Multiple recurrence theorem for measure preserving actions of a nilpotent group

A. Leibman

Abstract

The multidimensional ergodic Szemerédi theorem of Furstenberg and Katznelson, which deals with commuting transformations, is extended to the case where the transformations generate a nilpotent group:

Theorem. *Let $(X, \mathfrak{B}, \mu)$ be a measure space with $\mu(X) < \infty$ and let $T_1, \ldots, T_k$ be measure preserving transformations of X generating a nilpotent group. Then for any $A \in \mathfrak{B}$ with $\mu(A) > 0$,*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu\Big(T_1^{-n}A \cap \ldots \cap T_k^{-n}A\Big) > 0.$$

Our main result also generalizes the polynomial Szemerédi Theorem in [BL1]. In the course of the proof we describe a relatively simple form to which any unitary action of a finitely generated nilpotent group on a Hilbert space and any measure preserving action of a finitely generated nilpotent group on a probability space can be reduced.

0. Introduction

0.1. We were inspired by the following two theorems:

Theorem T. ([FW]) *Let (X, ρ) be a compact metric space and let $T_1, \ldots, T_k$ be commuting homeomorphisms of X. Then for any $\varepsilon > 0$ there exist $x \in X$ and $n \in \mathbb{N}$ such that*

$$\rho\Big(T_i^n x, x\Big) < \varepsilon \text{ for every } i = 1, \ldots, k.$$

Theorem M. ([FK1]) *Let $(X, \mathfrak{B}, \mu)$ be a measure space with $\mu(X) < \infty$, let $T_1, \ldots, T_k$ be commuting measure preserving transformations of X and let $A \in \mathfrak{B}$ with $\mu(A) > 0$. Then*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu\Big(T_1^{-n}A \cap \ldots \cap T_k^{-n}A\Big) > 0.$$

1991 *Mathematics Subject Classification:* Primary 28D15, 47A35; Secondary 05A17, 46C50.
Key words: nilpotent group, measure preserving transformations, Szemerédi theorem.

Theorems T and M are both statements about recurrence. Theorem T asserts that, given a finite collection of commuting transformations of a compact metric space X, a point $x \in X$ returns to any its neighborhood simultaneously under the action of these transformations (topological recurrence). It follows from Theorem M that, given a finite collection of commuting transformations of a probability space X, every set $A \subseteq X$ of positive measure contains a subset of positive measure returning into A simultaneously under the action of these transformations (measurable recurrence). Since every compact metric space with an amenable group acting on it possesses an invariant Borel measure, Theorem T is a corollary of Theorem M. At the same time, Theorem T is an important component of the proof of Theorem M. Note also that both theorems are well known classical results in the "single recurrence" case $k = 1$.

0.2. A natural question is whether multiple recurrence theorems, Theorems T and M, hold true if the requirement that $T_1, \ldots, T_k$ commute is omitted. An elegant example of Furstenberg ([F2], Ch. 2) shows that, generally speaking, it is not so even if $T_1, \ldots, T_k$ generate a metabelian group (i.e. a group whose commutator is a commutative group). The class of nilpotent groups is in a sense closest to that of commutative groups. The following theorem, extending Theorem T to the case where $T_1, \ldots, T_k$, generate a nilpotent group, was proved in [L]:

Theorem NT. *Let (X, ρ) be a compact metric space, let G be a nilpotent group of homeomorphisms of X, let $T_1, \ldots, T_t \in G$ and let $p_{i,j}: \mathbb{Z} \longrightarrow \mathbb{Z}$ be polynomials satisfying $p_{i,j}(0) = 0$, $i = 1, \ldots, I$, $j = 1, \ldots, t$. For any $\varepsilon > 0$ there exist $x \in X$ and $n \in \mathbb{N}$ such that*

$$\rho\Big(T_t^{p_{i,t}(n)} \ldots T_1^{p_{i,1}(n)} x, x\Big) < \varepsilon \textit{ for every } i = 1, \ldots, I.$$

0.3. Our purpose is to obtain an analogous generalization of Theorem M. We prove:

Theorem NM′. *Let $(X, \mathfrak{B}, \mu)$ be a measure space with $\mu(X) < \infty$, let G be a nilpotent group of measure preserving transformations of X (acting on X from the right), let $T_1, \ldots, T_t \in G$, and let $p_{i,j}: \mathbb{Z} \longrightarrow \mathbb{Z}$ be polynomials satisfying $p_{i,j}(0) = 0$, $i = 1, \ldots, I$, $j = 1, \ldots, t$. Then for any $A \in \mathfrak{B}$ with $\mu(A) > 0$,*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu\Big(\bigcap_{i=1}^{I} A\big(T_t^{p_{i,t}(n)} \ldots T_1^{p_{i,1}(n)}\big)^{-1}\Big) > 0.$$

Remark. When dealing with an action of a noncommutative group G on a measure space X, we will follow the convention that G acts on X from the right. Such an action induces a natural action of G on functions on X from the left. Clearly, Theorem NM′ remains true if in the formulation of this theorem G is assumed to act on X from the left.

0.4. Some new, polynomial expressions of the form $T_t^{p_t(n)} \ldots T_1^{p_1(n)}$ arise in the formulations of Theorems NT and NM′ as compared with Theorems T and M. The "commutative polynomial" multiple recurrence theorems, where the transformations $T_1, \ldots, T_t$ commute and such polynomial expressions are present, that is Theorems NT and NM′ with commutative G, were obtained in [BL1].

0.5. As a matter of fact, we will prove the "uniform and multiparameter" version of Theorem NM$'$:

Theorem NM. *Let $(X,\mathfrak{B},\mu)$ be a measure space with $\mu(X)<\infty$, let G be a nilpotent group of measure preserving transformations of X (acting on X from the right), let $T_1,\dots,T_t\in G$, let $d\in\mathbb{N}$ and let $p_{i,j}\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ be polynomials satisfying $p_{i,j}(0)=0$, $i=1,\dots,I$, $j=1,\dots,t$. Then, for any $A\in\mathfrak{B}$ with $\mu(A)>0$, there exists $c>0$ such that the set*

$$S=\Bigl\{n\in\mathbb{Z}^d:\mu\Bigl(\bigcap_{i=1}^{I}A\bigl(T_t^{p_{i,t}(n)}\dots T_1^{p_{i,1}(n)}\bigr)^{-1}\Bigr)>c\Bigr\}$$

is syndetic (that is, has bounded gaps) in $\mathbb{Z}^d$.

The commutative version of Theorem NM (that is, Theorem NM with commutative G) was proved in [BM]. To derive Theorem NM$'$ from Theorem NM, note that in the case $d=1$ one has $s=\liminf\frac{\#(S\cap[0,N-1])}{N}>0$ and so,

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu\Bigl(\bigcap_{i=1}^{I}A\bigl(T_t^{p_{i,t}(n)}\dots T_1^{p_{i,1}(n)}\bigr)^{-1}\Bigr)>sc>0.$$

0.6. The proof of Theorem NM runs along the lines similar to those of the proof of Theorem M provided in [FK1] (and the proof of its uniform generalization in [BM]). Namely, given a measure space $\mathbf{X}=(X,\mathfrak{B},\mu)$ and a finitely generated nilpotent group G of measure preserving transformations of $\mathbf{X}$, we represent $\mathbf{X}$ as a transfinite tower of G-invariant factors:

$$\mathbf{X}_0=(X,\{\emptyset,X\},\mu)\longleftarrow\mathbf{X}_1=(X,\mathfrak{B}_1,\mu)\longleftarrow\dots\longleftarrow\mathbf{X}_\alpha=(X,\mathfrak{B}_\alpha,\mu)\longleftarrow\dots\longleftarrow\mathbf{X}$$

where for every limit ordinal α the σ-algebra $\mathfrak{B}_\alpha$ is generated by $\bigcup_{\beta<\alpha}\mathfrak{B}_\beta$ and for every ordinal α the extension $(\mathbf{X}_\alpha,G)\longleftarrow(\mathbf{X}_{\alpha+1},G)$ is *primitive* (see below). Theorem NM is proved by transfinite induction: we show that if Theorem NM holds for $(\mathbf{X}_\alpha,G)$ then it holds for $(\mathbf{X}_{\alpha+1},G)$. This allows us to deal with relatively simple primitive extensions of dynamical systems only. To clarify what a primitive extension is, let us first consider a unitary action of a group on a Hilbert space.

0.7. Let G be a finitely generated commutative group. Following [F2], we call a unitary action of G on a Hilbert space M *primitive* if a subgroup H of G acts compactly on M (that is, the orbit Hu of every $u\in M$ is precompact in M) and the action on M of every element of $G\setminus H$ is weakly mixing (that is, it has pure continuous spectrum). It is easy to see that for any unitary action of G on a Hilbert space M, M is decomposable into a direct sum of pairwise orthogonal G-invariant subspaces with a primitive action of G on each of them. In the case where G is generated by a single operator T, this decomposition is $M=M^c(T)\oplus M^w(T)$, where T has discrete spectrum on $M^c(T)$ and pure continuous spectrum on $M^w(T)$. In the general case, let H be the maximal subgroup of G with the property that the space $M^c(H)=\bigl\{u\in M:Hu\text{ is precompact}\bigr\}$ is nontrivial. Then $M^c(H)$ is closed and G-invariant, and every element T of $G\setminus H$ is weakly mixing on

$M^c(H)$: otherwise, we could pass to the nontrivial subspace $M^c(H) \cap M^c(T)$ and add T to H. So, the action of G on $M^c(H)$ is primitive; passing to $M^c(H)^\perp$ and using transfinite induction we can decompose M into the direct sum of such subspaces.

One could expect that the same holds for a nilpotent group G of unitary operators on a Hilbert space M. The ideal picture would be as follows: M contains a G-invariant subspace M' such that a normal subgroup $H \subseteq G$ acts compactly on M and for every $T \in G \setminus H$ the action of T on M' is weakly mixing. However, some difficulties arise here. First, it is not obvious that the elements of G acting compactly on a subspace M' of M form a group. (It is so, see Theorem 9.10.) Second, the space $M^c(T)$ of vectors on which an element $T \in G$ acts compactly is not, generally speaking, invariant with respect to G.

0.8. Nevertheless, the following theorem shows that the real situation is quite close to the ideal one described above (see Theorem 9.4):

Theorem. *Given a finitely generated nilpotent group G of unitary operators acting on a Hilbert space M, one can find a G-invariant subspace $M' \subseteq M$ on which the action of G is primitive in the following sense: there exists a (not necessarily G-invariant) closed subspace $L \subseteq M'$ such that the subgroup G_L of the elements of G preserving L contains a normal subgroup H so that*
1) H acts compactly on L,
2) every $T \in G_L \setminus H$ is weakly mixing on L,
3) the elements of the orbit of L under the action of G (they are in a natural one-to-one correspondence with the set of left cosets of G_L in G) are pairwise orthogonal and span M'.

The action of G on the orbit of L may have cycles. By passing to a normal subgroup G^* of finite index in G we can kill them; then the action on L of every element of $G^* \setminus G_L$ has pure Lebesgue spectrum (see Lemma 9.6).

0.9. Example. Let us give an example of "a primitive action" of a nilpotent group on a Hilbert space. It is a simple example, but Theorem 0.8 above says that it is quite typical.

Let L be a Hilbert space and let P be a unitary transformation of L with pure continuous spectrum. Define M as the Hilbert space generated by a family of pairwise orthogonal copies of L, indexed by $\mathbb{Z}$: $M = \overline{\bigoplus_{k \in \mathbb{Z}} L_k}$, $L_k \perp L_m$ for all $k \neq m \in \mathbb{Z}$, $\varphi_k \colon L_k \xrightarrow{\sim} L$, $k \in \mathbb{Z}$. Lift P to M by $P\big|_{L_k} = \varphi_k^{-1} \circ P \circ \varphi_k$, $k \in \mathbb{Z}$. Let S be "the coordinate shift" in M: $S\big|_{L_k} = \varphi_{k+1}^{-1} \circ \varphi_k$, $k \in \mathbb{Z}$, and let T be defined by $T\big|_{L_k} = P^k\big|_{L_k}$. Then the commutator $[T, S] = T^{-1}S^{-1}TS = P$ and T, S commute with P, so the group G generated by T, S is nilpotent of class 2. The action of G on M is primitive: the subgroup G_{L_0} of the elements of G preserving subspace L_0 is $\langle T, P\rangle$ (the subgroup generated by T and P),
1) its normal subgroup $H = \langle T\rangle$ acts identically on L_0,
2) every element of $\langle T, P\rangle$ outside of H coincides on L_0 with some P^n, $n \neq 0$, and hence, is weakly mixing on L_0,
3) and finally, for every element $S^kT^nP^m$ outside of $\langle T, P\rangle$ (that is, with $k \neq 0$) we have $S^kT^nP^m(L_0) = L_k \perp L_0$.

0.10. A similar picture can be observed if a nilpotent group G acts on a measure space X

as a group of measure preserving transformations (see Theorem 11.11):

Theorem. *Let G be a finitely generated nilpotent group. Given a measure preserving system $(X, \mathfrak{B}, \mu, G)$ with $\mu(X) < \infty$, one can find a nontrivial G-invariant factor X' of X on which the action of G is primitive in the following sense: X' is the product of a countable set $\mathcal{H}$ of independent factors, and G transitively acts on $\mathcal{H}$ as a group of permutations. For every $Z \in \mathcal{H}$, the subgroup G_Z of the elements of G preserving Z contains a normal subgroup H which acts compactly on Z, and every $T \in G_Z \setminus H$ is weakly mixing on Z.*

0.11. Example. The "measure-preserving" version of Example 0.9 is as follows. Let $(Z, \mathfrak{D}, \nu)$ be a measure space of finite measure and let P be a weakly mixing measure preserving transformation of Z. Let measure space $(X, \mathfrak{B}, \mu)$ be the direct product of a family of copies of Z, indexed by $\mathbb{Z}$: $X = \prod_{k \in \mathbb{Z}} Z_k$, $\varphi_k \colon Z_k \xrightarrow{\sim} Z$, $k \in \mathbb{Z}$. We will denote the k-th coordinate of $x \in X$ by x_k.

Lift P to X by $(xP)_k = \varphi_k^{-1}\big(\varphi_k(x_k)P\big)$, $k \in \mathbb{Z}$, and define $S, T \colon X \longrightarrow X$ by $(xS)_k = \varphi_k^{-1} \circ \varphi_{k-1}(x_{k-1})$ and $(xT)_k = x_k P^k$, $k \in \mathbb{Z}$. Then $[T, S] = P$ and T and S commute with P, so the group $G = \langle T, S \rangle$ is nilpotent of class 2. The action of G is primitive on X. Indeed, S shifts factors Z_k of X: $SZ_k = Z_{k+1}$, $k \in \mathbb{Z}$, and for every $k \in \mathbb{Z}$ the subgroup $G_{Z_k} = \langle T, P \rangle$ preserves Z_k. Its normal subgroup $H_k = \langle TP^{-k} \rangle$ is trivial on Z_k, so every element of $G_{Z_k} \setminus H_k$ coincides on Z_k with some P^n, $n \neq 0$, and thus is weakly mixing on Z_k.

0.12. *An extension* $\alpha \colon (X, \mathfrak{B}, \mu, G) \longrightarrow (Y, \mathfrak{D}, \nu, G)$ is a measurable mapping $\alpha \colon X \longrightarrow Y$ satisfying $\nu(B) = \mu(\alpha^{-1}(B))$ for all $B \in \mathfrak{D}$ and commuting with a measure preserving action of G on X and Y. Following the scheme of the proof of Theorem M, we deal with actions of G which are *relatively weakly mixing* and *relatively compact* with respect to Y (see [F2]). In fact, the notion of a primitive action is introduced and Theorem 0.10 is proved just in this context: Theorem 11.11 says that for any extension $\alpha \colon X \longrightarrow Y$ there exists an intermediate G-invariant factor X' of X such that X' is an extension of Y with a primitive action of G on it; we say that *X' is a primitive extension of Y*. As a corollary, every measure preserving system $(X, \mathfrak{B}, \mu, G)$ can be represented as a (transfinite) tower of primitive extensions.

0.13. Given an extension $\alpha \colon (X, \mathfrak{B}, \mu, G) \longrightarrow (Y, \mathfrak{D}, \nu, G)$, and under some natural mild assumptions of *regularity* on $(X, \mathfrak{B}, \mu)$, a family of measures μ_y, $y \in Y$, on X is defined (the measure μ_y is localized on the fiber $\alpha^{-1}(y)$, see [F2]). As a result, a family of inner products $\langle\,,\rangle_y$, $y \in Y$, measurable and compatible with multiplication by functions from $L^\infty(Y)$, is defined on the Hilbert space $L^2(X)$: $\langle f, g \rangle_y = \int_X f \bar{g} d\mu_y$. We define *a Y-pre-Hilbert space* as a module over $L^\infty(Y)$ equipped with such a system of inner products, and *a Y-Hilbert space* as a Y-pre-Hilbert space satisfying the natural conditions of nondegeneracy and completness (see Section 6). In particular, in the situation described above, any closed subspace of $L^2(X)$ invariant with respect to multiplication by functions from $L^\infty(Y)$ is a Y-Hilbert space. One of the important tools used in the proof of Theorem 0.5 is the "relativized" version of Theorem 0.8, which deals with the abstract Y-Hilbert spaces (Theorem 9.4). The reason why we do not confine ourselves to working with closed subspaces of $L^2(X)$ is that we want to minimize the assumptions under which our statements hold true, and that it is more covenient to deal with inner products than with integrals. We do not develop the theory of Y-Hilbert spaces, but prove (or, at least, formulate) everything we use in the sequel.

The idea of considering "generalized" Hilbert spaces is not new (see, for example, [I]). In fact, a Y-Hilbert space can be considered as the set of square integrable sections of *a Hilbert bundle* (a Hilbert bundle is a family of Hilbert spaces measurable parameterized by points of a measure space, see [R]). We have preferred the term "Y-Hilbert space" and a different approach to these spaces for the following reasons. First, we wished to emphasize the closeness of this construction to the construction of the conventional Hilbert space: a Y-Hilbert space is a "relative" Hilbert space, a Hilbert space for which the set of scalars is the set of measurable functions on a fixed measure space Y, and many constructions and facts known for conventional Hilbert spaces are transfered to Y-Hilbert spaces. In addition, it seems to be more natural for us to consider an element of a Y-Hilbert space (imagining it as a function from $L^2(X)$ for some extensioon X of Y) as an entire object, not as a set of vectors of distinct Hilbert spaces forming a Hilbert bundle over Y. For more details see [Z1], [Z2], [F1], [FK2].

0.14. We want to mention two additional difficulties arising in the proof of Theorem NM in comparison with Theorem M in [FK1]. The first one is that for a noncommutative group G, G-sequences of the form $g(n) = T^n$ with $T \in G$ do not form a group (with respect to the element-wise multiplication). If G is nilpotent, we are forced to deal with sequences of the form $g(n) = T_t^{p_t(n)} \ldots T_1^{p_1(n)}$, where $T_1, \ldots, T_t \in G$ and $p_1, \ldots, p_t$ are polynomials with rational coefficients taking on integer values on the integers; just these expressions arise in the formulations of Theorems NT and NM. We call such sequences *G-polynomials*. When G is a commutative group, ergodic theorems dealing with t sequences $T_1^n, \ldots, T_t^n$ in G are usually proved by induction on t. In the case where G is nilpotent, we have to involve G-polynomials and use a special induction process (so called *PET-induction*, see [B]), based on the fact that every G-polynomial is canceled if one applies to it the differential operator $Dg(n) = g(n)^{-1}g(n+1)$ several times.

0.15. The second difficulty relates to "the topological part" of the proof of Theorem NM. The proof of Theorem M uses Theorem T. To prove Theorem NM we need an abstract, stronger (Hales-Jewett's theorem type) version of Theorem NT. In [BL2] such a theorem is formulated and proved for the case of commutative G. We do not bring the most general theorem in this paper, confining ourselves to a rather technical result, formulated in the language of G-polynomials and necessary for our proof of Theorem NM (see Theorem 5.3). Theorem NT is a simple corollary of this result (see 5.5), and so, we demonstrate here another way of proving Theorem NT.

0.16. Multiple recurrence theorems, Theorems T and M, provide some important combinatorial facts as corollaries. Namely, Theorem T gives a multidimensional generalization of the known van der Waerden theorem about arithmetic progressions:

Theorem CT. ([FW]) *Let $d \in \mathbb{N}$ and let F be a finite set in $\mathbb{Z}^d$. For any finite coloring of $\mathbb{Z}^d$ there exist $n \in \mathbb{N}$ and $u \in \mathbb{Z}^d$ for which the set $u + n \cdot F$ is monochromatic.*

In other words, given a finite coloring of $\mathbb{Z}^d$, any finite subset of $\mathbb{Z}^d$ can be found in one color after being stretched and shifted in a suitable way. Theorem M implies a multidimensional generalization of the Szemerédi theorem ([Sz]), saying that such a subset can be found in any set of positive upper density in $\mathbb{Z}^d$:

Theorem CM. ([FK1]) *Let $d \in \mathbb{N}$, let F be a finite set in $\mathbb{Z}^d$, and let S be a subset of $\mathbb{Z}^d$ of positive upper Banach density (that is, for some sequence $\Pi = \{\Pi_k\}_{k \in \mathbb{N}}$ of parallelepipeds in $\mathbb{Z}^d$, $\Pi_k = \prod_{i=1}^d \{a_{k,i}, a_{k,i}+1, \ldots, b_{k,i}\}$ with $b_{k,i} - a_{k,i} \underset{k\to\infty}{\longrightarrow} \infty$, $i = 1, \ldots, d$, one has $\limsup_{k\to\infty} \frac{\#(S\cap\Pi_k)}{\#\Pi_k} > 0$). Then there exist $u \in \mathbb{Z}^d$ and $n \in \mathbb{N}$ such that $u + n \cdot F \subset S$.*

(Van der Waerden's and Szemerédi's theorems correspond to the case $d = 1$.)

0.17. A polynomial version of Theorem CM can be found in [BL1]. In this paper we establish analogous combinatorial facts, corresponding to Theorem NT and Theorem NM. Theorem NT gives the following nilpotent van der Waerden theorem (see Theorem 14.2):

Theorem NCT. *Let G be a nilpotent group, let $T_1, \ldots, T_t \in G$, and let $p_{i,j} : \mathbb{Z} \longrightarrow \mathbb{Z}$ be polynomials satisfying $p_{i,j}(0) = 0$, $i = 1, \ldots, I$, $j = 1, \ldots, t$. For any finite coloring of G there exist $n \in \mathbb{N}$ and $T \in G$ such that the set $\left\{T_t^{p_{i,t}(n)} \ldots T_1^{p_{i,1}(n)} T,\ i = 1, \ldots, l\right\}$ is monochromatic.*

(Theorem CT corresponds to $G = \mathbb{Z}^d$ and linear $p_{i,j}$.)

The nilpotent Szemerédi theorem, obtainable as a corollary of Theorem NM′, is

Theorem NCM. *Let G be a nilpotent group, let $T_1, \ldots, T_t \in G$, and let $p_{i,j} : \mathbb{Z} \longrightarrow \mathbb{Z}$ be polynomials satisfying $p_{i,j}(0) = 0$, $i = 1, \ldots, I$, $j = 1, \ldots, t$. and let S be a subset of G of positive upper density (that is with $\limsup_{k\to\infty} \frac{\#(S\cap\Phi_k)}{\#\Phi_k} > 0$ for some right Følner sequence $\{\Phi_k\}$ in G). There exist $n \in \mathbb{N}$ and $T \in G$ such that $T_t^{p_{i,t}(n)} \ldots T_1^{p_{i,1}(n)} T \in S$ for all $i = 1, \ldots, l$.*

(See Theorem 14.10.)

Taking as G the group of upper (or lower) triangular matrices with unit diagonal, one can also obtain some "pure combinatorial", number corollaries of the nilpotent van der Waerden and Szemerédi theorems (see Corollary 14.5 and Corollary 14.11).

0.18. Acknowledgments. This work is my doctoral thesis, carried out at the Department of Mathematics of Technion, Haifa, under supervision of Prof. Vitaly Bergelson; without his guidance and support this paper could not be written. I am very thankful also to Prof. H. Furstenberg for his attention to this work and to Prof. V. Ya. Lin for many useful contacts.

0.19. Sections 1 – 4 are preparatory. In particular, G-polynomials are introduced and studied in Section 3 and the PET-induction is described in Section 4. In Section 5 we formulate and prove an abstract and stronger version of Theorem NT, which is then used in the proof of Theorem NM. Y-Hilbert spaces and their transformations are defined and studied in Sections 6 and 7. Sections 8 – 11 are devoted to "the structure theory". The notions of a primitive action of a nilpotent group on a Y-Hilbert space and on an extension of a fixed measure space Y are introduced, and we prove that actions of a nilpotent group are reducible to primitive actions. Using this structure theory, we prove Theorem NM in Sections 12 and 13. In Section 14 we derive from Theorems NT and NM′ their combinatorial corollaries.

Notation:

$\mathbb{R}_+$	nonnegative real numbers
$\mathbb{Z}_+$	nonnegative integers
$\#A$	the cardinality of set A
$a_1 \dots \widehat{a_i} \dots a_k$	$a_1 \dots a_{i-1}a_{i+1} \dots a_k$ (the term a_i is left out)
$L(\Pi)$	the size of parallelepiped Π (subsection 1.1)
$d_*(S)$	the lower Banach density of $S \subseteq \mathbb{Z}^d$ (1.1)
$d^*(S)$	the upper Banach density of $S \subseteq \mathbb{Z}^d$ (1.1)
$d^*_\Phi(S)$	the upper density of set S with respect to Følner sequence $\Phi = \{\Phi_k\}$
ess-sup f	the essential supremum of function f: if f is defined on a measure space $(X, \mathfrak{B}, \mu)$, $\text{ess-sup}\, f = \inf\Big\{c : \mu\big(\{x \in X : f(x) > c\}\big) = 0\Big\}$
$\bigoplus_{\sigma\in\Sigma} L_\sigma$	the direct sum of a family of linear spaces
$L \bigoplus L'$	the sum of orthogonal subspaces L, L' of a Hilbert space
$\bigoplus_{\sigma\in\Sigma} L_\sigma$	the sum of a family of pairwise orthogonal linear subspaces of a Hilbert space
$L \ominus L'$	the orthogonal complement of a subspace L' of a Hilbert space L
Span(U)	the subspace spanned by subset U of a linear space
$\overline{U}$	the closure of subset U of a topological space
$\prod_{\sigma\in\Sigma} X_\sigma$	the product (or the relative product) of a family of factors of a measure space
$\coprod_{\sigma\in\Sigma} X_\sigma$	the product of a *relatively independent* family of factors (11.2)
$\mathcal{N}(u,v)$	$\int \lvert\langle u,v\rangle_y\rvert d\nu$ (6.10)
$\mathrm{Sel}((\mathbb{Z}^d)^p, \mathbb{Z}^d)$	the set of *selection* mappings from $(\mathbb{Z}^d)^p$ to $\mathbb{Z}^d$ (1.5)
$[T, P]$	the commutator of elements T, P of a group, $[T,P] = T^{-1}P^{-1}TP$
$N(H)$	the normalizer of subgroup H, $N(H) = \{T : THT^{-1} = H\}$
$\langle T_1, \dots T_k\rangle$	the group generated by $T_1, \dots, T_k$
$\text{D-}\lim\limits_n u(n)$	the limit of sequence $u(n)$ in density (1.3)
$u(n) \xrightarrow{D} u$	sequence $u(n)$ converges to u in density (1.3)
M^∞	the set of elements of Y-Hilbert space M with bounded norms
$M \otimes M'$	the tensor product of Y-Hilbert spaces M and M' (6.17)
$M^c(Q)$	the maximal subspace of Y-Hilbert space M on which set of transformations Q acts compactly (7.7)
$M^w(g)$	the maximal subspace of space M on which sequence of transformations $g(n)$, $n \in \mathbb{Z}^d$, is weakly mixing (7.4)
$\mathcal{X}$	the space $L^2(X)$ of square integrable functions on extension X of a measure space Y, considered as a Y-Hilbert space (11.1)
$\mathcal{X}^\infty$	the space of essentially bounded functions from $\mathcal{X}$ (11.3)
$X^c(Q)$	the factor of measure space X on which Q acts compactly (11.3)

$\mathcal{X}^c(Q)$	the space of functions from $\mathcal{X}$ on which Q acts compactly (11.3)
$D^m g$	the derivative of sequence g with step m, $D^m g(n) = g(n)^{-1} g(n+m)$ (3.13)
$\wp^d G$	the group of G-polynomials of d variables (3.1)
$\wp G$	$\wp^1 G$
$\wp_0^d G$	the subgroup of $\wp^d G$ consisting of G-polynomials vanishing at zero (3.1)
$\wp G(n,m)$	the group of G-polynomials of variables n, m (3.2)
$\wp_0 G(n,m)$	the subgroup of $\wp G(n,m)$ consisting of G-polynomials vanishing when $n = 0$ (3.2)
$w(g)$	*the weight* of G-polynomial g (3.8)
$w_{(n)}(g)$	the weight of g with respect to variable n (3.8)
$\mathcal{W}$	the set of weights (3.8)
$\omega(\mathcal{A})$	the weight of system $\mathcal{A}$ (4.3)
$\omega_{(n)}(\mathcal{A})$	the weight of $\mathcal{A}$ with respect to variable n (4.4)

1. Densities

1.1. *A parallelepiped* in $\mathbb{Z}^d$ is a set of the form

$$\Pi = \prod_{i=1}^{d} \{a_i, a_i+1, \ldots, b_i\},\ a_i \leq b_i \in \mathbb{Z},\ i = 1, \ldots, d.$$

The size $L(\Pi)$ of Π is the minimum of the length of its edges: $L(\Pi) = \min_{1 \leq i \leq d}(b_i - a_i + 1)$.

Given a set $S \subseteq \mathbb{Z}^d$, its *upper (Banach) density* is $d^*(S) = \limsup\limits_{L(\Pi)\to\infty} \dfrac{\#(S \cap \Pi)}{\#\Pi}$, its *lower (Banach) density* is $d_*(S) = \liminf\limits_{L(\Pi)\to\infty} \dfrac{\#(S \cap \Pi)}{\#\Pi}$, its *(Banach) density* is $\lim\limits_{L(\Pi)\to\infty} \dfrac{\#(S \cap \Pi)}{\#\Pi}$ (if exists).

1.2. A set $S \subseteq \mathbb{Z}^d$ is *syndetic* if it does not have arbitrarily large gaps: there exists $L \in \mathbb{N}$ such that every parallelepiped $\Pi \subset \mathbb{Z}^d$ with $L(\Pi) \geq L$ contains a point from S: $\Pi \cap S \neq \emptyset$. It is clear that $d_*(S) > 0$ if and only if S is syndetic.

A set $\Lambda \subseteq \mathbb{Z}^d$ is *thick* if it contains arbitrarily large parallelepipeds: for any $L \in \mathbb{N}$ there exists a parallelepiped $\Pi \subset \Lambda$ with $L(\Pi) = L$. It is clear that a set is syndetic if and only if it has a nonempty intersection with every thick set. It is also clear that a thick set remains thick after deleting any subset of zero density from it.

1.3. *A d-dimensional sequence $x(n)$, $n \in \mathbb{Z}^d$, in a set V* is a mapping $x : \mathbb{Z}^d \longrightarrow V$. We will often omit the word "d-dimensional".

Given a topological space V, a sequence $v(n)$, $n \in \mathbb{Z}^d$, in V *converges to $v \in V$ in density*, $\operatorname{D-lim}_n v(n) = v$ or $v(n) \xrightarrow{D} v$, if for every neighborhood U of v the set $\{n \in \mathbb{Z}^d : v(n) \notin U\}$ is of zero density (see [F2]).

Given a sequence $r(n)$, $n \in \mathbb{Z}^d$, in $\mathbb{R}$, *its upper limit in density* is

$$\operatorname{D-lim\,sup}_n r(n) = \sup\Bigl\{r \in \mathbb{R} : d^*\bigl(\{n \in \mathbb{Z}^d : r(n) > r\}\bigr) > 0\Bigr\}.$$

1.4. We will say that a statement holds true *for almost all points in* $\mathbb{Z}^d$ if it is true for all points in $\mathbb{Z}^d$ but a set of zeroes of a polynomial $\mathbb{Z}^d \longrightarrow \mathbb{Z}$. In particular, for $d = 1$ it means "for all but finitely many points".

Let us note that the set of zeroes of any nonzero polynomial defined on $\mathbb{Z}^d$ is of zero density. Thus, if a statement is true "for almost all $n \in \mathbb{Z}^d$", then it is also true "for all $n \in \mathbb{Z}^d$ but a set of zero density".

1.5. Given $d, p \in \mathbb{N}$, for any $P \subseteq \{1, \ldots, p\}$, $P \neq \emptyset$, we will call the linear surjective mapping $n\colon (\mathbb{Z}^d)^p \longrightarrow \mathbb{Z}^d$ defined by

$$n(m) = \sum_{i \in P} m_i,\ m = (m_1, \ldots, m_p),\ m_1, \ldots, m_p \in \mathbb{Z}^d,$$

a selection. The set of all selections $(\mathbb{Z}^d)^p \longrightarrow \mathbb{Z}^d$ will be denoted by $\mathrm{Sel}((\mathbb{Z}^d)^p, \mathbb{Z}^d)$, it is finite for any $d, p \in \mathbb{N}$.

The sum $(n_1 + n_2)\colon (\mathbb{Z}^d)^{p_1} \oplus (\mathbb{Z}^d)^{p_2} \longrightarrow \mathbb{Z}^d$ of selections $n_1 \in \mathrm{Sel}((\mathbb{Z}^d)^{p_1}, \mathbb{Z}^d)$, $n_2 \in \mathrm{Sel}((\mathbb{Z}^d)^{p_2}, \mathbb{Z}^d)$ is a selection as well: $n_1 + n_2 \in \mathrm{Sel}((\mathbb{Z}^d)^{p_1+p_2}, \mathbb{Z}^d)$.

1.6. Lemma. *Let $d, p \in \mathbb{N}$ and let $\Lambda \subseteq \mathbb{Z}^d$ be thick. Then*

$$\Lambda' = \bigcap_{n \in \mathrm{Sel}((\mathbb{Z}^d)^p, \mathbb{Z}^d)} n^{-1}(\Lambda) \subseteq \mathbb{Z}^{dp}$$

is thick.

