Non-Conventional Ergodic Averages And Nilmanifolds

Bernard Host, Université de Marne la Vallée
Bryna Kra, Pennsylvania State University

Preprints available at:
http://www.math.psu.edu/kra/

Theorem (Szeméredi). *If a set $E \subset \mathbb{Z}$ has positive upper density, then it contains arbitrarily long arithmetic progressions.*

Theorem (Furstenberg).
*Let $(X, \mathcal{X}, \mu, T)$ be a (p.m.p.t.) system, and $A \in \mathcal{X}$ with $\mu(A) > 0$.
Then for any $k \geq 1$ the* $\liminf$ *when $N \to \infty$ of*

$$\frac{1}{N}\sum_{n=0}^{N-1} \mu\Big(A \cap T^{-n}A \cap T^{-2n}A \cap \cdots \cap T^{-kn}A\Big)$$

is positive.

Question. *Is the* $\liminf$ *actually a* $\lim$*?*

Question. *Convergence in $L^2(\mu)$ of*

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n f_1 \cdot T^{2n} f_2 \cdot \ldots \cdot T^{kn} f_k \ . \qquad \text{(AP)}$$

THE RESULTS

History of the problem of convergence of (AP):
$k = 1$: Ergodic Theorem.
$k = 2$: Furstenberg.
k arbitrary, weak mixing system : Furstenberg.
$k = 3$: Conze & Lesigne (assuming total ergodicity; H&K in general)

Theorem 1. *Let $(X, \mathcal{X}, \mu, T)$ be a system, $k \geq 1$ be an integer and let f_j, $1 \leq j \leq k$, be k bounded measurable functions on X. Then the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n f_1 \cdot T^{2n} f_2 \cdot \ldots \cdot T^{kn} f_k \qquad \text{(AP)}$$

converge in $L^2(\mu)$ when $N \to +\infty$.

Open problem: Is a similar result valid for commuting transformations $T_1, T_2, \ldots, T_k$ substituted for $T, T^2, \ldots, T^k$?

Polynomial averages

Furstenberg's Theorem was generalized to polynomial averages by Bergelson and Leibman. It is natural to ask for similar generalizations of Theorem 1.

Furstenberg and Weiss : Convergence of

$$\frac{1}{N}\sum_{n=0}^{N-1} T^{n^2} f_1 \cdot T^n f_2 \text{ and}$$

$$\frac{1}{N}\sum_{n=0}^{N-1} T^{n^2} f_1 \cdot T^{n^2+n} f_2 \ .$$

Theorem 2. *Let* $(X, \mathcal{X}, \mu, T)$ *be a system,* $k \geq 1$ *and* $f_1, f_2, \ldots, f_k \in L^\infty(\mu)$. *Then for any integer polynomials* $p_1(\cdot), p_2(\cdot), \ldots, p_k(\cdot)$ *the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} T^{p_1(n)} f_1 \cdot T^{p_2(n)} f_2 \cdot \ldots \cdot T^{p_k(n)} f_k$$

converge in $L^2(\mu)$.

CHARACTERISTIC FACTORS

In the sequel, $(X, \mathcal{X}, \mu, T)$ is an ergodic system.

We use the word 'factor' with two different but equivalent meanings:

- A *factor* of X is a T-invariant sub-σ-algebra of $\mathcal{X}$.

- Let $(Y, \mathcal{Y}, \nu, S)$ be a system.
  A map $\pi : X \to Y$ is a *factor map* if $\pi^* \mu = \nu$ and $S \circ \pi = \pi \circ T$.
  In this case we also say that Y is a factor of X.

These definitions coincide up to the identification of $\mathcal{Y}$ with $\pi^{-1}(\mathcal{Y})$.

Let Y be a factor of X.
We say that Y is a *characteristic factor* for the convergence of the averages

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n f_1 \cdot T^{2n} f_2 \cdot \ldots \cdot T^{kn} f_k \qquad \text{(AP)}$$

if the difference between these averages and the same averages with $\mathbb{E}(f_1 \mid \mathcal{Y})$ substituted for f_1, . . . , $\mathbb{E}(f_k \mid \mathcal{Y})$ substituted for f_k:

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n \mathbb{E}(f_1 \mid \mathcal{Y}) \cdot \ldots \cdot T^{kn} \mathbb{E}(f_k \mid \mathcal{Y})$$

converge to 0 in $L^2(\mu)$.

This means that the averages (AP) converge to 0 whenever $\mathbb{E}(f_i \mid \mathcal{Y}) = 0$ for at least one i.

THE STRATEGY

1. Find a characteristic factor for the convergence of the averages (AP).

2. Give a ‘good’ description of this factor. This means, identify it with a system with a known ‘geometric’ or ‘algebraic’ structure.

3. Show convergence for this factor.

BUILDING CHARACTERISTIC FACTORS

A standard method for finding characteristic factors consists in using Van der Corput's Lemma several times.

Van der Corput's Lemma.
Let H be a Hilbert space and ξ_n, $n \geq 0$, in H with $\|\xi_n\| \leq 1$ for every n. For $h \geq 0$ define:

$$\gamma_h := \limsup_N \left| \frac{1}{N} \sum_{n=0}^{N-1} \langle \xi_{n+h}, \xi_n \rangle \right| .$$

Then

$$\limsup_N \left\| \frac{1}{N} \sum_{n=0}^{N-1} \xi_n \right\|^2 \leq \limsup_H \frac{1}{H} \sum_{h=0}^{H-1} \gamma_h .$$

First idea:
We 'forget' the original convergence problem and build a factor by a procedure that mimics successive uses of VdC's Lemma.

This factor will be 'automatically' characteristic.

Some seminorms

We build a sequence of seminorms on $L^\infty(\mu)$ such that for every $f \in L^\infty(\mu)$:

1. $|||f|||_1 = \left|\int f\,d\mu\right|$.

2. For every $k \geq 1$,

$$|||f|||_{k+1}^{2^{k+1}} = \lim_{H\to+\infty} \frac{1}{H} \sum_{h=0}^{H-1} |||f \cdot T^h f|||_k^{2^{k+1}} \ .$$

Furthermore:

- $|||f|||_1 \leq |||f|||_2 \leq \cdots \leq |||f|||_k \leq \cdots \leq \|f\|_\infty$.

- If X is weakly mixing then $|||f|||_k = |||f|||_1$ for every k.

Some factors

By using these seminorms we build factors $Z_k(X) = Z_k$, $k \geq 0$, of X with:

- For $f \in L^\infty(\mu)$, we have $\mathbb{E}(f \mid \mathcal{Z}_k) = 0$ if and only if $|||f|||_{k+1} = 0$.

Furthermore:

- $Z_0(X)$ is the trivial factor of X.

- The sequence of factors is increasing:

  $$Z_0 \leftarrow Z_1 \leftarrow \cdots \leftarrow Z_k \leftarrow Z_{k+1} \leftarrow \cdots \leftarrow X \ .$$

- If X is weakly mixing then $Z_k(X)$ is the trivial factor for every k.

