Weakly mixing PET

V. BERGELSON

Department of Mathematics, The Ohio State University, Columbus, Ohio 43210, *USA*

I. *Introduction*

Suppose that $(X, \mathscr{B}, \mu)$ is a probability measure space and T is an invertible measure perserving transformation of $(X, \mathscr{B}, \mu)$. T is called weakly mixing if for any two sets $A_1, A_2 \in \mathscr{B}$ one has:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} (\mu(A_1 \cap T^{-n} A_2) - \mu(A_1)\mu(A_2))^2 = 0.$$

In this case we will say that $(X, \mathscr{B}, \mu, T)$ is weakly mixing system. Furstenberg proved in [**F1**] that any weakly mixing system is 'weakly mixing of all orders': if T is weakly mixing then for any $k \in \mathbb{N}$ and $A_0, A_1, \ldots, A_k \in \mathscr{B}$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} (\mu(A_0 \cap T^{-n} A_1 \cap T^{-2n} A_2 \cap \cdots \cap T^{-kn} A_k)$$
$$- \mu(A_0)\mu(A_1) \cdots \mu(A_k))^2 = 0.$$

One can show (see [**FKO**, Theorem 3.1]) that the last statement is equivalent to the following:

THEOREM 1.1. *If* $(X, \mathscr{B}, \mu, T)$ *is a weakly mixing system, then for any* $k \in \mathbb{N}$ *and any* $f_1, f_2, \ldots, f_k \in L^\infty(X, \mathscr{B}, \mu)$

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} T^n f_1 T^{2n} f_2 \cdots T^{kn} f_k - \prod_{i=1}^{k} \int f_i \, d\mu \right\|_{L^2(X)} = 0.$$

The purpose of this paper is to prove an extension of Theorem 1.1 which we will call *weakly mixing* PET (PET stands for *polynomial ergodic theorem*). We will say that polynomials $p(t), q(t)$ are essentially distinct if $p(t) - q(t) \neq \text{const}$.

THEOREM 1.2 (weakly mixing PET). *Suppose that* $(X, \mathscr{B}, \mu, T)$ *is a weakly mixing system and let* $p_1(t), p_2(t), \ldots, p_k(t)$ *be pairwise essentially distinct polynomials with rational coefficients taking on integer values on the integers. Then for any* $f_1, f_2, \ldots, f_k \in L^\infty(X, \mathscr{B}, \mu)$ *one has*:

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} T^{p_1(n)} f_1 T^{p_2(n)} f_2 \cdots T^{p_k(n)} f_k - \prod_{i=1}^{k} \int f_i \, d\mu \right\|_{L^2(X)} = 0. \qquad (1.1)$$

Remark. As a matter of fact, one can prove that under the conditions of Theorem 1.2

$$\lim_{N-M\to\infty}\left\|\frac{1}{N-M}\sum_{n=M}^{N-1}T^{p_1(n)}f_1 T^{p_2(n)}f_2\cdots T^{p_k(n)}f_k-\prod_{i=1}^{k}\int f_i\,d\mu\right\|_{L^2(X)}=0.$$

This uniform version of the weakly mixing PET will be dealt with in § 3 where as an application of it we will show that any weakly mixing system is 'uniformly polynomially mixing of all orders'. Our weakly mixing PET is so to say a result of crossbreeding of Theorem 1.1 with the following PET of Furstenberg ([**F2**, p. 70]).

THEOREM 1.3. *Suppose that $p(t)$ is a polynomial with rational coefficients that takes on integer values on the integers and let U be a unitary operator on a Hilbert space $\mathcal{H}$. Assume that U has the following property: if for some $n \in \mathbb{N}$ and $f \in \mathcal{H}$, $U^n f = f$ then $f = 0$. Then:*

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=1}^{N}U^{p(n)}f\right\|_{\mathcal{H}}=0.$$

Furstenberg's proof of Theorem 1.3 makes use of the spectral theorem. We will give below quite a different proof (avoiding the use of the spectral theorem). A generalization of this method will lead us to the proof of weakly mixing PET. The main tools in the proof of Theorem 1.2 will be Theorem 1.4 which is an abstract version of classical van der Corput's difference theorem (cf. [**KN**, pp. 25-26]) and Theorem 1.5 which is a generalization of Theorem 1.4.

THEOREM 1.4. *Suppose that $\{x_n\}_{n=1}^{\infty}$ is a bounded sequence of vectors in Hilbert space and suppose that for any $h \geq h_0 > 0$*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\langle x_n, x_{n+h}\rangle=0. \tag{1.2}$$

Then

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=1}^{N}x_n\right\|=0. \tag{1.3}$$

Proof. By the classical van der Corput's fundamental inequality, if $x_1, x_2, \ldots, x_n$ are complex numbers and H is an integer with $1 \leq H \leq N$ then

$$H^2\left|\sum_{n=1}^{N}x_n\right|^2 \leq H(N+H-1)\sum_{n=1}^{N}|x_n|^2$$
$$+2(N+H-1)\sum_{h=1}^{H-1}(H-h)\,\mathrm{Re}\sum_{n=1}^{N-h}x_n\bar{x}_{n+h}, \tag{1.4}$$

where $\mathrm{Re}\,z$ denotes the real part of $z \in \mathbb{C}$.