Proof. First, note that for any thick set $\Lambda_0 \subseteq \mathbb{Z}^d$ there exist $m_1, m_2, \ldots \in \mathbb{Z}^d$ such that $\sum_{i \in P} m_i \in \Lambda_0$ for every finite nonempty $P \subset \mathbb{N}$. (That is, Λ_0 contains an IP-set in the terminology of [F2].) Indeed, take $m_1 \in \Lambda_0$, then find a parallelepiped $\Pi_1 \subset \Lambda_0$ with $L(\Pi_1) > 2|m_1|$ and take m_2 to be the center of Π_1 (or a point of $\mathbb{Z}^d$ nearest to the center). We have $m_2, m_1 + m_2 \in \Lambda_0$. Then find a parallelepiped $\Pi_2 \subset \Lambda_0$ with $L(\Pi_2) > 2(|m_1| + |m_2|)$, and choose m_3 to be the center of Π_2, and so on.

Let $L \in \mathbb{N}$ be given. Choose a thick set $\Lambda_0 \subseteq \Lambda$ satisfying

$$\Lambda_0 + p \cdot \{0, \ldots, L-1\}^d \subseteq \Lambda.$$

Find $m_1, \ldots, m_p \in \mathbb{Z}^d$ such that $\sum_{i \in P} m_i \in \Lambda_0$ for every $P \subseteq \{1, \ldots, p\}$, $P \neq \emptyset$, and put $m = (m_1, \ldots, m_p) \in (\mathbb{Z}^d)^p$. Then $n\Bigl(m + \bigl(\{0, \ldots, L-1\}^d\bigr)^p\Bigr) \subseteq \Lambda$ for any $n \in \mathrm{Sel}((\mathbb{Z}^d)^p, \mathbb{Z}^d)$, that is the parallelepiped $\Pi = m + \bigl(\{0, \ldots, L-1\}^d\bigr)^p$, whose size is L, is contained in Λ'. ■

2. Bases in a nilpotent group

We collect here, mostly without proofs, some simple assertions concerning nilpotent groups. For more details see, for example, [KM].

2.1. Let G be a group. For $Q_1, \dots, Q_k \subseteq G$, $\langle Q_1, \dots, Q_k \rangle$ denotes the subgroup of G generated by $\bigcup_{i=1}^k Q_i$. Given $T, P \in G$, *the commutator of* T, P is $[T, P] = T^{-1}P^{-1}TP$, the commutator of two subsets $Q_1, Q_2 \subseteq G$ is the subgroup $\big\langle \big\{ [T, P] : T \in Q_1,\ P \in Q_2 \big\} \big\rangle$.

2.2. A group G is called *nilpotent* if it has *a finite central series*, that is a finite sequence of normal subgroups $\{\mathbf{1}_G\} = G_0 \subseteq \ldots \subseteq G_t = G$ with $[G_i, G] \subseteq G_{i-1}$, $i = 1, \dots, t$.

It is easy to see that any finitely generated nilpotent group is a factor of a finitely generated torsion-free nilpotent group. Thus, every representation of a nilpotent group can be lifted to a representation of a torsion-free nilpotent group. Using this fact, we will deal in our considerations with torsion-free nilpotent groups only.

From now on G is a finitely generated torsion-free nilpotent group.

2.3. The first fact which we need is that G has a central series whose factors are infinite cyclic groups:

$$\{\mathbf{1}_G\} = G_0 \subseteq \ldots \subseteq G_t = G, \quad G_{i+1}/G_i \simeq \mathbb{Z},\ i = 1, \dots, t.$$

2.4. *An ordered basis* $(T_1, \dots T_t)$ *of* G is an ordered subset of G satisfying the following conditions:
1. every $P \in G$ can be **uniquely** represented in the form $P = T_t^{a_t} \ldots T_1^{a_1}$, $a_1, \dots, a_t \in \mathbb{Z}$,
2. for every $1 \le i < j \le t$ the commutator $[T_i, T_j] \in \langle T_1, \dots, T_{i-1} \rangle$.

Any nilpotent group G always has an ordered basis: one can take T_i, $i = 1, \dots, t$, such that for a central series $\{\mathbf{1}_G\} = G_0 \subseteq \ldots \subseteq G_t = G$ with infinite cyclic factors, $G_i = \langle T_i, G_{i-1} \rangle$.

2.5. Let $(T_1, \dots, T_t)$ be an ordered basis. Put $G_0 = \{\mathbf{1}_G\}$, $G_i = \langle T_1, \dots, T_i \rangle$, $i = 1, \dots, t$. Then $G_0 \subset \ldots \subset G_t = G$ are normal subgroups of G with $G_i/G_{i-1} \simeq \mathbb{Z}$, $i = 1, \dots, t$, $(T_1, \dots, T_i)$ is an ordered basis of G_i, $1 \le i \le t$, and $(T_{i+1}G_i, \dots, T_jG_i)$ is an ordered basis of G_j/G_i, $1 \le i < j \le t$.

For some $P_i \in G_{i-1}$ put $T_i' = T_iP_i$, $i = 1, \dots, t$. It is clear that $(T_1', \dots, T_t')$ is an ordered basis of G as well. In particular, since for all $R \in G$ one has $R^{-1}T_iR = T_i[T_i, R]$ with $[T_i, R] \in G_{i-1}$, $i = 1, \dots, t$, it follows that $(R_1^{-1}T_1R_1, \dots, R_t^{-1}T_tR_t)$ is an ordered basis of G for any $R_1, \dots, R_t \in G$.

2.6. Given an ordered basis $(T_1, \dots, T_t)$ of G, we have a coordinate mapping $a\colon G \longrightarrow \mathbb{Z}^t$ defined by

$$a(P) = \big(a_1(P), \dots, a_t(P)\big) \text{ for } P = T_t^{a_t(P)} \ldots T_1^{a_1(P)}.$$

This mapping is *polynomial* in the following sense:

Lemma. (See, for example, [KM].) *There exist polynomials* $f_i\colon \mathbb{Z}^{2(t-i)} \longrightarrow \mathbb{Z}$, $i = 1, \dots, t$, *and* $f'_i\colon \mathbb{Z}^{t-i+1} \longrightarrow \mathbb{Z}$, $i = 1, \ldots, t$, *such that*

$$a_i\bigl(P_1P_2\bigr) = a_i(P_1) + a_i(P_2) + f_i\Bigl(a_{i+1}(P_1), \ldots, a_t(P_1), a_{i+1}(P_2), \ldots, a_t(P_2)\Bigr),$$
$$a\bigl(P^n\bigr) = na_i(P) + f'_i\Bigl(a_{i+1}(P), \ldots, a_t(P), n\Bigr),$$

for every $P, P_1, P_2 \in G$, $n \in \mathbb{Z}$.

2.7. $(T_1, \ldots, T_t)$ *is a basis of* G if for some permutation σ of $\{1, \ldots, t\}$, $(T_{\sigma(1)}, \ldots, T_{\sigma(t)})$ is an ordered basis of G.

It follows from 2.5 that, given a basis $(T_1, \ldots, T_t)$ of G and elements $R_1, \ldots, R_t \in G$, $(R_1^{-1}T_1R_1, \ldots, R_t^{-1}T_tR_t)$ is a basis of G as well.

2.8. An important fact is that the multiplication in G remains polynomial with respect to any (not necessarily ordered) basis in G:

Proposition. *Let* $(T_1, \ldots, T_t)$ *be a basis of* G. *Then any* $P \in G$ *can be uniquely represented in the form* $P = T_t^{a_t} \ldots T_1^{a_1}$, $a_1, \ldots, a_t \in \mathbb{Z}$.

The coordinate mapping $a\colon G \longrightarrow \mathbb{Z}^t$ *defined by*

$$a(P) = \bigl(a_1(P), \ldots, a_t(P)\bigr) \quad \textit{for} \;\; P = T_t^{a_t(P)} \ldots T_1^{a_1(P)}$$

is polynomial: there exist polynomial mappings $F\colon \mathbb{Z}^{2t} \longrightarrow \mathbb{Z}^t$, $F'\colon \mathbb{Z}^{t+1} \longrightarrow \mathbb{Z}^t$ *such that*

$$a\bigl(P_1P_2\bigr) = F\bigl(a(P_1), a(P_2)\bigr), \quad a\bigl(P^n\bigr) = F'\bigl(a(P), n\bigr),$$

for every $P, P_1, P_2 \in G$, $n \in \mathbb{Z}$.

Proof. We already have this statement for every ordered basis by Lemma 2.6. Our task is to show that the statement holds true if we permute elements of an ordered basis. Let σ be the permutation of $\{1, \ldots, t\}$ for which $(T_{\sigma(1)}, \ldots, T_{\sigma(t)})$ is an ordered basis.

First of all, let us prove existence and uniqueness of the representation $P = T_t^{a_t} \ldots T_1^{a_1}$ for every $P \in G$. Using induction on t, we may assume that every $\tilde{P} \in G/\langle T_{\sigma(1)}\rangle$ can be uniquely represented in the form $\tilde{P} = T_t^{a_t} \ldots \widehat{T_{\sigma(1)}} \ldots T_1^{a_1}$ modulo $T_{\sigma(1)}$, that is every $P \in G$ can be uniquely represented in the form $P = T_t^{a_t} \ldots \widehat{T_{\sigma(1)}} \ldots T_1^{a_1} T_{\sigma(1)}^{a_{\sigma(1)}}$. Since $T_{\sigma(1)}$ commutes with each $T_1, \ldots, T_t$, we have $P = T_t^{a_t} \ldots T_1^{a_1}$.

We have to prove now that for $T_t^{c_t} \ldots T_1^{c_1} = T_t^{a_t} \ldots T_1^{a_1} T_t^{b_t} \ldots T_1^{b_1}$ and $T_t^{d_t} \ldots T_1^{d_1} = \bigl(T_t^{a_t} \ldots T_1^{a_1}\bigr)^n$, $c_1, \ldots, c_t$ are polynomials of $a_1, \ldots, a_t, b_1, \ldots, b_t$, and $d_1, \ldots, d_t$ are polynomials of $a_1, \ldots, a_t, n$. Of course, one could do this directly, utilizing the commutator calculus. We will use Lemma 2.6 instead.

Applying Lemma 2.6 several times, we reduce the problem to the following: for $P = T_t^{b_t} \ldots T_1^{b_1} = T_{\sigma(t)}^{a_t} \ldots T_{\sigma(1)}^{a_1}$, $b_1, \ldots, b_t$ are polynomials of $a_1, \ldots, a_t$. Using an induction on $k = t, \ldots, 1$, we may assume that in the representation $P = T_{i_t}^{b_{i_t}} \ldots T_{i_{k+1}}^{b_{i_{k+1}}} T_{\sigma(k)}^{c_k} \ldots T_{\sigma(1)}^{c_1}$ with $\{i_{k+1}, \ldots, i_t\} = \{\sigma(k+1), \ldots, \sigma(t)\}$ and $i_t > \ldots > i_{k+1}$, $b_{i_{k+1}}, \ldots, b_{i_t}, c_1, \ldots, c_k$ are polynomials of $a_1, \ldots, a_t$. We have now to move $T_{\sigma(k)}^{c_k}$ to its proper place.

Namely, let $i_l > \sigma(k) > i_{l-1}$. It is enough to prove that in the representation

$$P = T_{i_t}^{b_{i_t}} \dots T_{i_l}^{b_{i_l}} T_{\sigma(k)}^{c_k} T_{i_{l-1}}^{b_{i_{l-1}}} \dots T_{i_{k+1}}^{b_{i_{k+1}}} T_{\sigma(k-1)}^{d_{k-1}} \dots T_{\sigma(1)}^{d_1}$$

$d_1, \dots, d_{k-1}$ are polynomials of $b_{i_{k+1}}, \dots, b_{i_t}, c_1, \dots, c_k$. But

$$T_{\sigma(k-1)}^{d_{k-1}} \dots T_{\sigma(1)}^{d_1} = \left[T_{i_{l-1}}^{b_{i_{l-1}}} \dots T_{i_{k+1}}^{b_{i_{k+1}}}, T_{\sigma(k)}^{c_k}\right] T_{\sigma(k-1)}^{c_{k-1}} \dots T_{\sigma(1)}^{c_1} \in \langle T_{\sigma(1)}, \dots, T_{\sigma(k-1)} \rangle,$$

and the required fact follows from Lemma 2.6, applied a number of times. ■

2.9. We will say that an ordered subset $(S_1, \dots, S_s)$ of G is *a basis of G over H* if there exists a basis $(R_1, \dots, R_r)$ of H such that $(R_1, \dots, R_r, S_1, \dots, S_s)$ is a basis of G. When this is the case, by Proposition 2.8 every $P \in G$ can be **uniquely** represented in the form $P = S_1^{a_1} \dots S_s^{a_s} R$ with $R \in H$ and $a_i \in \mathbb{Z}$, $i = 1, \dots, s$.

We will say that H *is complete (in G)* if a basis of G over H exists. A one-to-one mapping from the set of left cosets of H in G onto $\mathbb{Z}^s$ is naturally defined in this case. The following lemma is trivial.

Lemma. *If H' is complete in H and H is complete in G, then H' is complete in G. If H is complete in G, then for every $T \in G$, THT^{-1} is complete in G. If H is a normal subgroup of G, then H is complete in G if and only if G/H is torsion-free.*

2.10. Proposition. *Let H be a subgroup of G. Then there exists a normal subgroup G^* of finite index in G such that $H \cap G^*$ is complete in G^*.*

Proof. Choose an ordered basis $(T_1, \dots, T_t)$ in G. We will look for a sequence $d_1, \dots, d_t \in \mathbb{N}$ such that, for every $k = 1, \dots, t$:
a) The group $G_k^* = \langle T_1^{d_1}, \dots, T_k^{d_k} \rangle$ is normal in G, and $(T_1^{d_1}, \dots, T_k^{d_k})$ is a basis of G_k^*.
b) Either there exists $P_{k-1} \in G_{k-1}^*$ such that $T_k^{d_k} P_{k-1} \in H$, or $T_k^{d_k n} P_{k-1} \notin H$ for all $n \in \mathbb{Z}$, $n \neq 0$, and all $P_{k-1} \in G_{k-1}^*$.

We will do it by induction on k; put $G_0^* = \{\mathbf{1}_G\}$. Assume that for some $0 \leq k < t$ the numbers $d_1, \dots, d_k$ have already been chosen.

Lemma. *There exists $c \in \mathbb{N}$ such that $\left[G, T_{k+1}^c\right] \subseteq G_k^*$.*

Proof. It is enough to find $c \in \mathbb{N}$ such that $[T, T_{k+1}^c] \subseteq G_k^*$ for every $T = T_1, \dots, T_t$.

Represent $[T, T_{k+1}^c]$ in the coordinate form:

$$[T, T_{k+1}^c] = T_k^{p_k(c)} \dots T_1^{p_1(c)}.$$

Since $[T, T_{k+1}^c] = \mathbf{1}_G$, the polynomials $p_1, \dots, p_k$ have no constant terms: $p_1(0) = \dots = p_k(0) = 0$, and their coefficients are rational numbers. Hence, if $c \in \mathbb{N}$ is divisible by d_i and by all denominators of the coefficients of p_i, $i = 1, \dots, k$, then d_i divides $p_i(c)$, $i = 1, \dots, k$. ■

Now we have two possibilities:

a) There exist $n \in \mathbb{N}$, $P_k \in G_k^*$ such that $T_{k+1}^{cn} P_k \in H$. We put $d_{k+1} = cn$ in this case.

b) $T_{k+1}^{cn} P_k \notin H$ for all $n \in \mathbb{N}$ and $P_k \in G_k^*$. Since $[T_{k+1}^c, G_k^*] \subseteq G_k^*$, it also implies $T_{k+1}^{-cn} P_k \notin H$ for all $n \in \mathbb{N}$ and $P_k \in G_k^*$. We put $d_{k+1} = c$.

In both cases the group $G_{k+1}^* = \langle T_1^{d_1}, \ldots, T_{k+1}^{d_{k+1}} \rangle$ is normal in G and $(T_1^{d_1}, \ldots, T_{k+1}^{d_{k+1}})$ is a basis of G_{k+1}^*. So, the step of the induction process has been done.

Now, put $G^* = G_t^* = \langle T_1^{d_1}, \ldots, T_t^{d_t} \rangle$. G^* is a normal subgroup of index $d_1 \ldots d_t$ in G. Denote

$$I = \Big\{ 1 \leq k \leq t : \text{there exists } P_{k-1} \in G_{k-1}^* \text{ such that } R_k = T_k^{d_k} P_{k-1} \in H \Big\}.$$

Then $\big(T_i^{d_i},\ i \notin I\big)$ is a basis of G^* over $H \cap G^*$. (And $\big(R_k = T_k^{d_k} P_{k-1},\ k \in I\big)$ is a basis of $H \cap G^*$.) ■

2.11. In conclusion, we introduce one more piece of notation. Given a subgroup $H \subseteq G$, $T \in G$ *normalizes* H if $THT^{-1} = H$. Elements normalizing H form a subgroup of G called *the normalizer of* H. We will denote it by $N(H)$:

$$N(H) = \Big\{ T \in G : THT^{-1} = H \Big\}.$$

H is normal in $N(H)$, and $N(H)$ is the maximal subgroup of G with this property.

The orbit of H under the left conjugation action of G is in a natural one-to-one correspondence with the set of left cosets of $N(H)$ in G: $THT^{-1} \leftrightarrow TN(H)$, $T \in G$.

3. G-polynomials

Sequences of the form $\{T^n\}_{n \in \mathbb{Z}^d}$, $T \in G$, do not form a group if G is a noncommutative group. The element-wise products of such "power" sequences are examples of what we will call *G-polynomials*. An operation of differentiation is defined on G-polynomials; a specific property of a nilpotent group G is that this differentiation cancels any G-polynomial after being applied finitely many times. So, an induction process on the "degree" (or *the weight*, see subsection 3.8) of a G-polynomial can be used.

We have a standard problem with notation, namely: given a mapping g, what is $g(n)$? Is it the mapping g itself or the value of g at a point n? We will more or less follow the rule: if n is not specified beforehand or immediately afterwards, then $g(n)$ is a mapping whose argument is n.

3.1. For $d \in \mathbb{N}$, *an integral polynomial of d variables* is a polynomial mapping $\mathbb{Z}^d \longrightarrow \mathbb{Z}$ (it may have rational coefficients). We denote the ring of integral polynomials of d variables by $\wp^d$.

The group of G-polynomials of d variables $\wp^d G$ is the minimal subgroup of the group $G^{\mathbb{Z}^d}$ of d-dimensional sequences $\mathbb{Z}^d \longrightarrow G$ which contains the constant sequences and is closed with respect to raising to integral polynomial powers: if $g, h \in \wp^d G$ and $p \in \wp^d$, then $gh, g^p \in \wp^d G$, where $gh(n) = g(n)h(n)$ and $g^p(n) = g(n)^{p(n)}$, $n \in \mathbb{Z}^d$. We will also denote $\wp = \wp^1$, $\wp G = \wp^1 G$. The elements of $\wp^d G$ are called *G-polynomials*. G itself is a subgroup of $\wp^d G$ and is represented by *constant G-polynomials*.

G-polynomials vanishing at zero form a subgroup of $\wp^d G$; we will denote this subgroup by $\wp_0^d G$:

$$\wp_0^d G = \Big\{ g \in \wp^d G : g(0) = \mathbf{1}_G \Big\}.$$

3.2. In order to specify arguments of G-polynomials we will sometimes use the notation $\wp(n_1, \ldots, n_k)$ and $\wp G(n_1, \ldots, n_k)$, $n_1 \in \mathbb{Z}^{d_1}$, ..., $n_k \in \mathbb{Z}^{d_k}$, for $\wp^{d_1+\ldots+d_k}$ and $\wp G^{d_1+\ldots+d_k}$ respectively. A polynomial $p \in \wp(n,m)$ and a G-polynomial $g \in \wp G(n,m)$, $n \in \mathbb{Z}^{d_1}$, $m \in \mathbb{Z}^{d_2}$, can be considered respectively as a polynomial and a G-polynomial of argument n with coefficients from $\wp(m)$. Fixing m we obtain a polynomial and a G-polynomial of n: for any $m \in \mathbb{Z}^{d_2}$ we have $p(n,m) \in \wp(n)$ and $g(n,m) \in \wp G(n)$.

By $\wp_0 G(n_1, \ldots, n_k)$ we will denote the subgroup of $\wp G(n_1, \ldots, n_k)$ consisting of G-polynomials vanishing at zero **at the first argument**:

$$\wp_0 G(n_1, \ldots, n_k) = \Big\{ g \in \wp G(n_1, \ldots, n_k) : g(0, n_2, \ldots, n_k) = \mathbf{1}_G \Big\}.$$

3.3. Let $(T_1, \ldots, T_t)$ be a basis of G.

Lemma. *Every $g \in \wp^d G$ can be uniquely represented in the form*

$$g = T_t^{p_t} \ldots T_1^{p_1} \tag{3.1}$$

with $p_i \in \wp^d$, $i = 1, \ldots, t$.

In particular, $g \in \wp_0^d G$ if and only if $p_1(0) = \ldots = p_t(0) = 0$.

Proof. Since G-polynomials are defined inductively, it is enough to check that g^{-1}, gh and g^p, $p \in \wp^d$, can be represented in the form (3.1) if $g, h \in \wp^d G$ can. But this is a corollary of the polynomiality of the multiplication in G (Proposition 2.8). Uniqueness of representation (3.1) follows from uniqueness of the decomposition of an element of G with respect to the basis. ∎

3.4. Corollary. *Let $g \in \wp G(n,m)$, $h \in \wp G(n)$, $n \in \mathbb{Z}^{d_1}$, $m \in \mathbb{Z}^{d_2}$. Then either $g(n,m) = h(n)$ for every $m \in \mathbb{Z}^{d_2}$, or $g(n,m) \neq h(n)$ for almost all $m \in \mathbb{Z}^{d_2}$.*

In particular, for $g \in \wp^{d_1} G$, $T \in G$, either $g \equiv T$ or $g(n) \neq T$ for almost all $n \in \mathbb{Z}^{d_1}$.

Proof. The polynomials $p_1, \ldots, p_t$ in the representation (3.1) are polynomials of n whose coefficients are polynomials of m. And each of these coefficients either is constant, or for every $a \in \mathbb{Z}$ it is not equal to a for almost all $m \in \mathbb{Z}^{d_2}$. ∎

3.5. Given an ordered subset $E = (S_1, \ldots, S_s)$ of G, we denote by $\wp^d E$ the set of G-polynomials of the form $S_s^{p_s} \ldots S_1^{p_1}$, $p_i \in \wp^d$, $i = 1, \ldots, s$.

Let H be a complete subgroup of G and let $E = (S_1, \ldots, S_s)$ be a basis of G over H. Then any $g \in \wp^d G$ can be uniquely represented in the form $g = g'h$ with $g' \in \wp^d E$, and h being an H-polynomial: $h \in \wp^d H$.

3.6. Proposition. *Let H be a complete subgroup of G. If $g \in \wp G(n,m)$, $n \in \mathbb{Z}^{d_1}$, $m \in \mathbb{Z}^{d_2}$, and $g \notin \wp H(n,m)$, then $g(n,m) \notin \wp H(n)$ for almost all $m \in \mathbb{Z}^{d_2}$.*

In particular, if $g \in \wp^d G$ and $g \notin \wp^d H$, then $g(n) \notin H$ for almost all $n \in \mathbb{Z}^d$.

Proof. Let E be a basis of G over H. Represent g in the form $g = g'h$ with $g' \in \wp E(n,m)$, $h \in \wp H(n,m)$. Since $g \notin \wp H(n,m)$, g' is nontrivial. By Corollary 3.4, g' is not equal to $\mathbf{1}_G$ for almost all $n \in \mathbb{Z}^d$. ■

3.7. The following simple technical corollary will be used below.

Corollary. *Let H be a complete subgroup of G and let $g_1, g_2 \in \wp_0^d G$.*
a) If $T_1, T_2 \in G$, $T_1^{-1}T_2 \notin H$, then $g_1(n)T_1H \neq g_2(n)T_2H$ for almost all $n \in \mathbb{Z}^d$.
b) If $T \in G$ and $g^{-1}g_2 \notin \wp(THT^{-1})$, then $g_1(n)TH \neq g_2(n)TH$ for almost all $n \in \mathbb{Z}^d$.

Proof.
a) The G-polynomial $g(n) = T_1^{-1}g_1(n)^{-1}g_2(n)T_2$ does not belong to $\wp^d H$ as $g(0) = T_1^{-1}T_2 \notin H$.
b) THT^{-1} is a complete subgroup of G as well, so $g_1^{-1}(n)g_2(n)(THT^{-1}) \neq THT^{-1}$ for almost all $n \in \mathbb{Z}^d$. ■

3.8. We fix from now on an ordered basis $(T_1, \dots, T_t)$ of G. We also denote $G_0 = \{\mathbf{1}_G\}$, $G_i = \langle T_1, \dots, T_i \rangle$, $i = 1, \dots, t$.

For $g \in \wp^d G$, $g \neq \mathbf{1}_G$, let $1 \leq k \leq t$ be such that $g \in \wp^d G_k$, $g \notin \wp^d G_{k-1}$. Then g can be uniquely represented in the form $g = T_k^p g'$, $p \in \wp^d$, $p \neq 0$, $g' \in \wp^d G_{k-1}$. We say that T_k is *the senior generator of g. The weight of g*, $w(g)$, is the pair $(k, \deg p)$; we put $w(\mathbf{1}_G) = (0,0) = 0$. If g is considered as a G-polynomial of several variables, $g \in \wp G(n,m)$, we also define *the weight of g with respect to n*, $w_{(n)}(g)$, as the pair $(k, \deg_{(n)} p)$ where $\deg_{(n)} p$ is the degree of the polynomial $p(n,m)$ with respect to the variable n.

Define an ordering on the set $\mathcal{W}$ of all weights, that is on the set of pairs (k,a), $0 \leq k \leq t$, $a \in \mathbb{Z}_+$, lexicographically: (k,a) is greater than (l,b) if either $k > l$, or $k = l$ and $a > b$. The set $\mathcal{W}$ becomes well ordered under this ordering.

3.9. Lemma. *Let $g \in \wp^d G$ and $w(g) = (k,a)$.*
If $r \in \wp^d$, then $w(g^r) = (k, a + \deg r)$.
If $g, h \in \wp^d G$ with $w(h) < w(g)$, then $w(gh) = w(hg) = w(g)$.
If $g \in \wp_0^d G$, then for any $n \in \mathbb{Z}^d$, $w(g(n)) < w(g)$.

Proof. Pass to the factor group $G' = G/G_{k-1}$. In G' we have $g = T_k^p$ with $p \neq 0$, $\deg p = a$, $h = T_k^q$ with $\deg q < a$. Hence $g^r = T_k^{pr}$, $gh = T_k^{p+q}$ in G', and the first two statements follow.

If additionally $g(0) = \mathbf{1}_G$, then $p(0) = 0$, and since p is nontrivial we have $a = \deg p > 0$. At the same time, for any $n \in \mathbb{Z}^d$, $p(n)$ is a constant and so $w(g(n)) \leq (k,0)$. ■

3.10. Let $g_1, g_2 \in \wp^d G$. Represent g_1 and g_2 in the form $g_j = T_{k_j}^{p_j} g_j'$ with $p_j \neq 0$, $g_j' \in \wp^d G_{k_j - 1}$, $j = 1, 2$. We will say that g_1 *is equivalent to* g_2 and write $g_1 \sim g_2$ if $k_1 = k_2$ and the leading terms of p_1 and p_2 coincide. We obtain an equivalence relation on $\wp^d G$, and $g_1 \sim g_2$ implies $w(g_1) = w(g_2)$. Thus, we may define *the weight of an equivalence class* in $\wp^d G$ as the weight of any of its members.

3.11. Lemma. *Let $g, h \in \wp^d G$.*
a) For any $m \in \mathbb{Z}^d$, $g(n+m) \sim g(n)$.
b) $gh \sim hg$.
c) $hgh^{-1} \sim g$.
d) If $w(h) < w(g)$, then $gh \sim g$.
e) If $w(h) = w(g)$ and $h \not\sim g$, then $w(gh) = w(g)$.
f) If $g' \in \wp^d G$ with $g' \sim g$ and $g \not\sim h$, then $h^{-1}g' \sim h^{-1}g$.
g) If $h \sim g$ and $h \neq \mathbf{1}_G$, then $w(h^{-1}g) < w(g)$.

Proof. Let $w(g) = (k, a)$. First of all, c) is a corollary of b), and we may assume $w(h) \leq w(g)$ in b). Thus, we may confine ourselves to $g, g', h \in \wp^d G_k$.

Passing to the factor group $G' = G_k/G_{k-1}$, we reduce the problem to that for polynomials: here the weight of a G-polynomial is defined by the degree of the corresponding polynomial, and two G-polynomials are equivalent if the leading terms of the corresponding polynomials coincide. In this context the statement of the lemma becomes clear. ∎

3.12. Given $g_1, g_2 \in \wp G(n, m)$, we say that g_1 *is equivalent to g_2 with respect to n* if $w_{(n)}(g_1^{-1}g_2) < w_{(n)}(g_1)$. All the considerations in subsection 3.10 are carried onto this "with respect to n" case. In particular, the weight $w_{(n)}$ of the class of G-polynomials equivalent with respect to n is defined.

3.13. *The differentiation with step $m \in \mathbb{Z}^d$* is the mapping $D^m : G^{\mathbb{Z}^d} \longrightarrow G^{\mathbb{Z}^d}$ acting by the rule $D^m(g)(n) = g(n)^{-1}g(n+m)$.

It is clear that the operation of differentiation preserves the group of G-polynomials: for $g \in \wp^d G$, $m \in \mathbb{Z}^d$, we have $D^m g \in \wp^d G$. A key property of G-polynomials is that they vanish after a finite number of differentiations. This follows from the following lemma.

Lemma. *For $g \in \wp^d G$, $g \neq \mathbf{1}_G$, and for any $m \in \mathbb{Z}^d$,*

$$w(D^m g) < w(g).$$

Proof. Indeed, by Lemma 3.11, $g(n+m) \sim g(n)$ and so, $w\Big(g(n)^{-1}g(n+m)\Big) < w\big(g(n)\big)$. ∎

4. Systems of G-polynomials and PET-induction

4.1. *A system* is a finite subset of $\wp^d G$.

4.2. Recall that as described in subsection 3.8, the set of weights $\mathcal{W}$ is well ordered. Denote by Ω the set of functions $\mathcal{W} \longrightarrow \mathbb{Z}_+$ having finite support. Ω is well ordered too by the rule

for $\omega_1, \omega_2 \in \Omega$, $\omega_1 \succ \omega_2$ if there exists $w \in \mathcal{W}$ such that
$\omega_1(w) > \omega_2(w)$ and $\omega_1(w') = \omega_2(w')$ for all $w' \in \mathcal{W}$ with $w' > w$.

To make this ordering clearer, let us write $\omega \in \Omega$ in the form of the list

$$\Big(\omega(w_1)w_1, \ldots, \omega(w_p)w_p\Big),$$

where $w_1 > \ldots > w_p$ and $\omega(w) = 0$ for all $w \notin \{w_1, \ldots, w_p\}$. Then, for $a_q \neq b_q$,

$$\Big(a_1 w_1, \ldots, a_{q-1} w_{q-1}, a_q w_q, \ldots, a_p w_p\Big) \succ \Big(a_1 w_1, \ldots, a_{q-1} w_{q-1}, b_q w_q, \ldots, b_p w_p\Big)$$

if and only if $a_q > b_q$.

4.3. For every system $\mathcal{A} \subset \wp^d G$ we define its *weight* $\omega(\mathcal{A}) \in \Omega$:

$$\omega(\mathcal{A})(w) = \begin{cases} \text{the number of equivalence classes of weight } w \text{ having a nonempty} \\ \text{intersection with } \mathcal{A}. \end{cases}$$

We will say that a system $\mathcal{A}' \subset \wp^d G$ *precedes* $\mathcal{A}$ if $\omega(\mathcal{A}) \succ \omega(\mathcal{A}')$.

Example. The weight of the system $\big\{S_1^{10n+10}, S_1^{5n^2+8n}, S_1^{7n^2+7n}, S_1^{7n^2+4}, S_2^{n^2-2} S_1^{4n^4-5}, S_2^{n^2+2n} S_1^{n^3+11}, S_2^{2n^3-2n^2} S_1^{9n^9-5n^2}, S_2^{n^3+4n^2+n} S_1^{5n^5+5}, S_2^{n^3+2n^2}, S_2^{n^3+n^2+n} S_1^{n^{11}-n}\big\}$ is $\Big(2(2,3), 1(2,2), 2(1,2), 1(1,1)\Big)$.

4.4. In the same way, *the weight of* $\mathcal{A} \subset \wp^d(n,m)$ *with respect to* n is defined by

$$\omega_{(n)}(\mathcal{A})(w) = \begin{cases} \text{the number of classes of elements of } \mathcal{A} \text{ equivalent with respect to} \\ n \text{ whose weight with respect to } n \text{ is } w. \end{cases}$$

We will say that a system $\mathcal{A}' \in \wp G(n,m)$ *precedes* $\mathcal{A}$ *with respect to* n if $\omega_{(n)}(\mathcal{A}') \prec \omega_{(n)}(\mathcal{A})$.

4.5. *The PET-induction* is induction on the well ordered set Ω. That is, if a statement is true for the system $\{\mathbf{1}_G\}$ and if one can deduce that the statement holds for a system $\mathcal{A}$ from the assumption that it is true for all systems preceding $\mathcal{A}$, then we can assert that it is true for all systems.

The following lemma is the main tool used in the PET-induction.

Lemma. *Let $\mathcal{A} \subset \wp^d G$ be a system.*
1) If $\mathcal{A}' \subset \wp^d G$ is a system consisting of G-polynomials of the form $g' = h^{-1} g h$ for $g \in \mathcal{A}$ and $h \in \wp^d G$, then $\omega(\mathcal{A}') \preceq \omega(\mathcal{A})$.
2) If $\mathcal{A}'$ is a system consisting of G-polynomials of the form $g' = hg$ and $g' = gh$ for $g \in \mathcal{A}$ and $h \in \wp^d G$ with $w(h) < w(g)$, then $\omega(\mathcal{A}') \preceq \omega(\mathcal{A})$. In particular, if $\mathcal{A} \subset \wp_0^d G$ and $\mathcal{A}'$ consists of G-polynomials of the form $g'(n) = g(n_0)^{-1} g(n+n_0)$ and $g'(n) = g(n+n_0) g(n_0)^{-1}$ with $g \in \mathcal{A}$ and $n_0 \in \mathbb{Z}^d$, then $\mathcal{A}' \subset \wp_0^d G$ and $\omega(\mathcal{A}') \preceq \omega(\mathcal{A})$.
3) Let $h \in \mathcal{A}$, $h \neq \mathbf{1}_G$, be a G-polynomial of weight minimal in $\mathcal{A}$: $w(h) \leq w(g)$ for all $g \in \mathcal{A}$. If $\mathcal{A}'$ is a system consisting of G-polynomials of the form $g' = h^{-1} g$ and $g' = g h^{-1}$ for $g \in \mathcal{A}$, then $\omega(\mathcal{A}') \prec \omega(\mathcal{A})$.