These factors are characteristic

Lemma. *If* $\|f_1\|_\infty, \ldots, \|f_k\|_\infty$ *are* ≤ 1*, then for* $\ell = 1, \ldots, k$ *we have*

$$\limsup_N \Big\| \frac{1}{N} \sum_{n=0}^{N-1} T^n f_1 \cdot T^{2n} f_2 \cdots T^{kn} f_k \Big\|_2 \leq \ell |||f_\ell|||_k \, .$$

Corollary. *The factor* Z_{k-1} *is characteristic for the convergence of the averages*

$$\frac{1}{N} \sum_{n=0}^{N-1} T^n f_1 \cdot T^{2n} f_2 \cdots T^{kn} f_k \qquad \text{(AP)}$$

Proof. By induction. We assume it is true for k. Define $\xi_n = T^n f_1 \cdot T^{2n} f_2 \cdots T^{(k+1)n} f_{k+1}$. We assume that $\ell > 1$ (the case $\ell = 1$ is similar). We have

$$\Big|\frac{1}{N}\sum_{n=0}^{N-1}\langle \xi_{n+h}, \xi_n\rangle\Big| =$$

$$\Big|\int (f_1 \cdot T^h f_1)\frac{1}{N}\sum_{n=0}^{N-1}\prod_{j=2}^{k+1} T^{(j-1)n}(f_j \cdot T^{jh} f_j)\, d\mu\Big|$$

$$\leq \Big\|\frac{1}{N}\sum_{n=0}^{N-1}\prod_{j=2}^{k+1} T^{(j-1)n}(f_j \cdot T^{jh} f_j)\Big\|_2$$

Induction hypothesis: $\gamma_h \leq \ell |\!|\!| f_\ell \cdot T^{\ell h} f_\ell |\!|\!|_k$.

$$\frac{1}{H}\sum_{h=0}^{H-1}\gamma_h \leq \ell\,\frac{\ell}{\ell H}\sum_{n=0}^{\ell H-1} |\!|\!| f_\ell \cdot T^n f_\ell |\!|\!|_k$$

and we conclude by using VdC and the definition of the seminorm $|\!|\!| \cdot |\!|\!|_{k+1}$. □

Building measures

By induction, we build a measure $\mu^{[k]}$ on $X^{[k]} := X^{2^k}$, invariant under $T^{[k]} := T\times\cdots\times T$ (2^k times).

- $\mu^{[0]} := \mu$.
- Let $\mathcal{I}^{[k]}$ be the invariant σ-algebra of $(X^{[k]}, \mu^{[k]}, T^{[k]})$.
  Then $\mu^{[k+1]}$ is the relatively independent square of $\mu^{[k]}$ over $\mathcal{I}^{[k]}$.

Explanation:
We identify $X^{[k+1]}$ with $X^{[k]}\times X^{[k]}$ and write

$$\mathbf{x} = (\mathbf{x}', \mathbf{x}'')$$

for a point of $X^{[k+1]}$, with $\mathbf{x}', \mathbf{x}'' \in X^{[k]}$.
When F, G are bounded functions on $X^{[k]}$,

$$\int_{X^{[k+1]}} F(\mathbf{x}')\, G(\mathbf{x}'')\, d\mu^{[k+1]}(\mathbf{x}) := \int_{X^{[k]}} \mathbb{E}(F \mid \mathcal{I}^{[k]})\, \mathbb{E}(G \mid \mathcal{I}^{[k]})\, d\mu^{[k]}$$

The points of $X^{[k]}$ are written

$$\mathbf{x} = \Big(x_\epsilon : \epsilon \in \{0,1\}^k\Big)$$

Lemma. *For* $f \in L^\infty(\mu)$*,*

$$|\!|\!| f |\!|\!|_k^{2^k} = \int_{X^{[k]}} \prod_{\epsilon\in\{0,1\}^k} f(x_\epsilon)\, d\mu^{[k]}(\mathbf{x}) \ .$$

Corollary. $|\!|\!| \cdot |\!|\!|_k$ *is a seminorm.*

By ergodicity:
$\mathcal{I}^{[0]}$ is trivial, $\mu^{[1]} = \mu\times\mu$ and $|\!|\!| f |\!|\!|_1 = \Big|\int f\, d\mu\Big|$.

The Kronecker factor

Let (Z, m, S) be the Kronecker factor of X:
It is the factor of X spanned by the eigenfunctions.
Z is a compact abelian group.
m is the Haar measure of Z.
The transformation $S : Z \to Z$ is a *rotation*:
$Sz = \alpha z$ for some fixed element α of Z.
$\{\alpha^n : n \in \mathbb{Z}\}$ is dense in Z.

For $f \in L^\infty(\mu)$ we write $\tilde{f} = \mathbb{E}(f \mid Z)$.
The ergodic decomposition of $\mu \times \mu$ for $T \times T$ is

$$\mu \times \mu = \int_Z \mu_s \, dm(s)$$

where

$$\int_{X \times X} f(x)g(y)\, d\mu_s(x, y) := \int_Z \tilde{f}(z)\tilde{g}(sz)\, dm(z) \ .$$

Computation of $\mu^{[2]}$, $|\!|\!| \cdot |\!|\!|_2$ and Z_1

We have $\mu^{[2]} = \int_Z \mu_s \times \mu_s \, dm(s)$.

$$|\!|\!| f |\!|\!|_2^4 := \int f \otimes f \otimes f \otimes f \, d\mu^{[2]}$$
$$= \int_{Z^3} \tilde{f}(z)\tilde{f}(sz)\tilde{f}(tz)\tilde{f}(stz) \, dm(z) \, dm(s) \, dm(t)$$

Corollary. $|\!|\!| f |\!|\!|_2$ *is the* ℓ^4*-norm of the Fourier Transform of* $\tilde{f}$.
$|\!|\!| \cdot |\!|\!|_2$ *is a seminorm.*
The factor Z_1 *is the Kronecker factor.*

A useful Lemma

Let

$$\mu^{[k]} = \int_{\Omega_k} \mu_\omega^{[k]} \, dP_k(\omega)$$

be the ergodic decomposition of $\mu^{[k]}$ for $T^{[k]}$. We can carry out the same construction with $(X^{[k]}, \mu_\omega^{[k]}, T^{[k]})$ substituted for (X, μ, T).

Lemma. *For any $\ell > 0$,*

$$\mu^{[k+\ell]} = \int_{\Omega_k} \Big(\mu_\omega^{[k]}\Big)^{[\ell]} \, dP_k(\omega) \ .$$

Proof. For $\ell = 1$ this is the definition of $\mu^{[k+1]}$.
Then by induction. □

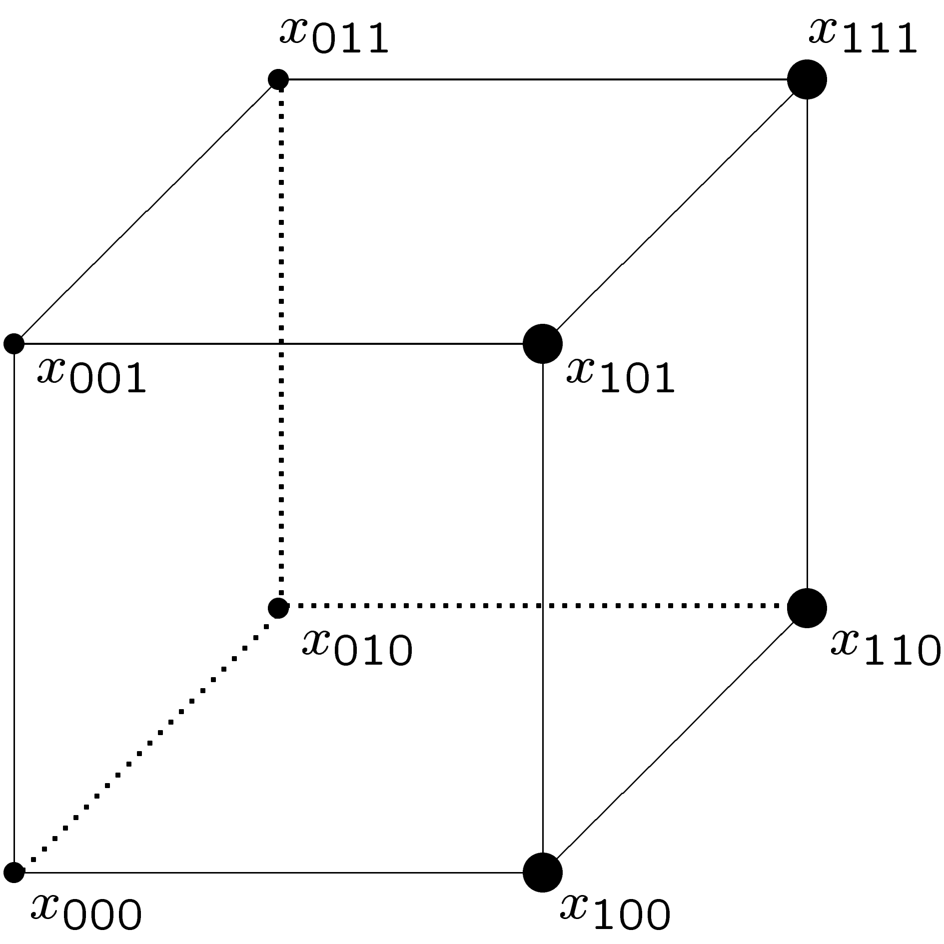
For $k = 1$ we have $\Omega_1 = Z$, $P_1 = m$ and

$$\mu^{[k+1]} = \int_Z \Big(\mu_s\Big)^{[k]} \, dm(s) \ .$$

Second idea:
Use a 'geometric' point of view.

