1. Cosine Sums

Conjecture. *Suppose $\alpha \in \mathbf{R} \setminus \mathbf{Q}$. Then*[1] $\frac{1}{N} \sum_{\{0,\ldots N-1\}} \cos n^2\alpha \xrightarrow{N} 0$.

Loosely speaking, this sum can be written in a way so that it resembles the 'time averages' of a certain 'measure preserving system'. Namely, as $\frac{1}{N} \sum_{\{0,\ldots N-1\}} f \circ S^n (0,0)$, where[2] $S : (x,y) \mapsto (\{x+\alpha\}, \{y+x\})$ and $f : (x,y) \mapsto \cos(x+2y)$. If our underlying measure preserving system is 'ergodic', we might be able to make use of Birkhoff's theorem, which tells us something about the convergence of time averages.

Now I'll introduce everything properly.

Definition 1.1. Let $(X, \mathcal{X}, \mu)$ be a probability space. If T is a measure space automorphism of X, (ie it is measurable, it has a measurable inverse, and μ is T-invariant: $T^*\mu = \mu$), call T a measure preserving transformation (mpt), and $(X, \mathcal{X}, \mu, T)$ a measure preserving system (mps).

Here is a somewhat useful characterisation:

Proposition 1.1. *With the above notation, μ is T-invariant iff $\forall f \in L^1$ $\mu f = \mu f \circ T$.*

Proof. Choosing f to be the characteristic function of a measurable set, we get the sufficient condition. The necessary condition comes from seeing that if the equality holds for characteristic functions, it also holds for simple functions by linearity. Then it holds for nonnegative functions by the monotone convergence theorem, and lastly integrable functions by linearity. $\square$

Proposition 1.2. *Suppose $\alpha \in \mathbf{R} \setminus \mathbf{Q}$. Then $([0,1), \mathcal{B}, \mathrm{leb}, x \mapsto \{x+\alpha\} : R_\alpha)$, the circle system, is an mps.*

Proof. R_α is measurable. Its inverse is $x \mapsto \{x - \alpha\}$, which is measurable. $R^* \mathrm{leb} = \mathrm{leb}$ because Lebesgue measure is translation invariant. $\square$

Definition 1.2. Let $(X, \mathcal{X}, \mu, T)$ be an mps. If $\{A \in \mathcal{X} : T^{-1}A = A\}$ is μ-trivial, call $(X, \mathcal{X}, \mu, T)$ ergodic.

This is some sort of irreducibility, for if $A \in \mathcal{X}$ and $\mu A = 1 - a \in (0,1)$, then calling $X \setminus A =: B$, $\left(B, \mathcal{X}|_B, \frac{1}{a}\mu|_B, T|_B\right)$ is an mps, a 'sub'-mps in a sense.

Result 1.3 (Birkhoff's Theorem)**.** *$(X, \mathcal{X}, \mu, T)$ is ergodic iff $\forall f \in L^1$*

$$\frac{1}{N} \sum_{\{0,\ldots,N-1\}} f \circ T^n \xrightarrow{N} \mu f \text{ as.}$$

So if we could show that the circle system is ergodic, and choose our f to be cos, then we could get $\sum_{\{0,\ldots,N-1\}} \cos(x + n\alpha) \to 0$ as. This is one step closer to the conjecture, but there is still a lot missing. Aside from there being an n instead of

[1] all trigonometric functions will have a 2π already in it, eg $\cos : t \mapsto \sum \frac{(-1)^n}{(2n)!} (2\pi t)^{2n}$

[2] $\{x\}$ is the fractional part of x

n^2, we can't even just choose $x = 0$ to get the analogous statement, since we only have almost sure convergence. Let me first fix the former problem.

Definition 1.3. Let $(X, \mathcal{X}, \mu, T)$ be an mps, $(Y, \mathcal{Y}, \nu)$ a probability space. If $(S_x)_X$ is a family of mpts of Y, call $(X \times Y, \mathcal{X} \otimes \mathcal{Y}, \mu \times \nu, (x, y) \mapsto (Tx, S_x y))$ the skew product of $(X, \mathcal{X}, \mu, T)$ and $(Y, \mathcal{Y}, \nu)$ with $(S_x)_X$.

Proposition 1.4. *With the above notation, suppose $X \times Y \ni (x, y) \mapsto S_x y \in Y$ is measurable. Then $(X \times Y, \mathcal{X} \otimes \mathcal{Y}, \mu \times \nu, (x, y) \mapsto (Tx, S_x y))$ is an mps.*

Proof. Since $X \times Y \ni (x, y) \mapsto S_x y \in Y$ is measurable, S is measurable. $(x, y) \mapsto \left(T^{-1}x, S_x^{-1}y\right)$ is its inverse, which is also measurable. To show $\mu \times \nu$ is S-invariant, let $E \in \mathcal{X} \otimes \mathcal{Y}$. Denoting by E_x the section $\{y : (x, y) \in E\}$, $(x, y) \in S^{-1}E$ iff $y \in S_x^{-1} E_{Tx}$. So $\left(S^{-1}E\right)_x = S_x^{-1} E_{Tx}$. By Fubini, $\mu \times \nu S^{-1} E = \int \nu S_x^{-1} E_{Tx} \mathrm{d}\mu = \int \nu E_{Tx} \mathrm{d}\mu$, because S_x is an mpt. Since $x \mapsto \nu E_x$ is integrable, by the previous proposition $\int \nu E_{Tx} \mathrm{d}\mu = \int \nu E_x \mathrm{d}\mu = \mu \times \nu E$. So $\mu \times \nu$ is S-invariant. $\square$