Proof. All this is a corollary of Lemma 3.9 and Lemma 3.11. In both 1) and 2), every element of $\mathcal{A}'$ is equivalent to some element of $\mathcal{A}$ and hence, the number of equivalence classes of each weight can not be greater in $\mathcal{A}'$ than in $\mathcal{A}$.

In 3), the equivalence classes in $\mathcal{A}$ change when we pass to $\mathcal{A}'$, but the equivalence of elements is preserved and their weights remain the same. The only exception is the equivalence class containing h; it is replaced by equivalence classes having smaller weights.

$\blacksquare$

4.6. An analogous fact holds if we deal with "weights with respect to n". The following lemma can be proved in the same way as Lemma 4.5.

Lemma. *Let $\mathcal{A} \subset \wp G(n,m)$ be a system.*
1) If $\mathcal{A}' \subset \wp G(n,m)$ is a system consisting of G-polynomials of the form $g' = h^{-1}gh$ for $g \in \mathcal{A}$ and $h \in \wp G(n,m)$, then $\omega_{(n)}(\mathcal{A}') \preceq \omega_{(n)}(\mathcal{A})$.
2) If $\mathcal{A}'$ is a system consisting of G-polynomials of the form $g' = hg$ and $g' = gh$ for $g \in \mathcal{A}$ and $h \in \wp^d G$ with $w_{(n)}(h) < w_{(n)}(g)$, then $\omega_{(n)}(\mathcal{A}') \preceq \omega_{(n)}(\mathcal{A})$. In particular, if $\mathcal{A} \subset \wp_0 G(n,m)$ and $\mathcal{A}'$ consists of G-polynomials of the form $g'(n) = g(n_0(m))^{-1}g(n+n_0(m))$ and $g'(n) = g(n+n_0(m))g(n_0(m))^{-1}$ with $g \in \mathcal{A}$ and $n_0(m)$ being a polynomial mapping, then $\mathcal{A}' \subset \wp_0 G(n,m)$ and $\omega_{(n)}(\mathcal{A}') \preceq \omega_{(n)}(\mathcal{A})$.
3) Let $h \in \mathcal{A}$, $h \neq \mathbf{1}_G$, be a G-polynomial whose weight with respect to n is minimal in $\mathcal{A}$: $w_{(n)}(h) \leq w_{(n)}(g)$ for all $g \in \mathcal{A}$. If $\mathcal{A}'$ is a system consisting of G-polynomials of the form $g' = h^{-1}g$ and $g' = gh^{-1}$ with $g \in \mathcal{A}$, then $\omega_{(n)}(\mathcal{A}') \prec \omega_{(n)}(\mathcal{A})$.

5. An abstract topological recurrence theorem

In the course of the proof of Theorem NT in [L] no properties of natural numbers as arguments of G-polynomials were used, except the fact that one can add them together. It follows that Theorem NT can be generalized and formulated for G-polynomials whose arguments are elements of a commutative semigroup. We will not do this here because of the too extensive preparatory work required (see [BL2] for the case of commutative G). However, in order to prove Theorem NM, we need a stronger statement than Theorem NT. This section is devoted to obtaining such an intermediate theorem.

5.1. First of all, we describe the environment in which we will formulate and prove Theorem 5.3 below. As before, G denotes a finitely generated torsion-free nilpotent group. An ordered basis of G is assumed to be fixed and so, the weights of G-polynomials and systems are assumed to be defined with respect to this basis.

We fix $d \in \mathbb{N}$ and a group $Z \simeq \mathbb{Z}^d$. We fix letter "n" as a formal variable taking values in Z: "a mapping $F(n)$" will mean "a mapping F from Z". In particular, groups of G-polynomials of argument n, $\wp G(n) \simeq \wp^d G$ and $\wp_0 G(n) \simeq \wp_0^d G$ are defined.

We also fix an infinite alphabet M. Denote by $\mathcal{F}(M)$ the set of all finite nonempty subsets of M. For every $a \in M$ let $Z_a \simeq \mathbb{Z}^d$ be a group. (To be more formal, we assume that Z, Z_a ($a \in M$) and all their products are pairwise disjoint sets.) For every $m \in \mathcal{F}(M)$ define $Z_m = \bigoplus_{a \in m} Z_a$, $Z_m \simeq (\mathbb{Z}^d)^{\# m}$. As a formal variable in Z_m we will use the same symbol m: "a mapping $F(m)$" will mean "a mapping F from Z_m". In particular, for $m \in \mathcal{F}(M)$ the group of G-polynomials $\wp G(m)$ is defined, for pairwise disjoint $m_1, \ldots, m_k \in \mathcal{F}(M)$ the group $\wp G(m_1, \ldots, m_k)$ is defined.

5.2. For $m_1, m_2 \in \mathcal{F}(M)$ with $m_1 \subseteq m_2$ there is a natural embedding $\wp G(m_1) \subseteq \wp G(m_2)$. Thus the limit group $\wp G(M) = \bigcup_{m \in \mathcal{F}(M)} \wp G(m)$ is defined.

Let $N, m \in \mathcal{F}(M)$, $N \cap m = \emptyset$, let $g \in \wp G(n, N)$. Then for any polynomial mapping $n(m)\colon Z_m \longrightarrow Z$, in particular, for any selection $n(m) \in \mathrm{Sel}(Z_m, Z)$, one has $g(n(m), N) \in \wp G(m, N)$.

5.3. Let (X, ρ) be a compact metric space and let the group $\wp G(M)$ act on X.

Theorem. *For any system $\mathcal{A} \subset \wp_0 G(n, N)$, $N \in \mathcal{F}(M)$, and any $\varepsilon > 0$ there exist $m \in \mathcal{F}(M)$, $m \cap N = \emptyset$, and a system $\mathcal{B} \subset \wp_0 G(m, N)$ such that for any $x \in X$ there exist $h \in \mathcal{B}$ and a selection $n(m)\colon Z_m \longrightarrow Z$ for which*

$$\rho\bigl(g(n)hx, hx\bigr) < \varepsilon \textit{ for every } g \in \mathcal{A}.$$

(To simplify the notation we write $g(n)$ instead of $g(n, N)$ or $g(n(m), N)$.)

5.4. Proof. Let $\mathcal{A} \in \wp_0 G(n, N)$ and $\varepsilon > 0$ be given. We will use the PET-induction on the weight of $\mathcal{A}$ with respect to n. The statement is trivial for the system $\{\mathbf{1}_G\}$, and so, we have the beginning of the induction process.

5.4.1. Let $k \in \mathbb{N}$ be such that there are two points at a distance less than $\varepsilon/2$ among any $k + 1$ points of X. Let $g_0 \in \mathcal{A}$ be a G-polynomial whose weight with respect to n is minimal in $\mathcal{A}$. We may assume that $\mathcal{A}$ does not contain constant G-polynomials and, thus, $g_0 \neq \mathbf{1}_G$.

5.4.2. Let

$$\mathcal{A}_0 = \Bigl\{g(n)g_0(n)^{-1},\ g \in \mathcal{A}\Bigr\}, \quad \varepsilon_0 = \frac{\varepsilon}{2k}.$$

By Lemma 4.6, $\mathcal{A}_0$ precedes $\mathcal{A}$ with respect to n. Thus there exist $m_0 \in \mathcal{F}(M)$ with $m_0 \cap N = \emptyset$ and a system $\mathcal{B}_0 \subset \wp_0 G(m_0, N)$ such that for any $x \in X$ there exist $h_0 \in \mathcal{B}_0$ and $n_0 \in \mathrm{Sel}(Z_{m_0}, Z)$ such that

$$\rho\bigl(f(n_0)h_0x, h_0x\bigr) < \varepsilon_0 \text{ for every } f \in \mathcal{A}_0.$$

Denote

$$\begin{aligned}\mathcal{A}_1 = \Bigl\{\bigl(g_0(n_0)^{-1}h_0\bigr)^{-1}g(n_0)^{-1}g(n+n_0)g_0(n)^{-1}\bigl(g_0(n_0)^{-1}h_0\bigr),\\ g \in \mathcal{A},\ h_0 \in \mathcal{B}_0,\ n_0 \in \mathrm{Sel}(Z_{m_0}, Z)\Bigr\} \subset \wp G_0(n, m_0, N),\end{aligned}$$

and choose $0 < \varepsilon_1 < \varepsilon/2k$ such that the inequality $\rho(y_1, y_2) < \varepsilon_1$ implies $\rho(ey_1, ey_2) < \dfrac{\varepsilon}{2k}$ for all

$$e = g(n_0)g_0(n_0)^{-1}h_0 \in \wp G(m_0, N),\ g \in \mathcal{A},\ h_0 \in \mathcal{B}_0,\ n_0 \in \mathrm{Sel}(Z_{m_0}, Z).$$

By Lemma 4.6, $\mathcal{A}_1$ precedes $\mathcal{A}$ with respect to n. Thus there exist $m_1 \in \mathcal{F}(M)$ with $m_1 \cap (m_0 \cup N) = \emptyset$ and a system $\mathcal{B}_1 \subset \wp_0 G(m_1, m_0, N)$ such that for any $x \in X$ there exist $h_1 \in \mathcal{B}_1$ and $n_1 \in \mathrm{Sel}(Z_{m_1}, Z)$ such that

$$\rho\big(f(n_1)h_1 x, h_1 x\big) < \varepsilon_1 \text{ for all } f \in \mathcal{A}_1.$$

Continue this process: assume that numbers $\varepsilon_0, \dots, \varepsilon_{j-1}$, sets $m_0, \dots, m_{j-1} \in \mathcal{F}(M)$ and systems $\mathcal{B}_l \in \wp G(m_l \cup \dots \cup m_0 \cup N)$, $l = 1, \dots, j-1$, have already been chosen.

Denote

$$\begin{aligned}\mathcal{A}_j = \Big\{ &\big(g_0(n_{j-1})^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_{j-1}\big)^{-1} \\ &\qquad g(n_{j-1} + \dots + n_i)^{-1} g(n + n_{j-1} + \dots + n_i) g_0(n)^{-1} \\ &\qquad\qquad \big(g_0(n_{j-1})^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_{j-1}\big), \\ &\text{for } \ i = 0, \dots, j-1,\ g \in \mathcal{A},\ h_l \in \mathcal{B}_l,\ n_l \in \mathrm{Sel}(Z_{m_l}, Z),\ l = 0, \dots, j-1 \Big\} \\ &\qquad \subset \wp G_0(n, m_{j-1}, \dots, m_0, N),\end{aligned}$$

and choose $0 < \varepsilon_j < \varepsilon/2k$ such that the inequality $\rho(y_1, y_2) < \varepsilon_j$ implies $\rho(ey_1, ey_2) < \dfrac{\varepsilon}{2k}$ for every

$$\begin{aligned} e = g\big(n_{j-1} + \dots + n_i\big)\Big(g_0(n_{j-1})^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_{j-1}\Big) &\in \wp G(m_{j-1}, \dots, m_0, N), \\ i = 0, \dots, j-1,\ g \in \mathcal{A},\ h_l \in \mathcal{B}_l,\ n_l \in \mathrm{Sel}(Z_{m_l}, Z),&\ l = 0, \dots, j-1.\end{aligned}$$

By Lemma 4.6, $\mathcal{A}_j$ precedes $\mathcal{A}$ with respect to n. Thus there exist $m_j \in \mathcal{F}(M)$ with $m_j \cap (m_{j-1} \cup \dots \cup m_0 \cup N) = \emptyset$, and a system $\mathcal{B}_j \subset \wp_0 G(m_j, \dots, m_0, N)$ such that for any $x \in X$ there exist $h_j \in \mathcal{B}_j$ and $n_j \in \mathrm{Sel}(Z_{m_j}, Z)$ such that

$$\rho\big(f(n_j)h_j x, h_j x\big) < \varepsilon_j \text{ for all } f \in \mathcal{A}_j.$$

We continue this process up to $j = k$. Then we put

$$\begin{gathered} m = m_k \cup \dots \cup m_0 \in \mathcal{F}(M), \\ \mathcal{B} = \Big\{ g_0(n_j)^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_k,\ h_i \in \mathcal{B}_i,\ n_i \in \mathrm{Sel}(Z_{m_i}, Z),\ i = 0, \dots, k,\ j = 1, \dots, k \Big\} \\ \subset \wp_0 G(m, N). \end{gathered}$$

5.4.3. Let $x \in X$ be given.

Put $y_k = x$. There exist $h_k \in \mathcal{B}_k$ and $n_k \in \mathrm{Sel}(Z_{m_k}, Z)$ such that

$$\rho\big(f(n_k)h_k y_k, h_k y_k\big) < \varepsilon_k \text{ for all } f \in \mathcal{A}_k.$$

Put $y_{k-1} = h_k y_k$. There exist $h_{k-1} \in \mathcal{B}_{k-1}$ and $n_{k-1} \in \mathrm{Sel}(Z_{m_{k-1}}, Z)$ such that

$$\rho\big(f(n_{k-1})h_{k-1} y_{k-1}, h_{k-1} y_{k-1}\big) < \varepsilon_{k-1} \text{ for all } f \in \mathcal{A}_{k-1}.$$

Continue the process of choosing y_j, h_j and m_j up to $j = 0$. We have in particular $y_0 = h_0 \dots h_k x$.

5.4.4. Put
$$x_j = g_0(n_j)^{-1} \dots g_0(n_0)^{-1} y_0,\ j = 0, \dots, i.$$

Lemma. *For any* $0 \le i < j \le k$
$$\rho\big(g(n_j + \dots n_{i+1})x_j, x_i\big) < \frac{\varepsilon}{2k}(j-i),\ g \in \mathcal{A}. \tag{5.1}$$

Proof. Indeed,
$$\begin{aligned}&\rho\Big(h_{j-1}^{-1} \dots h_0^{-1} g_0(n_0) \dots g_0(n_{j-1})\\ &g\big(n_{j-1} + \dots + n_{i+1}\big)^{-1} g\big(n_j + \dots + n_{i+1}\big) g_0(n_j)^{-1} g_0(n_{j-1})^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_{j-1} h_j y_j,\\ &\qquad\qquad h_j y_j\Big) < \varepsilon_j.\end{aligned}$$

By the choice of ε_j,
$$\begin{aligned}&\rho\Big(\big(g(n_{j-1} + \dots + n_{i+1}) g_0(n_{j-1})^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_{j-1}\big) h_{j-1}^{-1} \dots h_0^{-1} g_0(n_0) \dots g_0(n_{j-1})\\ &\quad g(n_{j-1} + \dots + n_{i+1})^{-1} g(n_j + \dots + n_{i+1}) g_0(n_j)^{-1} g_0(n_{j-1})^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_j y_j,\\ &\qquad \big(g(n_{j-1} + \dots + n_{i+1}) g_0(n_{j-1})^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_{j-1}\big) h_j y_j\Big) < \frac{\varepsilon}{2k}.\end{aligned}$$

We have consequently
$$\begin{aligned}\rho\Big(&g(n_j + \dots + n_{i+1}) g_0(n_j)^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_k x,\\ &g(n_{j-1} + \dots + n_{i+1}) g_0(n_{j-1})^{-1} \dots g_0(n_0)^{-1} h_0 \dots h_k x\Big) < \frac{\varepsilon}{2k},\end{aligned}$$

that is
$$\rho\Big(g(n_j + \dots + n_{i+1})x_j, g\big(n_{j-1} + \dots + n_{i+1}\big)x_{j-1}\Big) < \frac{\varepsilon}{2k}.$$

(In particular,
$$\rho\big(g(n_{i+1})x_j, x_{j-1}\big) < \frac{\varepsilon}{2k}.)$$

Clearly, this implies (5.1). ■

5.4.5. By the choice of k, there are $0 \le i < j \le k$ for which $\rho(x_j, x_i) < \frac{\varepsilon}{2}$. Coupled with (5.1) this implies
$$\rho\big(g(n_j + \dots n_{i+1})x_j, x_j\big) < \varepsilon,$$

that is
$$\rho\big(g(n)hx, hx\big) < \varepsilon,$$

where
$$\begin{aligned}n(m) &= n_j(m_j) + \dots n_{i+1}(m_{i+1}) \in \mathrm{Sel}(Z_m, Z),\\ h(m) &= g_0(n_j(m_j))^{-1} \dots g_0(n_0(m_0))^{-1} h_0(m_0) \dots h_k(m_k) \in \mathcal{B}.\end{aligned}$$

■

5.5. Proof of Theorem NT. Now let us show how Theorem 5.3 implies Theorem NT. The statement of Theorem NT concerns the part of a nilpotent group G generated by the finite set $\{T_1, \dots, T_t\}$ only. Since any finitely generated nilpotent group is a factor of a finitely generated torsion-free nilpotent group, we may assume as before that G is a finitely generated torsion-free nilpotent group.

Define a mapping $\varphi: \wp G(M) \longrightarrow G$ by

$$\varphi\Big(S_t^{p_t(m_1,\dots,m_k)} \dots S_1^{p_1(m_1,\dots,m_k)}\Big) = S_t^{p_t(1,\dots,1)} \dots S_1^{p_1(1,\dots,1)},$$

where $S_1, \dots, S_t \in G$ and $p_1, \dots, p_t$ are integral polynomials of variables $m_1, \dots, m_k \in M$. That is, φ is defined by putting all variables from M to 1. We obtain an action of $\wp G(M)$ on X by $gx = \varphi(g)x$. Define G-polynomials $g_i(n) = T_1^{p_{i,1}(n)} \dots T_t^{p_{i,t}(n)}$, $i = 1, \dots, I$, and a system $\mathcal{A} = \{g_1, \dots, g_I\} \subset \wp_0 G$. Given $\varepsilon > 0$, find $m \in \mathcal{F}(M)$ and a system $\mathcal{B} \subset \wp_0 G(m)$ satisfying the conclusion of Theorem 5.3. Fix any $x \in X$ and find $h \in \mathcal{B}$ and a selection $n(m): Z_m \longrightarrow Z$ such that $\rho\big(g_i(n)hx, hx\big) < \varepsilon$, $i = 1, \dots, I$. Then the point $x' = hx \in X$ and the number $n' = n(1, \dots, 1) \in \mathbb{N}$ satisfy the conclusion of Theorem NT. ■

5.6. Now we will derive a "coloring" corollary from Theorem 5.3; this corollary (Theorem 5.9 below) will be used in the proof of Theorem NM.

Given a set H and $r \in \mathbb{N}$, *an r-coloring of H* is a mapping from H into an r-element set. We fix in this section such a set Q, $\#Q = r$.

The set of all r-colorings of H, Q^H, is compact in the product topology. If H is countable, $H = \{h_i\}_{i \in \mathbb{N}}$, then Q^H can be metrized by

$$\rho(\chi_1, \chi_2) = \Big(\min\{i \in \mathbb{N} : \chi_1(h_i) \neq \chi_2(h_i)\}\Big)^{-1}, \ \chi_1, \chi_2 \in Q^H.$$

If H is a group, then the (right) action of H on itself induces a continuous (left) action of H on Q^H:

$$(h\chi)(h') = \chi(h'h), \ h, h' \in H, \ \chi \in Q^H.$$

5.7. Applying Theorem 5.3 to the compact metric space of r-colorings of $\wp G(M)$ (under the notation of Section 5), we obtain the following statement.

Corollary. *Let $r \in \mathbb{N}$. For any system $\mathcal{A} \subset \wp_0 G(n)$ and any $\varepsilon > 0$ there exist $m \in \mathcal{F}(M)$ and a system $\mathcal{B} \subset \wp_0 G(m)$ such that for any r-coloring χ of $\wp G(M)$ there exist $h \in \mathcal{B}$ and a selection $n(m): Z_m \longrightarrow Z$ for which χ is constant on the set $\{g(n)h : g \in \mathcal{A}\}$.*

Proof. Indeed, if $\rho\big(g(n)h\chi, h\chi\big)$ is small enough, then $\chi(g(n)h) = g(n)h\chi(\mathbf{1}_G) = h\chi(\mathbf{1}_G) = \chi(h)$, for all $g \in \mathcal{A}$. ■

5.8. Certainly, h in the formulation of Corollary 5.7 can be placed on the left of g as well:

Corollary. *Let $r \in \mathbb{N}$. For any system $\mathcal{A} \subset \wp_0 G(n)$ and any $\varepsilon > 0$ there exist $m \in \mathcal{F}(M)$ and a system $\mathcal{B} \subset \wp_0 G(m)$ such that for any r-coloring χ of $\wp G(M)$ there exist $h \in \mathcal{B}$ and a selection $n(m): Z_m \longrightarrow Z$ for which χ is constant on the set $\{hg(n) : g \in \mathcal{A}\}$.*

Proof. Put $\mathcal{A}' = \{g^{-1} : g \in \mathcal{A}\}$, find m and $\mathcal{B}'$ satisfying the statement of Corollary 5.7, applied to the system $\mathcal{A}'$, and put $\mathcal{B} = \{h'^{-1} : h' \in \mathcal{B}'\}$.

Let a coloring χ of $\wp G(M)$ be given. Define a coloring χ' of $\wp G(M)$ by $\chi'(f) = \chi(f^{-1})$, $f \in \wp G(M)$. By Corollary 5.7, there exist $h' \in \mathcal{B}'$ and a selection $n(m)\colon Z_m \longrightarrow Z$ such that χ is constant on the set $\{g'(n)h' : g' \in \mathcal{A}'\}$. Put $h = h'^{-1}$. Since $\chi(hg(n)) = \chi'(g(n)^{-1}h')$, χ is constant on $\{hg(n) : g \in \mathcal{A}\}$. ■

5.9. Now note that the formulation of Corollary 5.8 deals with colorings not of all $\wp G(M)$ but of $\wp G(m)$ only, and even of a finite subset of $\wp G(m)$. Consequently, we have:

Theorem. *Let $r \in \mathbb{N}$. For any system $\mathcal{A} \subset \wp_0 G(n)$ and any $\varepsilon > 0$ there exist $m \in \mathcal{F}(M)$ and a system $\mathcal{B} \subset \wp_0 G(m)$ such that for any r-coloring χ of the set*

$$\Big\{hg(n) : g \in \mathcal{A},\ h \in \mathcal{B},\ n \in \mathrm{Sel}(Z_m, Z)\Big\}$$

there exist a G-polynomial $h \in \mathcal{B}$ and a selection $n(m)\colon Z_m \longrightarrow Z$ for which χ is constant on the set $\{hg(n) : g \in \mathcal{A}\}$.

6. Y-Hilbert spaces

From now on, $(Y, \mathfrak{D}, \nu)$ will be a measure space with $\nu(Y) = 1$. In this section we introduce the notion of *a Y-Hilbert space*: it is, so to say, a relative Hilbert space, a Hilbert space over the ring of measurable functions on Y (cf. [I], *inner product modules*, and [R], [Z1], *Hilbert bundles*). We also consider the simplest notions and constructions related to Y-Hilbert spaces.

6.1. *A Y-pre-Hilbert space M* is a module over the ring $L^\infty(Y)$ equipped with a nonnegative inner product $\langle\,,\rangle\colon M^2 \longrightarrow L^1(Y)$: for $u, v \in M$ a function (more exactly, an equivalence class of functions) $\langle u, v\rangle \in L^1(Y)$ is defined and, for every $u, v, v' \in M$, $\varphi, \psi \in L^\infty(Y)$, one has
1) $\langle u + u', v\rangle = \langle u, v\rangle + \langle u', v\rangle$,
2) $\langle \varphi u, \psi v\rangle = \varphi\overline{\psi}\langle u, v\rangle$,
3) $\langle u, v\rangle = \overline{\langle v, u\rangle}$,
4) $\langle u, u\rangle \geq 0$.
The value of the function $\langle u, v\rangle$ at a point $y \in Y$ (defined modulo sets of zero measure in Y) will be denoted by $\langle u, v\rangle_y$. The norm $\|u\| \in L^2(Y)$ for $u \in M$ is defined as $\|u\| = \langle u, u\rangle^{\frac{1}{2}}$, its value at $y \in Y$ is $\|u\|_y = \langle u, u\rangle_y^{\frac{1}{2}}$.

6.2. Let M be a Y-pre-Hilbert space. Every set $A \in \mathfrak{D}$ defines a nonnegative inner product $\langle\,,\rangle_A$ on M by $\langle u, v\rangle_A = \int_A \langle u, v\rangle_y d\nu$. We put $\|u\|_A = \langle u, u\rangle_A^{\frac{1}{2}}$. In particular, $\langle u, v\rangle_Y = \int_Y \langle u, v\rangle_y d\nu$, $\|u\|_Y = \left(\int_Y \|u\|_y^2 d\nu\right)^{\frac{1}{2}} = \big\|\|u\|\big\|_{L^2(Y)}$.

6.3. The metric on a Y-pre-Hilbert space M is defined by the norm $\|\ \|_Y$. It is clear that multiplication by functions from $L^\infty(Y)$ is continuous in this metric.

A Y-pre-Hilbert space M is *a Y-Hilbert space* if $\langle\,,\rangle_Y$ defines on M the structure of a Hilbert space, that is if M is complete with respect to $\|\ \|_Y$, and $\|u\|_Y = 0$ implies $u = 0$.

6.4. Examples.

6.4.1. Let Y be a single-element set, $Y = \{y\}$, Then $L^1(Y) \simeq L^\infty(Y) \simeq \mathbb{C}$, and thus, any Y-Hilbert space is a conventional Hilbert space.

6.4.2. $M = L^2(Y)$, $\langle u, v\rangle = u\bar{v}$.

6.4.3. Let $(Z, \mathfrak{C}, \eta)$ be a measure space, $\eta(Z) < \infty$, let $(X, \mathfrak{B}, \mu)$ be the product $(Y, \mathfrak{D}, \nu) \times (Z, \mathfrak{C}, \eta)$. Take $M = L^2(X)$. For $u, v \in M$ choose their representatives $\tilde{u}, \tilde{v}$, and for $\varphi \in L^\infty(Y)$, $y \in Y$, $z \in Z$ put $\langle u, v\rangle_y = \int_Z \tilde{u}(y, z)\overline{\tilde{v}(y, z)} d\eta$ and $(\varphi u)(y, z) = \varphi(y)\tilde{u}(y, z)$. Since $u\bar{v} \in L^1(X)$, $\langle u, v\rangle_y$ is defined for almost all $y \in Y$ and $\langle u, v\rangle \in L^1(Y)$. Clearly, M is a Y-Hilbert space. In fact, it is easy to see that any Y-Hilbert space is a "disjoint union" of Y-Hilbert spaces of this special form.

6.5. Let M be a Y-pre-Hilbert space. We say that a sequence $u_1, u_2, \ldots \in M$ *converges to $u \in M$ pointwise* if the sequence of the functions $\|u_1 - u\|, \|u_2 - u\|, \ldots$ pointwise converges to 0. If a sequence $u_1, u_2, \ldots$ converges to $u \in M$, then some its subsequence $u_{k_1}, u_{k_2}, \ldots$ converges to u pointwise. Moreover, it is easy to see that every sequence fundamental in M contains a pointwise fundamental subsequence. It is also clear that if two pointwise fundamental sequences $u_1, u_2, \ldots$ and $v_1, v_2, \ldots$ are *equivalent*, that is if they define the same point in *the topological completion* of M, then the sequence $u_1, v_1, u_2, v_2, \ldots$ is also pointwise fundamental. This shows that the completion of M, whose elements are equivalence classes $[\{u_k\}]$ of fundamental sequences $\{u_k\}$ in M, inherits the structure of a Y-Hilbert space: to define $\langle u, v\rangle$ for $u = [\{u_k\}]$, $v = [\{v_k\}]$, pick up pointwise fundamental subsequences $\{u_{k_i}\}$ and $\{v_{k_i}\}$ and put $\langle u, v\rangle$ to be the pointwise limit of $\langle u_{k_i}, v_{k_i}\rangle$ while $i \to \infty$.

6.6. From now on, M will be a Y-Hilbert space.

If M is $L^2(Y)$ (see example 6.4.2), its elements, remaining in M, can be multiplied not only by functions from $L^\infty(Y)$, but also by suitable unbounded functions. This suggests the following definition:

Let $u, v \in M$, let φ be a measurable function on Y. We say that $v = \varphi u$ if for all $c \in \mathbb{R}$, $1_{A_c} v = (1_{A_c}\varphi)u$, where $A_c = \{y \in Y : \varphi(y) < c\}$.

Lemma. *Let $u \in M$ and a measurable function φ on Y be such that $\varphi\|u\| \in L^2(Y)$. Then $\varphi u \in M$ (that is, there is $v \in M$ such that $v = \varphi u$), and $\|\varphi u\| = |\varphi|\|u\|$.*

Proof. Let $A_k = \{y \in Y : \varphi(y) < k\}$, put $v_k = (1_{A_k}\varphi)u$, $k = 1, 2, \ldots$. Then for $l > k$, $v_l - v_k = (1_{A_l \setminus A_k}\varphi)u$, and so

$$\|v_l - v_k\|_Y = \int_{A_l \setminus A_k} |\varphi(y)|^2 \|u\|_y^2 d\nu \le \int_{Y \setminus A_k} |\varphi(y)|^2 \|u\|_y^2 d\nu \underset{k\to\infty}{\longrightarrow} 0.$$

Hence, the sequence $v_1, v_2, \ldots$ is fundamental in M, and its limit v satisfies the conclusion of the lemma. ■

6.7. Corollary. *Let $u \in M$, let $A = \{y \in Y : \|u\|_y \neq 0\}$. Then $u/\|u\| = (1_A\|u\|^{-1})u \in M$ is defined and $\big\|u/\|u\|\big\| = 1_A$.*

6.8. Corollary. *Let $u, v \in M$ and a measurable function φ on Y satisfy $v = \varphi u$. Then $1_A u = (1_A\varphi^{-1})v$, where $A = \{y \in Y : \varphi(y) \neq 0\}$.*

Proof. Indeed, $u' = (1_A\varphi^{-1})v$ exists by Lemma 6.6, and $\|\varphi(1_A u - u')\| = |\varphi|\|1_A v - 1_A v\| = 0$. It follows that $\|1_A u - u'\| = 0$ on A; since also $1_{Y\setminus A}1_A u = 1_{Y\setminus A}u' = 0$, we have $\|1_A u - u'\| = 0$, and so $1_A u = u'$. ■

6.9. A submodule $N \subseteq M$, that is a subset of M invariant with respect to multiplication by functions from $L^\infty(Y)$, will be called *a subspace of M*. The sum of two subspaces and the closure of a subspace of M are subspaces of M as well. A closed subspace of M is a Y-Hilbert space.

6.10. For $u, v \in M$ we denote by $\mathcal{N}(u, v)$ the L^1-norm of $\langle u, v\rangle$:

$$\mathcal{N}(u, v) = \big\|\langle u, v\rangle\big\|_{L^1(Y)} = \int |\langle u, v\rangle_y| d\nu.$$

The proof of the following proposition is immediate:

Proposition. *For any $u, u', v \in M$, $\varphi \in L^\infty(Y)$*
a) $\|u\|_Y^2 = \mathcal{N}(u, u)$.
b) $\mathcal{N}(u, v) \geq 0$,
c) $\mathcal{N}(u + u', v) \leq \mathcal{N}(u, v) + \mathcal{N}(u', v)$,
d) $\mathcal{N}(u, v) = \mathcal{N}(v, u)$,
e) $\mathcal{N}(u, v) \leq \|u\|_Y\|v\|_Y$,
f) $\mathcal{N}(\varphi u, v) \leq \text{ess-sup}(|\varphi|)\mathcal{N}(u, v)$.

In particular, $\mathcal{N}$ is continuous on $M \times M$.

6.11. Two vectors $u, v \in M$ are *Y-orthogonal*, $u \perp_Y v$, if $\langle u, v\rangle = 0$, that is if $\mathcal{N}(u, v) = 0$. The Y-orthogonal complement $N^{\perp_Y} = \Big\{v \in M : v \perp_Y u \text{ for any } u \in N\Big\}$ of a subset $N \subseteq M$ is a closed subspace of M.

Example. Let $A \in \mathfrak{D}$, let $N = 1_A \cdot M = \{1_A u : u \in M\}$. Then $N^{\perp_Y} = 1_{Y\setminus A} \cdot M$.

6.12. The following lemma shows that for subspaces of M the Y-orthogonality coincides with the conventional orthogonality; this allows us not to distinguish between these notions.

Lemma. *Let N be a subspace of M and let $u \in M$ be orthogonal to N: $\langle u, v\rangle_Y = 0$ for all $v \in N$. Then $u \perp_Y N$.*

Proof. Let $v \in N$. Denote

$$\varphi(y) = \begin{cases} \dfrac{\langle u, v\rangle_y}{|\langle u, v\rangle_y|}, & \text{if } \langle u, v\rangle_y \neq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Then $\varphi \in L^\infty(Y)$ and

$$\mathcal{N}(u,v) = \int |\langle u,v\rangle_y| d\nu = \int \overline{\varphi(y)} \langle u,v\rangle_y d\nu = \int \langle u, \varphi v\rangle_y d\nu = \langle u, \varphi v\rangle_Y = 0.$$

■

6.13. Corollary. *Let N be a closed subspace of M. Then $M = N \bigoplus N^{\perp_Y}$.*

6.14. We say that a system $\mathcal{S}$ of vectors in M is *orthonormal* if $\langle v, v'\rangle = 0$ for all $v, v' \in \mathcal{S}$, $v \neq v'$, and for every $v \in \mathcal{S}$, $\|v\| = 1_{A_v}$ for some $A_v \in \mathfrak{D}$ with $\nu(A_v) > 0$. By the Zorn lemma, there is a maximal (in the sense of inclusion) orthonormal system in M; we call every such system *an orthonormal basis of M*. If $\mathcal{B}$ is an orthonormal basis of M, then the set of finite linear combinations $L = \Big\{\varphi_1 v_1 + \ldots + \varphi_k v_k : \varphi_1, \ldots, \varphi_k \in L^\infty(Y),\ v_1, \ldots, v_k \in \mathcal{B}\Big\}$ is dense in M. Indeed, if u were a nonzero vector in $L^\perp$, we could add the vector $u/\|u\|$ (see Corollary 6.7) to $\mathcal{B}$. Moreover,

Lemma. *Let $\mathcal{B}$ be an orthonormal basis of M. Then every $u \in M$ is uniquely representable in the form $u = \sum_{k=1}^\infty \psi_k v_k$, where $v_1, v_2, \ldots \in \mathcal{B}$ are pairwise distinct and nonzero functions $\psi_1, \psi_2, \ldots \in L^2(Y)$ satisfy $\psi\big|_{Y \setminus A_k} = 0$, where $A_k = \{y \in Y : \|v_k\|_y \neq 0\}$.*

Proof. Uniqueness of such an expansion for u is evident from the equality $\langle u, v_m\rangle = \sum_{k=1}^\infty \psi_k \langle v_k, v_m\rangle = 1_{A_m} \psi_m$.