We identify $\{0,1\}^k$ with the set of vertices of the Euclidian cube of dimension k.
We use geometric words like *side, edge, vertex...* for subsets of $\{0,1\}^k$

A point $\mathbf{x}$ of $X^{[3]}$
and the side $\{100, 101, 110, 111\}$



The *symmetry group of the k-cube* is the group of permutations of $\{0,1\}^k$ arising from the isometries of $\mathbb{R}^k$ preserving the cube.

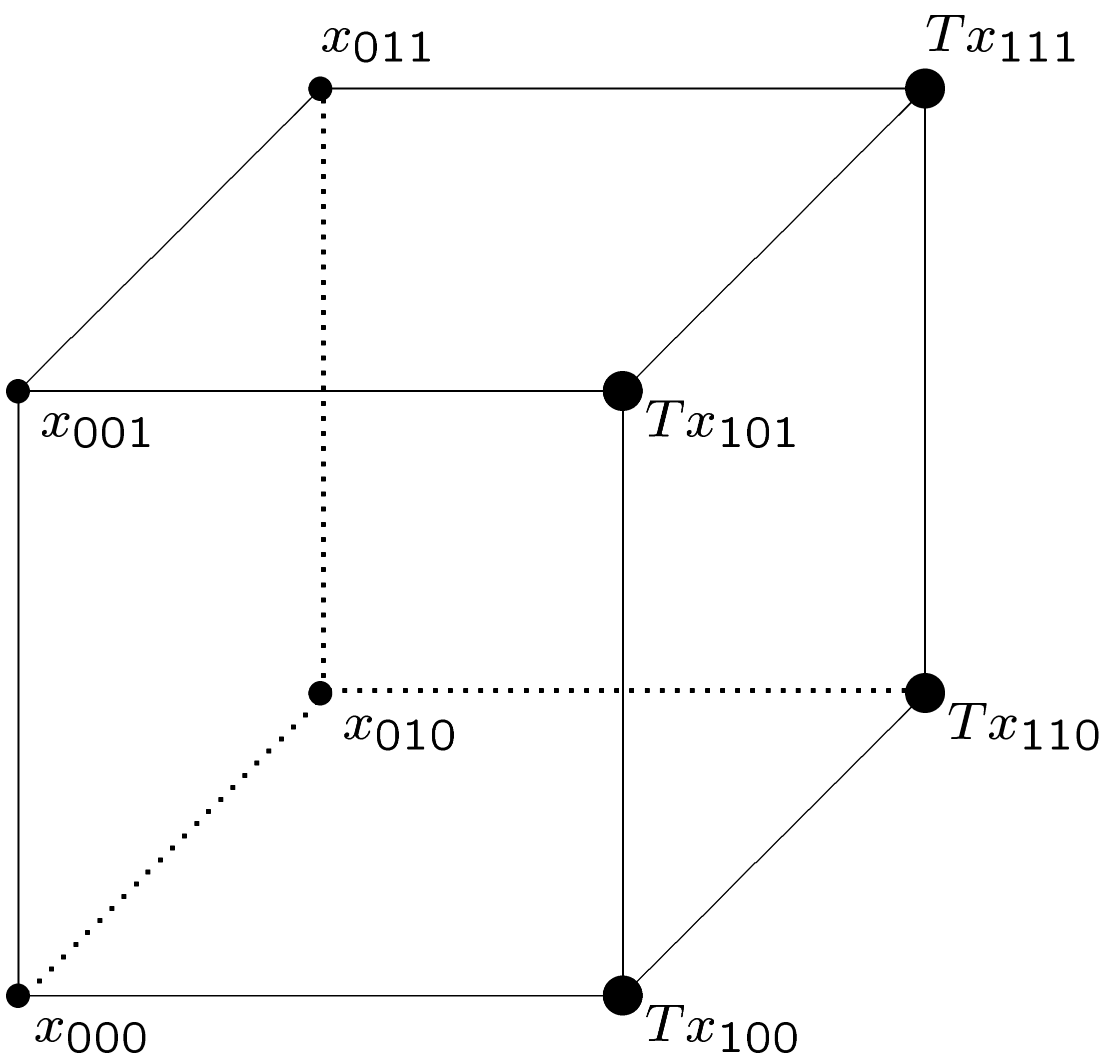
Lemma. *The measure $\mu^{[k]}$ is invariant under the symmetry group of the k-cube.*

Let α be a *face* of the cube.
The *side transformation* $T_\alpha^{[k]} : X^{[k]} \to X^{[k]}$ is given by:

$$(T_\alpha^{[k]}\mathbf{x})_\epsilon = \begin{cases} Tx_\epsilon & \text{if } \epsilon \in \alpha \\ x_\epsilon & \text{if } \epsilon \notin \alpha \end{cases}$$

Lemma. *The measure $\mu^{[k]}$ is invariant under the side transformations.*
It is ergodic under the joint action spanned by these transformations.

The point $T_\alpha^{[3]}\mathbf{x}$ for $\alpha = \{100, 101, 110, 111\}$



Building the factors

Let $\mathbf{0} = 00\cdots0 \in \{0,1\}^k$. $X^{[k]} = X \times X^{2^k-1}$. We consider a point $\mathbf{x} \in X^{[k]}$ as a pair:

$$\mathbf{x} = (x_\mathbf{0}, \tilde{\mathbf{x}}) \text{ with } x_\mathbf{0} \in X,\ \tilde{\mathbf{x}} \in X^{2^k-1} .$$

Consider the transformations $T_\alpha^{[k]}$, for all faces α not containing $\mathbf{0}$.

Lemma. *A function on $X^{[k]}$ is invariant under these transformations if and only if it depends on $x_\mathbf{0}$ only.*

Let $B \subset X^{2^k-1}$ be invariant under these transformations. Then there exists $A \subset X$ with

$$\mathbf{1}_A(x_\mathbf{0}) = \mathbf{1}_B(\tilde{\mathbf{x}}) \quad \text{for } \mu^{[k]}\text{-a.e. } \mathbf{x} \in X^{[k]} .$$

Definition.
A subset A of X belongs to $\mathcal{Z}_{k-1}$ if and only if there exists $B \subset X^{2^k-1}$ with

$$\mathbf{1}_A(x_{\mathbf{0}}) = \mathbf{1}_B(\tilde{\mathbf{x}}) \quad \text{for } \mu^{[k]}\text{-a.e. } \mathbf{x} \in X^{[k]} \ .$$

Lemma. *The 2^k-joining $\mu^{[k]}$ of X is relatively independent with respect to its projection on $Z_{k-1}^{[k]}$.*

Corollary. *For $f \in L^\infty(\mu)$,*

$$|||f|||_k = 0 \Longleftrightarrow \mathbb{E}(f \mid Z_{k-1}) = 0 \ .$$

The factors are built!

Review

We want a factor Y so that the average converges to 0 in $L^2(\mu)$ when $\mathbb{E}(f_i \mid \mathcal{Y}) = 0$ for some i.

We constructed factors

$$Z_0 \leftarrow Z_1 \leftarrow \cdots \leftarrow Z_{k-1} \leftarrow Z_k \leftarrow \cdots \leftarrow X$$

of X that work for arithmetic progressions.

Z_0 is the trivial factor of X.
Z_1 is the Kronecker factor of X.

Definition.
X is a system of level *k if $Z_k(X) = X$.*

For every X, $Z_k(X)$ is of level k.
Therefore, we only have to describe the systems of level k.