Proposition 1.5. *The skew product of the circle system and itself with the translation maps, $(y \mapsto \{y + x\})_{x \in [0,1)}$, is an mps.*

Proof. If $x \in [0, 1)$, then $y \mapsto \{y + x\}$ is an mpt (similar to Proposition 1.2). By the previous proposition, it suffices to prove that $(x, y) \mapsto \{y + x\}$ is measurable. But it is continuous, so this is done. $\square$

Now, as I had vaguely outlined at the beginning, if we call the mpt on the skew product space S, we can now use the integrable function $f : (x, y) \mapsto \cos(x + 2y)$ to get that $\frac{1}{N} \sum_{\{0,\dots N-1\}} \cos n^2 \alpha = \frac{1}{N} \sum_{\{0,\dots N-1\}} f \circ S^n (0, 0)$. So if we knew our skew product was ergodic, and somehow upgraded the almost sure convergence in Birkhoff's theorem to pointwise everywhere, then we would have our result.

For the latter, I'm going to have to start paying attention to the topology that we have lying in the background. This is because letting the test functions be L^1 really gives us no hope of retrieving pointwise everywhere convergence. To see this, take the circle system, $x \in [0, 1)$. If t is in the orbit of x, then take $f : t = x + n\alpha \mapsto n+1$, zero otherwise. Then $\mu f = 0$ but $\frac{1}{N} \sum_{\{0,\dots,N-1\}} f \circ T^n (x) \geq \frac{N-1}{2}$. Of course, I haven't yet proven the circle system is ergodic, but this is coming. Importantly though, we have a compact topology on the circle system (and so the Borel σ-algebra), so we can relax the test function requirement by only insisting that they be continuous.

So in a bit more generality, suppose $(X, \mathcal{X}, \mu, T)$ is an mps where X has a compact topology and $\mathcal{X}$ is the Borel σ-algebra. Since we have a topology, we ought to also ask that T respect it, ie by being continuous. I'll call this a topological mps from now on. By writing $\frac{1}{N} \sum_{\{0,\dots N-1\}} \delta_{T^n x} =: \mu_x^N$, the conclusion of Birkhoff's theorem in this context looks like[3] $\mu_x^N \overset{*}{\rightharpoonup} \mu$ as. I think you might want to interpret the measures μ_x^N as 'empirical', and so the previous sentence can be interpreted as the empirical measures weakly converging to μ almost surely. Now suppose further that that this weak convergence happens to hold not just almost surely, but

[3] $\overset{*}{\rightharpoonup}$ denotes weak convergence of probability measures

pointwise everywhere. In this case, it turns out that μ is unique: in fact it is not just the only ergodic measure for T, but the only T-invariant measure at all.

To see this, suppose ν is a T-invariant measure, and let $f \in \mathcal{C}X$. By Proposition 1.1, $\forall N \int \frac{1}{N}\sum_{\{0,\ldots,N-1\}} f \circ T^n \mathrm{d}\nu = \int f \mathrm{d}\nu$. Similarly, and by the triangle inequality, $\left|\frac{1}{N}\sum_{\{0,\ldots,N-1\}} f \circ T^n(x)\right| \leq \nu|f|$. So by the dominated convergence theorem and our pw everywhere hypothesis, $\int f \mathrm{d}\nu = \lim_N \int \frac{1}{N}\sum_{\{0,\ldots,N-1\}} f \circ T^n \mathrm{d}\nu = \int f \mathrm{d}\mu$. This gives us $\nu = \mu$ (take a characteristic function of a measurable set, approximate it piecewise linearly, apply the dominated convergence theorem).

It turns out that this also goes in the other direction too:

Result 1.6. *Suppose $(X, \mathcal{X}, \mu, T)$ is a topological mps. Then $\forall x \in X$ $\mu_x^N \overset{*}{\rightharpoonup} \mu$ iff μ is the only T-invariant measure. In this case, μ is automatically ergodic.*

If $(X, \mathcal{X}, \mu, T)$ is such a topological mps, let's call it uniquely ergodic.

It will be useful to point out that if one wants to verify the uniqueness of μ, it suffices to only check the convergence of integrals in weak convergence when f belongs to a dense subset of $\mathcal{C}X$. This is because, following similar steps to earlier, it is enough to know $\int f \mathrm{d}\nu = \int f \mathrm{d}\mu$ for f in a dense subset to conclude that $\mu = \nu$.

Proposition 1.7. *The circle system is uniquely ergodic.*

Proof. R_α is continuous, so it now suffices to prove that $\frac{1}{N}\sum_{\{0,\ldots,N-1\}} f \circ T^n(x) \to \operatorname{leb} f$ for f a trigonometric polynomial, and all $x \in [0,1)$. Let $x \in [0,1)$ and take $g : x \mapsto \exp ikx$, for some $k \neq 0$. Then because it is a geometric sum, $\left|\frac{1}{N}\sum_{\{0,\ldots,N-1\}} g \circ T^n(x)\right| \leq K\frac{1}{N}$, for some $K > 0$ by the triangle inequality. So $\frac{1}{N}\sum_{\{0,\ldots,N-1\}} g \circ T^n(x) \to 0 = \operatorname{leb} g$. Since trigonometric polynomials are linear combinations of gs, and possibly a constant term (which is fine), this is done. $\square$

Question 1.8. *Is the skew product of the circle system and itself with the translation maps uniquely ergodic?*

Answering this proves the conjecture, since $(x, y) \mapsto \cos(x + 2y)$ is continuous.