For $\varepsilon > 0$, let distinct $v_1, \ldots, v_m \in \mathcal{B}$ be such that $\big\|u - \sum_{k=1}^m \varphi_k v_k\big\|_Y < \varepsilon$ for some $\varphi_1, \ldots, \varphi_m \in L^\infty(Y)$. Let N be the closure of the subspace of M generated by $v_1, \ldots, v_m$. Put $\psi_k = \langle u, v_k\rangle$, $k = 1, \ldots, m$. Then $\|\psi_k\| \leq \|u\| \|v_k\| \leq \|u\|$, and so, $\psi_k \in L^2(Y)$, $k = 1, \ldots, m$, and by Lemma 6.6, $u' = \sum_{k=1}^m \psi_k v_k$ is defined. Since $u - u' \perp v_k$, $k = 1, \ldots, m$, u' is the orthogonal projection of u onto N, and so, $\|u - u'\|_Y < \varepsilon$. Decreasing ε and increasing N, we obtain a series $\sum_{k=1}^\infty \langle u, v_k\rangle v_k$ converging to u. ■

6.15. Corollary. *Under the conditions of Lemma 6.14, $\psi_k = \langle u, v_k\rangle$, $k = 1, 2, \ldots$, and $\|u\|^2 = \sum_{k=1}^\infty |\psi_k|^2$ in $L^1(Y)$.*

Proof. $\|u\|^2 = \sum_{k,l=1}^\infty \psi_k \overline{\psi_l} \langle v_k, v_l\rangle = \sum_{k=1}^\infty |\psi_k|^2$. ■

6.16. Now we will describe two constructions of Y-Hilbert spaces. The first one is given by the operation of *complex conjugation*: $\overline{M} = \{\bar{u} : u \in M\}$ with $\langle \bar{u}, \bar{v}\rangle_y = \overline{\langle u, v\rangle_y}$ and $\varphi \bar{u} = \overline{\overline{\varphi} u}$ is a Y-Hilbert space.

6.17. The second construction is *the tensor product of Y-Hilbert spaces*. We denote by M^∞ the subspace of M consisting of vectors whose norms are essentially bounded: $M^\infty = \Big\{u \in M : \|u\| \in L^\infty(Y)\Big\}$. It is clear that M^∞ is dense in M. Let M' be another Y-Hilbert space. Then $M^\infty \otimes M'^\infty = M^\infty \otimes_{L^\infty(Y)} M'^\infty$ with the inner product given by $\langle u_1 \otimes u_1', u_2 \otimes u_2'\rangle = \langle u_1, u_2\rangle \langle u_1', u_2'\rangle$ is a well defined Y-pre-Hilbert space.

Lemma. *The introduced inner product is nondegenerate on $M^\infty \otimes M'^\infty$: for nonzero $w \in M^\infty \otimes M'^\infty$ one has $\|w\| \neq 0$.*

Proof. Let $w=\sum_{k=1}^{l} u_i\otimes u'_k$, $u_1,\ldots,u_l\in M^\infty$, $u'_1,\ldots,u'_l\in M'^\infty$. Let N and N' be the closed subspaces of M and M' spanned by $u_1,\ldots,u_l$ and $u'_1,\ldots,u'_l$ respectively. Choose orthonormal bases $v_1,\ldots,v_m$ in N and $v'_1,\ldots,v'_{m'}$ in N', then $u_k=\sum_{i=1}^{m}\varphi_{k,i}v_i$, $u'_k=\sum_{i=1}^{m'}\varphi'_{k,j}v'_j$, $k=1,\ldots,l$, where $\varphi_{k,i}=\langle u_k,v_i\rangle$, $\varphi'_{k,j}=\langle u'_k,v_j\rangle$, and, since $\|u_k\|,\|u'_k\|\in L^\infty(Y)$, all $\varphi_{k,i},\varphi'_{k,j}\in L^\infty(Y)$, $k=1,\ldots,l$, $i=1,\ldots,m$, $j=1,\ldots,m'$. We have then $w=\sum_{\substack{1\le i\le m\\ 1\le j\le m'}}\varphi_{k,i}\varphi'_{k,j}v_i\otimes v'_j$, and $\|w\|^2=\sum_{\substack{1\le i\le m\\ 1\le j\le m'}}|\varphi_{k,i}|^2|\varphi'_{k,j}|^2$. ∎

We will call the Y-Hilbert space obtained by the completion of $M^\infty\otimes M'^\infty$ *the tensor product of M and M'* and denote by $M\otimes M'$. If $\mathcal{B}$ and $\mathcal{B}'$ are orthonormal bases in M and M' respectively, then $\big\{v\otimes v' : v\in\mathcal{B},\ v'\in\mathcal{B}' \text{ with } v\otimes v'\neq 0\big\}$ is an orthonormal basis in $M\otimes M'$. It follows that every element w of $M\otimes M'$ is representable in the form $w=\sum_{i,j=1}^{\infty}\psi_{i,j}v_i\otimes v'_j$, $v_1,v_2,\ldots,\mathcal{B}$, $v'_1,v'_2,\ldots,\mathcal{B}'$, and if we put $u_i=\sum_{j=1}^{\infty}\psi_{i,j}v'_j$, $i=1,2,\ldots$, in the form $w=\sum_{i=1}^{\infty}v_i\otimes u_i$.

Examples.

1. If Y is single-element and so, M and M' are conventional Hilbert spaces, then $M\otimes M'$ is the completion of the tensor product $M\otimes_{\mathbb{C}} M'$.

2. If $M=L^2(Y\times Z)$ and $M'=L^2(Y\times Z')$ (see example 6.4.3), then $M\otimes M'=L^2(Y\times Z\times Z')$, it corresponds to *the relative product* $Y\times Z\times Z'$ of the measure spaces $Y\times Z$ and $Y\times Z'$.

6.18. For $u\in M$, $b\in\mathbb{R}$, denote $B_{u,b}=\big\{y\in Y : \|u\|_y>b\big\}$. It is clear that for every $u\in M$, $\|u\|_{B_{u,b}}\longrightarrow 0$ when $b\longrightarrow\infty$. We say that a set $U\in M$ has *uniformly bounded growth* if for any $\varepsilon>0$ there exists $b\in\mathbb{R}$ such that $\|u\|_{B_{u,b}}<\varepsilon$ for all $u\in U$.

The following lemma is one of the main technical tools in our further considerations.

Lemma. *Let $u(n)$, $n\in\mathbb{Z}^d$, be a sequence in M with uniformly bounded growth and satisfying*

$$\operatorname*{D-lim}_{m}\operatorname*{D-lim\,sup}_{n}\mathcal{N}(u(n+m),u(n))=0.$$

Then, for any $v\in M$, $\mathcal{N}(u(n),v)\xrightarrow{D}0$.

Proof. The proof is based on an analogous proposition taking place in the "absolute" case (that is, where M is a Hilbert space), Lemma 4.9 in [F2]. This lemma is formulated for the case of one-dimensional sequences only, $d=1$, but its proof can be verbatim transferred to the case $d\ge 2$.

Let us assume first that $u(n)\in M^\infty$, $n\in\mathbb{Z}^d$, and there exists $b\in\mathbb{R}$ such that ess-sup $\|u(n)\|<b$, $n\in\mathbb{Z}^d$. Then for $n,m\in\mathbb{Z}^d$,

$$\begin{aligned}\mathcal{N}(u(n),u(n+m))&=\int\big|\langle u(n),u(n+m)\rangle_y\big|d\nu>\frac{1}{b^2}\int\big|\langle u(n),u(n+m)\rangle_y\big|^2d\nu\\&=\frac{1}{b^2}\int\langle u(n),u(n+m)\rangle_y\overline{\langle u(n),u(n+m)\rangle_y}d\nu\\&=\frac{1}{b^2}\int\Big\langle u(n)\otimes\bar u(n),u(n+m)\otimes\bar u(n+m)\Big\rangle_y d\nu=\frac{1}{b^2}\big\langle w(n),w(n+m)\big\rangle_Y,\end{aligned}$$

where $w(n) = u(n) \otimes \bar{u}(n)$ is a bounded sequence of elements of $M \otimes \overline{M}$, considered as a Hilbert space with the inner product $\langle\, ,\rangle_Y$. By Lemma 4.9 in [F2], the equality $\underset{m}{\text{D-lim}}\, \underset{n}{\text{D-lim}\sup} \big|\langle w(n), w(n+m)\rangle_Y\big| = 0$ implies $\underset{n}{\text{D-lim}} \langle w(n), w'\rangle_Y = 0$ for all $w' \in M \otimes \overline{M}$. Since

$$\langle w(n), v \otimes \bar{v}\rangle_Y = \int \langle u(n) \otimes \bar{u}(n), v \otimes \bar{v}\rangle_y d\nu = \int \big|\langle u(n), v\rangle_y\big|^2 d\nu$$
$$\geq \Big(\int \big|\langle u(n), v\rangle_y\big| d\nu\Big)^2 = \mathcal{N}(u(n), v)^2,$$

we obtain that $\underset{n}{\text{D-lim}}\, \mathcal{N}(u(n), v) = 0$.

Now, for general v and $u(n)$, $n \in \mathbb{Z}^d$, the condition that $u(n)$, $n \in \mathbb{Z}^d$, have uniformly bounded growth implies that for every $\varepsilon > 0$ there are $v' \in M^\infty$ and $u'(n) \in M^\infty$, $n \in \mathbb{Z}^d$, with $\|v - v'\|_Y < \varepsilon$, $\|u(n) - u'(n)\|_Y < \varepsilon$ and $\text{ess-sup}\, \|u'(n)\| < b$, $n \in \mathbb{Z}^d$, for some $b \in \mathbb{R}$. ■

6.19. Given $u \in M$ and $\varepsilon > 0$, a set $V \subseteq M$ is *an ε-Y-net for u* if $\int \min_{v \in V} \|u - v\|_y^2 d\nu < \varepsilon^2$.

Lemma. *Let $V \subseteq M$ be an ε-Y-net for $u \in M$ and let $\|u' - u\|_Y < \varepsilon$. Then V is a 2ε-Y-net for u'.*

Proof. Since for any $f, g \in L^2(Y)$ one has $\int |f+g|^2 d\nu \leq 2\Big(\int |f|^2 d\nu + \int |g|^2 d\nu\Big)$,

$$\int \min_{v \in V} \|u' - v\|_y^2 d\nu \leq \int \big(\|u' - u\|_y + \min_{v \in V} \|u - v\|_y\big)^2 d\nu$$
$$\leq 2\Big(\int \|u' - u\|_y^2 d\nu + \int \min_{v \in V} \|u - v\|_y^2 d\nu\Big) < 4\varepsilon^2.$$

■

6.20. Lemma. *Let a finite set $\{v_1, \ldots, v_k\}$ be an ε-Y-net for $u \in M$. Then*

$$\|u\|_Y^2 < \sum_{i=1}^k \Big(\|v_i\|_Y^2 + 2\varepsilon \|v_i\|_Y\Big) + \varepsilon^2$$

and, for all $w \in M$,

$$\mathcal{N}(u, w) < \sum_{i=1}^k \mathcal{N}(v_i, w) + \varepsilon \|w\|_Y.$$

Proof.

$$\|u\|_Y^2 = \int \|u\|_y^2 d\nu \leq \int \min_{1 \leq i \leq k} \Big(\|u - v_i\|_y^2 + 2\big|\langle u - v_i, v_i\rangle_y\big| + \|v_i\|_y^2\Big) d\nu$$
$$\leq \int \min_{1 \leq i \leq k} \Big(\|u - v_i\|_y^2 + 2\|u - v_i\|_y \sum_{j=1}^k \|v_j\|_y\Big) d\nu + \sum_{j=1}^k \int \|v_j\|_y^2 d\nu$$
$$\leq \int \min_{1 \leq i \leq k} \|u - v_i\|_y^2 d\nu + 2 \sum_{j=1}^k \Big(\int \min_{1 \leq i \leq k} \|u - v_i\|_y^2 d\nu\Big)^{1/2} \Big(\int \|v_j\|_y^2 d\nu\Big)^{1/2} + \sum_{j=1}^k \int \|v_j\|_y^2 d\nu$$
$$< \varepsilon^2 + \sum_{j=1}^k \Big(2\varepsilon \|v_j\|_Y + \|v_j\|_Y^2\Big).$$

$$\begin{aligned}\mathcal{N}(u,w) &= \int \big|\langle u,w\rangle_y\big| d\nu \le \int \min_{1\le i\le k}\Big(\big|\langle u-v_i,w\rangle_y\big| + \big|\langle v_i,w\rangle_y\big|\Big)d\nu \\ &\le \int \min_{1\le i\le k}\|u-v_i\|_y\|w\|_y d\nu + \int \sum_{j=1}^{k}\big|\langle v_j,w\rangle_y\big| d\nu \\ &\le \Big(\int \min_{1\le i\le k}\|u-v_i\|_y^2 d\nu\Big)^{1/2}\Big(\int \|w\|_y^2 d\nu\Big)^{1/2} + \sum_{j=1}^{k}\int \big|\langle v_j,w\rangle_y\big| d\nu \\ &< \varepsilon\|w\|_Y + \sum_{j=1}^{k}\mathcal{N}(v_j,w).\end{aligned}$$

■

6.21. Given a set $U \subseteq M$ and $\varepsilon > 0$, a set $V \subseteq M$ is *an ε-Y-net for U* if V is an ε-Y-net for every $u \in U$.

A set $U \subset M$ is *Y-precompact* if for any $\varepsilon > 0$ there exists a finite ε-Y-net for U. Y-precompactness is weaker than precompactness: it is clear that any precompact subset of M is Y-precompact(in particular, any finite subset of M is Y-precompact). The inverse is not true generally speaking; it follows from Lemma 6.20 however that any Y-precompact set is bounded.

6.22. We can even say more.

Lemma. *Let a set $U \subset M$ be Y-precompact. Then for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $B \in \mathfrak{D}$ with $\nu(B) < \delta$ and any $u \in U$ one has $\|u\|_B < \varepsilon$.*

Proof. For $\varepsilon > 0$, let $\{v_1, \dots, v_k\}$ be an ε-Y-net for U, and let $\delta > 0$ be such that $\|v_i\|_B < \varepsilon/k$ and $\|v_i\|_B^2 < \varepsilon^2/k$ for any $B \in \mathfrak{D}$ satisfying $\nu(B) < \delta$. Then by a modification of Lemma 6.20, for any $u \in U$

$$\|u\|_B^2 < \sum_{i=1}^{k}\Big(\|v_i\|_B^2 + 2\varepsilon\|v_i\|_B\Big) + \varepsilon^2 < 4\varepsilon^2.$$

■

6.23. The following lemma demonstrates that Y-precompactness is "an inner property":

Lemma. *Let M be a separable Y-Hilbert space and let a set $U \subset M$ be Y-precompact. Then for any $\varepsilon > 0$ there exists a finite ε-Y-net for U, contained in U.*

Proof. (It is unexpectedly long.) Given $\varepsilon > 0$, using Lemma 6.22 find $0 < \delta < \varepsilon^2$ such that $B \in \mathfrak{D}$, $\nu(B) < 2\delta$ implies $\|u\|_B < \varepsilon$ for all $u \in U$. Let $\{v_1, \dots, v_k\}$ be a δ-Y-net for U. For each $u \in U$, $1 \le i \le k$ denote $A_{u,i} = \Big\{y \in Y : \min_{1\le i\le k}\|u - v_i\|_y < \sqrt{\delta}\Big\}$, $A_i = \bigcup_{u\in U} A_{u,i}$, and $A_u = \bigcup_{i=1}^{k} A_{u,i}$. We have $\nu(A_u) > 1 - \delta$.

Choose a countable family $\{u_j\}_{j\in\mathbb{N}}$ dense in U. Then we also have $A_i = \bigcup_{j=1}^{\infty} A_{u_j,i}$ for each $i = 1, \dots, k$. Choose $K \in \mathbb{N}$ so big that for $A_i' = \bigcup_{j=1}^{K} A_{u_j,i}$ one has $\nu(A_i \setminus A_i') < \delta/k$.

Let now $u \in U$. Denote $A_{u,i}' = A_{u,i} \cap A_i'$, $i = 1, \dots, k$. Then for any $y \in A_{u,i}'$ we have

$$\min_{1\le j\le K}\|u - u_j\|_y \le \|u - v_i\|_y + \min_{1\le j\le K}\|v_i - u_j\|_y < 2\sqrt{\delta}.$$

On the other hand, since $A_u \subseteq \bigcup_{i=1}^k A_i$ we have

$$B_u = Y \setminus \bigcup_{i=1}^k A'_{u,i} \subseteq (Y \setminus A_u) \cup \bigcup_{i=1}^k \Big(A_i \setminus A'_i\Big),$$

and so, $\nu(B_u) < 2\delta$. Hence,

$$\int \min_{1\le j\le K} \|u-u_j\|_y^2 d\nu \le \int_{\bigcup_{i=1}^k A'_{u,i}} \min_{1\le j\le K} \|u-u_j\|_y^2 d\nu + \Big(\|u\|_{B_u} + \|u_1\|_{B_u}\Big)^2$$
$$< 4\delta + 4\varepsilon^2 < 8\varepsilon^2. \qquad \blacksquare$$

7. Weakly mixing and compact actions on a Y-Hilbert space

7.1. *A transformation* T of a Y-Hilbert space M is a pair (T_Y, T_M) where T_Y is a measure preserving transformation of Y, $y \to yT$, and T_M is a linear self-mapping of M, $u \to Tu$, that satisfy:
a) $\langle Tu, Tv\rangle = T\langle u, v\rangle$, $u, v \in M$,
b) $T(\varphi u) = T\varphi Tu$, $\varphi \in L^\infty(Y)$, $u \in M$,
where, for a function f on Y, the function Tf is defined by $(Tf)(y) = f(yT)$. We will follow the convention that transformations of M act on Y from the right and on elements of M from the left.

It is clear that any transformation T preserves the forms $\langle\,,\rangle_Y$ and $\mathcal{N}$: $\langle Tu, Tv\rangle_Y = \langle u, v\rangle_Y$, $\mathcal{N}(Tu, Tv) = \mathcal{N}(u, v)$ for $u, v \in M$.

We say that a set Q *acts on* M if a mapping from Q into the set of transformations of M is given. If a one-element set $\{T\}$ acts on M, we say that T *acts on* M. *A group* G *acts on* M if a homomorphism of G into the group of invertible transformations of M is given.

Remark. In the terminology of [R], [Z1], the transformations we defined are *cocycle representations of* $(Y, \mathbb{Z})$ *on* M. We have preferred the term "transformation" as a neutral one.

7.2. Examples.
7.2.1. If Y is a single-element set, $Y = \{y\}$, then M is a Hilbert space and its transformations are its isometries.

7.2.2. If $M = L^2(Y)$, its transformations are the isometries of this Hilbert space induced by measure preserving transformations of Y.

7.2.3. Let $(X, \mathfrak{B}, \mu) = (Y, \mathfrak{D}, \nu) \times (Z, \mathfrak{C}, \eta)$, let T be a measure preserving transformation of X which also preserves the fibers of the projection $\pi: X \longrightarrow Y$: $\pi(x_1) = \pi(x_2)$ implies $\pi(x_1T) = \pi(x_2T)$. Then T induces a measure preserving transformation of Y by $(\pi(x))T = \pi(xT)$, and the natural action of T on $L^2(X)$ is a transformation of this Y-Hilbert space. Note that not all transformations of $L^2(X)$ are obtainable in this way (compare with the absolute case: though any Hilbert space is isomorphic to $L^2(X)$ for a suitable measure space X, not all isometries of $L^2(X)$ result from measure preserving transformations of X).

7.3. From now on, M will be a separable Y-Hilbert space.

Given a sequence $g(n)$, $n \in \mathbb{Z}^d$, of transformations of M and a vector $u \in M$, we say that *the action of g on u is weakly mixing* or *g is weakly mixing on u* if $g(n)u$ "weakly converges to zero in density": $\mathcal{N}(g(n)u, v) \xrightarrow{D} 0$ for all $v \in M$. We say that *$g(n)$ is weakly mixing on $N \subseteq M$* if $g(n)$ is weakly mixing on every $u \in N$. If T is an invertible transformation of M, we say that *T is weakly mixing on $u \in M$ (on $N \subseteq M$)* if the sequence $g(n) = T^n$, $n \in \mathbb{Z}$, is weakly mixing on u (on N).

Examples.

1. If Y is single-element and so, M is a conventional Hilbert space and its invertible transformation T is a unitary operator on M, then T is weakly mixing on M if it is weakly mixing on M in the usual sense, that is if T has pure continuous spectrum.

2. Let $M = L^2(Y \times Z)$, let T be an invertible measure preserving transformation of Y and S be a unitary operator on $L^2(Z)$. Lift T to M by $T(\varphi f) = T\varphi Sf$ for $\varphi \in L^{\infty(Y)}$ and $f \in L^2(Z)$. Then T is weakly mixing on M if and only if S is weakly mixing on the Hilbert space $L^2(Z)$, that is if T is weakly mixing in the usual sense "on the fibers" of M (in the terminology of [Z2] and [F2], T is *relatively* weakly mixing).

Note that our weak mixing is stronger than the "absolute" weak mixing: $\langle g(n)u, v\rangle \xrightarrow{D} 0$ in $L^1(Y)$ implies $\langle g(n)u, v\rangle_Y \xrightarrow{D} 0$, and so, if a sequence $g(n)$ is weakly mixing on M, then $g(n)$ is weakly mixing on the Hilbert space M with the inner product $\langle\, ,\rangle_Y$.

7.4. For a sequence $g(n)$, $n \in \mathbb{Z}^d$, of transformations of M, we define

$$M^w(g) = \big\{u \in M : g \text{ is weakly mixing on } u\big\}.$$

It is clear that $M^w(g)$ is a closed subspace of M. If T is an invertible transformation of M, let $M^w(T) = M^w\big(T^n,\ n \in \mathbb{Z}\big)$.

The following elementary lemma shows that the property "a sequence g is weakly mixing on $u \in N$" is an "inner" property of the vector u.

Lemma. *Let $g(n)$, $n \in \mathbb{Z}^d$, be a sequence of transformations of M and let N be a closed subspace of M invariant with respect to all $g(n)$. Then $N^w(g) = M^w(g) \cap N$.*

Proof. Let $u \in N^w(g)$. For $v \in M$, let $w \in N$ be the orthogonal projection of v onto N. Then by Lemma 6.12, $\mathcal{N}(g(n)u, v) = \mathcal{N}(g(n)u, w)$ for all $n \in \mathbb{Z}^d$. ∎

7.5. The following lemma is the main tool when we deal with "polynomial" weakly mixing sequences.

Lemma. *Let $g(n)$, $n \in \mathbb{Z}^d$, be a sequence of invertible transformations of M and let $u \in M$. If for almost all $m \in \mathbb{Z}^d$ the derivative sequence $D^m g = g(n)^{-1}g(n+m)$ is weakly mixing on $u \in M$, then g is weakly mixing on u.*

Proof. Denote $u(n) = g(n)u$, $n \in \mathbb{Z}^d$. Then, for $n, m \in \mathbb{Z}^d$,

$$\mathcal{N}\big(u(n+m), u(n)\big) = \mathcal{N}\big(g(n+m)u, g(n)u\big) = \mathcal{N}\big(g(n)^{-1}g(n+m)u, u\big) = \mathcal{N}\big(D^m g(n)u, u\big).$$

For all m but a set of zero density in $\mathbb{Z}^d$ we have $\operatorname*{D\text{-}lim}_n \mathcal{N}\big(D^m g(n)u, u\big) = 0$, hence

$$\operatorname*{D\text{-}lim}_m \operatorname*{D\text{-}lim\,sup}_n \mathcal{N}\big(u(n+m), u(n)\big) = 0.$$

It is also clear that $u(n)$, $n \in \mathbb{Z}^d$, have uniformly bounded growth. By Lemma 6.18 it follows that $\operatorname*{D\text{-}lim}_n \mathcal{N}(u(n), v) = 0$ for all $v \in M$. ∎

7.6. Given a set Q acting on M, we will say that Q *acts compactly on* $u \in M$ if the set $Qu = \big\{Tu,\ T \in Q\big\}$ is Y-precompact, and that Q *acts compactly on* $N \subseteq M$ if Q acts compactly on every $u \in N$. If T is an invertible transformation of M, T *acts compactly on* $u \in M$ *(on* $N \subseteq M$*)* if $\{T^n,\ n \in \mathbb{Z}\}$ does.

7.7. For a set Q acting on M, we will denote

$$M^c(Q) = \big\{u \in M : Q \text{ acts compactly on } u\big\}.$$

By Lemma 6.19, $M^c(Q)$ it is a closed subspace of M. When T is an invertible transformation of M, we will $M^c(\{T^n,\ n \in \mathbb{Z}\})$ by $M^c(T)$.

Since the property to be Y-precompact is an "inner" property of a set (see Lemma 6.23), we have

Lemma. *Let a set Q act on M and and let N be a subspace of M invariant with respect to the action of Q. Then $N^c(Q) = M^c(Q) \cap N$.*

7.8. Proposition. *Let a subset $U \subset M$ be Y-precompact and let a set Q act compactly on U. Then the set $QU = \{Tu,\ T \in Q,\ u \in U\}$ is Y-precompact.*

Proof. Let $\{v_1, \dots, v_k\} \subseteq U$ be an ε-Y-net for U, and let $w_1, \dots, w_l$ be an ε/k-Y-net for $\bigcup_{i=1}^k Qv_i$. Then, for any $u \in U$, $T \in Q$,

$$\int \min_{1\le i\le l} \|Tu - w_i\|_y^2 d\nu \le \int \min_{1\le j\le k} \|Tu - Tv_j\|_y^2 d\nu + \sum_{j=1}^k \int \min_{1\le i\le l} \|Tv_j - w_i\|_y^2 d\nu$$
$$+2\Big(\int \min_{1\le j\le k} \|Tu - Tv_j\|_y^2 d\mu_y\Big)^{1/2} \sum_{j=1}^k \Big(\int \min_{1\le i\le l} \|Tv_j - w_i\|_y^2 d\nu\Big)^{1/2} < 4\varepsilon^2.$$

∎

7.9. Corollary. *Let a set Q_1 act compactly on $u \in M$ and let a set Q_2 act compactly on M. Then $Q_2Q_1 = \{T_2T_1,\ T_1 \in Q_1,\ T_2 \in Q_2\}$ acts compactly on u.*

7.10. Corollary. *Let G be a group acting on M and let H be a subgroup of G of finite index. Then $M^c(G) = M^c(H)$.*

Proof. Let $Q \subseteq G$ be any finite set containing representatives of all left cosets of H in G. Then Q, being finite, acts compactly on M, and $G = QH$. ∎

7.11. Corollary. *For $m \in \mathbb{Z}$, $m \ne 0$, $M^c(T^m) = M^c(T)$.*

7.12. Corollary. *Let G be a group acting on M, let H_1 and H_2 be subgroups of G such that H_2 acts compactly on M and H_1 acts compactly on$u \in M$ and normalizes H_2: $H_1 \subseteq N(H_2)$. Then the group H generated by H_1 and H_2 acts compactly on u.*

Proof. Any element T of H can be written in the form $T = T_2T_1$ with $T_1 \in H_1$ and $T_2 \in H_2$. ■

7.13. Lemma. *Let G be a group acting on M, let Q be a subset of G and let $T \in G$. Then $TM^c(Q) = M^c(TQT^{-1})$.*

Proof. For $u \in M^c(Q)$, $TQT^{-1}(Tu) = TQu$ is Y-precompact. ■

7.14. We are going now to establish some facts concerning relations between weakly mixing and compact actions.

Proposition. *Let a sequence $g(n)$, $n \in \mathbb{Z}^d$, of transformations of M be weakly mixing on M and let a sequence $h(n)$, $n \in \mathbb{Z}^d$, of transformations of M act compactly on $u \in M$. Then the sequence $g(n)h(n)$, $n \in \mathbb{Z}^d$, is weakly mixing on u.*

Proof. For $\varepsilon > 0$, choose an ε-Y-net $\{w_1, \ldots, w_k\}$ for $\{h(n)u,\ n \in \mathbb{Z}^d\}$. Then, for any $n \in \mathbb{Z}^d$, the set $\{g(n)w_1, \ldots, g(n)w_k\}$ is an ε-Y-net for $g(n)h(n)u$. So, by Lemma 6.20, for any $v \in M$ and any $n \in \mathbb{Z}^d$

$$\mathcal{N}\big(g(n)h(n)u, v\big) < \sum_{i=1}^{k} \mathcal{N}\big(g(n)w_i, v\big) + \varepsilon\|v\|_Y.$$

Since $\mathcal{N}\big(g(n)w_i, v\big) \xrightarrow{D} 0$, $i = 1, \ldots, k$, we have $\operatorname{D-lim\,sup}_n \mathcal{N}\big(g(n)h(n)u, v\big) < \varepsilon\|v\|_Y$. Since ε is arbitrary, $\operatorname{D-lim}_n \mathcal{N}\big(g(n)h(n)u, v\big) = 0$. ■

7.15. Proposition. *Let $g(n)$, $n \in \mathbb{Z}^d$, be a sequence of transformations of M. Then $M^w(g) \perp M^c(g)$.*

Proof. Let $v \in M^w(g)$, $u \in M^c(g)$; we have to prove that $\mathcal{N}(u, v) = 0$.

For $\varepsilon > 0$, let $\{w_1, \ldots, w_k\}$ be an ε-Y-net for gu. Then, for any $n \in \mathbb{Z}^d$, by Lemma 6.20

$$\mathcal{N}(u, v) = \mathcal{N}\big(g(n)u, g(n)v\big) < \sum_{i=1}^{k} \mathcal{N}\big(w_i, g(n)v\big) + \varepsilon\|v\|_Y.$$

Since $\mathcal{N}\big(w_i, g(n)v\big) \xrightarrow{D} 0$, $i = 1, \ldots, k$, this proves that $\mathcal{N}(u, v)$ is smaller than any positive number, and hence is zero. ■

8. Weakly mixing action of an amenable group

In this section we bring a theorem (Theorem 8.4 below) saying that, for actions of a countable amenable group G on a Y-Hilbert space, the notions of compactness and weakly mixing are complementary. We will use this fact in the cases where G is commutative or nilpotent, but its proof remains the same in the general amenable case.

In this section, let M be a separable Y-Hilbert space.

8.1. Let G be a countable group. A sequence $\Phi_1, \Phi_2, \ldots$ of finite subsets of G is called *a (right) Følner sequence* if $\dfrac{\#(\Phi_k T \triangle \Phi_k)}{\#\Phi_k} \underset{k\to\infty}{\longrightarrow} 0$ for every $T \in G$. Groups having a Følner sequence are called *amenable*.

8.2. Commutative and nilpotent groups are amenable: a sequence of parallelepipeds $\Pi_1, \Pi_2, \ldots \subset \mathbb{Z}^t$ with $L(\Pi_k) \underset{k\to\infty}{\longrightarrow} \infty$ provides us with an example of a Følner sequence in $\mathbb{Z}^t$. Now, let G be a finitely generated torsion-free nilpotent group, let $(T_1, \ldots, T_t)$ be an ordered basis in G. Then G is identified with $\mathbb{Z}^t$ (as a set, not as a group of course) by the coordinate mapping $\big(T_t^{n_t} \ldots T_1^{n_1}\big) \mapsto (n_1, \ldots, n_t)$ (see subsection 2.6). Let numbers $a_{k,i} \le b_{k,i} \in \mathbb{Z}$, $i = 1, \ldots, d$, $k \in \mathbb{N}$, satisfy the conditions

$$b_{k,t} - a_{k,t} \underset{k\to\infty}{\longrightarrow} \infty, \ \frac{b_{k,i} - a_{k,i}}{|a_{k,i+1}|^p + |b_{k,i+1}|^p + 1} \underset{k\to\infty}{\longrightarrow} \infty \text{ for all } p > 0, \ i = t-1, \ldots, 1, \tag{8.1}$$

then the sequence of parallelepipeds

$$\Pi_k = \prod_{i=1}^{t} \{a_{k,i}, a_{k,i} + 1, \ldots, b_{k,i}\} \subset \mathbb{Z}^t, \ k \in \mathbb{N},$$

is a Følner sequence in G. This easily follows from the fact that the multiplication in G in coordinates $(n_1, \ldots, n_t)$ is polynomial. Indeed, fix $T \in G$. The mapping $M_T\colon G \longrightarrow G$, $P \mapsto PT$, acts on points of $\mathbb{Z}^t$ by the rule

$$M_T\Big((n_1, \ldots, n_t)\Big) = \big(n_1 + f_1(n_2, \ldots, n_t), \ldots, n_t + f_t\big)$$

for some polynomials $f_i(n_{i+1}, \ldots, n_t)$, $i = 1, \ldots, t$, (see subsection 2.6). Let $C \in \mathbb{R}$, $p \in \mathbb{N}$ be such that

$$\big|f_i(n_{i+1}, \ldots, n_t)\big| < C\big(|n_{i+1}|^p + \ldots + |n_t|^p + 1\big) \text{ for any } n_1, \ldots, n_t \in \mathbb{Z}, \, i = 1, \ldots, t.$$

A point $n = (n_1, \ldots, n_t) \in \Pi_k$ can leave Π_k under the action of M_T only if for some $1 \le i \le t$, either $0 \le n_i - a_{k,i} < \big|f_i(n_{i+1}, \ldots, n_t)\big|$ or $0 \le b_{k,i} - n_i < \big|f_i(n_{i+1}, \ldots, n_t)\big|$. Hence

$$\begin{aligned} \frac{\#(M_T\Pi_k \triangle \Pi_k)}{\#\Pi_k} = 2\frac{\#(M_T\Pi_k \setminus \Pi_k)}{\#\Pi_k} &\le 4\sum_{i=1}^{t} \frac{\max_{n\in\Pi_k} \big|f_i(n_{i+1}, \ldots, n_t)\big|}{b_{k,i} - a_{k,i}} \\ &< 4\sum_{i=1}^{t} \frac{\max_{n\in\Pi_k} C\Big(|n_{i+1}|^p + \ldots + |n_t|^p + 1\Big)}{b_{k,i} - a_{k,i}} \\ \le 4C\sum_{i=1}^{t} \frac{|a_{k,i+1}|^p + \ldots + |a_{k,t}|^p + |b_{k,i+1}|^p + \ldots + |b_{k,t}|^p + 1}{b_{k,i} - a_{k,i}} &\underset{k\to\infty}{\longrightarrow} 0. \end{aligned}$$

Remark. The condition (8.1) is far from being necessary.