We know:
X is of level 0 $\iff$ X is trivial.
Systems of level 1 are ergodic rotations.

Later we show that systems of level 2 are Conze-Lesigne systems.

Nilpotent groups

For g, h in a group G, write $[g, h]$ for the commutator $g^{-1}h^{-1}gh$ of g and h.
If $A, B \subset G$, write $[A, B]$ for the subgroup of G generated by $\{[a, b] : a \in A, b \in B\}$.

Define
$G^{(1)} = G$ and $G^{j+1} = [G, G^{(j)}]$ for $j \geq 1$.

G is nilpotent of level k if $G^{(k+1)} = \{1\}$.

Definition.
*If G is a k-step nilpotent Lie group and Λ is a discrete cocompact subgroup, then the compact space $X = G/\Lambda$ is a k-*step nilmanifold*.*

G acts on G/Λ by left translation

$$T_a(x\Lambda) = (ax)\Lambda \ .$$

Definition.
There is a unique probability measure μ *on* X *that is invariant under the action of* G *by left translations, called the* Haar measure.

Fix an element $a \in G$.
The system $(G/\Lambda, \mathcal{G}/\Lambda, T_a, \mu)$ *is a* k-step nilsystem *and* T_a *is a* nilrotation.

In Lecture 3, we show:

Theorem. *A system of level* k *is an inverse limit of* k*-step nilsystems.*

Theorem. *Assume that* G *is spanned by the connected component of* 0 *and* a.
Then the nilsystem G/Λ *is ergodic if and only if the action of* a *on* $G/G^{(2)}\Lambda$ *is ergodic. In this case, the Kronecker factor of* X *is* $G/G^{(2)}\Lambda$.

Follows from results of Lesigne.
See also Leibman.

Typical example of a non-abelian two-step ergodic nilsystem.

G is the Heisenberg group

$$G = \left\{ \begin{pmatrix} 1 & x_1 & x_2 \\ 0 & 1 & x_3 \\ 0 & 0 & 1 \end{pmatrix} : x_i \in \mathbb{R} \right\}$$

with matrix multiplication.
It is two-step nilpotent.

$$\Lambda = \left\{ \begin{pmatrix} 1 & z_1 & z_2 \\ 0 & 1 & z_3 \\ 0 & 0 & 1 \end{pmatrix} : z_i \in \mathbb{Z} \right\} \; ; \; a = \begin{pmatrix} 1 & a_1 & a_2 \\ 0 & 1 & a_3 \\ 0 & 0 & 1 \end{pmatrix}$$

where $a_1, a_3 \in \mathbb{R}$ are rationally independent and $a_2 \in \mathbb{R}$.

G/Λ is compact and T_a is a nilrotation on G/Λ.

We have $G/G^{(2)}\Lambda \simeq \mathbb{T}^2$ and rotation on $\mathbb{T}^2$ by (a_1, a_3) is ergodic. Previous theorem gives T_a is ergodic and Kronecker is factor induced by functions on x_1, x_3.

Convergence on nilmanifolds

Theorem. *Let $X = G/\Lambda$ be a nilmanifold with Haar measure μ and let $a_1, \ldots, a_\ell$ be commuting elements of G.*
If the group spanned by the translations $a_1, \ldots, a_\ell$ acts ergodically on (X, μ), then X is uniquely ergodic for this group.

Parry proved this for one translation. Same proof works for commuting. Proof based on argument of Furstenberg. Also Leibman.

Theorem. *Let $X = G/\Lambda$ be a nilmanifold and $T : X \to X$ be a nilrotation. Then for any continuous function f on X, the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$$

converge everywhere on X as $N \to \infty$.

Special case of theorem of Shah.
Lesigne proved this when G is connected.
Leibman proved general case.

Idea: For each $x \in X$, the closed orbit $Y = \overline{\{T^n x : n \in \mathbb{Z}\}}$ is minimal by distality. Give Y the structure of a nilmanifold and use Parry result that a minimal nilmanifold is uniquely ergodic.

Corollary. *For any* $k \in \mathbb{N}$, $a_1, \ldots, a_k \in G$, $x_1, \ldots,$ $x_k \in X$ *and continuous functions* $f_1, \ldots, f_k$,

$$\lim_{N \to +\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(a_1^n x_1) \ldots f_k(a_k^n x_k)$$

exists.

Proof. Apply theorem to the element $a = (a_1, \ldots, a_k)$ of the group G^k at the point $x = (x_1, \ldots, x_k) \in X^k = G^k/\Lambda^k$ with the continuous function $f_1(x_1) \ldots f_k(x_k)$. □

Corollary. *Convergence for arithmetic progressions in* $L^2(\mu)$.

Proof. Characteristic factor is an inverse limit of $k-1$-step nilsystems. Use previous corollary with $a_i = a^i$ for $i = 1, \ldots, k$ to prove this for a nilsystem and by density of continuous functions, have result for the inverse limit. □

Characteristic factors for polynomials

We consider

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{n=0}^{N-1} T^{p_1(n)}f_1\cdot T^{p_2(n)}f_2\cdot\ldots\cdot T^{p_\ell(n)}f_\ell\ .$$

Same factors work for polynomial averages:

Theorem. *Assume that $\{p_1,\ldots,p_\ell\}$ are non-constant integer valued polynomials so that p_i-p_j is not constant for all $i\neq j$.*
There exists an integer $k\geq 0$ so that for any ergodic system $(X,\mathcal{X},\mu,T)$ and functions $f_1,\ldots,f_\ell\in L^\infty(\mu)$, $Z_k(X)$ is characteristic for the associated polynomial average.

This means: if $\mathbb{E}\Big(f_m|\mathcal{Z}_k(X)\Big)=0$, then

$$\sup_M\Big\|\frac{1}{N}\sum_{n=M}^{M+N-1} T^{p_1(n)}f_1\cdot\ldots\cdot T^{p_\ell(n)}f_\ell\Big\|_{L^2(\mu)}\to 0$$

as $N\to+\infty$.

Convergence for polynomials

Theorem (Leibman).
If (X, μ, T) *is a nilsystem,* $p_1, \ldots, p_\ell$ *are integer valued polynomials and* $f_1, \ldots, f_\ell$ *are continuous functions on* X*, then for all sequences* $\{M_i\}$ *and* $\{N_i\}$ *with* $N_i \to +\infty$*, the averages*

$$\frac{1}{N_i} \sum_{n=M_i}^{M_i+N_i-1} f_1(T^{p_1(n)}x) \ldots f_\ell(T^{p_\ell(n)}x)$$

converge for every $x \in X$*.*

Corollary.
Convergence for polynomial averages.

Proof. The characteristic factor for a polynomial average is an inverse limit of k-step nilsystems for some k and Leibman's Theorem gives result for nilsystems. By density, pass to the inverse limit. $\square$

Averages along cubes

2-dimensional cube:

$$f(x)f(T^m x)f(T^n x)f(T^{m+n} x)$$

Theorem (Bergelson). *Let* $(X, \mathcal{X}, \mu, T)$ *be an invertible measure preserving system. Then if* $f_1, f_2, f_3 \in L^\infty(\mu)$*,*

$$\lim_{N\to+\infty} \frac{1}{N^2} \sum_{n,m=0}^{N-1} f_1(T^n x) f_2(T^m x) f_3(T^{n+m} x)$$

exists in $L^2(\mu)$*.*

Also same result over interval $[M, N]$, as $N - M \to +\infty$.

3-dimensional cube:

$$f(x)f(T^m x)f(T^n x)f(T^{m+n} x)$$
$$f(T^p x)f(T^{m+p} x)f(T^{n+p} x)f(T^{m+n+p} x)$$

Theorem. *Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving probability system and let f_j, $1 \leq j \leq 7$, be 7 bounded measurable functions on X. Then the averages*

$$\frac{1}{N^3} \sum_{m,n,p=0}^{N-1} f_1(T^m x) f_2(T^n x) f_3(T^{m+n} x)$$
$$f_4(T^p x) f_5(T^{m+p} x) f_6(T^{n+p} x) f_7(T^{m+n+p} x)$$

converge in $L^2(\mu)$ as $N \to +\infty$.