Examining the proof of (1.4) one sees that the same inequality holds for vectors in Hilbert space, absolute values in (1.4) being replaced by the norm and the products $x_n\bar{x}_{n+h}$ by the scalar products $\langle x_n, x_{n+h}\rangle$. We get:

$$H^2\left\|\sum_{n=1}^{N}x_n\right\|^2 \leq H(N+H-1)\sum_{n=1}^{N}\|x_n\|^2$$
$$+2(N+H-1)\sum_{h=1}^{H-1}(H-h)\,\mathrm{Re}\sum_{n=1}^{N-h}\langle x_n, x_{n+h}\rangle. \tag{1.5}$$

Assuming (without the loss of generality) that $\|x_n\| \leq 1$ for $n = 1, 2, \ldots, N$ we get from (1.5):

$$\left\| \frac{1}{N} \sum_{n=1}^{N} x_n \right\|^2 \leq \frac{N+H-1}{HN} + 2 \sum_{h=1}^{H-1} \frac{(N+H-1)(H-h)(N-h)}{H^2 N^2} \left| \frac{1}{N-h} \sum_{n=1}^{N-h} \langle x_n, x_{n+h} \rangle \right|.$$

Using the assumption (1.2) we obtain:

$$\overline{\lim_{N \to \infty}} \left\| \frac{1}{N} \sum_{n=1}^{N} x_n \right\|^2 \leq \frac{1}{H}. \tag{1.6}$$

As H is arbitrary, (1.6) gives us (1.3). Theorem 1.4 is proved.

It is clear from the proof of Theorem 1.4 that the conclusion remains true if (1.2) is replaced by

$$\lim_{h \to \infty} \left(\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h} \rangle \right) = 0.$$

As a matter of fact it is enough even to assume that

$$\text{D-}\lim_{h} \left(\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h} \rangle \right) = 0;$$

(D-$\lim_h a_h = 0$ means that $\lim_{h \to \infty, h \notin P} a_h = 0$, where $P \subset \mathbb{N}$ is a set of zero density). This leads to the following:

THEOREM 1.5 (cf. Lemma 4.9 in [F2]). *Suppose that* $\{x_n\}_{n=1}^{\infty}$ *is a bounded sequence of elements in a Hilbert space. If*

$$\text{D-}\lim_{h} \left(\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h} \rangle \right) = 0,$$

then

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} x_n \right\| = 0.$$

Before giving the formal proof of Theorem 1.2 (in § 2) we shall give two illustrative examples.

To begin with, let us prove Furstenberg's PET (Theorem 1.3) with the help of Theorem 1.4. We will use the induction on the degree d_p of polynomial $p(t)$. If $d_p = 1$ then the result in question follows from the von Neumann's classical mean ergodic theorem. Suppose that $d_p > 1$. Writing $x_n = U^{p(n)}f$, we have:

$$\langle x_n, x_{n+h} \rangle = \langle U^{p(n)}f, U^{p(n+h)}f \rangle$$
$$= \langle f, U^{p(n+h)-p(n)}f \rangle.$$

Notice that for any $h \in \mathbb{N}$ the degree of $p(n+h) - p(n)$ is equal to $d_p - 1$. By the induction hypothesis

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h} \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle f, U^{p(n+h)-p(n)}f \rangle = 0,$$

and the result follows from Theorem 1.4.

Remark. Suppose that T is a weakly mixing invertible transformation of $(X, \mathcal{B}, \mu)$. Then $T \times T$ is also a weakly mixing transformation (of the product space). Applying

Theorem 1.3 to the unitary operator induced by $T \times T$ and performing routine computations we obtain that for any $f_1, f_2 \in L^2(X, \mathcal{B}, \mu)$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left(\int f_1 T^{p(n)} f_2 \, d\mu - \int f_1 \, d\mu \int f_2 \, d\mu \right)^2 = 0.$$

This gives

$$\underset{n}{\text{D-lim}} \int f_1 T^{p(n)} f_2 \, d\mu = \int f_1 \, d\mu \int f_2 \, d\mu. \tag{1.7}$$

The last formula will be used in the proof of implication (ii)$\Rightarrow$(iii) below.

Suppose now that $(X, \mathcal{B}, \mu, T)$ is a weakly mixing system. Let $\mathcal{H} = L^2(X, \mathcal{B}, \mu)$, $f_1, f_2 \in L^\infty(X, \mathcal{B}, \mu)$ and let us try to prove that

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} T^{n^2} f_1 T^{2n^2} f_2 - \int f_1 \, d\mu \int f_2 \, d\mu \right\|_{\mathcal{H}} = 0. \tag{1.8}$$

One might try to apply the same idea as above to the sequence $x_n = T^{n^2} f_1 T^{2n^2} f_2$. Unfortunately, the expressions $\langle x_n, x_{n+h} \rangle$ do not look too simple in this case. Nevertheless one can give the following multi-stage proof to (1.8). Consider the following statements.

(i) If $a_1, a_2, \ldots, a_7$ are pairwise distinct non-zero integers then for any $f_1, f_2, \ldots, f_7 \in L^\infty(X, \mathcal{B}, \mu)$

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{7} T^{a_i n} f - \prod_{i=1}^{7} \int f_i \, d\mu \right\|_{\mathcal{H}} = 0.$$

Remark. Statement (i) is 'almost' a partial case of Theorem 1.1. We say 'almost' because Theorem 1.1 deals with positive powers of T only. One checks readily that this restriction is insignificant in any of the proofs of Theorem 1.1, given in [F1], [F2], [FKO]. The reader is also referred to § 3 of the present paper where a slight generalization of Theorem 1.1 is proved (formula $(3.1)_k$).