8.3. Let G be a countable amenable group acting on M. A set $\Gamma \subseteq G$ is said to be *of zero density in* G if for every Følner sequence $\Phi_1, \Phi_2, \ldots$ in G, $\dfrac{\#(\Gamma \cap \Phi_k)}{\#\Phi_k} \underset{k\to\infty}{\longrightarrow} 0$.

We say that G *is weakly mixing on* $u \in M$ if for any Følner sequence $\Phi_1, \Phi_2, \ldots$ in G, any $v \in M$ and any $c > 0$, the set $\Big\{T \in G : \mathcal{N}(Tu, v) > c\Big\}$ is of zero density in G. We say that G is weakly mixing on $N \subseteq M$ if G is weakly mixing on every $u \in N$; the maximal subspace of M on which G is weakly mixing we will denote by $M^w(G)$.

If G is commutative, say $G = \mathbb{Z}^d$, then G can be considered as a d-dimensional sequence. It is easy to see then that G is weakly mixing on $u \in M$ as an amenable group if and only if G is weakly mixing on u in the sense of subsection 7.3, that is as a sequence. In particular, if T is an invertible transformation of M, then T is weakly mixing on $u \in M$ if and only if the group generated by T is weakly mixing on u.

8.4. Theorem. *Let G be a countable amenable group acting on M. If G is not weakly mixing on M, then there exists nonzero $u \in M$ such that G acts compactly on u.*

As the matter of fact, it is the key point: the notions of weakly mixing and compact actions of an amenable group are complementary.

8.5. The proof of Theorem 8.4 we bring here (as well as the proof of Theorem 8.11 below) can be found in [F2], Ch. 6; we only have to adapt it to the more abstract situation we deal with. It is based on the (well known) fact that the Mean Ergodic Theorem holds for unitary actions of amenable groups:

Lemma. *Let H be a Hilbert space, let G be a countable amenable group of unitary operators on H and let $\Phi_1, \Phi_2, \ldots$ be a Følner sequence in G. Then for every $w \in H$ the limit $\lim\limits_{k\to\infty} \dfrac{1}{\#\Phi_k} \sum\limits_{T\in\Phi_k} Tw$ exists and equals the orthogonal projection of w onto the space of G-invariant elements of H.*

Proof. Let $w \in H$ be of the form $w = w' - Sw'$ for some $w' \in H$, $S \in G$. Then

$$\Big\|\sum_{T\in\Phi_k} Tw\Big\| = \Big\|\sum_{T\in\Phi_k} Tw' - \sum_{T\in\Phi_k} TSw'\Big\| = \Big\|\sum_{T\in\Phi_k\setminus\Phi_k S} Tw' - \sum_{T\in\Phi_k S\setminus\Phi_k} Tw'\Big\| \leq \#(\Phi_k S \triangle \Phi_k)\|w'\|$$

and so, $\lim\limits_{k\to\infty} \dfrac{1}{\#\Phi_k} \sum\limits_{T\in\Phi_k} Tw = 0$.

Now, let $w \in H$ be orthogonal to all vectors of the form $w' - Sw'$, $w' \in H$, $S \in G$. Then, in particular, for any $S \in G$, $\langle w, w - Sw\rangle = \langle w, w - S^{-1}w\rangle = 0$. Thus

$$\|w - Sw\|^2 = \langle w - Sw, w - Sw\rangle = \langle w, w - Sw\rangle - \langle w, S^{-1}w - w\rangle = 0,$$

so $w = Sw$, that is w is G-invariant and $\dfrac{1}{\#\Phi_k} \sum\limits_{T\in\Phi_k} Tw = w$ for all k. ∎

8.6. Proof of Theorem 8.4. Consider the Y-Hilbert space $M \otimes \overline{M}$ (see 6.16 and 6.17). Let $v \in \overline{M}^\infty$, define a linear mapping $*v\colon M^\infty \otimes \overline{M}^\infty \longrightarrow M$, $w \mapsto w * v$, by

$$(u \otimes v') * v = \langle v', v\rangle u.$$

Let $\mathcal{B}$ be an orthonormal basis in M. For $w = \sum_{k=1}^{l} u_k \otimes v_k \in M^\infty \otimes \overline{M}^\infty$ with $u_1, \ldots, u_l \in \mathcal{B}$ we have

$$\|w * v\| = \Big\|\sum_{k=1}^{l} \langle v_k, v\rangle u_k\Big\| = \Big(\sum_{k=1}^{l} |\langle v_k, v\rangle|^2\Big)^{\frac{1}{2}} \leq \|v\| \Big(\sum_{k=1}^{l} \|v_k\|^2\Big)^{\frac{1}{2}} \leq \text{ess-sup}(\|v\|)\|w\| \tag{8.2}$$

(as elements of $L^2(Y)$, that is for almost all $y \in Y$), so

$$\|w * v\|_Y \leq \text{ess-sup}(\|v\|)\|w\|_Y. \tag{8.3}$$

Hence, the mapping $*v$ is bounded and can be lifted to $M \otimes \overline{M}$: for $w = \sum_{k=1}^{\infty} u_k \otimes v_k$ with $u_1, u_2, \ldots \in \mathcal{B}$, one has $w * v = \sum_{k=1}^{\infty} \langle v_k, v\rangle u_k \in M$.

Moreover, for every nonzero $w \in M \otimes \overline{M}$ there is $v \in \overline{M}^\infty$ for which $w * v \neq 0$. Indeed, if $w = \sum_{k=1}^{\infty} u_k \otimes v_k$ with $u_1, u_2, \ldots \in \mathcal{B}$ and $v_1 \neq 0$, let $A \in \mathfrak{D}$ be such that $v = 1_A v_1 \in \overline{M}^\infty$ and $v \neq 0$. Then

$$\|w * v\|^2 = \sum_{k=1}^{\infty} \big|\langle v_k, v\rangle\big|^2 \geq \big|\langle v_1, v\rangle\big|^2 = \|v\|^2 > 0.$$

Now, let G be a countable amenable group acting on M. The action of G is naturally lifted to $M \otimes \overline{M}$ by $T(u \otimes \bar{v}) = Tu \otimes \overline{Tv}$, $T \in G$. Let $v \in \overline{M}^\infty$, then for $w = u \otimes v' \in M^\infty \otimes \overline{M}^\infty$, $T \in G$ we have

$$T(w * v) = T\big(\langle v', v\rangle u\big) = T\langle v', v\rangle Tu = \langle Tv', Tv\rangle Tu = Tw * Tv.$$

It follows that the equality $T(w * v) = Tw * Tv$ holds for all $w \in M \otimes \overline{M}$.

Let us assume now that G is not weakly mixing on M, let $u, v \in M$ and $c > 0$ be such that the set $\Gamma = \{T \in G : \mathcal{N}(Tu, v) > c\}$ is not of zero density in G: for some Følner sequence $\Phi_1, \Phi_2, \ldots$ in G and some $e > 0$, $\dfrac{\#(\Gamma \cap \Phi_k)}{\#\Phi_k} > e$ for all $k \in \mathbb{N}$. Slightly changing u and v, we may assume that $u, v \in M^\infty$. Then, for all $k \in \mathbb{N}$,

$$\begin{aligned} \frac{1}{\#\Phi_k} \sum_{T \in \Phi_k} \langle T(u \otimes \bar{u}), v \otimes \bar{v}\rangle_Y &= \frac{1}{\#\Phi_k} \sum_{T \in \Phi_k} \int \big|\langle Tu, v\rangle_y\big|^2 d\nu \\ &\geq \frac{1}{\#\Phi_k} \sum_{T \in \Phi_k} \mathcal{N}(Tu, v)^2 > c^2 e > 0, \end{aligned}$$

and so, $\lim_{k\to\infty} \sum_{T \in \Phi_k} T(u \otimes \bar{u}) \neq 0$ in the Hilbert space $M \otimes \overline{M}$ with the inner product $\langle\,,\rangle_Y$. It follows that the subspace of G-invariant elements in $M \otimes \overline{M}$ is nontrivial; let

$w \in M \otimes \overline{M}$, $w \neq 0$, be such that $Tw = w$ for all $T \in G$. Choose $v \in \overline{M}^{\infty}$ for which $u = w * v \neq 0$.

We claim that G acts compactly on u. Indeed, for $\varepsilon > 0$, let $w' = \sum_{k=1}^{l} u_k \otimes v_k \in M^{\infty} \otimes \overline{M}^{\infty}$ be such that $\|w - w'\|_Y < \varepsilon/(2\,\text{ess-sup}\,\|v\|)$. Then for any $T \in G$, by (8.3)

$$\big\|Tu - w' * Tv\big\|_Y = \big\|Tw * Tv - w' * Tv\big\|_Y = \big\|w * Tv - w' * Tv\big\|_Y \leq \|w - w'\|_Y \cdot \text{ess-sup}\,\|v\| < \frac{\varepsilon}{2}.$$

Let $C \in \mathbb{R}$ be such that $\big|\langle v_k, v\rangle\big| < C$, $k = 1, \ldots, l$, and let E be a finite $\frac{\varepsilon}{2}$-net for the disc $|z| < C$ in $\mathbb{C}$. Since $w' * Tv = \sum_{k=1}^{l} \langle v_k, Tv\rangle u_k$, the set $\Big\{\sum_{k=1}^{l} c_k u_k : c_k \in E,\ k = 1, \ldots, l\Big\}$ is an $\frac{\varepsilon}{2}$-Y-net for $w' * Tv$ and so, by Lemma 6.19, an ε-Y-net for Tu. ∎

8.7. Corollary. *Let G be a countable amenable group acting on M. Then $M = M^w(G) \oplus M^c(G)$. In particular, if T is an invertible transformation of M, then $M = M^w(T) \oplus M^c(T)$.*

Proof. By ("the amenable" version of) Proposition 7.15, $M^w(G) \perp M^c(G)$. But every G-invariant closed subspace of M which properly contains $M^w(G)$ has a nonempty intersection with $M^c(G)$. ∎

8.8. Remark. It is seen from the proof of Theorem 8.4 that, for an amenable group G, G is weakly mixing on M if and only if G is ergodic on the Hilbert space $M \otimes \overline{M}$. Usually this is used as the definition of weak mixing (which definition is suited for actions of any, not necessarily amenable group). It follows that the complementary to $M^w(G)$ subspace $M^c(G)$ is exactly the subspace of M on which G has *relatively discrete spectrum*, that is the space spanned by finite-dimensional G-invariant subspaces of M (see [Z1] and [Z2]).

8.9. In fact, some strengthening of Theorem 8.4 holds true, it can be obtained if one replaces the notion of Y-precompactness for a stronger one. Let $\varepsilon > 0$. Following [F2], we say that a set $V \subseteq L^2(X)$ *is ε-spanning for $U \subseteq M$ on $A \in \mathfrak{D}$* if for every $u \in U$ and almost all $y \in A$ there exists $v \in V$ with $\|u - v\|_y < \varepsilon$.

Let Q be a set of transformations of M. A vector $u \in M$ is called *almost periodic (with respect to Q)* if the set Qu possesses a finite ε-spanning set on Y for every $\varepsilon > 0$. Denote by $M^s(Q)$ *the closure* of the subspace of M consisting of vectors almost periodic with respect to Q. Clearly, $M^s(Q) \subseteq M^c(Q)$.

8.10. Lemma. *$u \in M^s(Q)$ if and only if for every $\delta > 0$ there is $B \in \mathfrak{D}$ with $\nu(B) > 1 - \delta$ such that $1_B u$ is almost periodic.*

Proof. Clearly, if $u \in M$ satisfies the condition above, it belongs to $M^s(Q)$. Now let u be in $M^s(Q)$, and let $\delta > 0$ be given. For every $k = 1, 2, \ldots$ pick an almost periodic $v \in M$ with $\|u - v\| < \dfrac{\delta}{k2^k}$. Let $B_k = \big\{y \in Y : \|u - v\|_y \leq 1/k\big\}$, then $\nu(B_k) > 1 - \delta/2^k$. Let V be a $\frac{1}{k}$-spanning set for Qv on Y, then $V \cup \{0\}$ is a $\frac{2}{k}$-spanning set for $Q(1_{B_k} u)$ on Y. It follows that for $B = \bigcup_{k=1}^{\infty} B_k$, $\nu(B) > 1 - \delta$, the vector $1_B u$ is almost periodic. ∎

8.11. Theorem. *Let G be a countable amenable group of transformations of M. If G is not weakly mixing on M, then there is $u \in M$ almost periodic with respect to G.*

Proof. The proof extends the proof of Theorem 8.4. Let $w_1, w_2, \ldots \in M^\infty \otimes \overline{M}^\infty$ be a sequence converging to a G-invariant element $w \in M \otimes \overline{M}$ pointwise (see 6.5). Choose a subset $A \in \mathfrak{D}$ with $\nu(A) > 0$ such that $1_A w \neq 0$ and $w_1, w_2, \ldots$ converge to w uniformly on A (that is, the sequence $\|w_1 - w\|, \|w_2 - w\|, \ldots$ converges to 0 uniformly on A). Find $v \in \overline{M}^\infty$ for which $u = 1_A w * v \neq 0$.

We claim that u is almost periodic with respect to G. Let $\varepsilon > 0$ be given, let $w_j = \sum_{k=1}^{l} u_k \otimes v_k$ be such that $\|w - w_j\| < \varepsilon/(2\,\text{ess-sup}\,\|v\|)$ on A. Then, for every $T \in G$, by (8.2)

$$\big\|Tu - w_j * Tv\big\|_y = \big\|w * Tv - w_j * Tv\big\|_y \leq \|w - w_j\|_y \cdot \text{ess-sup}\,\|v\| < \frac{\varepsilon}{2}$$

for almost all $y \in A$. Again, let $C \in \mathbb{R}$ be such that $\big|\langle v_k, v\rangle\big| < C$, $k = 1, \ldots, l$, and let E be a finite $\frac{\varepsilon}{2}$-net for $\{z \in \mathbb{C} : |z| < C\}$. Since $w' * Tv = \sum_{k=1}^{l} \langle v_k, Tv\rangle u_k$, the set $V = \Big\{\sum_{k=1}^{l} c_k u_k : c_k \in E,\ k = 1, \ldots, l\Big\}$ is, in fact, an $\frac{\varepsilon}{2}$-spanning set for $w_j * Gv = \{w_j * Tv : T \in G\}$ on Y and so, ε-spanning for Gu on A.

We will change V in order to obtain an ε-spanning set for Gu on whole Y. Since for $T, S \in G$, $v \in V$ and $y \in Y$, $\|STu - Sv\|_{yS^{-1}} = \|Tu - v\|_y$, the set SV is ε-spanning for Gu on AS^{-1}. Count the elements of G: $G = \{S_1 = \mathbf{1}_G, S_2, \ldots\}$, put $A_k = AS_k^{-1} \setminus \bigcup_{i=1}^{k-1} A_i$, and, for $v \in V$, define $\tilde{v} = \sum_{k=1}^{\infty} 1_{A_k} S_k v$ (it exists by Lemma 6.6). Then the set $\tilde{V} = \{\tilde{v} : v \in V\}$ is ε-spanning for Gu on $B = \bigcup_{k=1}^{\infty} A_k = \bigcup_{T \in G} AT^{-1}$, and, since $1_{Y \setminus B} Tu = 1_{Y \setminus B} 1_{AT^{-1}} T(w * v) = 0$ for all $T \in G$, the set $\tilde{V} \cup \{0\}$ is ε-spanning for Gu on Y. ■

8.12. Corollary. *If G is a countable amenable group acting on M, then $M^s(G) = M^c(G)$.*

8.13. And Lemma 8.10 gives

Corollary. *Let G be a countable amenable group acting compactly on M. Then for any $u \in M$ and any $\delta > 0$ there exists $B \in \mathfrak{D}$ with $\nu(B) > 1 - \delta$ such that $G(1_B u)$ is almost periodic.*

9. Primitive action of a nilpotent group on a Y-Hilbert space

We fix a finitely generated torsion-free nilpotent group G acting on a separable Y-Hilbert space M. We also fix an ordered basis in G, and thus the weights of G-polynomials and systems are assumed to be defined.

9.1. We begin with a special case of the action of G.

Proposition. *Let H be a normal subgroup of G such that H acts compactly on M and every $T \notin H$ is weakly mixing on M. Then every $g \in \wp G$ with $gg(0)^{-1} \notin \wp H$ is weakly mixing on M.*

Proof. We will prove this by induction on the weight of g. First of all, replacing g by $gg(0)^{-1}$, we may assume that $g(0) = \mathbf{1}_G$. It follows from Corollary 7.11 that G/H is torsion-free, and so, H is complete in G.

If $D^m g(n)(D^m g(0))^{-1} \notin \wp H(n,m)$, then, by Proposition 3.6, for almost all $m \in \mathbb{Z}$ we have $D^m g(n)(D^m g(0))^{-1} \notin \wp H(n)$ and we may assume by the induction hypothesis that $D^m g(n)$ is weakly mixing on M (see Lemma 3.13). Lemma 7.5 says that g is weakly mixing on M in this case.

Let now $D^m g(n) g(m)^{-1} \in \wp H(n,m)$. Since H is normal we have $g(m)^{-1} g(n)^{-1} g(n+m) \in \wp H(n,m)$. Denote $S = g(1)$. Then $h(n) = S^{-n} g(n) \in H$ for any $n \in \mathbb{Z}$. Hence by Proposition 3.6, $h \in \wp H$. Since $g \notin \wp H$, so $S \notin H$. Hence, S is weakly mixing on M. Since h acts compactly on M, by Proposition 7.14 $g(n) = S^n h(n)$ is weakly mixing on M. ■

9.2. We return to the general case now. The following proposition describes a way to find a subspace of M with a "primitive" action of G on it. Recall that $N(H)$ denotes the normalizer of H in G.

Proposition. *G contains a subgroup H satisfying:*
0) $M^c(H)$ is nontrivial,
1) every $T \in N(H) \setminus H$ is weakly mixing on $M^c(H)$,
2) for every $T \notin N(H)$, $TM^c(H) \perp M^c(H)$.

Proof. Let $\{\mathbf{1}_G\} = G_0 \subseteq G_1 \subseteq \ldots \subseteq G_l = G$ be a central series. We use induction on $k = 0, \ldots, l$ to find a subgroup $H_k \subseteq G_k$ satisfying:
0) $M^c(H_k)$ is nontrivial,
1) every $T \in (N(H_k) \cap G_k) \setminus H_k$ is weakly mixing on $M^c(H_k)$,
2) for every $T \notin N(H_k)$, $TM^c(H_k) \perp M^c(H_k)$.
Then we put $H = H_l$.

Let $H_0 = \{\mathbf{1}_G\}$, then conditions 0)–2) are trivially satisfied for $k = 0$. Assume that H_k has been already found.

The group $N(H_k) \cap G_{k+1}$ preserves $M^c(H_k)$. Choose in this group a maximal subgroup $H_{k+1} \supseteq H_k$ such that $M^c(H_{k+1}) \subseteq M^c(H_k)$ is nontrivial.

1) For $T \notin N(H_k)$, $TM^c(H_k) \perp M^c(H_k)$, so $TM^c(H_{k+1}) \perp M^c(H_{k+1})$. But for any $T \in N(H_{k+1})$, $TM^c(H_{k+1}) = M^c(TH_{k+1}T^{-1}) = M^c(H_{k+1})$. Hence, $N(H_{k+1})$ preserves $M^c(H_{k+1})$ and $N(H_{k+1}) \subseteq N(H_k)$.

In particular, $N(H_{k+1}) \cap G_{k+1} \subseteq N(H_k) \cap G_{k+1}$. But every element of $\big(N(H_{k+1}) \cap N(H_k) \cap G_{k+1}\big) \setminus H_{k+1}$ is weakly mixing on $M^c(H_{k+1})$: if it were not so, we could add this element to H_{k+1} by Corollary 8.7 and Corollary 7.12.

2) Let $T \notin N(H_{k+1})$. We have to prove that $M^c(H_{k+1}) \perp TM^c(H_{k+1})$. If $T \notin N(H_k)$, then even $TM^c(H_k) \perp M^c(H_k)$. So, let $T \in N(H_k)$.

Let $P \in H_{k+1}$ be such that $TPT^{-1} \notin H_{k+1}$. Define $g(n) = (TPT^{-1})^n P^{-n} \in \wp(N(H_k) \cap G_k)$, then $g(0) = \mathbf{1}_G$. Since $g(1) = TPT^{-1}P^{-1} \notin H_{k+1} \supseteq H_k$, $g \notin \wp H_k$. So, we can use Proposition 9.1 to see that g is weakly mixing on $M^c(H_k)$. But then $g(n)P^n = (TPT^{-1})^n$ is weakly mixing on $M^c(H_{k+1})$ by Proposition 7.14, and at the same time acts compactly on $TM^c(H_{k+1})$ by Lemma 7.13. By Proposition 7.15, $M^c(H_{k+1}) \perp$

$TM^c(H_{k+1})$. ■

9.3. Definition. *An action of G on M is primitive (or G acts on M primitively) if a subgroup $H \subseteq G$ and a subspace $M(H) \subseteq M$ exist such that:*
1) H preserves $M(H)$ and acts compactly on $M(H)$,
2) $N(H)$ preserves $M(H)$ and every $T \in N(H) \setminus H$ is weakly mixing on $M(H)$,
3) $TM(H) \perp M(H)$ for every $T \notin N(H)$,
4) $M = \overline{\bigoplus_{T \in G/N(H)} TM(H)}$, where $G/N(H)$ denotes the set of left cosets of $N(H)$ in G.

9.4. The main structure theorem about an action of a nilpotent group G on a Y-Hilbert space M is the following:

Theorem. *M is decomposable into a direct sum of G-invariant subspaces on each of which the action of G is primitive.*

Proof. It is enough to point at a G-invariant subspace of M on which G acts primitively.

Choose $H \subseteq G$ satisfying the conclusion of Proposition 9.2 and put $M(H) = M^c(H)$, $M' = \overline{\sum_{T \in G} TM(H)}$. Then M' is a nontrivial G-invariant Y-Hilbert space, and G acts primitively on it. ■

9.5. From now on let G act on M primitively.

Denote the orbit of H under the conjugation action of G by $\mathcal{H}$:

$$\mathcal{H} = \{THT^{-1},\ T \in G\}.$$

$\mathcal{H}$ is in one-to-one correspondence with the set $G/N(H)$ of left cosets of $N(H)$ in G, and with the orbit of $M(H)$ under the action of T: $THT^{-1} \leftrightarrow TN(H) \leftrightarrow TM(H)$, $T \in G$.

For $F \in \mathcal{H}$, $F = THT^{-1}$, denote $M(F) = TM(H)$. The action of G on M remains primitive if we change $H \mapsto F$, $M(H) \mapsto M(F)$.

9.6. The action of G on $\mathcal{H}$ may have cycles; we want now to reduce G slightly in order to remove them. Then every element outside of $N(H)$ will act on vectors of $M(H)$ like a coordinate shift and, in particular, be weakly mixing on $M(H)$.

By Proposition 2.10, G contains a normal subgroup G^* of finite index such that $N(H) \cap G^*$ is complete in G^*.

Lemma. *Let $g \in \wp G^*$, $g \notin \wp N(H)$. Then $g(n)M(H) \perp M(H)$ for almost all $n \in \mathbb{Z}$.*

Proof. Proposition 3.6 applied to $N(H) \cap G^*$ says that $g(n) \notin N(H)$ for almost all $n \in \mathbb{Z}$. ■

9.7. Proposition. *Let $g \in \wp_0 G^*$ and $g \notin \wp H$. Then g is weakly mixing on $M(H)$.*

Proof. If $g \notin \wp N(H)$, Lemma 9.6 even gives more than we need. Otherwise, g preserves $M(H)$ and we may apply Proposition 9.1 to the space $M(H)$ and the normal subgroup H of $N(H)$. ■

9.8. Corollary. *Let $g \in \wp G^*$ be such that $gg(0)^{-1} \notin \wp F$ for all $F \in \mathcal{H}$. Then g is weakly mixing on M.*

Proof. Every $F \in \mathcal{H}$ is complete in G^* as well, so, by Proposition 9.7 applied to F, $gg(0)^{-1}$ is weakly mixing on $M(F)$. Hence, $gg(0)^{-1}$ and so, g itself are weakly mixing on M. ■

9.9. Remark. It follows from Corollary 7.11 that $N(H)/H$ is torsion-free and so, H is complete in G^*. Choose a basis E of G^* over H. Every $g \in \wp G^*$ whose senior generator is from E satisfy the condition of Corollary 9.8 and so, is weakly mixing on M. This follows from the fact that the conjugation action of G preserves the senior generator of every G-polynomial.

9.10. The following theorem demonstrates that nilpotent groups are sometimes not worse than commutative those.

Theorem. *Let $\mathcal{G}$ be a nilpotent group acting on M and let $u \in M$. Then elements of $\mathcal{G}$ acting compactly on u form a group.*

Proof. Let $T_1, T_2 \in \mathcal{G}$ act compactly on u, let $G = \langle T_1, T_2 \rangle$. We may assume that G is torsion-free and that its action on M is primitive. Let H, $\mathcal{H}$ and G^* be as above, and let $c \in \mathbb{N}$ be such that $T_1^c, T_2^c \in G^*$. Decompose $u = \sum_{F \in \mathcal{H}} u_F$, $u_F \in M(F)$. Since T_1^c and T_2^c act compactly on u, it must be that $T_1^c, T_2^c \in F$ for every $F \in \mathcal{H}$ with $u_F \neq 0$. But then the group $G' = \langle T_1^c, T_2^c \rangle \subseteq F$ for each such F and so G' acts compactly on u. Since G' is of finite index in G, G acts compactly on u by Corollary 7.10. ■

10. Weak mixing of G-polynomials of several variables

The technical statements obtained in the end of Section 9 are not enough for our further aims. The goal of this section is to establish generalizations of Proposition 9.7, Corollary 9.8 and, mostly, of Proposition 9.1 for G-polynomials from $\wp^d G$ with $d \geq 2$.

We preserve all notation of Section 9.

10.1. Proposition. *Let $d \in \mathbb{N}$, let $g \in \wp_0^d G^*$ and $g \notin \wp^d H$. Then g is weakly mixing on $M(H)$.*

10.2. Corollary. *Let $g \in \wp^d G^*$ be such that $gg(0)^{-1} \notin \wp^d F$ for all $F \in \mathcal{H}$. Then g is weakly mixing on M.*

10.3. The proofs of Proposition 10.1 and Corollary 10.2 copy the proofs of Proposition 9.7 and Corollary 9.8 respectively; however, instead of Proposition 9.1, we use the following its multiparameter version:

Proposition. *Let H be a normal subgroup of G such that H acts compactly on M and every $T \notin H$ is weakly mixing on M. Then any $g \in \wp^d G$ with $gg(0)^{-1} \notin \wp^d H$ is weakly mixing on M.*

Proof. Replacing g by $gg(0)^{-1}$, we may assume that $g(0) = \mathbf{1}_G$. It follows from Corollary 7.11 that G/H is torsion-free, and so, H is complete in G.

If $D^m g(n)\big(D^m g(0)\big)^{-1} \notin \wp^d H(n,m)$, then, by Proposition 3.6, for almost all $m \in \mathbb{Z}$ we have $D^m g(n)\big(D^m g(0)\big)^{-1} \notin \wp^d H(n)$ and we may assume by induction on the weight of g that $D^m g(n)$ is weakly mixing on M (see Lemma 3.13). Lemma 7.5 says that g is weakly mixing on M in this case.

Now let $D^m g(n) g(m)^{-1} \in \wp^d H(n,m)$. Since H is normal in G, we have

$$g(m)^{-1} g(n)^{-1} g(n+m) \in \wp^d H(n,m). \tag{10.1}$$

Denote $S_1 = g(1,0,\ldots,0)$, $S_2 = g(0,1,0,\ldots,0)$, $\ldots$, $S_d = g(0,\ldots,0,1)$ and define $G' = \langle S_1,\ldots,S_d,H\rangle$. It follows from (10.1) that $[S_i,S_j] \in H$, $1 \le i,j \le d$, that is G'/H is commutative. Furthermore, $h(n) = \big(S_d^{n_d}\ldots S_1^{n_1}\big)^{-1} g(n) \in H$ for any $n = (n_1,\ldots,n_d) \in \mathbb{Z}^d$. By Proposition 3.6, $h \in \wp^d H$. By Proposition 7.14, it is only to prove that the G-polynomial $s(n) = S_d^{n_d}\ldots S_1^{n_1} \in \wp^d G$ is weakly mixing on M.

Since $g \notin \wp^d H$, not all of $S_1,\ldots,S_d$ are in H. We may assume that $(S_1,\ldots,S_d)$ is a basis of G'/H.

Assume that $s(n)$ is not weakly mixing on M. We will prove in this assumption that G' is not weakly mixing on M. If this is the case, Theorem 8.4 says that G' acts compactly on some $w \in M$, and this contradicts the fact that one of $S_1,\ldots,S_d$ is weakly mixing on M.

10.4. For $\Lambda \subseteq \mathbb{N}$ denote $\underline{d}(\Lambda) = \liminf\limits_{N\to\infty} \frac{1}{N}\#\Big(\Lambda \cap \{1,\ldots,N\}\Big)$.

Lemma. *Let $u \in M$ and let R act compactly on u. Then for any $a,e > 0$ there exist $b,c > 0$ such that for any $v \in M^\infty$ satisfying* ess-sup $\|v\|_y \le e$ *and $\mathcal{N}(u,v) > a$, one has $\underline{d}\big(\{n \in \mathbb{N} : \mathcal{N}(R^n u,v) > b\}\big) > c$.*

Proof. Changing u slightly, we may assume that $u \in M^\infty$; put $e' = \text{ess-sup}\,\|u\|_y$.

Define $B_1 = \Big\{y \in Y : \big|\langle u,v\rangle_y\big| > \frac{a}{2}\Big\}$, then $\nu(B_1) \ge \frac{a}{2ee'}$. Put $\varepsilon = a/4e$. Now, if for some $y \in B_1$ and $u' \in M$ one has $\|u'-u\|_y < \varepsilon$, then

$$\big|\langle u',v\rangle_y\big| \ge \big|\langle u,v\rangle_y\big| - \varepsilon e > \frac{a}{4}. \tag{10.2}$$

Using Corollary 8.13, find $B_2 \subseteq B_1$ with $\nu(B_2) > a/4ee'$ such that $1_{B_2}u$ is almost periodic with respect to $\{R^n : n \in \mathbb{Z}\}$. Let $\{w_1,\ldots,w_{Q-1}\}$ be an $\frac{\varepsilon}{2}$-spanning set for $\big\{R^n(1_{B_2}u),\ n \in \mathbb{Z}\big\}$ on Y.

It follows from the "classical" measurable recurrence theorem, Theorem M, that there exist $b',c' > 0$ depending only on Q and $\nu(B_2)$ such that for

$$\Lambda' = \Big\{m \in \mathbb{N} : \nu\Big(\bigcap_{q=1}^{Q} B_2 R^{-qm}\Big) > b'\Big\}$$

one has $\underline{d}(\Lambda') > c'$ (see also subsection 14.7).

Fix $m \in \Lambda'$ and put $B_3 = \bigcap_{q=1}^{Q} B_2 R^{-qm}$, then $\nu(B_3) > b'$. For every $y \in B_3$ and $q = 1, \ldots, Q$,

$$\left\| R^{qm} u - w_r \right\|_y = \left\| 1_{B_2 R^{-qm}} R^{qm} u - w_r \right\|_y = \left\| R^{qm} (1_{B_2} u) - w_r \right\|_y < \frac{\varepsilon}{2}$$

for some r. Define a mapping $\chi_y \colon \{1, \ldots, Q\} \longrightarrow \{1, \ldots, Q-1\}$ by the rule

$$\chi_y(q) = r \quad \text{if} \quad \left\| R^{qm} u - w_r \right\|_y < \frac{\varepsilon}{2}.$$

For any $y \in B_3$, there exist $1 \leq q_1(y) < q_2(y) \leq Q$ such that $\chi_y(q_1(y)) = \chi_y(q_2(y)) = r$ for some $1 \leq r \leq Q-1$, that is

$$\left\| R^{q_1(y)m} u - w_r \right\|_y < \frac{\varepsilon}{2}, \quad \left\| R^{q_2(y)m} u - w_r \right\|_y < \frac{\varepsilon}{2},$$

and consequently,

$$\left\| R^{q_2(y)m} u - R^{q_1(y)m} u \right\|_y = \left\| R^{q_1(y)m} (R^{(q_2(y)-q_1(y))m} u - u) \right\|_y < \varepsilon.$$

Find $B_4 \subseteq B_3$, $\nu(B_4) > b'/Q^2$, such that $q_1(y) = q_1$, $q_2(y) = q_2$ are constant on B_4.

Put $B = B_4 R^{q_1 m}$, $q(m) = q_2 - q_1$. Then $\nu(B) > b'/Q^2$ and for any $y \in B$ we have $\left\| R^{q(m)m} u - u \right\|_y < \varepsilon$ and therefore, by (10.2), $\left| \langle R^{q(m)m} u, v \rangle_y \right| > a/4$.

Hence, for any $m \in \Lambda'$ there exists $1 \leq q(m) \leq Q$ such that $\mathcal{N}(R^{q(m)m} u, v) > \dfrac{a}{4} \dfrac{b'}{Q^2}$. It is also clear that for $\Lambda = \left\{ q(m)m,\ m \in \Lambda' \right\}$ we have $\underline{d}(\Lambda) > \dfrac{c'}{Q^2}$. ■

10.5. Let us return to the proof of Proposition 10.3.

Assume that for $u, v \in M$ and $a > 0$ the set

$$\Gamma \left\{ (n_1, \ldots, n_d) \in \mathbb{Z}^d : \mathcal{N}\big(S_d^{n_d} \ldots S_1^{n_1} u, v\big) > a \right\}$$

is not of zero density in $\mathbb{Z}^d$, that is there exist $\alpha > 0$ and a sequence of parallelepipeds $\Pi_1, \Pi_2, \ldots \subset \mathbb{Z}^d$ with $L(\Pi_k) \underset{k \to \infty}{\longrightarrow} \infty$ such that $\#(\Gamma \cap \Pi_k)/\#\Pi_k > \alpha$, $k \in \mathbb{N}$. Let $(R_1, \ldots, R_r)$ be an ordered basis of H, then $(R_1, \ldots, R_r, S_1, \ldots, S_d)$ is an ordered basis of G'. Changing a and v slightly, we may assume that $v \in M^\infty$. Put $e = \text{ess-sup} \|v\|_y$.