Again, same result holds over intervals $[M, M']$, $[N, N']$, $[P, P']$, as $M' - M, N' - N, P' - P \to +\infty$.

More generally, this theorem holds for cubes of $2^k - 1$ functions.

General averages along cubes

For $\epsilon = \epsilon_1 \dots \epsilon_k \in \{0,1\}^k$ and $\mathbf{n} = (n_1, \dots, n_k) \in \mathbb{Z}^k$,

$$\epsilon \cdot \mathbf{n} = \epsilon_1 n_1 + \epsilon_2 n_2 + \cdots + \epsilon_k n_k$$

$\mathbf{0}$ denotes the element $00\dots0$ of $\{0,1\}^k$.

Theorem. *Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving probability system, let $k \geq 1$ be an integer and let f_ϵ, $\epsilon \in \{0,1\}^k \setminus \{\mathbf{0}\}$, be $2^k - 1$ bounded functions on X. Then the averages*

$$\frac{1}{N^k} \cdot \sum_{\mathbf{n} \in [0,N)^k} \prod_{\substack{\epsilon \in \{0,1\}^k \\ \epsilon \neq \mathbf{0}}} f_\epsilon(T^{\epsilon \cdot \mathbf{n}} x)$$

converge in $L^2(\mu)$ as N tends to $+\infty$.

Same result for averages with $\mathbf{n} \in [M_1, N_1) \times \cdots \times [M_k, N_k)$, as $N_1 - M_1, \dots, N_k - M_k \to +\infty$.

Lower bounds

Restricting to the indicator function of a measurable set:

Theorem. *Let* $(X, \mathcal{X}, \mu, T)$ *be an invertible measure preserving probability system and let* $A \in \mathcal{X}$*. Then the limit of the averages*

$$\prod_{i=1}^{k} \frac{1}{N_i - M_i} \cdot \sum_{\substack{n_1 \in [M_1, N_1) \\ \cdots\cdots\cdots \\ n_k \in [M_k, N_k)}} \mu\Big(\bigcap_{\epsilon \in \{0,1\}^k} T^{\epsilon \cdot \mathbf{n}} A \Big)$$

exists and is greater than or equal to $\mu(A)^{2^k}$ *when* $N_1 - M_1, N_2 - M_2, \ldots, N_k - M_k$ *tend to* $+\infty$*.*

Corollary. *For every* $\delta > 0$ *the subset*

$$\Big\{ \mathbf{n} \in \mathbb{Z}^k : \mu\Big(\bigcap_{\epsilon \in \{0,1\}^k} T^{\epsilon \cdot \mathbf{n}} A \Big) > \mu(A)^{2^k} - \delta \Big\}$$

of $\mathbb{Z}^k$ *is syndetic.*

Characteristic factors for cubes

Theorem.
The factor $Z_{k-1}(X)$ is characteristic for the $L^2(\mu)$ convergence of the k-dimensional cubic averages.

This means:
This average converges to 0 in $L^2(\mu)$ if $\mathbb{E}\Big(f_\epsilon | \mathcal{Z}_{k-1}(X)\Big) = 0$ for some $\epsilon \in \{0,1\}^k \setminus \mathbf{0}$.

Proof of L^2 convergence relies on first understanding convergence for numerical averages:

$$\frac{1}{N^k} \sum_{n_1,\ldots,n_k=0}^{N-1} \int \prod_{\epsilon \in \{0,1\}^k} f_\epsilon \circ T^{\epsilon \cdot n} \, d\mu \ .$$

Limit for integral of 2^k terms

We have reduced to the case that:
$X = G/\Lambda$ is a k-step nilmanifold where G is k-step nilpotent Lie group; Λ is discrete cocompact subgroup.

Recall definition of the measure $\mu^{[k]}$ on $X^{[k]} = X^{2^k}$. Defined inductively with

- $\mu^{[0]} := \mu$.
- $\mu^{[k+1]}$ is the relatively independent square of $\mu^{[k]}$ over $\mathcal{I}^{[k]}$, where $\mathcal{I}^{[k]}$ is the invariant σ-algebra of $(X^{[k]}, \mu^{[k]}, T^{[k]})$.

For this set up, the measure $\mu^{[k]}$ on $X^{[k]}$ has a simple description: it can be identified with the Haar measure on some submanifold of X^{2^k}.

Definition. *Let* $G^{[k]}_{k-1}$ *be the subgroup of* $G^{[k]}$ *spanned by*

$\{g^{[k]}_\alpha : g \in G$ *and* α *is a* $k-1$ *face of* $\{0,1\}^k\}$.

Called the side subgroup.

Define X_k to be the nilmanifold

$$X_k = G^{[k]}_{k-1}/(\Lambda^{[k]} \cap G^{[k]}_{k-1})$$

embedded in $X^{[k]}$ in the natural way.

Proposition. $\mu^{[k]}$ *is the Haar measure of* X_k.

The transformations $T^{[k]}_\alpha$, where α is a side of $\{0,1\}^k$, span an ergodic action on $(X^{[k]}, \mu^{[k]})$. Thus this action on $X^{[k]}$ is uniquely ergodic.

Theorem. *The averages*

$$\frac{1}{N^k} \sum_{n_1,\ldots,n_k=0}^{N-1} \int \prod_{\epsilon \in \{0,1\}^k} f_\epsilon \circ T^{\epsilon \cdot \mathbf{n}} \, d\mu \ .$$

converge to

$$\int_{X^{2^k}} \prod_{\epsilon \in \{0,1\}^k} f_\epsilon(x_\epsilon) \, d\mu^{[k]}(\mathbf{x}) \ .$$

L^2 Convergence for cubic averages

Every $\mathbf{x} \in X^{[k]}$ is written $\mathbf{x} = (x_{\mathbf{0}}, \tilde{\mathbf{x}})$ with $x_{\mathbf{0}} \in X$ and $\tilde{\mathbf{x}} \in X^{2^k-1}$. We partition X_k according to the coordinate $\mathbf{0}$. For $x \in X$,

$$M_x = \left\{ \tilde{\mathbf{x}} \in X^{2^k-1} : (x, \tilde{\mathbf{x}}) \in X_k \right\} .$$

For almost every $x \in X$, M_x is a nilmanifold uniquely ergodic for the action spanned by the transformations $T_\alpha^{[k]}$ where α is a side of $\{0,1\}^k$ not containing $\mathbf{0}$. We write μ_x for its Haar measure

We deduce the L^2 convergence of the cubic averages with $2^k - 1$ terms to

$$\int_{X^{[k]}} \prod_{\epsilon \in \{0,1\}^k \setminus \{\mathbf{0}\}} f_\epsilon(x_\epsilon) \, d\mu_x^{[k]}(\mathbf{y}) \ .$$

We recall some notation

Let (X, μ, T) be an ergodic system.

$X^{[k]} := X^{2^k}$. Points of $X^{[k]}$ are written

$$\mathbf{x} = \left(x_\epsilon : \epsilon \in \{0,1\}^k\right) .$$

$\mu^{[k]}$ is a measure on $X^{[k]}$, invariant under $T^{[k]}$.
$\mathcal{I}^{[k]}$ is the invariant σ-algebra of $(X^{[k]}, \mu^{[k]}, T^{[k]})$.
For $f \in L^\infty(\mu)$,

$$|\!|\!| f |\!|\!|_k^{2^k} := \int_{X^{[k]}} \prod_{\epsilon \in \{0,1\}^k} f(x_\epsilon)\, d\mu^{[k]}(\mathbf{x}) .$$

The factor Z_k of X is characterized by: for $f \in L^\infty(\mu)$,

$$\mathbb{E}(f \mid \mathcal{Z}_k) = 0 \iff |\!|\!| f |\!|\!|_{k+1} = 0 .$$

The factors Z_k form an increasing sequence:

$$Z_0 \leftarrow Z_1 \leftarrow Z_k \leftarrow Z_{k+1} \cdots \leftarrow X .$$

Z_0 is the trivial factor,
Z_1 is the Kronecker factor of X.

Definition.
X is a system of level k *if* $Z_k(X) = X$.