(ii) For any non-zero integer c, any pairwise distinct integers a_1, a_2, a_3, a_4 and any $f_1, f_2, f_3, f_4 \in L^\infty(X, \mathcal{B}, \mu)$

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{N} T^{cn^2 + a_i n} f_i - \prod_{i=1}^{4} \int f_i \, d\mu \right\|_{\mathcal{H}} = 0.$$

(iii) For any $a_1, a_2, b, c \in \mathbb{Z}$ with $a_1 \neq a_2$; $b, c \neq 0$ and for any $f_0, f_1, f_2 \in L^\infty(X, \mathcal{B}, \mu)$

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} T^{bn} f_0 T^{cn^2 + a_1 n} f_1 T^{cn^2 + a_2 n} f_2 - \int f_1 \, d\mu \int f_2 \, d\mu \int f_3 \, d\mu \right\|_{\mathcal{H}} = 0.$$

(iv) For any $f_1, f_2 \in L^\infty(X, \mathcal{B}, \mu)$

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} T^{n^2} f_1 T^{2n^2} f_2 - \int f_1 \, d\mu \int f_2 \, d\mu \right\|_{\mathcal{H}} = 0.$$

We will prove (taking – as we may – (i) for granted) that

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(iv)}.$$

Before giving the proof of (1.8) we want to make two agreements (which will be kept in § 2 as well). First of all, we will confine ourself to the case of real scalars; it is completely clear that this can be done without the loss of generality. Second,

if (s) denotes a statement of the form: 'f_n is strongly convergent to f' we will denote by $(s)_w$ the (weaker) statement 'f_n is weakly convergent to f'. We will prove actually the following:

$$(i)_w \Rightarrow (ii); \quad (ii)_w \Rightarrow (iii); \quad (iii)_w \Rightarrow (iv).$$

$(i)_w \Rightarrow (ii)$. Assume without the loss of generality that $\int f_1 \, d\mu = 0$ and denote $\prod_{i=1}^{4} T^{cn^2 + a_i n} f_i$ by x_n. We have to prove that

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} x_n \right\|_{\mathcal{H}} = 0.$$

By Theorem 1.4 it is enough to check that for $h \geq h_0 > 0$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h} \rangle = 0.$$

We have:

$$\langle x_n, x_{n+h} \rangle = \int \prod_{i=1}^{4} T^{cn^2 + a_i n} f_i \prod_{i=1}^{4} T^{c(n+h)^2 + a_i(n+h)} f_i \, d\mu$$

$$= \int T^{cn^2 + a_1 n} \left(\prod_{i=1}^{4} T^{(a_i - a_1)n} f_i \prod_{i=1}^{4} T^{2chn + (a_i - a_1)n + ch^2 + a_i h} f_i \right) d\mu$$

$$= \int f_1 \prod_{i=2}^{4} T^{(a_i - a_1)n} f_i \prod_{i=1}^{4} T^{(2ch + a_i - a_1)n} \tilde{f}_i \, d\mu, \tag{1.9}$$

where $\tilde{f}_i = T^{ch^2 + a_i h} f_i$, $i = 1, 2, 3, 4$.

It is clear that for sufficiently large h the powers of T appearing in (1.9) are pairwise distinct. Thus by $(i)_w$ (and by assumption $\int f_1 \, d\mu = 0$) we get:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h} \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int f_1 \prod_{i=2}^{4} T^{(a_i - a_1)n} f_i \prod_{i=1}^{4} T^{(2ch + a_i - a_1)n} \tilde{f}_i \, d\mu$$

$$= \prod_{i=1}^{4} \int f_i \, d\mu \prod_{i=1}^{4} \int T^{h^2 + a_i h} f_i \, d\mu = \prod_{i=1}^{4} \left(\int f_i \, d\mu \right)^2 = 0.$$

The application of Theorem 1.4 now gives (ii).

$(ii)_w \Rightarrow (iii)$. For the proof of this implication Theorem 1.5 is needed. Assuming without the loss of generality that $\int f_0 \, d\mu = 0$ and writing

$$x_n = T^{bn} f_0 T^{cn^2 + a_1 n} f_1 T^{cn^2 + a_2 n} f_2,$$

we get:

$$\langle x_n, x_{n+h} \rangle = \int T^{bn} f_0 T^{cn^2 + a_1 n} f_1 T^{cn^2 + a_2 n} f_2 T^{b(n+h)} f_0 T^{c(n+h)^2 + a_1(n+h)} f_1$$

$$\times T^{c(n+h)^2 + a_2(n+h)} f_2 \, d\mu$$

$$= \int T^{bn} (f_0 T^{bh} f_0 T^{cn^2 + (a_1 - b)n} f_1 T^{cn^2 + (a_2 - b)n} f_2 T^{cn^2 + (2ch + a_1 - b)n + ch^2 + a_1 h} f_1$$

$$\times T^{cn^2 + (2ch + a_2 - b)n + ch^2 + a_2 h} f_2) \, d\mu$$

$$= \int \tilde{f}_0 T^{cn^2 + (a_1 - b)n} f_1 T^{cn^2 + (a_2 - b)} f_2 T^{cn^2 + (2ch + a_1 - b)n} \tilde{f}_1 T^{cn^2 + (2ch^2 + a_2 - b)n} \tilde{f}_2 \, d\mu,$$

where $\tilde{f}_0 = f_0 T^{bh} f_0$, $\tilde{f}_1 = T^{ch^2 + a_1 h} f_1$, $\tilde{f}_2 = T^{ch^2 + a_2 h} f_2$.

It follows from (ii)$_w$ that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h}\rangle = \int f_0 T^{bh} f_0 \, d\mu \int f_1 \, d\mu \int f_2 \, d\mu \int T^{ch^2+a_1 h} f_1 \, d\mu$$

$$\times \int T^{ch^2+a_2 h} f_2 \, d\mu$$

$$= \left(\int f_0 T^{bh} f_0 \, d\mu\right)\left(\int f_1 \, d\mu\right)^2 \left(\int f_2 \, d\mu\right)^2.$$

By (1.7) and by assumption $\int f_0 \, d\mu = 0$ we get:

$$\text{D-}\lim_{h} \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h}\rangle = \left(\int f_0 \, d\mu\right)^2 \left(\int f_1 \, d\mu\right)^2 \left(\int f_2 \, d\mu\right)^2 = 0.$$

This proves (iii).