Applying Lemma 10.4 to u, R_r, a, e, find $b_r, c_r > 0$ such that for every $(n_1, \ldots, n_d) \in \Gamma$,

$$\underline{d} \left\{ m_r \in \mathbb{N} : \mathcal{N}\big(R_r^{m_r} u, (S_d^{n_d} \ldots S_1^{n_1} R_r^{m_r})^{-1} v\big) > b_r \right\} > c_r.$$

Choose a sequence $q_{r,k} \in \mathbb{N}$, $k \in \mathbb{N}$, such that

$$\frac{1}{q_{r,k}} \# \left\{ 1 \leq m_r \leq q_{r,k} : \mathcal{N}\big(S_d^{n_d} \ldots S_1^{n_1} R_r^{m_r} u, v\big) > b_r \right\} > c_r$$

for every $(n_1,\dots,n_d)\in\Pi_k$, $k=1,2,\dots$, and $q_{r,k}/\max_{n\in\Pi_k}\|n\|^p \underset{k\to\infty}{\longrightarrow}\infty$ for every $p\in\mathbb{N}$. Define $\Pi_{r,k}=\{1,\dots,q_{r,k}\}\times\Pi_k\subset\mathbb{Z}^{d+1}$, $k=1,2,\dots$, and

$$\Gamma_r=\Big\{(m_r,n_1,\dots,n_d)\in\mathbb{Z}^{d+1}:\mathcal{N}\big(S_d^{n_d}\dots S_1^{n_1}R_r^{m_r}u,v\big)>b_r\Big\}.$$

Then $\#(\Gamma_r\cap\Pi_{r,k})/\#\Pi_{r,k}>\alpha c_r$ for every $k=1,2,\dots$.

Now apply Lemma 10.4 to u,R_{r-1},b_r,e to find $b_{r-1},c_{r-1}>0$ such that for every $(m_r,n_1,\dots,n_d)\in\Gamma_r$,

$$\underline{d}\Big\{m_{r-1}\in\mathbb{N}:\mathcal{N}\big(R_{r-1}^{m_{r-1}}u,(S_d^{n_d}\dots S_1^{n_1})^{-1}v\big)>b_{r-1}\Big\}>c_{r-1}.$$

Choose a sequence $q_{r-1,k}\in\mathbb{N}$, $k\in\mathbb{N}$, such that

$$\frac{1}{q_{r-1,k}}\#\Big\{1\le m_{r-1}\le q_{r-1,k}:\mathcal{N}\big(S_d^{n_d}\dots S_1^{n_1}R_r^{m_r}R_{r-1}^{m_{r-1}}u,v\big)>b_{r-1}\Big\}>c_{r-1}$$

for every $(m_r,n_1,\dots,n_d)\in\Pi_{r,k}$, $k=1,2,\dots$, and $q_{r-1,k}/q_{r,k}^p\underset{k\to\infty}{\longrightarrow}\infty$ for every $p\in\mathbb{N}$. Define $\Pi_{r-1,k}=\{1,\dots,q_{r-1,k}\}\times\Pi_{r,k}\subset\mathbb{Z}^{d+2}$, $k=1,2,\dots$, and

$$\Gamma_{r-1}=\Big\{(m_{r-1},m_r,n_1,\dots,n_d)\in\mathbb{Z}^{d+2}:\mathcal{N}\big(S_d^{n_d}\dots S_1^{n_1}R_r^{m_r}R_{r-1}^{m_{r-1}}u,v\big)>b_{r-1}\Big\}.$$

Then $\#(\Gamma_{r-1}\cap\Pi_{r-1,k})/\#\Pi_{r-1,k}>\alpha c_r c_{r-1}$ for every $k=1,2,\dots$.

Continue this process and find $b_l,c_l\in\mathbb{R}$, $q_{l,k}\in\mathbb{N}$, $l=r-2,\dots,1$, $k=1,2,\dots$, such that $q_{l,k}/q_{l-1,k}^p\underset{k\to\infty}{\longrightarrow}\infty$ for every $p\in\mathbb{N}$, and for

$$\begin{aligned}&\Gamma_1=\Big\{(m_1,\dots,m_r,n_1,\dots,n_d)\in\mathbb{Z}^{d+r}:\mathcal{N}\big(S_d^{n_d}\dots S_1^{n_1}R_r^{m_r}\dots R_1^{m_1}u,v\big)>b_1\Big\},\\&\Pi_{1,k}=\{1,\dots,q_{1,k}\}\times\dots\times\{1,\dots,q_{r,k}\}\times\Pi_k,\ k=1,2,\dots,\end{aligned}$$

we have

$$\frac{\#(\Gamma_1\cap\Pi_{1,k})}{\#\Pi_{1,k}}>\alpha c_r\dots c_1,\ k=1,2,\dots. \tag{10.3}$$

Define $\Phi_k=\Big\{S_d^{n_d}\dots S_1^{n_1}R_r^{m_r}\dots R_1^{m_1}:(m_1,\dots,m_r,n_1,\dots,n_d)\in\Pi_{1,k}\Big\}$, $k=1,2,\dots$. Then $\Phi_1,\Phi_2,\dots$ is a Følner sequence in G' (see subsection 8.2), and it follows from (10.3) that G' is not weakly mixing on u. ■

11. Primitive action of a nilpotent group on an extension

We will pass now from an abstract Y-Hilbert space to an extension of a measure space Y, provided with an action of a nilpotent group G. Our purpose is to find a nontrivial G-invariant factor of such an extension on which the induced action of G is "primitive". We also establish a "multi-weak mixing" relation, generalizing Corollary 9.8 (see Theorem 11.15). All this can be considered as a generalization of the "bilinear" propositions of Section 9 for the "multilinear" case.

As before, let G be a finitely generated torsion-free nilpotent group, let an ordered basis of G be fixed and so, the weights of G-polynomials and systems be defined.

11.1. Through this section $\alpha\colon \big(X,\mathfrak{B},\mu,G\big) \longrightarrow \big(Y,\mathfrak{D},\nu,G\big)$ will be a fixed *extension*, that is a mapping $\alpha\colon X \longrightarrow Y$ of measure spaces $\big(X,\mathfrak{B},\mu\big)$ and $\big(Y,\mathfrak{D},\nu\big)$ with $\mu(X)=\nu(Y)=1$, satisfying $\alpha^{-1}(B)\in\mathfrak{B}$ and $\mu\big(\alpha^{-1}(B)\big)=\nu(B)$ for $B\in\mathfrak{D}$, and commuting with a measure preserving action of G on X and Y. We will say that X is *an extension* (of Y). We will follow the convention that G acts on points of X and Y from the right; then G acts on the set of functions on X and Y from the left by $(Tf)(x)=f(xT)$. Under the assumption that $\big(X,\mathfrak{B},\mu\big)$ is regular (see [F2]), we have a decomposition of the measure μ, that is an (almost everywhere) uniquely defined system of measures μ_y, $y\in Y$, on $\big(X,\mathfrak{B}\big)$ satisfying
a) $\int u d\mu = \int\big(\int u d\mu_y\big)d\nu$ for $u\in L^1(X)$,
b) $\int \varphi u d\mu_y = \varphi(y)\int u d\mu_y$ a.e. for $u\in L^1(X)$ and $\varphi\in L^\infty(Y)$,
c) $\int Tu d\mu_y = \int u d\mu_{yT}$ a.e. for $u\in L^1(X)$ and $T\in G$.
This decomposition turns $L^2(X)$ into a Y-Hilbert space by

$$\langle u,v\rangle_y = \int u\bar{v}d\mu_y, \quad u,v\in L^2(X),\ y\in Y,$$

with G acting on it. We will denote this Y-Hilbert space by $\mathcal{X}$, and assume $\mathcal{X}$ to be separable.

11.2. A sub-σ-algebra $\mathfrak{C}$ of σ-algebra $\mathfrak{B}$ defines *a factor* $X'=(X,\mathfrak{C},\mu)$ of the measure space X. If $\mathfrak{C}$ contains $\alpha^{-1}(\mathfrak{D})$, then X' is *a factor over* Y. All factors of X which we deal with will be over Y and we will not mention this specifically. A factor *is nontrivial* if it is distinct from Y. We will identify a factor and the corresponding measure space. The space $L^1(X')$ of measurable functions on a factor X' of X is a subspace of $L^1(X)$.

G acts on the set of factors of X: if X' is the factor corresponding to a sub-σ-algebra $\mathfrak{C}$ of $\mathfrak{B}$, then TX', $T\in G$, corresponds to the sub-σ-algebra $\Big\{AT^{-1} : A\in\mathfrak{C}\Big\}$. A G-invariant factor X' of X with the induced action of G on X' is *a factor of the measure preserving system* (X,G). If, additionally, X' is a factor over Y, we will say that X' is *a subextension of* X (over Y).

Given a system of factors $\{X_\xi\}_{\xi\in\Xi}$ of X, one can define *the product* $\prod_{\xi\in\Xi}X_\xi$ as the factor corresponding to the σ-algebra of subsets of X generated by the σ-algebras corresponding to X_ξ, $\xi\in\Xi$. We will say that the system of factors is *relatively independent* (over Y) if for any pairwise distinct $\xi_1,\dots,\xi_k\in\Xi$ and any $u_i\in L^\infty(X_{\xi_i})$, $i=1,\dots,k$, one has $\int\prod_{i=1}^k u_i d\mu_y = \prod_{i=1}^k\int u_i d\mu_y$ for almost all $y\in Y$. When this is the case, we write $\coprod_{\xi\in\Xi}X_\xi$ instead of $\prod_{\xi\in\Xi}X_\xi$.

11.3. We do not distinguish between $L^2(Y)$ and its image $\alpha^*\big(L^2(Y)\big)$ in $\mathcal{X}$. The dense subspace $L^\infty(X) \subset \mathcal{X}$ of (essentially) bounded functions will be denoted by $\mathcal{X}^\infty$. For every subset U of $\mathcal{X}$ we denote $U^\infty = U \cap \mathcal{X}^\infty$. We also denote the Y-Hilbert space $\mathcal{X} \ominus L^1(Y)$ of functions orthogonal to "the constants" by M: $u \in M$ if and only if $\int u d\mu_y = 0$ for almost all $y \in Y$.

We say that a sequence $g(n) \in G$, $n \in \mathbb{Z}^d$, *is weakly mixing on* X if $g(n)$ is weakly mixing on M.

Given a set $Q \subseteq G$, the subspace $\mathcal{X}^c(Q)$ of functions from $\mathcal{X}$ on which Q acts compactly contains $L^2(Y)$, is closed under truncations and complex conjugations of its elements, and $\mathcal{X}^c(Q) \cap L^\infty(X)$ is a subalgebra of $L^\infty(X)$. Hence, $\mathcal{X}^c(Q)$ defines a factor of X (see [F2]); we denote it by $X^c(Q)$: $L^2(X^c(Q)) = \mathcal{X}^c(Q)$. If $X = X^c(Q)$, we say that Q *acts compactly on* X.

11.4. We begin with technical lemmas.

A set $U \subset \mathcal{X}$ is *uniformly bounded* if there exists $a \in \mathbb{R}$ such that for all $u \in U$, ess-sup $|u(x)| < a$.

Lemma. *Let $U_1, U_2 \subset \mathcal{X}^\infty$ be Y-precompact and uniformly bounded. Then $U_1 U_2 = \big\{u_1 u_2,\ u_1 \in U_1,\ u_2 \in U_2\big\}$ is Y-precompact.*

Proof. Let ess-sup $|u_1| < a$, ess-sup $|u_2| < a$ for all $u_1 \in U_1$ and $u_2 \in U_2$. Then first of all, ess-sup $|u_1 u_2| < a^2$ for $u_1 \in U_1$ and $u_2 \in U_2$.

Let $V_1 \subseteq U_1$ be a finite $\frac{\varepsilon}{a}$-Y-net for U_1 and let $V_2 \subseteq U_2$ be a finite $\frac{\varepsilon}{a}$-Y-net for U_2. Since for any $u_1 \in U_1$, $u_2 \in U_2$, $v_1 \in V_1$, $v_2 \in V_2$,

$$\|u_1 u_2 - v_1 v_2\|_y \leq a\big(\|u_1 - v_1\|_y + \|u_2 - v_2\|_y\big) \text{ a.e.,}$$

we have

$$\int \min_{\substack{v_1 \in V_1 \\ v_2 \in V_2}} \|u_1 u_2 - v_1 v_2\|_y^2 d\nu \leq 4\varepsilon^2.$$

It follows that $V_1 V_2$ is a finite 2ε-Y-net for $U_1 U_2$. ■

11.5. The following lemma is a "multilinear" version of Lemma 6.20.

Lemma. *Let $U_1, \ldots, U_k$ be Y-precompact uniformly bounded subsets of $\mathcal{X}^\infty$. For any $\varepsilon > 0$ there exists a set of vectors $v_{i,j} \in U_i$, $i = 1, \ldots, k$, $j = 1, \ldots, s_i$, such that for any $g_1, \ldots, g_k \in G$ and any vectors $u_i \in U_i$, $i = 1, \ldots, k$, one has*

$$\int \Big| \int \prod g_i u_i d\mu_y \Big| d\nu < \sum_{j_1=1}^{s_1} \cdots \sum_{j_k=1}^{s_k} \int \Big| \int \prod g_i v_{i,j_i} d\mu_y \Big| d\nu + \varepsilon.$$

Proof. Let ess-sup $|u| < a$ for all $u \in U$. Choose an $\frac{\varepsilon}{ka^{k-1}}$-$Y$-net $\{v_{1,1}, \ldots, v_{1,s_1}\} \subseteq U_1$ for U_1, then choose an $\frac{\varepsilon}{ka^{k-1}s_1}$-$Y$-net $\{v_{2,1}, \ldots, v_{2,s_2}\} \subseteq U_2$ for U_2, …, and an $\frac{\varepsilon}{ka^{k-1}s_1 \ldots s_{k-1}}$-$Y$-net $\{v_{k,1}, \ldots, v_{k,s_k}\} \subseteq U_k$ for U_k (see Lemma 6.23). Then $\big\{g_i v_{i,1}, \ldots, g_i v_{i,s_1}\big\}$ is an

$\frac{\varepsilon}{ka^{k-1}s_1\dots s_{i-1}}$-$Y$-net for g_iU_i, $i=1,\dots,k$. By Lemma 6.20,

$$\begin{aligned}\int\Big|\int\prod g_iu_id\mu_y\Big|d\nu&=\mathcal{N}\Big(g_1\bar{u}_1,\prod_{i=2}^k g_iu_i\Big)\\&<\sum_{j_1=1}^{s_1}\mathcal{N}\Big(g_1\bar{v}_{1,j_1},\prod_{i=2}^k g_iu_i\Big)+\frac{\varepsilon}{ka^{k-1}}\Big\|\prod_{i=2}^k g_iu_i\Big\|\\<\sum_{j_1=1}^{s_1}\int\Big|\int g_1v_{1,j_1}\prod_{i=2}^k g_iu_id\nu\Big|d\nu+\frac{\varepsilon}{k}&=\sum_{j_1=1}^{s_1}\mathcal{N}\Big(g_2\bar{u}_2,g_1v_{1,j_1}\prod_{i=3}^k g_iu_i\Big)+\frac{\varepsilon}{k}<\dots\\<\sum_{j_1=1}^{s_1}\dots\sum_{j_k=1}^{s_k}\int\Big|\int\prod_{i=1}^k g_iv_{i,j_i}d\mu_y\Big|d\nu&+\frac{\varepsilon}{k}+s_1\frac{\varepsilon}{ks_1}+\dots+(s_1\dots s_{k-1})\frac{\varepsilon}{ks_1\dots s_{k-1}}.\end{aligned}$$

■

11.6. Lemma. *Let $M_1,\dots,M_k\subseteq\mathcal{X}$ be subspaces of $\mathcal{X}$ and let $g_1(n),\dots,g_k(n)\in G$, $n\in\mathbb{Z}^d$, be sequences in G satisfying*

$$\int\prod_{i=1}^k g_i(n)u_id\mu_y-\prod_{i=1}^k\int g_i(n)u_id\mu_y\xrightarrow{D}0\ \textit{in}\ L^1(Y)$$

for all $u_i\in M_i^\infty$, $i=1,\dots,k$. Then for any Y-precompact uniformly bounded sequences $u_i(n)\in M_i^\infty$, $n\in\mathbb{Z}^d$, $i=1,\dots,k$,

$$\int\prod_{i=1}^k g_i(n)u_i(n)d\mu_y-\prod_{i=1}^k\int g_i(n)u_i(n)d\mu_y\xrightarrow{D}0\ \textit{in}\ L^1(Y).\tag{11.1}$$

Proof. By multilinearity of (11.1) we may replace $u_i(n)$ by $u_i(n)-\int u_i(n)d\mu_y$, $i=1,\dots,k$, and assume that $M_i\subseteq M$, $i=1,\dots,k$. We have to prove under this assumption that

$$\int\Big|\int\prod_{i=1}^k g_i(n)u_i(n)d\mu_y\Big|d\nu\xrightarrow{D}0.$$

But in this form it is a corollary of Lemma 11.5. ■

11.7. The following proposition deals with an extension of a special type.

Proposition. *Let a normal subgroup $H\subseteq G$ have the property that H acts compactly on X and every $T\in G\setminus H$ is weakly mixing on X. Let $g_1,\dots,g_k\in\wp G$ be such that $g_ig_i(0)^{-1}\notin\wp H$ and $\big(g_ig_i(0)^{-1}\big)^{-1}g_jg_j(0)^{-1}\notin\wp H$ for $i,j=1,\dots,k$, $i\neq j$. Then for any $u_1,\dots,u_k\in\mathcal{X}^\infty$, $v\in\mathcal{X}^\infty$,*

$$\int\prod_{i=1}^k g_i(n)u_ivd\mu_y-\prod_{i=1}^k\int g_i(n)u_id\mu_y\int vd\mu_y\xrightarrow{D}0\ \textit{in}\ L^1(Y).\tag{11.2}$$

Proof. We will prove this using PET-induction, applied to the system of G-polynomials $\mathcal{A} = \{g_1, \ldots, g_k\}$. Since $g_i g_i(0)^{-1} \notin \wp H$, $i = 1, \ldots, k$, $\mathcal{A}$ does not contain constant polynomials.

Replacing u_i by $g_i(0)u_i$, we may assume that $g_i(0) = \mathbf{1}_G$, $i = 1, \ldots, k$. By multilinearity of (11.2), we may also assume that $u_i \in M^\infty$, $i = 1, \ldots, k$, and prove

$$\mathcal{N}\big(\prod_{i=1}^{k} g_i(n)u_i, \bar{v}\big) \xrightarrow{D} 0.$$

By Lemma 6.18 it follows from

$$\operatorname*{D-lim}_m \operatorname*{D-lim\,sup}_n \mathcal{N}\Big(\prod_{i=1}^{k} g_i(n)u_i, \prod_{i=1}^{k} g_i(n+m)u_i\Big) = 0,$$

that is

$$\operatorname*{D-lim}_m \operatorname*{D-lim\,sup}_n \int\Big|\int \prod_{i=1}^{k} g_i(n)u_i g_i(n+m)\bar{u}_i d\mu_y\Big| d\nu = 0.$$

It follows from Corollary 7.11 that G/H is torsion-free, and thus, H is complete in G. Note the following:

a) $$g_i(n)^{-1} g_j(n) \notin \wp H(n) \text{ for } i, j = 1, \ldots, k,\ i \neq j; \tag{11.3}$$

b) $\big(g_i(n+m)g_i(m)^{-1}\big)^{-1} g_j(n+m)g_j(m)^{-1} \notin \wp H(n,m)$ for $i, j = 1, \ldots, k,\ i \neq j$, as it is so for $m = 0$. Hence by Proposition 3.6, for almost all $m \in \mathbb{Z}$

$$\big(g_i(n+m)g_i(m)^{-1}\big)^{-1} g_j(n+m)g_j(m)^{-1} \notin \wp H(n),\ i, j = 1, \ldots, k,\ i \neq j; \tag{11.4}$$

c) By the same reason, for almost all $m \in \mathbb{Z}$

$$g_i(n)^{-1} g_j(n+m)g_j(m)^{-1} \notin \wp H(n),\ i, j = 1, \ldots, k,\ i \neq j; \tag{11.5}$$

d) We assume that $g_i(n)^{-1} g_i(n+m)g_i(m)^{-1} \notin \wp H(n,m)$ for $i = 1, \ldots, r$, and that $g_i(n)^{-1} g_i(n+m)g_i(m)^{-1} \in \wp H(n,m)$ for $i = r+1, \ldots, k$. Then, for almost all $m \in \mathbb{Z}$,

$$g_i(n)^{-1} g_i(n+m)g_i(m)^{-1} \notin \wp H(n),\ i = 1, \ldots, r. \tag{11.6}$$

Since H is normal, we also have

$$h_i(n,m) = g_i(m)^{-1} g_i(n)^{-1} g_i(n+m) \in \wp H(n,m),\ i = r+1, \ldots, k,$$

and so, h_i, $i = r+1, \ldots, k$, acts compactly on M.

By Lemma 11.5, for any $\varepsilon>0$ there exist $v_{i,j}\in M^\infty$, $i=r+1,\dots,k$, $j=1,\dots,s$, such that for any $n,m\in\mathbb{Z}$

$$\begin{aligned}&\int\Big|\int\prod_{i=1}^k g_i(n)u_ig_i(n+m)\bar{u}_id\mu_y\Big|d\nu\\ &=\int\Big|\int\prod_{i=1}^r g_i(n)u_ig_i(n+m)\bar{u}_i\prod_{i=r+1}^k g_i(n)u_ig_i(n)g_i(m)h_i(n,m)\bar{u}_id\mu_y\Big|d\nu\\ &<\sum_{j_1,\dots,j_k=1}^s\int\Big|\int\prod_{i=1}^r g_i(n)u_ig_i(n+m)\bar{u}_i\prod_{i=r+1}^k g_i(n)u_ig_i(n)g_i(m)v_{i,j_i}d\mu_y\Big|d\nu+\varepsilon.\end{aligned}$$

Thus, it is enough to prove that for any $v_i\in M^\infty$, $i=r+1,\dots,k$,

$$\operatorname*{D-lim}_m\operatorname*{D-lim\,sup}_n\int\Big|\int\prod_{i=1}^r g_i(n)u_i\prod_{i=1}^r g_i(n+m)\bar{u}_i\prod_{i=r+1}^k g_i(n)\big(u_ig_i(m)v_id\mu_y\big)\Big|d\nu=0.$$

For $m\in\mathbb{Z}$, denote

$$g_{i,m}(n)=\begin{cases}g_i(n),\ 1\le i\le k\\ g_{i-k}(n+m),\ k+1\le i\le k+r\end{cases},\quad u_{i,m}=\begin{cases}u_i,\ 1\le i\le r\\ u_ig_i(m)v_i,\ r+1\le i\le k\\ \bar{u}_{i-k},\ k+1\le i\le k+r\end{cases}.$$

We have to prove that

$$\operatorname*{D-lim}_m\operatorname*{D-lim\,sup}_n\int\Big|\int\prod_{i=1}^{k+r}g_{i,m}(n)u_{i,m}d\mu_y\Big|d\nu=0.$$

Let $g_{t,m}$ be a G-polynomial of the minimal weight in the system $\{g_{i,m},\ i=1,\dots,k+r\}$. After a rearrangement, we may assume that $t=1$. Since $\mathcal{A}=\{g_1,\dots,g_k\}$ does not contain constant polynomials, $g_{t,m}$ is not constant and so, the system $\mathcal{A}_m=\{g_{1,m}^{-1}g_{i,m},\ i=2,\dots,k+r\}$ precedes $\mathcal{A}$ (see Lemma 4.5). Since by (11.3) – (11.6) $\mathcal{A}_m$ satisfy the condition of the proposition for almost all $m\in\mathbb{Z}$, we may apply a PET-induction hypothesis and assume that

$$\operatorname*{D-lim}_n\Big(\int u_{1,m}\prod_{i=2}^{k+r}g_{1,m}^{-1}(n)g_{i,m}(n)u_{i,m}d\mu_y-\int u_{1,m}d\mu_y\prod_{i=2}^{k+r}\int g_{1,m}^{-1}(n)g_{i,m}(n)u_{i,m}d\mu_y\Big)=0$$

in $L^1(Y)$ for almost all $m\in\mathbb{Z}$. It follows that

$$\operatorname*{D-lim}_n\Big(\int\prod_{i=1}^{k+r}g_{i,m}(n)u_{i,m}d\mu_y-\prod_{i=1}^{k+r}\int g_{i,m}(n)u_{i,m}d\mu_y\Big)=0$$

in $L^1(Y)$ for almost all $m \in \mathbb{Z}$. Hence, it is enough to prove that

$$\operatorname*{D\text{-}lim}_m \operatorname*{D\text{-}lim\,sup}_n \int \Big| \prod_{i=1}^{k+r} \int g_{i,m}(n) u_{i,m} d\mu_y \Big| d\nu = 0.$$

Let $a > 1$ be such that $\text{ess-sup}\,|u_i| < a$, $i = 1, \ldots, k$, and $\text{ess-sup}\,|v_i| < a$, $i = r+1, \ldots, k$. Then, by the definition of $u_{i,m}$, $\left|\prod_{i=2}^{k+r} \int g_{i,m} u_{i,m} d\mu_y\right| < a^{2k}$ for almost all $y \in Y$, $i = 1, \ldots, k+r$, $m \in \mathbb{Z}$. So,

$$\int \Big| \prod_{i=1}^{k+r} \int g_{i,m}(n) u_{i,m} d\mu_y \Big| d\nu < a^{2k} \int \Big| \int g_{1,m}(n) u_{1,m} d\mu_y \Big| d\nu = a^{2k} \int \Big| \int u_{1,m} d\mu_y \Big| d\nu$$

Remind that either $u_{1,m} = u_1$ or $u_{1,m} = u_1 g_1(m) v_1$. In the first case, $\int u_{1,m} d\mu_y = 0$ in $L^1(Y)$, in the second case

$$\operatorname*{D\text{-}lim}_m \int \Big| \int u_{1,m} d\mu_y \Big| d\nu = \operatorname*{D\text{-}lim}_m \mathcal{N}\big(u_1, g_1(m) v_1\big) = 0$$

by Proposition 9.1. ■

11.8. Corollary. *Assume that a normal subgroup H of G acts compactly on X and that every $T \in G \setminus H$ is weakly mixing on X.*

Let $g_1, \ldots, g_k \in \wp G$ be such that $\big(g_i g_i(0)^{-1}\big)^{-1} g_j g_j(0)^{-1} \notin \wp H$ for $i \neq j = 1, \ldots, k$. Then for any Y-precompact uniformly bounded sequences $u_i(n) \in U_i$, $n \in \mathbb{Z}$, $i = 1, \ldots, k$,

$$\int \prod_{i=1}^{k} g_i(n) u_i(n) d\mu_y - \prod_{i=1}^{k} \int g_i(n) u_i(n) d\mu_y \xrightarrow{D} 0 \text{ in } L^1(Y).$$

Proof. Applying Lemma 11.6, we reduce the problem to the case where sequences $u_i(n)$, $i = 1, \ldots, k$, are constant. It is enough then to apply Proposition 11.7 to $g_i' = g_1^{-1} g_i$, $i = 2, \ldots, k$, ■

11.9. The following proposition describes a procedure of building a "primitive" subextension.

Proposition. *G contains a subgroup H such that:*
0) $X^c(H)$ is nontrivial,
1) every $T \in N(H) \setminus H$ is weakly mixing on $X^c(H)$,
2) the factors $TX^c(H)$ for T running through the set $G/N(H)$ of left cosets of $N(H)$ in G are relatively independent: for any $T_1, \ldots, T_t \in G$ with $T_i^{-1} T_j \notin N(H)$ for $1 \leq i < j \leq t$ and any $u_i \in T_i X^c(H)^\infty$,

$$\int \prod_{i=1}^{t} u_i d\mu_y = \prod_{i=1}^{t} \int u_i d\mu_y \text{ in } L^1(Y).$$

Proof. Let $\{\mathbf{1}_G\} = G_0 \subseteq G_1 \subseteq \ldots \subseteq G_l = G$ be a central series in G. We will use induction on $k = 0, \ldots, l$ to find a subgroup $H_k \subseteq G_k$ such that:
0) $X^c(H_k)$ is nontrivial,
1) every $T \in (N(H_k) \cap G_k) \setminus H_k$ is weakly mixing on $X^c(H_k)$,
2) factors $TX^c(H_k)$ for T running through the set $G/N(H_k)$ are relatively independent.
Then we put $H = H_l$.

Put $H_0 = \{\mathbf{1}_G\}$, then conditions 0)–2) are trivially satisfied for $k = 0$. Assume that H_k has been already found.

The group $N(H_k) \cap G_{k+1}$ preserves $X^c(H_k)$. Choose in this group a maximal subgroup $H_{k+1} \supseteq H_k$ such that the factor $X^c(H_{k+1})$ of $X^c(H_k)$ is nontrivial. Denote $M^c(H_k) = L^2(X^c(H_k)) \ominus L^2(Y)$, $M^c(H_{k+1}) = L^2(X^c(H_{k+1})) \ominus L^2(Y)$.

1) For $T \notin N(H_k)$, $TM^c(H_k) \perp M^c(H_k)$, thus $TM^c(H_{k+1}) \perp M^c(H_{k+1})$. But for any $T \in N(H_{k+1})$, $TM^c(H_{k+1}) = M^c(TH_{k+1}T^{-1}) = M^c(H_{k+1})$. Hence, $N(H_{k+1})$ preserves $M^c(H_{k+1})$, and $N(H_{k+1}) \subseteq N(H_k)$.

In particular, $N(H_{k+1}) \cap G_{k+1} \subseteq N(H_k) \cap G_{k+1}$. But every element of $\Big(N(H_{k+1}) \cap N(H_k) \cap G_{k+1}\Big) \setminus H_{k+1}$ is weakly mixing on $M^c(H_{k+1})$: if it were not so, we could add this element to H_{k+1} by Corollary 8.7 and Corollary 7.12.

2) Let $T_1, \ldots, T_t \in G$ satisfy $T_i^{-1}T_j \notin N(H_{k+1})$ for $1 \leq i < j \leq t$, let $u_i \in T_i X^c(H_{k+1})^\infty$, $i = 1, \ldots, t$. We have to prove that

$$\int \prod_{i=1}^{t} u_i d\mu_y = \prod_{i=1}^{t} \int u_i d\mu_y \text{ in } L^1(Y). \tag{11.7}$$

We will do it by induction on t.

a) If there are $1 \leq i, j \leq t$ such that $T_i^{-1}T_j \notin N(H_k)$, then by the induction hypothesis on k the set $\{u_i,\ i = 1, \ldots, t\}$ is nontrivially subdivided into subsets $\{u_{l,i},\ i = 1, \ldots, t_l\}$, $l = 1, \ldots, s$, belonging to (L^2 of) distinct independent spaces. So,

$$\int \prod_{i=1}^{t} u_i d\mu_y = \prod_{l=1}^{s} \int \prod_{i=1}^{t_l} u_{l,i} d\mu_y \text{ in } L^1(Y)$$

with $t_1 + \ldots + t_s = t$, and induction on t finishes the proof.

b) Now let all T_i be in the same left coset of $N(H_k)$ in G. We may assume that $T_i \in N(H_k)$, $i = 1, \ldots, t$, and consequently, $T_i M(H_k) \subseteq M(H_k)$. By multilinearity of (11.7) we may assume that $u_i \in T_i M^c(H_{k+1})^\infty$, $i = 1, \ldots, t$, and prove under this assumption that $\int \prod_{i=1}^{t} u_i d\mu_y = 0$ in $L^1(Y)$.

We may and will also assume that $T_1 = \mathbf{1}_G$. It is given that $T_i H_{k+1} T_i^{-1}$, $i = 1, \ldots, t$, are all pairwise distinct; choose $P \in H_{k+1}$ such that $T_2 P T_2^{-1} \notin H_{k+1}$. Then, for any $n \in \mathbb{Z}$,

$$\begin{aligned} \int \Big| \int \prod_{i=1}^{t} u_i d\mu_y \Big| d\nu &= \int \Big| \int P^{-n} \prod_{i=1}^{t} u_i d\mu_y \Big| d\nu \\ &= \int \Big| \int \prod_{i=1}^{t} P^{-n} (T_i P T_i^{-1})^n (T_i P T_i^{-1})^{-n} u_i d\mu_y \Big| d\nu. \end{aligned} \tag{11.8}$$

We are going to prove that for any $\varepsilon > 0$ there exists $n \in \mathbb{Z}$ for which the expression in (11.8) is smaller than ε.

Let $u_i(n) = (T_i P T_i^{-1})^{-n} u_i$, $i = 1, \ldots, t$, $n \in \mathbb{Z}$. Since $u_i \in T_i M^c(H_{k+1})$, these sequences are Y-precompact. Define $g_i(n) = P^{-n}\big(T_i P T_i^{-1}\big)^n$, $i = 1, \ldots, t$, $n \in \mathbb{Z}$. Then $g_i \in \wp\big(N(H_k) \cap G_k\big)$. Subdivide the set of indices $\{1, \ldots, t\} = I_1 \cup \ldots \cup I_s$ in such a way that $g_i^{-1} g_j \in \wp H_k$ if and only if $i, j \in I_q$ for some $1 \le q \le s$. Since $g_1(1)^{-1} g_2(1) = P^{-1} T_2 P T_2^{-1} \notin H_{k+1} \supseteq H_k$, this partition is not trivial.

For each $q = 1, \ldots, s$ choose $i_q \in I_q$. The sequences $w_q(n) = \prod_{i \in I_q} g_{i_q}(n)^{-1} g_i(n) u_i(n)$, $q = 1, \ldots, s$, are Y-precompact and uniformly bounded, so we may apply Corollary 11.8 to obtain

$$\int \Big| \int \prod_{q=1}^{s} g_{i_q}(n) w_q(n) d\mu_y \Big| d\nu - \int \Big| \prod_{q=1}^{s} \int g_{i_q}(n) w_q(n) d\mu_y \Big| d\nu \xrightarrow{D} 0.$$

Since

$$g_{i_q}(n) w_q(n) = \prod_{i \in I_q} P^{-n} (T_i P T_i^{-1})^n (T_i P T_i^{-1})^{-n} u_i = P^{-n} \prod_{i \in I_q} u_i,$$

it is enough to have $\int | \int \prod_{i \in I_q} u_i d\mu_y | d\nu = 0$. But this is given by the induction hypothesis on t. ■

11.10. Definition. *An action of G on X is primitive (or G acts on X primitively) if a subgroup $H \subseteq G$ and a factor $X(H)$ exist such that:*
1) H preserves $X(H)$ and acts compactly on $X(H)$,
2) $N(H)$ preserves $X(H)$ and every $T \in N(H) \setminus H$ is weakly mixing on $X(H)$,
3) the factors $TX(H)$ for T passing on the set $G/N(H)$ of left cosets of $N(H)$ in G are relatively independent,
4) and $X = \coprod_{T \in G/N(H)} TX(H)$.