For every X, $Z_k(X)$ is a system of level k.
Thus we have to describe systems of level k.

X is of level 0 $\iff$ X is trivial.
Systems of level 1 are ergodic rotations.
We shall see later that systems of level 2 are Conze-Lesigne algebras.

Theorem. *A system of level k is an inverse limit of k-steps nilsystems.*

How to find the group?

Idea:
We consider the group of transformations of X preserving the 'cubic structure'.

Notation.

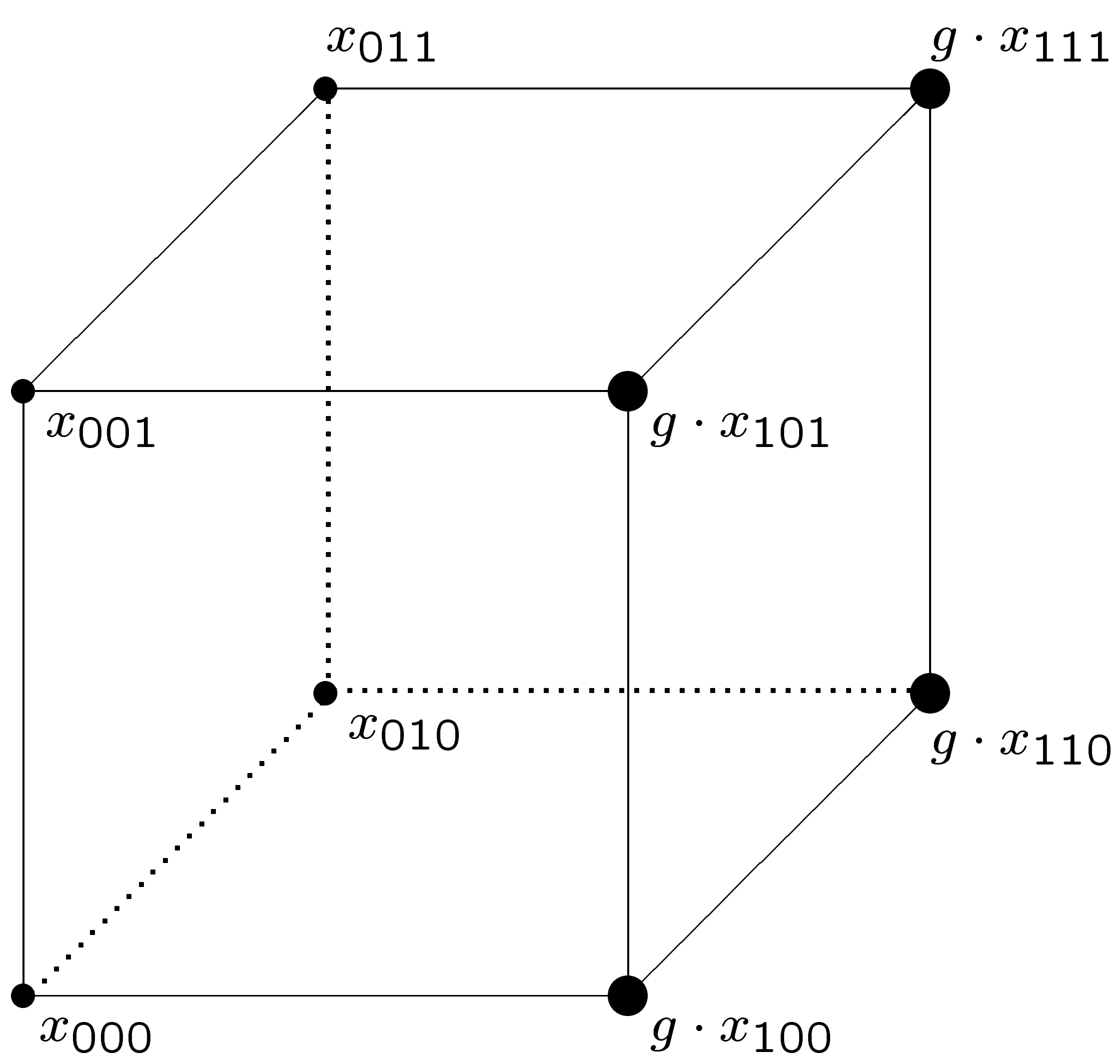
When $x \mapsto g \cdot x$ is a transformation of X and $\alpha \subset \{0,1\}^k$, we write $g_\alpha^{[k]} : X^{[k]} \to X^{[k]}$ for the transformation defined by

$$(g_\alpha^{[k]} \cdot \mathbf{x})_\epsilon = \begin{cases} g \cdot x_\epsilon & \text{if } \epsilon \in \alpha, \\ x_\epsilon & \text{else.} \end{cases}$$

We use this notation for 'sides', vertices',... of $\{0,1\}^k$.

We know that for every k and every side α of $\{0,1\}^k$ the measure $\mu^{[k]}$ is invariant under $T_\alpha^{[k]}$.

The transformation $g_\alpha^{[3]}\mathbf{x}$ for $\alpha = \{100, 101, 110, 111\}$



We recall that $\mathcal{I}^{[k]}$ is the invariant σ-algebra of $(X^{[k]}, \mu^{[k]}, T^{[k]})$.

Lemma. *For a measure preserving transformation $x \mapsto g \cdot x$ of X, TFAE:*

1. *For any k and any side α of $\{0,1\}^k$, the transformation $g_\alpha^{[k]}$ of $X^{[k]}$ leaves the measure $\mu^{[k]}$ invariant.*
2. *Same as 1. plus: and $g_\alpha^{[k]}$ maps the σ-algebra $\mathcal{I}^{[k]}$ to itself.*
3. *For every k, the transformation $g^{[k]}$ of $X^{[k]}$ leaves the measure $\mu^{[k]}$ invariant and acts trivially on the σ-algebra $\mathcal{I}^{[k]}$.*

Remark. By symmetry, we can substitute 'some side' for 'any side' in 1. and 2.

Definition. $\mathcal{G} = \mathcal{G}(X)$ *is the group of measure preserving transformations of* X *with these properties.*

Properties.

- $\mathcal{G}$ is a Polish group.
- $T \in \mathcal{G}$
- If $TS = ST$ then $S \in \mathcal{G}$.
- If X is a compact abelian group and T an ergodic rotation, then $\mathcal{G} = X$ acting on itself by translations.
- If $g \in \mathcal{G}$ then for every k, it leaves $\mathcal{Z}_k$ invariant.
- If $X \to Y$ is a factor map and $g \in \mathcal{G}(X)$ leaves $\mathcal{Y}$ invariant, then the transformation induced on Y belongs to $\mathcal{G}(Y)$.

The inverse problem: Can a given $h \in \mathcal{G}(Y)$ be lifted to an element of $\mathcal{G}(X)$?

Commutators

We recall that $[g,h] := g^{-1}h^{-1}gh$.

Remark. For $g,h : X \to X$ and $\alpha,\beta \in \{0,1\}^k$,

$$\left[g_\alpha^{[k]}, h_\beta^{[k]}\right] = [g,h]_{\alpha\cap\beta}^{[k]} \ .$$

We recall that $\mathcal{G}^{(1)} = \mathcal{G}$ and $\mathcal{G}^{(i+1)} = [\mathcal{G}^{(i)}, G]$.