(iii)$_w \Rightarrow$(iv). Assuming $\int f_1 \, d\mu = 0$ and denoting $x_n = T^{n^2} f_1 T^{2n^2} f_2$ check (1.4):

$$\langle x_n, x_{n+h}\rangle = \int T^{n^2} f_1 T^{2n^2} f_2 T^{n^2+2nh+h^2} f_1 T^{2n^2+4nh+2h^2} f_2 \, d\mu$$

$$= \int f_1 T^{2nh} \tilde{f}_1 T^{n^2} f_2 T^{n^2+4nh} \tilde{f}_2 \, d\mu,$$

where

$$\tilde{f}_1 = T^{h^2} f_1, \qquad \tilde{f}_2 = T^{2h^2} f_2.$$

It follows from (iii)$_w$ that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h}\rangle = \int f_1 \, d\mu \int \tilde{f}_1 \, d\mu \int f_2 \, d\mu \int \tilde{f}_2 \, d\mu = 0.$$

Application of Theorem 1.4 now gives (iv).

2. Proof of Theorem 1.2

In this section we will assume that all the polynomials involved have zero constan
term. It is clear that such an assumption can be made without the loss of generality
On the other hand, this assumption will allow us to read in the formulation o
Theorem 1.2 'pairwise distinct polynomials' instead of 'pairwise essentially distinc
polynomials'. Let $F = \{p_1(t), \ldots, p_k(t)\}$ be a finite set of polynomials and assum
that the largest of the degrees of $p_i(t)$ equals d. In such a case we will write $\deg F = a$
We have: $F = \bigcup_{i=1}^{d} F_i$ where F_i denotes the subset of F consisting of polynomial
having degree i, $1 \le i \le d$. By our assumption $F_d \ne \varnothing$. Each F_i can be furthe
decomposed into the disjoint union

$$F_i = \bigcup_{j=1}^{n_i} F_j^{(i)},$$

where $F_j^{(i)}, j = 1, 2, \ldots, n_i$ have the following property: if $p(t) \in F_r^{(i)}$ and $q(t) \in F_s^{(i}$
then the leading coefficients of $p(t)$ and $q(t)$ are equal if and only if $r = s$. We hav
thus

$$F = \bigcup_{i=1}^{d} \bigcup_{j=1}^{n_i} F_j^{(i)}.$$

It follows that to any finite set of polynomials F one can attach a vector $(n_1, n_2, \ldots, n_d)$ where $d = \deg F$, $n_i \geq 0$ for $i = 1, 2, \ldots, d - 1$ and $n_d > 0$. The ith coordinate of this vector is equal to the number of different groups of polynomials of degree i (the groups are defined by the demand that polynomials $p(t)$ and $q(t)$ belong to the same group if and only if their degrees are equal and their leading coefficients coincide). We will say that $(n_1, n_2, \ldots, n_d)$ is the characteristic vector of F.

Let $\mathscr{F}(n_1, n_2, \ldots, n_d)$ be the family of all finite sets of polynomials having characteristic vector $(n_1, n_2, \ldots, n_d)$. Emphasizing on sets of polynomials for which Theorem 1.2 is valid, consider the following two statements:

$T(n_1, n_2, \ldots, n_d)$: 'Theorem 1.2 is valid for any $F \in \mathscr{F}(n_1, n_2, \ldots, n_d)$'.

$T(\overline{n_1, \ldots, n_i}, n_{i+1}, \ldots, n_d)$: '$T(n_1, n_2, \ldots, n_d)$ is valid for any $n_1, n_2, \ldots, n_i$'.

Noting that Theorem 1.2 is equivalent to the statement: '$T(\overline{n_1, n_2, \ldots, n_d})$ is valid for any d' and bearing in mind that for $d = 1$ and $n_1 = 1$, $T(n_1, n_2, \ldots, n_d)$ is a partial case of von Neumann's theorem, we see that Theorem 1.2 follows from the following implications:

$$\left. \begin{array}{cc} T(n_1) \Rightarrow T(n_1 + 1), & n_1 \geq 1 \\[2mm] T(\overline{n_1, n_2, \ldots, n_d}) \Rightarrow T(\overline{n_1, n_2, \ldots, n_{d+1}}), & d \geq 1 \end{array} \right\} . \tag{2.1}$$

We claim that in order to prove (2.1) it is enough to prove the following implications:

$$T(n_1, n_2, \ldots, n_d) \Rightarrow T(n_1 + 1, n_2, \ldots, n_d); \qquad n_1, \ldots, n_{d-1} \geq 0, n_d \geq 1, d \geq 1 \quad (2.2)_d$$

$$T(\overline{n_1, \ldots, n_{i-1}}, n_i, \ldots, n_d) \Rightarrow T(\underbrace{0, \ldots, 0}_{i-1 \text{ zeros}}, n_i + 1, n_{i+1}, \ldots, n_d),$$
$$n_1, \ldots, n_{d-1} \geq 0, n_d \geq 1, d \geq i > 1 \tag{$2.3)_{d,i}$}$$

$$T(\overline{n_1, \ldots, n_d}) \Rightarrow T(\underbrace{0, \ldots, 0}_{d \text{ zeros}}, 1), \qquad d \geq 1 \tag{$2.4)_d$}$$