11.11. The main structure theorem is the following:

Theorem. *There exists a nontrivial subextension X' of X such that the action of G on X' is primitive.*

Proof. Choose $H \subseteq G$ satisfying the conclusion of Proposition 9.2 and put $X(H) = X^c(H)$, $X' = \coprod_{T \in G} TX(H)$. Then X' is nontrivial and G-invariant, and G acts primitively on it. ■

11.12. From now on let G act on X primitively.

Denote by $\mathcal{H}$ the orbit of H under the conjugation action of G:

$$\mathcal{H} = \{THT^{-1},\ T \in G\}.$$

$\mathcal{H}$ is in one-to-one correspondence with the set $G/N(H)$ of left cosets of $N(H)$ in G, and with the orbit of $X(H)$ under the action of T: $THT^{-1} \leftrightarrow TN(H) \leftrightarrow TX(H)$, $T \in G$.

For $F \in \mathcal{H}$, $F = THT^{-1}$, denote $X(F) = TX(H)$; the action of G on X remains primitive if we change $H \mapsto F$, $X(H) \mapsto X(F)$. We also denote $\mathcal{X}(F) = L^2(X(F))$, $M(F) = \mathcal{X}(F) \ominus L^2(Y)$, $M(\emptyset) = L^2(Y)$ and, for $\mathcal{Q} \subseteq \mathcal{H}$ with $\#\mathcal{Q} < \infty$, $M(\mathcal{Q}) = \overline{\operatorname{Span} \prod_{F \in \mathcal{Q}} M(F)^\infty}$.

11.13. Proposition. *The action of G on the Y-Hilbert space $\overline{\sum_{F\in\mathcal{H}} M(F)}$ is primitive. $M(\mathcal{Q}_1)\perp M(\mathcal{Q}_2)$ for any $\mathcal{Q}_1\neq\mathcal{Q}_2\subseteq\mathcal{H}$ with $\#\mathcal{Q}_1,\#\mathcal{Q}_2<\infty$, and $\mathcal{X}=\bigoplus_{\substack{\mathcal{Q}\subseteq\mathcal{H}\\ \#\mathcal{Q}<\infty}} M(\mathcal{Q})$.*

Proof. The first statement is clear by definition. It is also evident that spaces $M(\mathcal{Q})$ with $\mathcal{Q}\subseteq\mathcal{H}$, $\#\mathcal{Q}<\infty$, span $\mathcal{X}$. We have to check the orthogonality of distinct $M(\mathcal{Q})$ only. Let $\mathcal{Q}_1=\{F_{1,1},\ldots,F_{1,k_1}\}$, $\mathcal{Q}_2=\{F_{2,1},\ldots,F_{2,k_2}\}$, and let $F_{1,1}\notin\mathcal{Q}_2$. Then for any $u_{i,j}\in M(F_{i,j})^\infty$, $i=1,2$, $j=1,\ldots,k_i$,

$$\int\prod_{j=1}^{k_1}u_{1,j}\prod_{j=1}^{k_2}u_{2,j}\,d\mu_y=\int u_{1,1}\,d\mu_y\int\prod_{j=2}^{k_1}u_{1,j}\prod_{j=1}^{k_2}u_{2,j}\,d\mu_y=0\text{ in }L^1(Y).$$

■

11.14. Let G^* be a normal subgroup of G of finite index such that $N(H)\cap G^*$ is complete in G^*. Then, by Lemma 2.9, for any $F\in\mathcal{H}$ the subgroup $N(F)\cap G^*$ is complete in G^* as well.

The following proposition is the main "technical" result of this section.

Proposition. *Let $F_1,\ldots,F_k\in\mathcal{H}$ be pairwise distinct subgroups. Let $g_{j,i}\in\wp_0G^*$, $i=1,\ldots,k$, $j=1,\ldots,k_i$, be such that $g_{j_1,i}^{-1}g_{j_2,i}\notin\wp F_i$ for every $1\leq i\leq k$ and every $1\leq j_1<j_2\leq k_i$. Let $u_{j,i}(n)\in\mathcal{X}(F_i)^\infty$, $n\in\mathbb{Z}$, $i=1,\ldots,k$, $j=1,\ldots,k_i$, be Y-precompact uniformly bounded sequences. Then*

$$\int\prod_{i=1}^{k}\prod_{j=1}^{k_i}g_{j,i}(n)u_{j,i}(n)\,d\mu_y-\prod_{i=1}^{k}\prod_{j=1}^{k_i}\int g_{j,i}(n)u_{j,i}(n)\,d\mu_y\xrightarrow{D}0\textit{ in }L^1(Y).$$

Proof. For each $i=1,\ldots,k$, subdivide the set of indices $\{1,\ldots,k_i\}=J_1\cup\ldots\cup J_{L_i}$ in such a way that $g_{j_1,i}^{-1}g_{j_2,i}\in\wp N(F_i)$ if and only if j_1 and j_2 belong to the same J_l, $1\leq l\leq L_i$. By Corollary 3.7, applied to the subgroup $N(H)\cap G^*$ of G^*, $g_{j_1,i_1}(n)F_{i_1}$ and $g_{j_2,i_2}(n)F_{i_2}$ do not coincide for almost all $n\in\mathbb{Z}$ if either $i_1\neq i_2$, or $i_1=i_2=i$ but j_1 and j_2 are not in the same J_l, $1\leq l\leq L_i$. Since the factors $X(F)$, $F\in\mathcal{H}$, are relatively independent, we have

$$\int\prod_{i=1}^{k}\prod_{j=1}^{k_i}g_{j,i}(n)u_{j,i}(n)\,d\mu_y=\prod_{i=1}^{k}\prod_{l=1}^{L_i}\int\prod_{j\in J_l}g_{j,i}(n)u_{j,i}(n)\,d\mu_y\text{ in }L^1(Y)$$

for almost all $n\in\mathbb{Z}$. And, by Corollary 11.8,

$$\int\prod_{j\in J_l}g_{j,i}(n)u_{j,i}(n)\,d\mu_y-\prod_{j\in J_l}\int g_{j,i}(n)u_{j,i}(n)\,d\mu_y\xrightarrow{D}0\text{ in }L^1(Y)$$

for any $i=1,\ldots,k$ and $l=1,\ldots,L_i$. ■

11.15. Theorem. *Let $g_1,\dots,g_k \in \wp G^*$ be such that $\big(g_i g_i(0)^{-1}\big)^{-1} g_j g_j(0)^{-1} \notin \wp F$ for $i,j = 1,\dots,k$, $i \neq j$, and all $F \in \mathcal{H}$. Then for any Y-precompact uniformly bounded sequences $u_i(n) \in \mathcal{X}^\infty$, $n \in \mathbb{Z}$, $i = 1,\dots,k$,*

$$\int \prod_{i=1}^{k} g_i(n)u_i(n)d\mu_y - \prod_{i=1}^{k} \int g_i(n)u_i(n)d\mu_y \xrightarrow{D} 0 \text{ in } L^1(Y).$$

Proof. Replacing $u_i(n)$ by $g_i(0)u_i(n)$ we may assume that $g_i(0) = \mathbf{1}_G$, $i = 1,\dots,k$. Lemma 11.6 allows us to reduce the problem to the case in which the sequences $u_i(n)$ are constant, $u_i(n) = u_i$, $n \in \mathbb{Z}$. By Proposition 11.13 we may assume that $u_i \in M(\mathcal{Q}_i)$ for some $\mathcal{Q}_i \subseteq \mathcal{H}$ with $\#\mathcal{Q}_i < \infty$, $i = 1,\dots,k$. Corollary 3.7, applied to the complete subgroup $N(H) \cap G^*$ of G^* reduces the problem to the case $u_i \in M(F)^\infty$, $g_i \in \wp N(F)$, $i = 1,\dots,k$, for some $F \in \mathcal{H}$. And then Corollary 11.8 applied to the subgroup F finishes the proof. ■

11.16. Remark. It follows from Corollary 7.11 that $N(H)/H$ is torsion-free and so, H is complete in G^*. Choose a basis E of G^* over H. Now any pairwise distinct $g_1,\dots,g_k \in \wp E$ with $g(0) = \mathbf{1}_G$ satisfy the condition of Theorem 11.15. It simply follows from the fact that any conjugation in G preserves the senior generator of any G-polynomial.

12. NSZ-property

We pass to the proof of Theorem NM. The preceding part of the paper can be considered as preparatory for this concluding proof.

Let $(X,\mathfrak{B},\mu)$ be a measure space with $\mu(X) < \infty$, let G be a nilpotent group of right measure preserving transformations of X. Scaling μ, we may assume that $\mu(X) = 1$. Similarly to Theorem NT, the statement of Theorem NM deals with the part of G generated by a finite set $\{T_1,\dots,T_t\}$. Thus, we may and will assume that, as before, G is a finitely generated torsion-free nilpotent group.

12.1. We will say that a dynamical system $(Y,\mathfrak{D},\nu,G)$ *has the NSZ-property* if for every $A \in \mathfrak{D}$ with $\nu(A) > 0$, for every $d \in \mathbb{N}$ and for every system $\mathcal{A} \subset \wp_0^d G$ there exists $c > 0$ such that for every thick set $\Lambda \subseteq \mathbb{Z}^d$ there exists $n \in \Lambda$ for which

$$\nu\Big(\bigcap_{g\in\mathcal{A}} Ag(n)^{-1}\Big) > c.$$

12.2. Assume that $(X,\mathfrak{B},\mu,G)$ possesses the NSZ-property. In the notation of Theorem NM, the set of G-polynomials

$$\mathcal{A} = \Big\{g_i = T_t^{p_{i,t}} \dots T_1^{p_{i,1}},\ i = 1,\dots,I\Big\}$$

is a system in $\wp_0^d G$. We have consequently, for a suitable $c > 0$, that the set

$$S = \Big\{n \in \mathbb{Z}^d : \mu\Big(\bigcap_{i=1}^{I} Ag_i(n)^{-1}\Big) > c\Big\}$$

has a nonempty intersection with every thick subset in $\mathbb{Z}^d$. Hence, S is syndetic and the conclusion of Theorem NM holds in this case.

12.3. The set of G-invariant factors of X is naturally ordered by the relation "to be a factor": $(X,\mathfrak{B}_1,\mu) < (X,\mathfrak{B}_2,\mu)$ if $\mathfrak{B}_1 \subset \mathfrak{B}_2$. This is the case, the identity map id_X turns $(X,\mathfrak{B}_2,\mu,G)$ into an extension of $(X,\mathfrak{B}_1,\mu,G)$. We have:

Proposition. *The family of G-invariant factors of X possessing the NSZ-property has a maximal element.*

Proof. We copy the proof of Proposition 3.3 in [FK1]. In light of Zorn's lemma, it suffices to prove that if $\{\mathfrak{B}_s\}$ is a linearly ordered family of G-invariant sub-σ-algebras of $\mathfrak{B}$ such that $(X,\mathfrak{B}_s,\mu,G)$ has the NSZ-property for all s, then $(X,\bigcup\mathfrak{B}_s,\mu,G)$ has the NSZ-property.

Let a set $A \in \bigcup\mathfrak{B}_s$, $\mu(A)>0$, and a system $\mathcal{A} \subset \wp_0^d G$, $I=\#\mathcal{A}\geq 1$, be given. Find $A' \in \mathfrak{B}_s$ for some s with $\mu(A'\triangle A) < \dfrac{\mu(A)}{4I}$. Let μ_x, $x\in X$, be the decomposition of μ with respect to the factor $(X,\mathfrak{B}_s,\mu)$. Define $B=\Big\{x\in X : \mu_x(A)\geq 1-\dfrac{1}{2I}\Big\}$. Then $B\in\mathfrak{B}_s$, and B is of positive measure: since $\mu_x(A')=1$ for $x\in A'$, one would have

$$\mu(A'\setminus A) \geq \frac{1}{2I}\mu(A') > \frac{1}{2I}\cdot\frac{\mu(A)}{2} = \frac{\mu(A)}{4I}$$

otherwise.

Now, let $c>0$ and $n\in\mathbb{Z}^d$ be such that $\mu\Big(\bigcap_{g\in\mathcal{A}} Bg(n)^{-1}\Big) > c$. At every point $x\in\bigcap_{g\in\mathcal{A}} Bg(n)^{-1}$ we have

$$\mu_x\Big(\bigcap_{g\in\mathcal{A}} Ag(n)^{-1}\Big) \geq 1-\sum_{g\in\mathcal{A}}\big(1-\mu_x(Ag(n)^{-1})\big) = 1-\sum_{g\in\mathcal{A}}\big(1-\mu_{xg(n)}(A)\big) \geq 1-\frac{I}{2I}=\frac{1}{2}.$$

So, $\mu\left(\bigcap_{g\in\mathcal{A}} Ag(n)^{-1}\right) > c/2$ for such n. ■

12.4. It follows that, in order to prove Theorem NM, it suffices to establish the following fact: if Y is a proper G-invariant factor of X which has the NSZ-property, then there exists a nontrivial subextension of X (with respect to Y) having the NSZ-property too. Indeed, then the maximal element in the family of G-invariant factors of X having the NSZ-property must coincide with X itself.

We say that an extension $(X,\mathfrak{B},\mu,G)\longrightarrow(Y,\mathfrak{D},\nu,G)$ *is primitive* if the action of G on X as an extension of Y is primitive (see Definition 11.10). Since, by Theorem 11.11, X contains a nontrivial primitive subextension, it remains to prove the following proposition.

12.5. Proposition. *Let $(X,\mathfrak{B},\mu,G)\longrightarrow(Y,\mathfrak{D},\nu,G)$ be a primitive extension and let $(Y,\mathfrak{D},\nu,G)$ have the NSZ-property. Then $(X,\mathfrak{B},\mu,G)$ has the NSZ-property as well.*

12.6. We assume from now on that X is a primitive extension of Y. That is, a system $\mathcal{H}$ of conjugated subgroups of G and a system $X(H)=(X,\mathfrak{B}(H),\mu)$, $H\in\mathcal{H}$, of factors of X is fixed and is such that $X=\coprod_{H\in\mathcal{H}} X(H)$, and for every $H\in\mathcal{H}$ the following holds: H preserves $X(H)$, acts compactly on $X(H)$, and every $T\in N(H)\setminus H$ is weakly mixing on $X(H)$ (see subsection 11.10).

We also fix a normal subgroup G^* of finite index in G such that $N(H)\cap G^*$ is complete in G^* for every $H\in\mathcal{H}$ (see subsection 11.14).

12.7. The first ingredient of the proof of Proposition 12.5 is Theorem 5.9, the second one has to be the multiparameter version of Proposition 11.14:

Proposition. *Let $F_1,\ldots,F_K\in\mathcal{H}$ be pairwise distinct subgroups. Let $d\in\mathbb{N}$ and let $g_{l,k}\in\wp_0^d G^*$, $k=1,\ldots,K$, $l=1,\ldots,L_k$, be such that $g_{l_1,k}^{-1}g_{l_2,k}\notin\wp^d F_k$ for every $1\le k\le K$ and every $1\le l_1<l_2\le L_k$. Let $u_{l,k}(n)\in L^\infty\bigl(X(F_k)\bigr)$, $n\in\mathbb{Z}^d$, $k=1,\ldots,K$, $l=1,\ldots,L_k$, be Y-precompact uniformly bounded sequences. Then*

$$\int\prod_{k=1}^K\prod_{l=1}^{L_k}g_{l,k}(n)u_{l,k}(n)d\mu_y-\prod_{k=1}^K\prod_{l=1}^{L_k}\int g_{l,k}(n)u_{l,k}(n)d\mu_y\xrightarrow{D}0\ \textit{in}\ L^1(Y).$$

The proof of Proposition 12.7 is analogous to that of Proposition 11.14, the only difference is that the reference to Proposition 9.1 in Proposition 11.7 has to be replaced by the reference to its multiparameter version, Proposition 10.3.

13. Proof of the measure recurrence theorem

We continue the proof of Proposition 12.5 and preserve all notation of Section 12. Let $A\in\mathfrak{B}$ with $\mu(A)=4a>0$, $d\in\mathbb{N}$, a thick set $\Lambda\subseteq\mathbb{Z}^d$, and a system $\mathcal{A}=\{g_1,\ldots,g_I\}\subset\wp_0^d G$ be given.

13.1. First of all, let $N\in\mathbb{N}$ be such that, if we write $g_i'(n)=g_i(Nn)$, then $g_i'\in\wp_0^d G^*$, $i=1,\ldots,I$. Such N exists as $g_i(0)=\mathbf{1}_G$, $i=1,\ldots,I$, and G^* is of finite index in G. Since $\Lambda\cap(N\mathbb{Z}^d)$ is thick, we may pass to the subgroup $N\cdot\mathbb{Z}^d$ of $\mathbb{Z}^d$ and to the system $\mathcal{A}=\{g_1',\ldots,g_I'\}$, and identify G with G^*.

13.2. We need an inequality in our proof. It is a very cumbersome algebraic inequality of a "convex" type, and we have not been able to prove it by algebraic methods. We are forced to use an ergodic-theoretical trick.

Lemma. *For any $a>0$ and $I\in\mathbb{N}$ there exists $C(a,I)>0$ with the following property. Let $K,J\in\mathbb{N}$ and let $\Omega\subseteq\{1,\ldots,J\}^K$. Let a family $\{a_{j,k}\ge 0,\ k=1,\ldots,K,\ j=1,\ldots,J\}$ satisfy*

$$\sum_{j=1}^J a_{j,k}\le 1,\ k=1,\ldots,K,\quad\textit{and}\quad\sum_{(j_1,\ldots,j_K)\in\Omega}\prod_{k=1}^K a_{j_k,k}\ge a.\tag{13.1}$$

Let for each $k=1,\ldots,K$ the set $\{1,\ldots,I\}$ be partitioned into L_k subsets:

$$\{1,\ldots,I\}=\bigcup_{l=1}^{L_k}I_{l,k},\quad I_{l_1,k}\cap I_{l_2,k}=\emptyset,\ 1\le l_1\ne l_2\le L_k.\tag{13.2}$$

Denote by Θ the subset of Ω^I consisting of the I-tuples of elements of Ω which are constant on each $I_{l,k}$, $k=1,\ldots,K$, $l=1,\ldots,L_k$:

$$\Theta=\Bigl\{\theta=(j_k^i)\in\Omega^I: j_k^i=t_k^l(\theta)\ \textit{for all}\ i\in I_{l,k},\ k=1,\ldots,K,\ l=1,\ldots,L_k\Bigr\}.$$

Thus the collection $\{t_k^l(\theta),\ k=1,\dots,K,\ l=1,\dots,L_k\}$ *is well defined for all* $\theta\in\Theta$.

Then

$$\sum_{\theta\in\Theta}\prod_{k=1}^{K}\prod_{l=1}^{L_k}a_{t_k^l(\theta),k}\geq C(a,I). \tag{13.3}$$

Example. For $a_{1,1},a_{2,1},a_{1,2},a_{2,2}\geq 0$, $a_{1,1}+a_{2,1}\leq 1$, $a_{1,2}+a_{2,2}\leq 1$, and $I\in\mathbb{N}$ one has

$$a_{1,1}a_{1,2}^I+a_{2,1}a_{2,2}^I\geq C(a_{1,1}a_{1,2}+a_{2,1}a_{2,2},I)$$

(one can put $C(a,I)=a^I$ in this case). This inequality corresponds to the case $K=2$, $J=2$, $\Omega=\{(1,1),(2,2)\}$, $L_1=1$, $I_{1,1}=\{1,\dots,I\}$, $L_2=I$, $I_{l,2}=\{l\}$, $l=1,\dots,I$.

Proof. It follows from the "classical" measurable recurrence theorem, Theorem M, that for any $a>0$ and any $I\in\mathbb{N}$ there exists a constant $C(a,I)>0$ with the following property: for any family of pairwise commuting measure preserving transformations $T_1,\dots,T_I$ of a measure space $(Z,\mathfrak{C},\eta)$ with $\eta(Z)=1$ and for any $D\in\mathfrak{C}$ with $\eta(D)\geq a$ there exists arbitrarily big $n\in\mathbb{N}$ for which $\eta\Big(\bigcap_{i=1}^I DT_i^{-n}\Big)\geq C(a,I)$ (compare with subsection 14.7). This is just the constant we need.

Let $a>0$ and $I\in\mathbb{N}$ be given. We construct a dynamical system, which will give us (13.3). Denote by η the standard measure on $[0,1]$ and put $Z=[0,1]^K$. Let T be a measure preserving transformation of $[0,1]$ which is strong mixing of all orders; one can take for instance $T(x)=2x\,(\mathrm{mod}\,1)$.

Let $K,J\in\mathbb{N}$, $\Omega\in\{1,\dots,J\}^K$, numbers $a_{j,k}\geq 0$, $j=1,\dots,J$, $k=1,\dots,K$, satisfying (13.1), numbers $L_1,\dots,L_K$ and sets $I_{l,k}\subseteq\{1,\dots,I\}$, $k=1,\dots,K$, $l=1,\dots,L_k$, satisfying (13.2) be given. For each $k=1,\dots,K$, choose pairwise distinct intervals $D_{1,k},\dots,D_{L_k,k}\subset[0,1]$ with $\eta(D_{j,k})=a_{j,k}$, and put $D=\bigcup_{(j_1,\dots,j_K)\in\Omega}\prod_{k=1}^K D_{j_k,k}\subseteq Z$. Then by (13.1), $\eta^K(D)\geq a$.

For each $k=1,\dots,K$, denote $\lambda_k(i)=l$ if $i\in I_{l,k}$. Define $T_1,\dots,T_I\colon Z\longrightarrow Z$ by $T_i=(T^{\lambda_1(i)},\dots,T^{\lambda_K(i)})$, $i=1,\dots,I$. Since T is strong mixing of all orders, we have

$$\begin{aligned}\eta^K\Big(\bigcap_{i=1}^I DT_i^{-n}\Big)&=\sum_{(j_k^i)\in\Omega^I}\eta^K\Big(\bigcap_{i=1}^I(\prod_{k=1}^K D_{j_k^i,k})T_i^{-n}\Big)=\sum_{(j_k^i)\in\Omega^I}\prod_{k=1}^K\eta\Big(\bigcap_{i=1}^I D_{j_k^i,k}T^{-\lambda_k(i)n}\Big)\\&=\sum_{(j_k^i)\in\Omega^I}\prod_{k=1}^K\eta\Big(\bigcap_{l=1}^{L_k}(\bigcap_{i\in I_{l,k}}D_{j_k^i,k})T^{-ln}\Big)\underset{n}{\longrightarrow}\sum_{(j_k^i)\in\Omega^I}\prod_{k=1}^K\prod_{l=1}^{L_k}\eta(\bigcap_{i\in I_{l,k}}D_{j_k^i,k}).\end{aligned}$$

Note that $\bigcap_{i\in I_{l,k}}D_{j_k^i,k}=\emptyset$ if j_k^i for $i\in I_{l,k}$ do not all coincide. Hence

$$\eta^K\Big(\bigcap_{i=1}^I DT_i^{-n}\Big)\underset{n}{\longrightarrow}\sum_{\theta\in\Theta}\prod_{k=1}^K\prod_{l=1}^{L_k}\eta\big(D_{t_k^l(\theta),k}\big)=\sum_{\theta\in\Theta}\prod_{k=1}^K\prod_{l=1}^{L_k}a_{t_k^l(\theta),k}.$$

By the choice of $C(a,I)$ this implies (13.3). ∎

13.3. Put $C = C(a, I)$ and $\varepsilon = \dfrac{aC}{4I}$. It is clear that $C \leq a$ and thus $\varepsilon < a^2$ (anyway, we may assume this).

13.4. Find $K, J \in \mathbb{N}$, $H_1, \ldots, H_K \in \mathcal{H}$, $A_{j,k} \in \mathfrak{B}(H_k)$, $j = 1, \ldots, J$, $k = 1, \ldots, K$, and a subset $\Omega \subseteq \{1, \ldots, J\}^K$ such that:
a) for all $1 \leq k \leq K$, $A_{j_1,k} \cap A_{j_2,k} = \emptyset$ for $1 \leq j_1 \neq j_2 \leq J$, and $X = \bigcup_{j=1}^J A_{j,k}$;
b) the set $A' = \bigcup_{(j_1,\ldots,j_K)\in\Omega} \prod_{k=1}^K A_{j_k,k}$ satisfies $\mu(A \triangle A') < \varepsilon$.
Existence of such $K, J, H_k, A_{j,k}, \Omega$ easily follows from the fact that the σ-algebra $\mathfrak{B}$ is generated by $\{\mathfrak{B}(H),\ H \in \mathcal{H}\}$.

13.5. Define $D = \big\{y \in Y : \mu_y(A \triangle A') > \varepsilon/a\big\}$. Then $\nu(D) < a$ (it would be

$$\mu(A \triangle A') \geq \int_D \mu_y(A \triangle A') d\nu \geq \frac{\varepsilon}{a} a = \varepsilon$$

otherwise).

Put $B_1 = \big\{y \in Y : \mu_y(A) > 2a\big\}$. Then $\nu(B_1) \geq 2a$ (it would be

$$\mu(A) = \int_{B_1} \mu_y(A) d\nu + \int_{Y \setminus B_1} \mu_y(A) d\nu \leq \nu(B_1) + 2a\nu(Y \setminus B_1) < 4a$$

otherwise).

13.6. Put $B_2 = B_1 \setminus D$. Then $\nu(B_2) > a$ and for every $y \in B_2$ we have $\mu_y(A) > 2a$ and $\mu_y(A \triangle A') \leq \varepsilon/a$. Since $\varepsilon/a < a$ by 13.3, we also have $\mu_y(A') > a$ for $y \in B_2$.

13.7. Thus, if for some $y \in Y$ and $n \in \mathbb{Z}^d$ we have:
1) $yg_i(n) \in B_2$, $i = 1, \ldots, I$,
2) $\mu_y\Big(\bigcap_{i=1}^I A' g_i(n)\Big) > \dfrac{C}{2}$,
then

$$\mu_y\Big(\bigcap_{i=1}^{I} A g_i(n)\Big) > \frac{C}{2} - I\frac{\varepsilon}{a} = \frac{C}{2} - I\frac{1}{a}\frac{aC}{4I} = \frac{C}{4}.$$

13.8. Consider the inequality (13.3). Its left part is a continuous function of argument $(a_{j,k},\ k = 1, \ldots, K,\ j = 1, \ldots, J)$, running through the closed subset of the cube $[0,1]^{KJ}$ defined by (13.1). It is thus uniformly continuous on this subset and so, we can find $\delta > 0$ with the following property:
Let a collection $a_{j,k} \geq 0$, $j = 1, \ldots, J$, $k = 1, \ldots, K$, satisfy

$$\sum_{j=1}^{J} a_{j,k} \leq 1,\ k = 1, \ldots, K, \quad \textit{and} \quad \sum_{(j_1,\ldots,j_k)\in\Omega} \prod_{k=1}^{K} a_{j_k,k} \geq a. \tag{13.4}$$

Let $1 \leq L_1, \ldots, L_K \leq I$ and partitions $\{1, \ldots, I\} = \bigcup_{l=1}^{L_k} I_{l,k}$, $k = 1, \ldots, K$, be given. Define

$$\Theta = \Big\{\theta = (j_k^i) \in \Omega^I : j_k^i = t_k^l(\theta) \textit{ for all } i \in I_{l,k},\ k = 1, \ldots, K,\ l = 1, \ldots, L_k\Big\}. \tag{13.5}$$

Let a collection $a'_{j,k,l} \geq 0$, $j = 1,\dots,J$, $k = 1,\dots,K$, $l = 1,\dots,L_k$, *satisfy*

$$|a'_{j,k,l} - a_{j,k}| < \delta, \quad j = 1,\dots,J,\ k = 1,\dots,K,\ l = 1,\dots,L_k. \tag{13.6}$$

Then

$$\sum_{\theta\in\Theta} \prod_{k=1}^{K} \prod_{l=1}^{L_k} a'_{t_k^l(\theta),k,l} \geq \frac{3C}{4}. \tag{13.7}$$

We fix such a δ.

13.9. Define $u_{j,k} = 1_{A_{j,k}}$, $j = 1,\dots,J$, $k = 1,\dots,K$. Since H_k acts compactly on $u_{j,k}$, $j = 1,\dots,J$, $k = 1,\dots,K$, by Corollary 8.13 there exist $B_{j,k} \in \mathfrak{D}$ with $\nu(B_{j,k}) > 1 - \dfrac{a}{2KJ}$ such that $1_{B_{j,k}} u_{j,k}$ is almost periodic with respect to H_k. Let $\{v_{j,k,q},\ q = 1,\dots,Q\}$ be a $\frac{\delta}{4I}$-spanning set for $H_k(1_{B_{j,k}} u_{j,k})$ on Y, $k = 1,\dots,K$, $j = 1,\dots,J$.

Put $B_3 = B_2 \cap \Big(\bigcap_{\substack{1\leq j\leq J\\ 1\leq k\leq K}} B_{j,k}\Big)$. We have $\nu(B_3) > \dfrac{a}{2} > 0$. Now, for any $1 \leq k \leq K$, any $1 \leq j \leq J$, any $R_k \in H_k$ and any $y \in B_3 R_k^{-1}$ there exists $1 \leq q \leq Q$ such that

$$\big\|R_k u_{j,k} - v_{j,k,q}\big\|_y = \big\|1_{B_{j,k}R_k^{-1}} R_k u_{j,k} - v_{j,k,q}\big\|_y = \big\|R_k(1_{B_{j,k}} u_{j,k}) - v_{j,k,q}\big\|_y < \frac{\delta}{4I}.$$

13.10. B_3 contains a subset B_4 with $\nu(B_4) > 0$, such that $\mu_y(A_{j,k})$, $j = 1,\dots,J$, $k = 1,\dots,K$, are constant on B_4 with exactness up to $\delta/2$: there exists a collection of numbers $a_{j,k} \geq 0$, $j = 1,\dots,J$, $k = 1,\dots,K$, such that

$$\big|\mu_y(A_{j,k}) - a_{j,k}\big| < \frac{\delta}{2},\ j = 1,\dots,J,\ k = 1,\dots,K,$$

for all $y \in B_4$. Since by 13.6 for $y \in B_4$

$$\mu_y(A') = \sum_{(j_1,\dots,j_K)\in\Omega} \prod_{k=1}^{K} \mu_y(A_{j_k,k}) > a,$$

we may assume that $a_{j,k}$, $j = 1,\dots,J$, $k = 1,\dots,K$, satisfy (13.4).

13.11. Using Theorem 5.9, choose $p \in \mathbb{N}$ and a system $\mathcal{B} \subset \wp_0^{dp} G$ corresponding to the system $\mathcal{A} = \{g_1,\dots,g_I\}$ and Q^{JK} colors: for any mapping $\chi\colon \wp^{dp} G \longrightarrow \{1,\dots,Q^{JK}\}$ there exist $h_0 \in \mathcal{B}$ and a selection $n \in \mathrm{Sel}(\mathbb{Z}^{dp},\mathbb{Z}^d)$ such that χ is constant on the set $\big\{h_0(m) g_i(n(m)),\ i = 1,\dots,I\big\}$.

Denote

$$\mathcal{A}' = \Big\{h(m) g(n(m)),\ h \in \mathcal{B},\ g \in \mathcal{A},\ n \in \mathrm{Sel}(\mathbb{Z}^{dp},\mathbb{Z}^d)\Big\} \subset \wp_0^{dp} G.$$

For each $k = 1, \dots, K$ choose a set $\mathcal{S}_k \subset \wp_0^{dp} G$ of representatives of those left cosets of $\wp_0^{dp} H_k$ in $\wp_0^{dp} G$ which contain elements of $\mathcal{A}'$:

$$\text{for every } f \in \mathcal{A}' \text{ there is } s \in \mathcal{S}_k \text{ such that } r = s^{-1} f \in \wp_0^{dp} H_k,$$
$$\text{and } s_1^{-1} s_2 \notin \wp_0^{dp} H_k \text{ for all } s_1, s_2 \in \mathcal{S}_k,\ s_1 \neq s_2.$$

Denote by $\mathcal{R}_k$ the set of all "compact residues" of elements of $\mathcal{A}'$:

$$\mathcal{R}_k = \left\{ s^{-1} f \in \wp_0^{dp} H_k,\ f \in \mathcal{A}',\ s \in \mathcal{S}_k \right\},$$

$k = 1, \dots, K$.

13.12. Using the assumption that $(Y, \mathfrak{D}, \nu, G)$ is a NSZ-system, find $b > 0$ such that for every thick set $\Lambda' \subseteq \mathbb{Z}^{dp}$ there exists $m \in \Lambda'$ for which $\nu\Big(\bigcap_{f \in \mathcal{A}'} B_4 f(m)^{-1} \Big) > b$.

13.13. Let $\Delta = Q^{JK\#\mathcal{A}'}$; Δ is the number of all possible colorings of $\mathcal{A}'$ by Q^{JK} colors. Note that b and Δ have been chosen independently on Λ.

13.14. For $m \in \mathbb{Z}^{dp}$ denote by D_m the set of $y \in Y$ for which there exist
1. $(j_k^i,\ k = 1, \dots, K,\ i = 1, \dots, I) \in \Omega^I$,
2. $1 \leq L_1, \dots, L_K \leq I$ and disjoint partitions $\{1, \dots, I\} = \bigcup_{l=1}^{L_k} I_{l,k}$, $k = 1, \dots, K$,
3. $s_{l,k} \in \mathcal{S}_k$, $l = 1, \dots, L_k$, pairwise distinct for every $k = 1, \dots, K$,
4. $r_{i,k} \in \mathcal{R}_k$, $i = 1, \dots, I$, $k = 1, \dots, K$,

such that

$$\left| \int \prod_{k=1}^{K} \prod_{l=1}^{L_k} s_{l,k}(m) \prod_{i \in I_{l,k}} r_{i,k}(m) u_{j_k^i,k} d\mu_y - \prod_{k=1}^{K} \prod_{l=1}^{L_k} \int s_{l,k}(m) \prod_{i \in I_{l,k}} r_{i,k}(m) u_{j_k^i,k} d\mu_y \right| > \frac{C}{4J^{KI}}.$$

By the definition of $\mathcal{S}_k, \mathcal{R}_k$, $k = 1, \dots, K$, in 13.11 and by Proposition 12.7, the set

$$\Gamma = \left\{ m \in \mathbb{Z}^{dp} : \nu(D_m) > \frac{b}{2\Delta} \right\}$$

is of zero density.