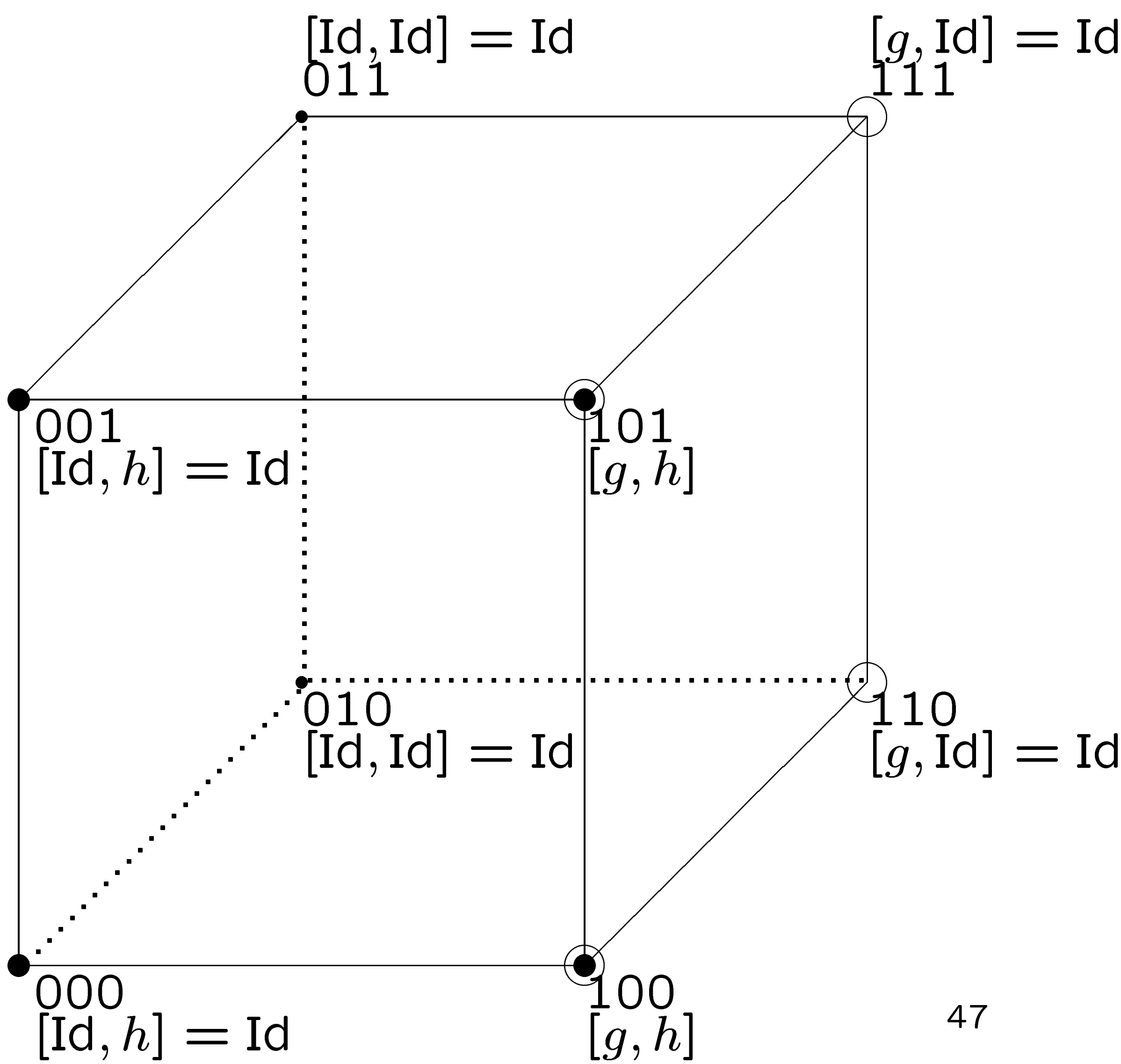
Lemma. *Let $g \in \mathcal{G}^{(j)}$ and α a face of dimension $k-j$ of $\{0,1\}^k$.*
Then the measure $\mu^{[k]}$ is invariant under $g_\alpha^{[k]}$. Moreover this transformation maps the σ-algebra $\mathcal{I}^{[k]}$ to itself.

Proposition. *If X is a system of level k then $\mathcal{G}$ is a k-step nilpotent group.*

Proof of the remark.

$$\alpha = \{100, 101, 110, 111\}$$
$$\beta = \{000, 100, 001, 101\}$$
$$\gamma = \{100, 101\} = \alpha \cap \beta \ .$$

We check that $\left[g_\alpha^{[3]}, h_\beta^{[3]}\right] = [g,h]_\gamma^{[3]}$



Notation.
Let U be a group and $\rho : Y \to U$ a map. Define $\Delta^k\rho : Y^{[k]} \to U$ by

$$\Delta^k\rho(\mathbf{y}) = \prod_{\epsilon\in\{0,1\}^k} \rho(y_\epsilon)^{(-1)^{|\epsilon|}},$$

where $|\epsilon| = \epsilon_1 + \cdots + \epsilon_k$.

Examples: $\Delta\rho(\mathbf{y}) = \rho(y_0) \cdot \rho(y_1)^{-1}$.

$$\Delta^2(\rho)(\mathbf{y}) = \rho(y_{00})\rho(y_{01})^{-1}\rho(y_{10})^{-1}\rho(y_{11})$$

With additive notation, $\Delta\rho(\mathbf{y}) = \rho(y_0) - \rho(y_1)$.

Definition. *Let (Y, ν, S) be an ergodic system, U a group, and $\rho : Y \to U$ a map.*

We say that ρ is a cocycle of type k *if $\Delta^k\rho : Y^{[k]} \to U$ is a coboundary of $(Y^{[k]}, \nu^{[k]}, S^{[k]})$.*

This means that there exists $F : Y^{[k]} \to U$ with

$$\Delta^k\rho(\mathbf{y}) = F(S^{[k]}\mathbf{y}) \cdot F(\mathbf{y})^{-1} \ (\nu^{[k]}\text{-a.e.})$$

A cocycle of type 0 is a coboundary ($\rho = f \circ T \cdot f^{-1}$ for some f).
A cocycle of type 1 with values in $\mathcal{S}^1$ is a quasi-coboundary (the product of a constant and a coboundary)

Relations between consecutive factors

Proposition. *Let X be a system of level k. We write (Y, ν, S) for the factor Z_{k-1}. Then X is an extension of Y by a compact abelian group U:*

$$X = Y \times U \ ; \ T(y, u) = (Sy, \rho(y)u) \ .$$

This extension is given by a cocycle $\rho : Y \to U$ of type k.

Steps of the proof.
1) *X is an isometric extension of Y.*

We can write $X = Y \times G/H$, where G is a compact group and H a compact subgroup. There is an obvious action of G on X.

2) *For $g \in G$ and α an edge of $\{0,1\}^k$, $g_\alpha^{[k]}$ acts trivially on $\mathcal{I}^{[k]}$.*

3) *For $g \in G$, the associated transformation of X belongs to the center of $\mathcal{G}(X)$.*
In particular G is abelian.

4) For $\chi \in \widehat{U}$, we define $f(y,u) = \chi(u)$ and

$$F(\mathbf{x}) = \prod_{\epsilon \in \{0,1\}^k} f(x_\epsilon)^{(-1)^{|\epsilon|}} .$$

We have $|||f|||_{k+1} \neq 0$, thus $\mathbb{E}(F \mid \mathcal{I}^{[k]}) \neq 0$.

Cocycles of type 2 and the Conze-Lesigne Equation

Proposition.
Let (Z, m, T) be an ergodic rotation and let $\rho : Z \to \mathbb{T}^d$ be a cocycle of type 2.
Then for every $s \in Z$ there exists $f : Z \to \mathbb{T}^d$ and $c \in \mathbb{T}^d$ with

$$\rho(sz) - \rho(z) = f(Tz) - f(z) + c \ . \qquad \text{(CL)}$$

The transformations of $Z \times \mathbb{T}^d$ of the form

$$(z, u) \mapsto \Big(sz, u + f(z)\Big)$$

where s, f, c satisfy (CL) form a group $\mathcal{G}_\rho$.

- $\mathcal{G}_\rho$ is 2-step nilpotent and locally compact.
- The action of $\mathcal{G}_\rho$ on $Z \times \mathbb{T}^d$ is transitive, ergodic and $Z \times \mathbb{T}^d = \mathcal{G}_\rho / \mathbb{Z}^d$
- $\mathcal{G}_\rho$ is a Lie group if Z is a Lie group.