Indeed, suppose that for some $d \geq 1$ $T(\overline{n_1, \ldots, n_d})$ is given. Then by $(2.4)_d$ and $(2.2)_{d+1}$ we see that $T(\underbrace{1, 0, \ldots, 0, 1}_{d+1 \text{ symbols}})$ is true. Recursive application of $(2.2)_{d+1}$ gives then $T(\underbrace{n_1, 0, \ldots, 0, 1}_{d+1 \text{ symbols}})$ for any n_1. Then by $(2.3)_{d+1,2}$ we get $T(\underbrace{0, 1, 0, \ldots, 0, 1}_{d+1 \text{ symbols}})$. Applying $(2.2)_{d+1}$ once again, we get $T(\underbrace{\bar{n}_1, 1, 0, \ldots, 1}_{d+1 \text{ symbols}})$ which in its turn (by $(2.3)_{d+1,2}$) gives $T(\underbrace{0, 2, 0, \ldots, 0, 1}_{d+1 \text{ symbols}})$. Repeating this procedure, we see that $T(\underbrace{0, n_2, 0, \ldots, 0, 1}_{d+1 \text{ symbols}})$ is true for any n_2. It follows then from $(2.2)_{d+1}$ that $T(\underbrace{\overline{n_1, n_2}, 0, \ldots, 0, 1}_{d+1 \text{ symbols}})$ is true. By the same procedure one gets $T(\underbrace{\overline{n_1, \ldots, n_i}, 0, \ldots, 0, 1}_{d+1 \text{ symbols}})$ for any $1 \leq i \leq d$. In particular, one gets $T(\overline{n_1, \ldots, n_d}, 1)$. By $(2.3)_{d+1,d}$ this gives $T(\underbrace{0, \ldots, 0, 2}_{d+1 \text{ symbols}})$. Analogously to the foregoing one arrives at $T(\overline{n_1, \ldots, n_d}, 2)$ and more generally to $T(\overline{n_1, \ldots, n_d}, n_{d+1})$ for any $n_{d+1} \geq 1$. This is equivalent to $T(\overline{n_1, \ldots, n_d, n_{d+1}})$ and gives (2.1). Thus, to finish the argument we have to show $(2.2)_d$, $(2.3)_{d,i}$ and $(2.4)_d$. We will prove $(2.3)_{d,i}$; the proofs of the other two implications are left to the reader (the proof of $(2.4)_d$ is the same as that of $(2.3)_{d,i}$; the proof of $(2.2)_d$ is completely analogous to that of the implication (ii) $\Rightarrow$ (iii) which was proved in the first section).

Before embarking on the proof of $(2.3)_d$ we shall formulate a lemma, the easy proof of which is left to the reader.

LEMMA. *If the polynomials $p_1(t), \ldots, p_k(t)$ are pairwise distinct then for sufficiently large (and fixed) h the polynomials $p_1(t), \ldots, p_k(t), p_1(t+h), \ldots, p_k(t+h)$ are pairwise distinct.*

Another important remark is that in proving (1.1) we can assume without the loss of generality that one of the functions has zero integral. This can be shown with the help of the following elementary identity (see 3.8 in [**FKO**]):

$$\prod_{i=1}^{k} a_i - \prod_{i=1}^{k} b_i = \sum_{j=1}^{k} \left(\prod_{i=1}^{j-1} a_i \right) (a_j - b_j) \left(\prod_{i=j+1}^{k} b_i \right), \qquad (2.5)$$

(where by definition $\prod_{i=1}^{0} a_i = 1$ and $\prod_{i=k+1}^{k} b_i = 1$). Putting $a_i = T^{p_i(n)} f_i$, $b_i = \int f_i \, d\mu$, we obtain from (2.5):

$$\frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{k} T^{p_i(n)} f_i - \prod_{i=1}^{k} \int f_i \, d\mu = \frac{1}{N} \sum_{n=1}^{N} \left(\prod_{i=1}^{k} T^{p_i(n)} f_i - \prod_{i=1}^{k} \int f_i \, d\mu \right)$$

$$= \sum_{i=1}^{k} \frac{1}{N} \sum_{n=1}^{N} \left\{ \prod_{i=1}^{j-1} T^{p_i(n)} f_i \left(T^{p_j(n)} \left(f_j - \int f_j \, d\mu \right) \right) \prod_{i=j+1}^{k} T^{p_i(n)} f_i \right\}. \qquad (2.6)$$

We see that by (2.6) the general case is reduced to the sum of expressions satisfying the above condition.

Suppose now that F is a finite set of pairwise distinct polynomials and assume that the characteristic vector of F equals $(\underbrace{0, \ldots, 0}_{i-1 \text{ zeros}}, n_i + 1, n_{i+1}, \ldots, n_d)$. Fix any of the $n_i + 1$ groups of polynomials of degree i and denote its polynomials by $p_1(t), \ldots, p_k(t)$. Denote the rest of the polynomials in F by $q_1(t), \ldots, q_l(t)$. Let $x_n = T^{p_1(n)} f_1 \ldots T^{p_k(n)} f_k T^{q_1(n)} g_1 \ldots T^{q_l(n)} g_l$, where $f_1, \ldots, f_k, g_1, \ldots, g_l \in L^\infty(X, \mathcal{B}, \mu)$. By the remark above we may assume without the loss of generality that one of the functions $f_1, \ldots, f_k, g_1, \ldots, g_l$ has zero integral. With this assumption we have to show that

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} x_n \right\|_{L^2(X)} = 0.$$