13.15. Denote

$$\Lambda'' = \bigcap_{n \in \mathrm{Sel}(\mathbb{Z}^{dp}, \mathbb{Z}^d)} n^{-1}(\Lambda).$$

By Lemma 1.6, $\Lambda'' \subseteq \mathbb{Z}^{dp}$ is thick.

Put $\Lambda' = \Lambda'' \setminus \Gamma$, Λ' is thick as well. By 13.12 find $m \in \Lambda'$ such that for $B_5 = \bigcap_{f \in \mathcal{A}'} B_4 f(m)^{-1}$ one has $\nu(B_5) > b$.

We fix m from now on.

13.16. For every $y \in B_5$ we introduce a coloring of $\mathcal{A}'$ by Q^{JK} colors in the following way. For $f \in \mathcal{A}'$, $f = s_k r_k$ with $s_k \in \mathcal{S}_k$, $r_k \in \mathcal{R}_k$, we put $\chi_y^{j,k}(f) = q$ if

$$\left\| r_k(m) u_{j,k} - v_{j,k,q} \right\|_{y s_k(m)} < \frac{\delta}{4I}. \tag{13.8}$$

Since by the choice of B_5 in 13.15 we have $y s_k(m) r_k(m) = y f(m) \in B_4 \subseteq B_3$, (13.8) takes place for some $1 \le q \le Q$ (see subsection 13.9).

Then we define $\chi_y = \left(\chi_y^{j,k},\ j = 1, \ldots, J,\ k = 1, \ldots, K \right)$.

13.17. There is $B_6 \subseteq B_5$ with $\nu(B_6) > \dfrac{b}{\Delta}$ such that χ_y is constant on B_6: $\chi_y = \chi$ for all $y \in B_6$ (see 13.13 for the definition of Δ).

We put $B_7 = B_6 \setminus D_m$. Since by 13.15 $m \notin \Gamma$, we have $\nu(B_7) > \dfrac{b}{2\Delta}$ by 13.14.

13.18. Choose $h_0 \in \mathcal{B}$ and a selection $n \colon \mathbb{Z}^{dp} \longrightarrow \mathbb{Z}^d$ such that χ is constant on the set $\left\{ h_0(m) g_i(n(m)),\ i = 1, \ldots, I \right\}$ (see subsection 13.11). Then $n = n(m) \in \Lambda$ by 13.15.

We fix n from now on.

13.19. Define $B = B_7 h_0(m)$. Then $\nu(B) > \dfrac{b}{2\Delta}$.

Since $B_7 h_0(m) g_i(n) \subseteq B_4$ by 13.15 and the definition of $\mathcal{A}'$ in 13.11, we have $B g_i(n) \subseteq B_4 \subseteq B_2$, $i = 1, \ldots, I$.

13.20. Decompose

$$h_0(m) g_i(n(m)) = s_{l_k(i),k}(m) r_{i,k}(m), \quad i = 1, \ldots, I,\ k = 1, \ldots, K,$$

where $r_{i,k} \in \mathcal{R}_k$, $i = 1, \ldots, I$, $k = 1, \ldots, K$, and $s_{l,k} \in \mathcal{S}_k$, $k = 1, \ldots, K$, $l = 1, \ldots, L_k$, are pairwise distinct. Define

$$I_{l,k} = \left\{ i \in \{1, \ldots, I\} : l_k(i) = l \right\}, \quad k = 1, \ldots, K,\ l = 1, \ldots, L_k.$$

13.21. Let $y \in B$. Put $z = yh_0(m)^{-1} \in B_7$. Then

$$\begin{aligned}
\mu_y\Big(\bigcap_{i=1}^{I} A' g_i(n)^{-1}\Big) &= \mu_z\Big(\bigcap_{i=1}^{I} A'\big(h_0(m)g_i(n)\big)^{-1}\Big) \\
&= \int \prod_{i=1}^{I} \sum_{(j_k)\in\Omega} \prod_{k=1}^{K} s_{l_k(i),k}(m) r_{i,k}(m) u_{j_k,k} d\mu_z \\
&= \sum_{(j_k^i)\in\Omega^I} \int \prod_{k=1}^{K} \prod_{l=1}^{L_k} s_{l,k}(m) \Big(\prod_{i\in I_{l,k}} r_{i,k}(m) u_{j_k^i,k}\Big) d\mu_z \\
&> \sum_{(j_k^i)\in\Omega^I} \Big(\prod_{k=1}^{K} \prod_{l=1}^{L_k} \int s_{l,k}(m) \prod_{i\in I_{l,k}} r_{i,k}(m) u_{j_k^i,k} d\mu_z - \frac{C}{4J^{KI}}\Big) \\
&\qquad \text{(by 13.14, as } B_7 \cap D_m = \emptyset \text{ by 13.17)} \\
&\geq \sum_{(j_k^i)\in\Omega^I} \prod_{k=1}^{K} \prod_{l=1}^{L_k} \int \prod_{i\in I_{l,k}} r_{i,k}(m) u_{j_k^i,k} d\mu_{zs_{l,k}(m)} - \frac{C}{4} \\
&= \sum_{(j_k^i)\in\Omega^I} \bigcap_{k=1}^{K} \bigcap_{l=1}^{L_k} \mu_{zs_{l,k}(m)}\Big(\bigcap_{i\in I_{l,k}} A_{j_k^i,k} r_{i,k}(m)^{-1}\Big) - \frac{C}{4}.
\end{aligned}$$

13.22. Since $z \in B_7 \subseteq B_6$, we have by 13.16, 13.17 and 13.18

$$\Big\| r_{i,k}(m) u_{j,k} - v_{j,k,q(j,k)} \Big\|_{zs_{l_k(i),k}(m)} < \frac{\delta}{4I}$$

for some $1 \leq q(j,k) \leq Q$, $j = 1, \ldots, J$, $k = 1, \ldots, K$.

For every $k = 1, \ldots, K$, $l = 1, \ldots, L_k$ choose $i_{l,k} \in I_{l,k}$.

Fix now some $1 \leq j \leq J$, $1 \leq k \leq K$, $1 \leq l \leq L_k$. Then for any $i \in I_l$ we have

$$\mu_{zs_{l,k}(m)}\Big(A_{j,k} r_{i,k}(m)^{-1} \triangle A_{j,k} r_{i_{l,k},k}(m)^{-1}\Big) = \Big\| r_{i,k}(m) u_{j,k} - r_{i_{l,k},k}(m) u_{j,k} \Big\|_{zs_{l,k}(m)} < \frac{\delta}{2I}.$$

Thus,

$$\Big| \mu_{zs_{l,k}(m)}\Big(\bigcap_{i\in I_{l,k}} A_{j,k} r_{i,k}(m)^{-1}\Big) - \mu_{zs_{l,k}(m)}\Big(A_{j,k} r_{i_{l,k},k}(m)^{-1}\Big) \Big| < \frac{\delta}{2}.$$

Since $i_{l,k} \in I_{l,k}$, $s_{l,k}(m) r_{i_{l,k},k}(m) = h_0(m) g_{i_{l,k},k}(n)$ and so, $zs_{l,k}(m) r_{i_{l,k}k}(m) \in B_4$. By 13.10,

$$\Big| \mu_{zs_{l,k}(m)}\big(A_{j,k} r_{i_{l,k},k}(m)^{-1}\big) - a'_{j,k} \Big| = \Big| \mu_{zs_{l,k}(m) r_{i_{l,k},k}(m)}\big(A_{j,k}\big) - a'_{j,k} \Big| < \frac{\delta}{2}.$$

Hence

$$\Big| \mu_{zs_{l,k}(m)}\Big(\bigcap_{i\in I_{l,k}} A_{j,k} r_{i,k}(m)^{-1}\Big) - a'_{j,k} \Big| < \delta,$$

that is, the collection

$$a'_{j,k,l} = \mu_{zs_{l,k}(m)}\Big(\bigcap_{i\in I_{l,k}} A_{j,k} r_{i,k}(m)^{-1}\Big), \quad j=1,\dots,J,\ k=1,\dots,K,\ l=1,\dots,L_k,$$

satisfies the condition (13.6). Define $\Theta \subseteq \Omega^I$ by (13.5). Then by 13.8, the inequality (13.7) holds for these $a'_{j,k,l}$. Thus, by 13.21,

$$\begin{aligned}\mu_y\Big(\bigcap_{i=1}^{I} A' g_i(n)^{-1}\Big) &> \sum_{(j_k^i)\in\Theta} \bigcap_{k=1}^{K}\bigcap_{l=1}^{L_k} \mu_{zs_{l,k}(m)}\Big(\bigcap_{i\in I_{l,k}} A_{j_k^i,k} r_{i,k}(m)^{-1}\Big) - \frac{C}{4}\\ &= \sum_{\theta\in\Theta}\prod_{k=1}^{K}\prod_{l=1}^{L_k} a'_{t_k^l(\theta),k,l} \ge \frac{3C}{4} - \frac{C}{4} = \frac{C}{2}.\end{aligned}$$

13.23. By 13.18, $n\in\Lambda$. By 13.7, 13.22 and 13.19, for any $y\in B$ we have

$$\mu_y\Big(\bigcap_{i=1}^{I} A g_i(n)^{-1}\Big) > \frac{C}{4}.$$

Therefore, by 13.19,

$$\mu\Big(\bigcap_{i=1}^{I} A g_i(n)^{-1}\Big) > \frac{bC}{8\Delta}.$$

■

14. The nilpotent van der Waerden and Szemerédi Theorems

14.1. A "nilpotent" generalization of Theorem CT can be deduced directly from Theorem NT. However it is simpler to utilize an abstract version of this theorem, Theorem 5.3.

Let G be a finitely generated torsion-free nilpotent group, let $\mathcal{A}\subset\wp_0 G$ be a system and let $r\in\mathbb{N}$. By Corollary 5.7, there exist $p\in\mathbb{N}$ and a system $\mathcal{B}\subset\wp_0^p G$ such that for any r-coloring χ of $\wp^p G$ there exist $h\in\mathcal{B}$ and a selection $n\in\mathrm{Sel}(\mathbb{Z}^p,\mathbb{Z})$ such that χ is constant on the set $\{g(n)h,\ g\in\mathcal{A}\}$.

Let an r-coloring $\psi\colon G\longrightarrow\{1,\dots,r\}$ of G be given. Put $m=(1,\dots,1)\in\mathbb{Z}^p$ and define a coloring χ of $\wp^p G$ by $\chi(h)=\psi(h(m))$ for $h\in\wp^p G$. Find $h\in\mathcal{B}$ and $n\in\mathrm{Sel}(\mathbb{Z}^p,\mathbb{Z})$ corresponding to χ by Corollary 5.7; then ψ is constant on the set $\{g(n(m))h(m),\ g\in\mathcal{A}\}$ and $1\le n(m)\le p$.

14.2. Consequently, we have the following theorem.

Theorem. *Let $\mathcal{A}\subset\wp_0 G$ be a system and let $r\in\mathbb{N}$. Then there exist $p\in\mathbb{N}$ and a finite subset $Q\subset G$ such that for any r-coloring χ of G there exist $1\le n\le p$ and $T\in Q$ such that χ is constant on the set $\{g(n)T,\ g\in\mathcal{A}\}$.*

14.3. Theorem 14.2 is equivalent to Theorem NT. Moreover, with its help one can obtain a version of Theorem NT in the formulation of which the requirements that X is complete and the elements of G act continuously on X are omitted:

Corollary. *Let (X,ρ) be a totally bounded metric space, let G be a nilpotent group of (not necessarily continuous) transformations of X, let $T_1,\ldots,T_t \in G$ and let $p_{i,j}\colon \mathbb{Z} \longrightarrow \mathbb{Z}$ with $p_{i,j}(0) = 0$, $i = 1,\ldots,I$, $j = 1,\ldots,t$, be polynomials. Then for any $\varepsilon > 0$ there exist $p \in \mathbb{N}$ and a finite subset $Q \subset G$ such that for any $x \in X$ there exist $1 \le n \le p$ and $T \in Q$ such that*

$$\rho\Big(T_t^{p_{i,t}(n)} \ldots T_1^{p_{i,1}(n)} Tx, Tx\Big) < \varepsilon \text{ for each } i = 1,\ldots,I.$$

Proof. Let $\varepsilon > 0$ be given. Choose a finite $\varepsilon/2$-net $\{x_1,\ldots,x_r\}$ in X. Using Theorem 14.2, find $p \in \mathbb{N}$ and $Q \subset G$ corresponding to the system

$$\mathcal{A} = \Big\{\mathbf{1}_G, g_1(n) = T_t^{p_{1,t}(n)} \ldots T_1^{p_{1,1}(n)}, \ldots, g_I(n) = T_t^{p_{I,t}(n)} \ldots T_1^{p_{I,1}(n)}\Big\}$$

and r-colorings of G. Given a point $x \in X$, define a coloring χ of G by

$$\chi(T) = r \quad \text{if} \quad \rho(Tx, x_r) < \frac{\varepsilon}{2}$$

(if there is a number of possibilities, choose one of them). Then for some $1 \le n \le p$ and $T \in Q$ one has $\chi(g_i(n)T) = \chi(T)$, $i = 1,\ldots,I$, and so, $\rho\big(g_i(n)Tx, Tx\big) < \varepsilon$, $i = 1,\ldots,I$. ■

14.4. Now, let $d \in \mathbb{N}$ and let G be the multiplicative group of upper triangular $d \times d$ matrices over $\mathbb{Z}$ with unit diagonal entries. G is nilpotent and torsion-free (moreover, any finitely generated torsion-free nilpotent group is a subgroup of such G for d big enough (see, for example, [KM])). It is easy to see that the group $\wp G$ of G-polynomials is the multiplicative group of upper triangular matrices with unit diagonal entries over the ring of integral polynomials. Applying Theorem 14.2 to this case, we obtain the following corollary.

Corollary. *Let*

$$g_i(n) = \begin{pmatrix} 1 & p_{i,1,2}(n) & p_{i,1,3}(n) & \ldots & p_{i,1,d}(n) \\ 0 & 1 & p_{i,2,3}(n) & \ldots & p_{i,2,d}(n) \\ 0 & 0 & 1 & \ldots & p_{i,3,d}(n) \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix}, \quad i = 1,\ldots,I, \tag{14.1}$$

be $d \times d$ matrices, where $p_{i,j,l}\colon \mathbb{Z} \longrightarrow \mathbb{Z}$ are integral polynomials satisfying $p_{i,j,l}(0) = 0$, $i = 1,\ldots,I$, $1 \le j < l \le d$. For any finite coloring of the set G of $d \times d$ upper-triangular matrices over $\mathbb{Z}$ with unit diagonal entries there exists $T \in G$ and $n \in \mathbb{N}$ such that the set $\{g_i(n)T,\ i = 1,\ldots,I\}$ is monochromatic.

14.5. Corollary 14.4 implies the following "pure" combinatorial fact (compare with Theorem CT in 0.16):

Corollary. *Let $d, I \in \mathbb{N}$ and let $p_{i,j,l}: \mathbb{Z} \longrightarrow \mathbb{Z}$ be polynomials satisfying $p_{i,j,l}(0) = 0$, $i = 1, \ldots, I$, $1 \le j < l \le d$. For any finite coloring of $\mathbb{Z}^{d-1}$ there exist $n \in \mathbb{N}$ and $(u_1, \ldots, u_{d-1}) \in \mathbb{Z}^{d-1}$ for which the set*

$$\begin{aligned}\Big\{\big(u_1 + p_{i,1,2}(n)u_2 + \ldots + p_{i,1,d-1}(n)u_{d-1} + p_{i,1,d}(n),&\\ u_2 + \ldots + p_{i,2,d-1}(n)u_{d-1} + p_{i,2,d}(n),&\\ \vdots\qquad&\\ u_{d-1} + p_{i,d-1,d}(n)\big),\ i = 1, \ldots, I\Big\}&\end{aligned} \tag{14.2}$$

is monochromatic.

Proof. Let $\pi: \mathbb{Z}^d \longrightarrow \mathbb{Z}^{d-1}$ be the projection forgetting the last coordinate. Define $v = (0, \ldots, 0, 1) \in \mathbb{Z}^d$.

The matrix group G introduced in 14.4 acts on $\mathbb{Z}^d$ from the left in a natural way. Any finite coloring $\psi: \mathbb{Z}^{d-1} \longrightarrow \{1, \ldots, r\}$ induces a coloring $\chi: G \longrightarrow \{1, \ldots, r\}$ by the rule $\chi(T) = \psi\big(\pi(Tv)\big)$.

Put $g_1, \ldots, g_I \in \wp_0 G$ by (14.1). Find $T \in G$ and $n \in \mathbb{N}$ such that χ is constant on the set $\big\{g_i(n)T,\ i = 1, \ldots, I\big\}$. Then ψ is constant on the set $\big\{\pi \circ g_i(n)Tv,\ i = 1, \ldots, I\big\}$, and this gives (14.2) for $(u_1, \ldots, u_{d-1}) = \pi(Tv) \in \mathbb{Z}^{d-1}$. ∎

14.6. To obtain combinatorial corollaries of measurable multiple recurrence theorems one uses *the Furstenberg Correspondence Principle* (see, for example, [F2]). Now we will describe a possible way to establish this principle.

Let G be a countable semigroup. Let G^+ and G^- be two distinct copies of G; we denote by T^+ and T^- the elements corresponding to $T \in G$ in G^+ and G^- respectively. Let $\mathcal{G}$ be the set of nonempty finite subsets of $G^+ \cup G^-$. The semigroup G acts on $\mathcal{G}$ by the rule

$$T\{T_1^{s_1}, \ldots, T_l^{s_l}\} = \{(TT_1)^{s_1}, \ldots, (TT_l)^{s_l}\},$$

where $l \in \mathbb{N}$, $s_1, \ldots, s_l \in \{+,-\}$, $\{T_1^{s_1}, \ldots, T_l^{s_l}\} \in \mathcal{G}$ and $T \in G$. Define $\mathfrak{X} = [0,1]^{\mathcal{G}}$. Then $\mathfrak{X}$ is a compact metrizable space. A point $\omega \in \mathfrak{X}$ has coordinates $0 \le \omega_Q \le 1$, $Q \in \mathcal{G}$, and a sequence of points $\omega_1, \omega_2, \ldots \in \mathfrak{X}$ converges to $\omega \in \mathfrak{X}$ if and only if $(\omega_1)_Q, (\omega_2)_Q, \ldots$ converges to ω_Q for every $Q \in \mathcal{G}$.

Let $(X, \mathfrak{B}, \mu, G)$, $\mu(X) = 1$, be a measure preserving system, let $A \in \mathfrak{B}$. Denote $A^+ = A$, $A^- = X \setminus A$. Define a point $\Psi\Big((X, \mathfrak{B}, \mu, G), A\Big) \in \mathfrak{X}$ by the formula

$$\Psi\Big((X, \mathfrak{B}, \mu, G), A\Big)_{\{T_1^{s_1}, \ldots, T_l^{s_l}\}} = \mu\Big(\bigcap_{i=1}^{l} A^{s_i} T_i^{-1}\Big),\ l \in \mathbb{N},\ T_1, \ldots, T_l \in G,\ s_1, \ldots, s_l \in \{+,-\}.$$

Then, if we write $\omega = \Psi\Big((X, \mathfrak{B}, \mu, G), A\Big)$, for any $Q \in \mathcal{G}$ and any $T \in G$ we have:

a) $\omega_{\{T^+, T^-\}} = 0$;

b) $\omega_Q = \omega_{Q\cup\{T^+\}} + \omega_{Q\cup\{T^-\}}$;
c) $\omega_{TQ} = \omega_Q$.

Denote by $\mathfrak{M}$ the closed (and so, compact) subspace of $\mathfrak{X}$ defined by the equations a) – c). Then Ψ is a mapping from the class of measure preserving systems with a marked element of the corresponding σ-algebra to $\mathfrak{M}$. A key fact is that Ψ is surjective (and therefore the class of measure preserving systems "is compact" in the topology lifted from $\mathfrak{M}$).

Indeed, given $\omega \in \mathfrak{X}$ satisfying the conditions a) – c), define X to be the set of mappings from G to the two-element set $\{+,-\}$: $X = \{+,-\}^G$, and $\mathfrak{B}$ to be the σ-algebra generated by the "cylinders"

$$\mathcal{C}_{\{T_1^{s_1},\ldots,T_l^{s_l}\}} = \Big\{x \in X : x(T_i) = s_i,\ i = 1,\ldots,l\Big\}$$

for $l \in \mathbb{N}$, $T_1,\ldots,T_l \in G$, $s_1,\ldots,s_l \in \{+,-\}$. Define a measure μ on $\mathfrak{B}$ by

$$\mu\big(\mathcal{C}_{\{T_1^{s_1},\ldots,T_l^{s_l}\}}\big) = \omega_{\{T_1^{s_1},\ldots,T_l^{s_l}\}},\ l \in \mathbb{N},\ T_1,\ldots,T_l \in G,\ s_1,\ldots,s_l \in \{+,-\},$$

and an action of G on X by $(xT)(P) = x(TP)$, $x \in X$, $T,P \in G$. Then $(X,\mathfrak{B},\mu,G)$ is a measure preserving system and $\Psi\Big((X,\mathfrak{B},\mu,G),\mathcal{C}_{\{\mathbf{1}_G^+\}}\Big) = \omega$.

14.7. The first corollary which one can derive from the considerations in subsection 14.6 is the existence of "universal constants" in theorems on measurable recurrence. In application to Theorem NM$'$ it can be formulated in the following way:

Corollary. *Let G be a nilpotent group and let $\mathcal{A} \subset \wp_0 G$ be a system. For any $a > 0$ there exist $N \in \mathbb{N}$ and $C > 0$ such that, given a measure preserving system $(X,\mathfrak{B},\mu,G)$, $\mu(X) = 1$, and a set $A \in \mathfrak{B}$ with $\mu(A) \geq a$, there exists $1 \leq n \leq N$ for which*

$$\mu\Big(\bigcap_{g\in\mathcal{A}} Ag(n)^{-1}\Big) \geq C.$$

Proof. Assume that for some $a > 0$ there are no such N, C: let $\big((X_k,\mathfrak{B}_k,\mu_k,G),A_k\big)$, $k = 1,2,\ldots$, be a sequence of measure preserving systems and subsets $A_k \in \mathfrak{B}_k$ with $\mu_k(X_k) = 1$, $\mu_k(A_k) \geq a$ and such that

$$\mu_k\Big(\bigcap_{g\in\mathcal{A}} A_k g(n)^{-1}\Big) < \frac{1}{k}$$

for each $n = 1,\ldots,k$, $k = 1,2,\ldots$.

The sequence $\omega_k = \Psi\big((X_k,\mathfrak{B}_k,\mu_k,G),A_k\big)$, $k = 1,2,\ldots$, of points of the compact space $\mathfrak{M}$ has a limit point ω; let a sequence $k_1,k_2,\ldots \in \mathbb{N}$ be such that $\omega_{k_j} \underset{j\to\infty}{\longrightarrow} \omega$. Find a measure preserving system $(X,\mathfrak{B},\mu,G)$ with $\mu(X) = 1$ and $A \in \mathfrak{B}$ for which $\omega = \Psi\big((X,\mathfrak{B},\mu,G),A\big)$. We have then $\mu(A) = \lim_{j\to\infty} \mu_{k_j}(A_{k_j}) \geq a$, and, for any $n \in \mathbb{N}$,

$$\mu\Big(\bigcap_{g\in\mathcal{A}} Ag(n)^{-1}\Big) = \lim_{j\to\infty} \mu_{k_j}\Big(\bigcap_{g\in\mathcal{A}} A_{k_j} g(n)^{-1}\Big) = 0.$$

This contradicts Theorem NM$'$. ∎

14.8. Let a countable semigroup G act (from the right) on a set M. We say that a sequence $\Phi = \{\Phi_k\}_{k\in\mathbb{N}}$ of finite subsets of M *is a Følner sequence in M (with respect to the action of G)* if $\dfrac{\#(\Phi_k T \triangle \Phi_k)}{\#\Phi_k} \underset{k\to\infty}{\longrightarrow} 0$ for all $T \in G$. In particular, a Følner sequence in an amenable group G (see subsection 8.1) is a Følner sequence with respect to the action of G on itself by right multiplications.

Given a subset $S \in M$, *the upper density of S with respect to Φ* is

$$d^*_\Phi(S) = \limsup_{k\to\infty} \frac{\#(S\cap\Phi_k)}{\#\Phi_k}.$$

14.9. Now we formulate a version of Furstenberg's Correspondence Principle.

Proposition. *Let G be a countable semigroup acting on a set M, let S be a subset of M and let Φ be a Følner sequence in M. There exist a measure preserving system $(X,\mathfrak{B},\mu,G)$, $\mu(X)=1$, and a set $A\in\mathfrak{B}$ such that $\mu(A)=d^*_\Phi(S)$ and, for any $l\in\mathbb{N}$ and any $T_1,\ldots,T_l\in G$,*

$$\mu\Big(\bigcap_{i=1}^{l} AT_i^{-1}\Big) \le d^*_\Phi\Big(\bigcap_{i=1}^{l} ST_i^{-1}\Big).$$

Proof. Let $\Phi_1,\Phi_2,\ldots$ be a subsequence of Φ for which $\lim\limits_{k\to\infty}\dfrac{\#(S\cap\Phi_k)}{\#\Phi_k} = d^*_\Phi(S)$.

Denote $S^+ = S$, $S^- = M\setminus S$, and define a sequence of points $\omega_1,\omega_2,\ldots\in\mathfrak{X}$ by the formula

$$(\omega_k)_{\{T_1^{s_1},\ldots,T_l^{s_l}\}} = \frac{1}{\#\Phi_k}\#\Big(\Big(\bigcap_{i=1}^{l} S^{s_i}T_i^{-1}\Big)\cap\Phi_k\Big),$$
$$l\in\mathbb{N},\ T_1,\ldots,T_l\in G,\ s_1,\ldots,s_l\in\{+,-\},\ k=1,2,\ldots.$$

Find a point $\omega\in\mathfrak{X}$ and a sequence $k_1,k_2,\ldots\in\mathbb{N}$ such that $\omega_{k_j}\underset{j\to\infty}{\longrightarrow}\omega$.

Since $\Phi_1,\Phi_2,\ldots$ is a Følner sequence, it is easy to conclude that $\omega\in\mathfrak{M}$. Let a measure preserving system $(X,\mathfrak{B},\mu,G)$, $\mu(X)=1$, and a set $A\in\mathfrak{B}$ be such that $\Psi\Big((X,\mathfrak{B},\mu,G),A\Big)=\omega$. Then

$$\mu(A) = \omega_{\{\mathbf{1}_G^+\}} = \lim_{j\to\infty}(\omega_{k_j})_{\{\mathbf{1}_G^+\}} = \lim_{j\to\infty}\frac{\#(S\cap\Phi_{k_j})}{\#\Phi_{k_j}} = d^*_\Phi(S),$$

and for any $l\in\mathbb{N}$ and any $T_1,\ldots,T_l\in G$ we have

$$\mu\Big(\bigcap_{i=1}^{l} AT_i^{-1}\Big) = \omega_{\{T_1^+,\ldots,T_l^+\}} = \lim_{j\to\infty}(\omega_{k_j})_{\{T_1^+,\ldots,T_l^+\}}$$
$$= \lim_{j\to\infty}\frac{1}{\#\Phi_{k_j}}\#\Big(\Big(\bigcap_{i=1}^{l} ST_i^{-1}\Big)\cap\Phi_{k_j}\Big) \le d^*_\Phi\Big(\bigcap_{i=1}^{l} ST_i^{-1}\Big).$$

■

14.10. Now it is simple to obtain the following generalization of Szemerédi's Theorem.

Theorem. *Let G be a finitely generated nilpotent group and let $\mathcal{A} \subset \wp_0 G$ be a system. For any subset $S \subseteq G$ of positive upper density in G there exist $n \in \mathbb{N}$ and $T \in G$ such that $\big\{Tg(n),\ g \in \mathcal{A}\big\} \subset S$.*

Proof. G acts on itself by right multiplications. $S \subseteq G$ is of positive upper density in G means that $d^*_\Phi(S) > 0$ for some Følner sequence Φ in G with respect to this action. By Theorem NM$'$ and Proposition 14.9, $d^*_\Phi\Big(\bigcap_{g\in\mathcal{A}} Sg(n)^{-1}\Big) > 0$ for some $n \in \mathbb{N}$. In particular, $\bigcap_{g\in\mathcal{A}} Sg(n)^{-1}$ is nonempty. Choose $T \in \bigcap_{g\in\mathcal{A}} Sg(n)^{-1}$. Then $Tg(n) \in S$ for all $g \in \mathcal{A}$. ■

14.11. As an example of a pure combinatorial proposition deducible from Theorem 14.10, let us bring the following its corollary:

Corollary. *Let $d, I \in \mathbb{N}$ and let $p_{i,j,l}\colon \mathbb{Z} \longrightarrow \mathbb{Z}$ be polynomials satisfying $p_{i,j,l}(0) = 0$, $i = 1, \ldots, I$, $1 \le j < l \le d$. Let $\Pi = \{\Pi_k\}_{k\in\mathbb{N}}$ be a sequence of parallelepipeds in $\mathbb{Z}^{d-1}$, $\Pi_k = \prod_{i=1}^{d-1}\{a_{k,i}, a_{k,i}+1, \ldots, b_{k,i}\}$, where the integers $a_{k,i} \le b_{k,i} \in \mathbb{Z}$ satisfy*

$$b_{k,d-1} - a_{k,d-1} \underset{k\to\infty}{\longrightarrow} \infty, \qquad \frac{b_{k,i} - a_{k,i}}{|b_{k,i+1}| + |a_{k,i+1}| + 1} \underset{k\to\infty}{\longrightarrow} \infty,\ i = d-2, \ldots, 1,$$

and let S be a subset of $\mathbb{Z}^{d-1}$ whose upper density with respect to Π is positive: $d^*_\Pi(S) = \limsup_{k\to\infty} \frac{\#(S\cap\Pi_k)}{\#\Pi_k} > 0$. Then there exist $(u_1, \ldots, u_{d-1}) \in \mathbb{Z}^{d-1}$ and $n \in \mathbb{N}$ such that

$$\begin{aligned}\Big\{\big(u_1 + p_{i,1,2}(n)u_2 + \ldots + p_{i,1,d-1}(n)u_{d-1} + p_{i,1,d}(n),&\\ u_2 + \ldots + p_{i,2,d-1}(n)u_{d-1} + p_{i,2,d}(n),&\\ \vdots\qquad\qquad&\\ u_{d-1} + p_{i,d-1,d}(n)\big),\ i = 1, \ldots, I\Big\} \subset S.&\end{aligned} \tag{14.3}$$

Proof. Denote by G the nilpotent group of $d \times d$ lower triangular matrices over $\mathbb{Z}$ with unit diagonal entries. Put $\Phi_k = \Pi_k \times \{1\} \subset \mathbb{Z}^d$, $k \in \mathbb{N}$. It is easy to see that the sequence $\Phi = \{\Phi_k\}_{k\in\mathbb{N}}$ is a Følner sequence in $\mathbb{Z}^d$ with respect to the natural right action of G on $\mathbb{Z}^d$. Define $g_1, \ldots, g_I \in \wp_0 G$ by

$$g_i(n) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ p_{i,1,2}(n) & 1 & 0 & \ldots & 0 \\ p_{i,1,3}(n) & p_{i,2,3}(n) & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ p_{i,1,d}(n) & p_{i,2,d}(n) & p_{i,3,d}(n) & \ldots & 1 \end{pmatrix},\ i = 1, \ldots, I.$$

Define $S' = \big\{s \times \{1\},\ s \in S\big\}$, then $d^*_\Phi(S') > 0$. By Theorem NM$'$ and Proposition 14.9, $d^*_\Phi\Big(\bigcap_{i=1}^I S'g(n)^{-1}\Big) > 0$ for some $n \in \mathbb{N}$. In particular, $\bigcap_{i=1}^I S'g(n)^{-1}$ is nonempty. Choose $(u_1, \ldots, u_{d-1}, 1) \in \bigcap_{i=1}^I S'g(n)^{-1}$. We have then (14.3) for such $n \in \mathbb{N}$, $u_1, \ldots, u_{d-1} \in \mathbb{Z}$. ■

Bibliography

[B] V. Bergelson, Weakly mixing PET, *Ergod. Th. and Dynam. Sys.* **7** (1987), 337-349.

[BL1] V. Bergelson and A. Leibman, Polynomial extension of van der Waerden's and Szemerédi's Theorems, *Journal of AMS* **9** No. 3 (1996), 725-753.

[BL2] V. Bergelson and A. Leibman, Set-polynomials and a polynomial extension of Hales-Jewett Theorem, *submitted* .

[BM] V. Bergelson and R. McCutcheon, Uniformity in Polynomial Szemerédi Theorem, *Ergodic Theory of* $\mathbb{Z}^d$*-actions* (edited by M. Pollicott and K. Schmidt), London Math. Soc. Lecture Note Series **228** (1996), 273-296.

[F1] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d'Analyse Math.* **31** (1977), 204-256.

[F2] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.

[FK1] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations, *J. d'Analyse Math.* **34** (1978), 275-291.

[FK2] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem, *J. d'Analyse Math.* **57** (1991), 64-119.

[FW] H. Furstenberg and B. Weiss, Topological dynamics and combinatorial number theory, *J. d'Analyse Math.* **34** (1978), 61-85.

[I] V. I. Istrăţescu, *Inner Product Structures*, D. Reidel Publishing Company, Dordrecht, 1987.

[KM] M. Kargapolov and Ju. Merzljakov, *Fundamentals of the Theory of Groups*, Springer Verlag, New York, Heidelberg, Berlin, 1979.

[L] A. Leibman, Multiple recurrence theorem for nilpotent group actions, *Geom. and Funct. Anal.* **4**, No. 6 (1994), 648-659.

[R] A. Ramsay, Virtual groups and group actions, *Advances in Math.* **6** (1971), 253-322.

[Sz] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199-245.

[Z1] R. Zimmer, Extensions of ergodic group actions, *Illinois J. Math.* **20** (1976), 373-409.

[Z2] R. Zimmer, Ergodic actions with generalized discrete spectrum, *Illinois J. Math.* **20** (1976), 555-588.

A. Leibman
Department of Mathematics
The Ohio State University
Columbus, OH43210, USA
e-mail: leibman@math.ohio-state.edu