Lemma. *There exist a Lie quotient Z' of Z and a cocycle ρ' cohomologous to ρ and measurable wrt $\mathcal{Z}'$.*

The countability Lemma. *Up to the addition of quasi-coboundaries there exist only countably many cocycles of type 2 on Z.*

Systems of level 2

Let X be a system of level 2.
X is an extension of $Z = Z_1(X)$ by a compact abelian group U, given by a cocycle $\rho : Z \to U$ of type 2.
If U is a torus, then the groups $\mathcal{G}(X)$ and $\mathcal{G}_\rho$ are actually the same and thus $\mathcal{G}(X)$ acts transitively on X.

Lemma. *If Z is a Lie group and U is a torus, then X is a 2-step nilsystem.*

Lemma. *U is connected.*

Thus U is an inverse limit of tori.

Proposition. *Every system of level 2 is an inverse limit of 2-step nilsystems.*

We go back to the general case
and we summarize.

$$Z_1 \leftarrow Z_2 \leftarrow \cdots \leftarrow Z_{k-1} \leftarrow Z_k \leftarrow \cdots \leftarrow X \ .$$

• Z_1 is the Kronecker factor of X.
• Z_k is an extension of Z_{k-1} by a compact abelian group U_k.
• This extension is given by a cocycle
$\rho_k : Z_{k-1} \to U_k$ of type k.
This means that there exists $F_k : Z_{k-1}^{[k]} \to U_k$ with

$$\Delta^k \rho_k = F_k \circ T^{[k]} \cdot F_k^{-1} \ (\mu^{[k]}\text{-a.e.})$$

• U_2 is connected.
• Z_2 is an inverse limit of a sequence of 2-step nilsystems.
• When Z is an ergodic rotation there exist (up to quasi-coboundaries) only countably many cocycles $Z \to \mathbb{T}^d$ of type 2.

We generalize the last three properties for higher levels.
The proof by induction uses a lot of technology of cocycles plus a step ladder.

The step ladder

Let (X, μ, T) be a system of level k.
We recall the definition of the measure μ_s for $s \in Z$:
For $f_0, f_1 \in L^\infty(\mu)$,

$$\int_{X\times X} f_0(x_0) f_1(x_1)\, d\mu_s(x_0, x_1) := \int_Z \mathbb{E}(f_0 \mid \mathcal{Z})(sz)\, \mathbb{E}(f_1 \mid \mathcal{Z})(z)\, dm(z) \ .$$

The ergodic decomposition of $\mu \times \mu$ for $T \times T$ is

$$\mu \times \mu = \int_Z \mu_s\, dm(s) \ .$$

We write (Y, ν, S) for $Z_{k-1}(X)$.
X and Y have the same Kronecker factor Z.
We define the measures ν_s, $s \in Z$, in a similar way.
For almost every $s \in Z$ we write

$$X_s = (X \times X, \mu_s, T \times T)\ ;\ Y_s = (Y \times Y, \nu_s, S \times S)$$

and we study the relations between these systems.

We recall that X is an extension of Y by a compact abelian group U_k, given by a cocycle $\rho_k : Y \to U$ of type k.

Proposition (The step ladder).

• *X_s is a system of level k*

• *Y_s is a system of level $k-1$.*

• *X_s is an extension of Y_s by $U_k \times U_k$, given by the cocycle $(y_0, y_1) \mapsto (\rho_k(y_0), \rho_k(y_1))$ of type k.*

• *$Z_{k-1}(X_s)$ is an extension of Y_s by U_k, given by the cocycle $(y_0, y_1) \mapsto \rho_k(y_0)\rho_k(y_1)^{-1}$ of type $k-1$.*

$$Z_{k-2}(X_s) \overset{U_{k-1}}{\longleftarrow} Y_s \overset{U_k}{\longleftarrow} Z_{k-1}(X_s) \overset{U_k}{\longleftarrow} X_s$$

$$Z_1 \leftarrow Z_2 \leftarrow \cdots \leftarrow Z_{k-1} \leftarrow Z_k \leftarrow \cdots \leftarrow X \ .$$

Z_i is an extension of Z_{i-1} by U_i .

Lemma. *For each i, the compact abelian group U_i is connected.*

Definition. *A system X of level k is* toral *if the compact abelian group Z_1 is a Lie group (i.e. admits a torus as an open subgroup) and each Z_i is a torus.*

Proposition. *Every system of level k is an inverse limit of toral systems of level k.*

The countability Lemma can be extended only in a weaker form.

Lemma. *Let X be an ergodic system, $(\Omega, \mathcal{B}, P)$ a standard probability space, and $\omega \mapsto \rho_\omega$ a measurable map with values in the set of co-cycles $X \to \mathbb{T}^d$ of type k.*
Then there exists $\Omega_0 \in \mathcal{B}$, $P(\Omega_0) > 0$, such that ρ_ω is constant on Ω_0 (up to quasi - coboundaries).

Theorem.
Every toral system of level k is a nilsystem.

By induction. Assume it holds for $k-1$.
Let X be a toral system of level k.
We write (Y,ν,S) for $Z_{k-1}(X)$.
$Y=\mathcal{G}/\Gamma$, with $\mathcal{G}=\mathcal{G}(Y)$.

For $j=k,k-1,\dots,2,1$ we show that every element of $\mathcal{G}^{(j)}(Y)$ can be lifted as an element of $\mathcal{G}(X)$. The lift is built by solving a family of functional equations.

Recall that X is an extension of $Y=Z_{k-1}(X)$ by a torus $\mathbb{T}^d$. This extension is given by a cocycle $\rho : Y\to\mathbb{T}^d$ of type k.
There exists $F : Y^{[k]}\to\mathbb{T}^d$ with

$$\Delta^k\rho = F\circ T^{[k]} - F \qquad (*)$$

$\mu^{[k]}$-a.e.

The functional equations

For $\ell \leq k$, every face α of dimension ℓ of $\{0,1\}^k$ defines a projection $\pi_\alpha^{[k]} : X^{[k]} \to X^{[\ell]}$.

Lemma. *For $j \leq k$ and $g \in \mathcal{G}^{(j)}(Y)$ there exists $\phi_g : Y \to \mathbb{T}^d$ such that, for every face α of dimension $k+1-j$ of $\{0,1\}^k$,*

$$F \circ g_\alpha^{[k]} - g = \Delta^{k+1-j} \phi_g \circ \pi_\alpha^{[k]} \ . \qquad (**)$$

$\mu^{[k]}$-a.e.

The proof of this Lemma uses all the machinery. . .

Let ϕ_g be given by the Lemma.
The transformation $(y,u) \mapsto (g \cdot y, u + \phi_g(y))$ of $X = Y \times \mathbb{T}^d$ is a lift of g in $\mathcal{G}(X)$.

We deduce that $\mathcal{G}(X)$ acts transitively on X.
It is easy to check that $\mathcal{G}(X)$ is a Lie group.
The theorem follows.

Idea of the proof of the Lemma. Assume it holds for $j+1$. For $g \in \mathcal{G}^{(j)}$ we define

$$\theta_g = \phi_{[g^{-1},T^{-1}]} \circ Tg + \rho \circ g - \rho \ .$$

Let β be a face of dimension $k-j$ of $\{0,1\}^k$. From (*) and (**) we get

$$(F \circ g_\beta^{[k]} - F) \circ T^{[k]} - (F \circ g_\beta^{[k]} - F) = \Delta^{k-j} \theta_g \circ \pi_\beta^{[k]} \ .$$

It follows that θ_g is a cocycle of type $k-j$.

By the countability Lemma there exists $A \subset G$, of positive measure, such that $\theta_g - \theta_h$ is a quasi-coboundary for $g, h \in A$:
For $g, h \in A$ there exist $\theta : Y \to \mathbb{T}^d$ and $c \in \mathbb{T}^d$ with $\theta_g - \theta_h = \theta \circ T - \theta + c$.
By (**) the map

$$F \circ (gh^{-1})_\beta^{[k]} - F - \Delta^{k-j}(\theta \circ h^{-1}) \circ \pi_\beta^{[k]} \ .$$

is invariant under $T^{[k]}$.
It follows that the relation (**) is satisfied with gh^{-1} and $\theta \circ h^{-1}$.
The announced result holds for g in an open subset of $\mathcal{G}^{(j)}$. □