This will follow from Theorem 1.5 the conditions of which we will verify now. We have

$$\langle x_n, x_{n+h} \rangle = \int \prod_{j=1}^{k} T^{p_j(n)} f_j \prod_{j=1}^{l} T^{q_j(n)} g_j \prod_{j=1}^{k} T^{p_j(n+h)} f_j \prod_{j=1}^{l} T^{q_j(n+h)} g_j \, d\mu$$

$$= \int T^{p_1(n)} \left(f_1 \prod_{j=2}^{k} T^{p_j(n)-p_1(n)} f_j \prod_{j=1}^{l} T^{q_j(n)-p_1(n)} g_j \right.$$

$$\times \left. \prod_{j=1}^{k} T^{p_j(n+h)-p_1(n)} f_j \prod_{j=1}^{l} T^{q_j(n+h)-p_1(n)} g_j \right) d\mu$$

$$= \int f_1 \prod_{j=2}^{k} T^{\tilde{p}_j(n)} f_j \prod_{j=1}^{l} T^{\tilde{q}_j(n)} g_j \prod_{j=1}^{k} T^{\tilde{\tilde{p}}_j(n)} \tilde{f}_j \prod_{j=1}^{l} T^{\tilde{\tilde{q}}_j(n)} \tilde{g}_j \, d\mu \qquad (2.7)$$

where

$$\tilde{p}_j(n) = p_j(n) - p_1(n), j = 2, \ldots, k$$
$$\tilde{q}_j(n) = q_j(n) - p_1(n), j = 1, 2, \ldots, l$$
$$\bar{p}_j(n) = p_j(n+h) - p_j(h) - p_1(n), j = 1, 2, \ldots, k$$
$$\bar{q}_j(n) = q_j(n+h) - q_j(h) - p_1(n), j = 1, 2, \ldots, l$$
$$\tilde{f}_j = T^{p_j(h)}f_j, j = 1, 2, \ldots, k$$
$$\tilde{g}_j = T^{q_j(h)}g_j, j = 1, 2, \ldots, k.$$

By the lemma above for any sufficiently large (and fixed) h the polynomials appearing in (2.7) are pairwise distinct. The degrees of polynomials $\tilde{p}_j$ and $\bar{p}_j$ are smaller than i (and ≥ 1), whereas deg $\tilde{q}_j$ = deg $\bar{q}_j$ = deg q_j for $j = 1, 2, \ldots, l$; the leading coefficients of those of $\tilde{q}_j$ and $\bar{q}_j$ which have degree i, have the form $c_q - c_p$ where c_p is the leading coefficient of p_1 and are different from those of any of $\tilde{q}_j$ and $\bar{q}_j$.

Note also that if the degree of q_j is larger than i, then not only the degree but also the leading coefficients of $\tilde{q}_j$ and $\bar{q}_j$ are the same as those of q_j. It follows that the characteristic vector of the set $\{\bar{p}_2, \ldots, \bar{p}_k, \tilde{p}_1, \ldots, \bar{p}_k, \tilde{q}_1, \ldots, \tilde{q}_l, \bar{q}_1, \ldots, \bar{q}_l\}$ has the form $(a_1, a_2, \ldots, a_{i-1}, n_i, n_{i+1}, \ldots, n_d)$. Applying $(T(\overline{n_1, \ldots, n_{i-1}}, n_i, \ldots, n_d))_w$ and using the fact that one of the functions f_j, g_j has zero integral, we get

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h} \rangle$$

$$= \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int f_1 \prod_{j=2}^{k} T^{\bar{p}_j(n)}f_j \prod_{j=1}^{l} T^{\tilde{q}_j(n)}g_j \prod_{j=1}^{k} T^{\tilde{p}_j(n)}\tilde{f}_j \prod_{j=1}^{l} T^{\bar{q}_j(n)}\tilde{g}_j \, d\mu$$

$$= \prod_{j=1}^{k} \int f_j \, d\mu \prod_{j=1}^{l} \int g_j \, d\mu \prod_{j=1}^{k} \int \tilde{f}_j \, d\mu \prod_{j=1}^{l} \int \tilde{g}_j \, d\mu = 0.$$

The application of Theorem 1.4 completes the proof.

3. *Uniform weakly mixing PET and the 'polynomial mixing of all orders' in weakly mixing systems*

The following uniform version of weakly mixing PET bears the same relation to Theorem 1.2 as the notion of well distribution to that of uniform distribution.

THEOREM 3.1. *Suppose that* $(X, \mathcal{B}, \mu, T)$ *is a weakly mixing system and let* $p_1(t), \ldots, p_k(t)$ *be pairwise essentially distinct polynomials with rational coefficients taking on integer values on the integers. Then for any* $f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$ *one has*:

$$\lim_{N-M \to \infty} \left\| \frac{1}{N-M} \sum_{n=M}^{N-1} T^{p_1(n)}f_1 T^{p_2(n)}f_1 \cdots T^{p_k(n)}f_k - \prod_{i=1}^{k} \int f_i \, d\mu \right\|_{L^2(X)} = 0.$$

The proof of Theorem 3.1 is analogous to that of Theorem 1.2. One has just to check that the uniform limits can be pushed through the inductive steps. The following two statements, the proofs of which are left to the reader, play the role which was played by Theorem 1.4 and Theorem 1.5 in the proof of Theorem 1.2.

THEOREM 3.2. *Suppose that* $\{x_n\}_{n=-\infty}^{\infty}$ *is a bounded sequence of vectors in Hilbert space, and suppose that for any* $h \geq h_0 > 0$

$$\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \langle x_n, x_{n+h} \rangle = 0.$$

Then

$$\lim_{N-M \to \infty} \left\| \frac{1}{N-M} \sum_{n=M}^{N-1} x_n \right\| = 0.$$

THEOREM 3.3. *Suppose that* $\{x_n\}_{n=-\infty}^{\infty}$ *is a bounded sequence of vectors in Hilbert space. If*

$$\text{D-}\lim_{h} \lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \langle x_n, x_{n+h} \rangle = 0,$$

then

$$\lim_{N-M \to \infty} \left\| \frac{1}{N-M} \sum_{n=M}^{N-1} x_n \right\| = 0.$$

