu_\varphi$ for some homomorphism (isomorphism) φ.*

Definition: Say $\mathbf{X}$ has **discrete spectrum** if the eigenfunctions of T span $L^2_{\mathbb{C}}(\mu)$. (Example: $Tx = x + \alpha (\mathrm{mod} 1)$) Let $G(\mathbf{X})$ denote the set of eigenvalues.

Theorem (Halmos, von Neumann). *Suppose $\mathbf{X}$ and $\mathbf{Y}$ are ergodic systems with discrete spectrum and $G(\mathbf{X}) = G(\mathbf{Y})$. Then $\mathbf{X}$ and $\mathbf{Y}$ are isomorphic.*

Proof (Lemańczyk): Let λ be an ergodic joining of $\mathbf{X}$ and $\mathbf{Y}$. If c is an eigenvalue of T, f_c and g_c corresponding eigenfunctions of $\mathbf{X}$ and $\mathbf{Y}$, then $f \otimes 1$ and $1 \otimes g$ are both eigenfunctions with eigenvalue c for the ergodic system

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

This means that $f \otimes 1$ and $1 \otimes g$ are multiples of each other, since their ratio is an invariant function for this system, and hence constant. Since the eigenfunctions of the two systems span their respective L^2-spaces we conclude that

$$\mathcal{B} \times Y = X \times \mathcal{C} \quad \lambda - \text{a.e.}$$

$\square$

Simple systems and structure of factors

A **factor algebra** of $\mathbf{X}$ is a sub-sigma-algebra $\mathcal{A} \subset \mathcal{B}$ such that $T^{-1}\mathcal{A} = \mathcal{A}$. We identify factor algebras which agree up to null sets.

$$\bar{\mathbf{X}} = (X, \mathcal{A}, \mu, T)$$

may be viewed as a factor of $\mathbf{X}$ via the homomorphisms $\mathrm{id} : X \to X$. Conversely if $\varphi : \mathbf{X} \to \mathbf{Y}$ is a homomorphism then $\varphi^{-1}(\mathcal{C})$ is a factor algebra of $\mathbf{X}$ and, roughly speaking, every factor algebra arises in this way.

To each factor of $\mathbf{X}$ there corresponds a joining $\lambda \in J(\mathbf{X}, \mathbf{X})$ characterized by

$$\lambda(f \otimes g) = \langle P_{L^2(\mathcal{A})} f, g \rangle \ .$$

(This is the so-called **relative product** over the factor $\mathcal{A}$ which we will look at in more detail in talk # 2.)

Definition (Veech): $\mathbf{X}$ is **2-simple** if each $\lambda \in \mathcal{J}e(\mathbf{X}, \mathbf{X})$ is either $\mu \times \mu$ or a μ_φ for some $\varphi \in C(\mathbf{X}) = \{$ automorphisms of $\mathbf{X}\}$. Equivalently, for each $\lambda \in \mathcal{J}e(\mathbf{X}, \mathbf{X})$ either

$$X \times \mathcal{C} \perp \mathcal{B} \times X \quad (\lambda)$$

or

$$X \times \mathcal{C} = \mathcal{B} \times X \quad \lambda \text{ a.e. } .$$

$\mathbf{X}$ is **simple** (of all orders) if a similar condition holds for higher order self-joinings.

Remark: Note that $C(\mathbf{X}) \supset \{T^n : n \in \mathbb{Z}\}$ so graphic self-joinings always exist. If $\mathbf{X}$ is simple of all orders and $C(\mathbf{X}) = \{T^n : n \in \mathbb{Z}\}$ then we say $\mathbf{X}$ is **MSJ** (for minimal self-joinings). The notion of MSJ, due to Rudolph, came before simplicity. Rudolph's paper and Furstenberg's original paper on joinings were the seminal ones for the development of the theory of joinings.

Theorem (Veech). *A simple system is either weakly mixing or has discrete spectrum. If $\mathbf{X}$ is simple of order 2 and $\mathcal{A} \subset \mathcal{B}$ is any T-invariant sub-sigma-algebra then there is a compact subgroup K of $C(\mathbf{X})$ such that*

$$\mathcal{A} = \mathcal{F}(K) \equiv \{A \in \mathcal{B} : kA = A \text{ for all } k \in K\} \ .$$

In this situation we say that $\mathbf{X}$ is a **group extension** of it's factor $\bar{\mathbf{X}} = (X, \mathcal{A}, \mu, T)$, which we denote by $\mathbf{X}/K$.

Veech's theorem is an easy consequence of a more general result of Veech which characterizes group extensions in terms of joinings. We shall prove this result in talk # 2.

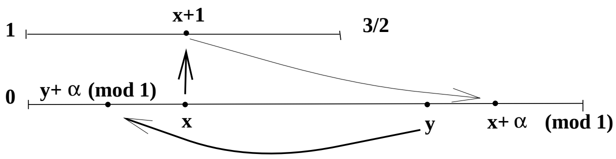
Theorem (Veech). *Suppose that $\mathcal{A}$ is a factor algebra of $\mathbf{X}$ with the property that each $\lambda \in J_e^2(\mathbf{X})$ which is diagonal on $\mathcal{A} \otimes \mathcal{A}$ is graphic. Then there is a compact subgroup K of $C(\mathbf{X})$ such that*

$$\mathcal{A} = \{A \in \mathcal{B} : kA = A \quad \forall k \in K\} \ .$$

(The converse is true as well.)

Veech's theorem gives us complete information about factors of $\mathbf{X}$, in terms of $C(\mathbf{X})$. In particular if $\mathbf{X}$ has MSJ, $C(\mathbf{X}) \sim \mathbb{Z}$ has no non-trivial compact subgroups so $\mathbf{X}$ must be **prime**, that is, the only factor algebras are the trivial one and $\mathcal{B}$ itself.

An example of MSJ



$|\alpha - p/q| > c/q^2$

MSJ, simplicity and primality are "exotic" properties. It is unknown if there is a smooth dynamical system with MSJ. Certain horocycle flows are known to be simple (Ratner). MSJ is known to be a meager property but it is unknown whether simplicity is meager or residual. The same goes for primality.

Thouvenot has shown that certain Gaussian systems, although never simple, allow a similar exhaustive analysis of factors via joinings. The factor structure of such systems is much richer than that of simple systems. More about this in Talk #3. The following result will also play a role in Talk #3. For $n > 0$ $\mathbf{X}^n = (\mathbf{X}, \mathcal{B}^n, \mu^n, T^{\otimes n})$ and $\mathbf{X}^{n\odot}$ denotes the factor system $\mathbf{X}/S_n$

Theorem (del Junco-Rudolph). *If $\mathbf{X}$ is simple and $\mathbf{Y}$ is any ergodic system which is not disjoint from $\mathbf{X}$ then there exist a compact subgroup K of $C(\mathbf{X})$ and an $n \geq 1$ such that $(\mathbf{X}/K)^{n\odot}$ is a factor of $\mathbf{Y}$.*

Corollary. *Simple systems without common factors are disjoint.*

$\mathbf{X}$ and $\mathbf{X}^{2\odot}$ are never disjoint because they sit as factors in $\mathbf{X}^2$ which are not independent. When $\mathbf{X}$ has MSJ it can be shown that $\mathbf{X}$ and $\mathbf{X}^{2\odot}$ have no common factors , in fact they are both prime and not isomorphic to each other. This was shown by Rudolph in order to give a negative answer to Furstenberg's question: if $\mathbf{X}$ and $\mathbf{Y}$ have no common factors must they be disjoint?