As an illustration let us prove the uniform version of Theorem 1.1 with the help of Theorem 3.3. We have to show that if $(X, \mathcal{B}, \mu, T)$ is a weakly mixing system, then for any $k \in \mathbb{N}$, any pairwise distinct non-zero integers $a_1, \ldots, a_k$ and any $f_1, \ldots, f_k \in L^{\infty}(X, \mathcal{B}, \mu)$ one has:

$$\lim_{N-M \to \infty} \left\| \frac{1}{N-M} \sum_{n=M}^{N-1} T^{a_1 n} f_1 T^{a_2 n} f_2 \cdots T^{a_k n} f_k - \prod_{i=1}^{k} \int f_i \, d\mu \right\|_{L^2(X)} = 0. \qquad (3.1)_k$$

Note that $(3.1)_1$ follows from the von Neumann's uniform ergodic theorem. Suppose that the assertion in question is valid for $k-1$. We shall establish it for k. Taking any k non-zero integers $a_1, \ldots, a_k$ and any $f_1, \ldots, f_k \in L^{\infty}(X, \mathcal{B}, \mu)$ and assuming without the loss of generality that $\int f_1 \, d\mu = 0$, let

$$x_n = T^{a_1 n} f_1 T^{a_2 n} f_2 \cdots T^{a_k n} f_k.$$

We have

$$\langle x_n, x_{n+h} \rangle = \int T^{a_1 n} f_1 T^{a_2 n} f_2 \cdots T^{a_k n} f_k T^{a_1(n+h)} f_1 T^{a_2(n+h)} f_2 \cdots T^{a_k(n+h)} f_k \, d\mu$$

$$= \int T^{a_1 n}(f_1 T^{a_1 h} f_1) T^{a_2 n}(f_2 T^{a_2 h} f_2) \cdots T^{a_k n}(f_k T^{a_k h} f_k) \, d\mu$$

$$= \int (f_1 T^{a_1 h} f_1) T^{(a_2 - a_1)n}(f_2 T^{a_2 h} f_2) \cdots T^{(a_k - a_1)n}(f_k T^{a_k h} f_k) \, d\mu.$$

By the induction hypothesis, we get:

$$\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \langle x_n, x_{n+h} \rangle = \left(\int f_1 T^{a_1 h} f_1 \, d\mu \right) \left(\prod_{i=2}^{k} \int f_i T^{(a_i - a_1)h} f_i \right).$$

Using the fact that T is weakly mixing, we finally get:

$$\text{D-}\lim_{h} \lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \langle x_n, x_{n+h} \rangle = \prod_{i=1}^{k} \left(\int f_i \, d\mu \right)^2 = 0.$$

Thus by Theorem 3.3 formula (3.1) holds for k. In similar fashion, one checks the uniform analogs of formulas (2.1), (2.2), (2.3), (2.4). We omit the details.

We are going now to apply Theorem 3.1 to the multiple recurrence in weakly mixing systems.

Definition 3.4. The uniform density $d_u(C)$ of a set $C \subset \mathbb{Z}$ is

$$\lim_{N-M \to \infty} \frac{|C \cap \{M, M+1, \ldots, N-1\}|}{N-M}$$

provided this limit exists.

Definition 3.5 (cf. Definition 4.2, page 84 in [**F2**]). A bilateral sequence $\{x_n\}_{n=-\infty}^{\infty}$ of points in a topological space X converges to a point $x \in X$ uniformly in density if for every neighbourhood V of x in X the set of n for which $x_n \in V$ has uniform density one. In this case we will write UD-$\lim_n x_n = x$. We will need the following technical lemma.

LEMMA 3.6. *If* $\{a_n\}_{n=-\infty}^{\infty}$ *is a bilateral sequence of real numbers with*

$$\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} a_n = \alpha, \qquad \lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} a_n^2 = \alpha^2,$$

then

$$\text{UD-}\lim_n a_n = \alpha. \tag{3.2}$$

Proof. It follows from the assertion of the lemma that

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

which is obviously equivalent to (3.2).

THEOREM 3.7. *If* $(X, \mathcal{B}, \mu, T)$ *is a weakly mixing system and* $p_1(t), \ldots, p_k(t)$ *are pairwise essentially distinct polynomials with rational coefficients taking on integer values on the integers, then for any* $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$

$$\text{UD-}\lim_n \int f_0 T^{p_1(n)} f_1 T^{p_2(n)} f_2 \cdots T^{p_k(n)} f_k \, d\mu = \prod_{i=0}^{k} \int f_i \, d\mu.$$

Proof. It follows from theorem 3.1 that

$$\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \int f_0 T^{p_1(n)} f_1 T^{p_2(n)} f_2 \cdots T^{p_k(n)} f_k \, d\mu = \prod_{i=0}^{k} \int f_i \, d\mu.$$

Applying Theorem 3.1 once again, this time to the product transformation $T \times T$ and to the functions $\tilde{f}_i(x, y) = f_i(x) f_i(y)$ we get

$$\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \left\{ \int f_0 T^{p_1(n)} f_1 T^{p_2(n)} f_2 \cdots T^{p_k(n)} f_k \, d\mu \right\}^2 = \prod_{i=0}^{k} \left(\int f_i \, d\mu \right)^2.$$

The desired result follows now from Lemma 3.5. Applying Theorem 3.6 to characteristic functions $f_i = 1_{A_i}$, we get the following result which might be regarded as the statement that weakly mixing systems are 'uniformly polynomially weakly mixing of all orders' (cf. Th. 4.11 in [**F2**]).

THEOREM 3.8. *If* $(X, \mathcal{B}, \mu, T)$ *is a weakly mixing system,* $A_0, A_2, \ldots, A_k$ *are* $k+1$ *sets in* $\mathcal{B}$ *and* $p_1(t), \ldots, p_k(t)$ *are pairwise distinct polynomials with rational coefficients taking on integer values on the integers, then for any* $\varepsilon > 0$ *the set*

$$\{n: |\mu(A_0 \cap T^{p_1(n)}A_1 \cap T^{p_2(n)}A_2 \cap \cdots \cap T^{p_k(n)}A_k) - \mu(A_0)\mu(A_1) \cdots \mu(A_k)| < \varepsilon\}$$

has uniform density one.

4. Relatives of weakly mixing PET

Theorem 1.2 is a representative of a large variety of PETs. In this section two additional examples will be given (without proof). The proofs will appear elsewhere.

(i) *Definition* 4.1 ([**F2**, p. 86]). Let G be a commutative group of measure preserving transformations of a measure space $(X, \mathcal{B}, \mu)$. G is called totally weakly mixing if each $T \in G$, $T \neq$ identity, is weakly mixing.

It is known ([**F2**, Th. 4.10]; see also [**FK1**, Th. 1.4]) that if G is a totally weakly mixing group of transformations of a probability space $(X, \mathcal{B}, \mu)$ then for any distinct $T_1, T_2, \ldots, T_k \in G$ and any $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$ one has:

$$\text{D-}\lim_n \int f_0 T_1^n f_1 T_2^n f_2 \cdots T_k^n f_k \, d\mu = \prod_{i=1}^k \int f_i \, d\mu. \tag{4.1}$$

The following theorem is a generalization of (4.1).

THEOREM 4.2. *Suppose that* $T_1, T_2, \ldots, T_k$, $T_i \neq$ *identity, are distinct elements of a totally weakly mixing group of transformations of a probability space* $(X, \mathcal{B}, \mu)$. *Then for any polynomial* $p(t)$ *which takes on integer values on the integers and for any* $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$ *one has:*

$$\text{UD-}\lim_n \int f_0 T_1^{p(n)} f_1 T_2^{p(n)} f_2 \cdots T_k^{p(n)} f_k \, d\mu = \prod_{i=0}^k \int f_i \, d\mu.$$

Remark. While generalizing (4.1), Theorem 4.2 is not a generalization of Theorem 1.2. However, it is possible to give an extension of Theorem 4.2 which will include Theorem 1.2 as a special case.

(ii) Our second example is mildly mixing PET. The notion of mild mixing was introduced in [**FW**] and plays the crucial role in the proof of the Furstenberg-Katznelson ergodic theorem for commuting IP-systems ([**FK2**]). We will give here only that minimal amount of facts about mild mixing which is necessary for the formulation of mildly mixing PET (Theorem 4.8 below). For a comprehensive account the reader is referred to [**FW**], [**F2**], [**F3**], [**FK2**].

Definition 4.3. Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. A function $f \in L^2(X, \mathcal{B}, \mu)$ is rigid for T if for some sequence $n_k \geq 1$, $T^{n_k}f \to f$ strongly in $L^2(X, \mathcal{B}, \mu)$.

Definition 4.4. A measure preserving system $(X, \mathcal{B}, \mu, T)$ is mildly mixing if there are no non-constant rigid functions in $L^2(X, \mathcal{B}, \mu)$.

Definition 4.5. A set of positive integers is called an IP-set if there exists a sequence $n_1, n_2, \ldots \in \mathbb{N}$ such that the set in question consists of the numbers n_i together with all finite sums $n_{i_1} + n_{i_2} + \cdots + n_{i_k}$ with $i_1 < i_2 < \cdots < i_k$.

Definition 4.6. A set $C \subset \mathbb{N}$ is called an IP*-set if it intersects each IP-set non-trivially.

Definition 4.7. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of points in a topological vector space X. A point $x \in X$ is said to be an IP*-limit of $\{x_n\}_{n=1}^{\infty}$ if for every neighbourhood V of x the set of n for which $x_n \in V$ is an IP*-set. In this case, we will write: $\text{IP*-lim}_n x_n = x$.

THEOREM 4.8. *Suppose* $(X, \mathcal{B}, \mu, T)$ *is a mildly mixing system and let* $p_1(t), \ldots, p_k(t)$ *be pairwise essentially distinct polynomials which take on integer values on the integers. Then for any* $f_0, f_1, \ldots, f_k \in L^{\infty}(X, \mathcal{B}, \mu)$ *one has*:

$$\text{IP*-lim}_n \int f_0 T^{p_1(n)} f_1 T^{p_2(n)} f_2 \cdots T^{p_k(n)} f_k \, d\mu = \prod_{i=0}^{k} \int f_i \, d\mu.$$

We leave to the reader the formulation of the mildly mixing analog of Theorem 4.2.

Many other PETs are beyond the scope of this paper. The author plans to take care of them in some other papers.

REFERENCES

[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, Princeton, New Jersey, 1981.

[F3] H. Furstenberg. IP-systems in ergodic theory. *Contemporary Mathematics* **26** (1984), 131–148.

[FK1] H. Furstenberg & Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. d'Analyse Math.* **34** (1978), 275–291.

[FK2] H. Furstenberg & Y. Katznelson. An ergodic Szemerédi theorem for IP-systems and combinatorial theory. To appear in *J. d'Analyse Math.*

[FKO] H. Furstenberg, Y. Katznelson & D. Ornstein. The ergodic theoretical proof of Szemerédi's theorem. *Bull. Amer. Math. Soc.* **7** (1982), 527–552.

[FW] H. Furstenberg & Weiss. The finite multipliers of infinite ergodic transformations. In *Structure of Attractors in Dynamical Systems*, Springer Lecture Notes #668, Springer (1978), 127–133.

[KN] L. Kuipers & H. Niederreiter. *Uniform Distribution of Sequences.* Wiley, New York, 1974.