Pointwise characteristic factors for the multidimensional return times theorem

Idris Assani

Department of Mathematics, UNC Chapel Hill, NC 27599

assani@math.unc.edu

Kimberly Presser

Department of Mathematics, Shippensburg University, PA 27517

kjpres@ship.edu

December 15, 2010

Abstract

This paper is an update and extension of a result the authors first proved in 2003. The goal of this paper is to study factors which are known to be L^2-characteristic for certain nonconventional averages and prove that these factors are pointwise characteristic for the multidimensional return times averages.

In memory of Dan Rudolph.

1 Introduction

A major result in ergodic theory in the late 1980's was the proof of the return times theorem by J. Bourgain [8] (which was later simplified by J. Bourgain, H. Furstenberg, Y. Katnzelson, D. Ornstein in [9]). This theorem created a key strengthening of the Birkhoff's Pointwise Ergodic Theorem [7].

Theorem 1. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic dynamical system of finite measure and $f \in L^\infty(\mu)$. Then there exists a set $X_f \subset X$ of full measure such that for any other ergodic dynamical system $(Y, \mathcal{G}, \nu, S)$ with $\nu(Y) < \infty$ and any $g \in L^\infty(\nu)$:*

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n y)$$

converges ν-a.e. for all $x \in X_f$.

In the BFKO proof [9] of the return times theorem, one of the keys to the argument was to decompose the given function using the Kronecker factor in order to prove the result independently for both the eigenfunctions and those functions in the orthocomplement of the Kronecker factor.

Using factors in convergence proofs in ergodic theory has long been a very useful tool. The notion of a characteristic factor is originally due to H. Furstenberg and is explicitly defined by H. Furstenberg and B. Weiss in [11].

Definition 1. *When the limiting behavior of a non-conventional ergodic average for $(X, \mathcal{F}, \mu, T)$ can be reduced to that of a factor system $(Y, \mathcal{G}, \nu, T)$, we shall say that the latter is a **characteristic factor** of the former.*

For each type of average under consideration, one will have to specify what is meant by reduced in the given case. In the case of H. Furstenberg and B. Weiss [11], they define the notion of characteristic factor for averages of the type

$$\frac{1}{N}\sum_{n=1}^{N}\left(f\circ T^{n}\right)\left(g\circ T^{n^2}\right).$$

Therefore their specific definition of characteristic factor is as follows.

Definition 2. *If $\{p_1(n), p_2(n), \ldots, p_k(n)\}$ are k integer-valued sequences, and $(Y, \mathcal{G}, \nu, T)$ is a factor of a system $(X, \mathcal{F}, \mu, T)$, we say that $\mathcal{G}$ is a* ***characteristic factor for the scheme*** *$\{p_1(n), p_2(n), \ldots, p_k(n)\}$, if for any $f_1, f_2, \ldots, f_k \in L^\infty(\mu)$ we have*

$$\frac{1}{N}\sum_{n=1}^{N}\left[f_1\circ T^{p_1(n)}\cdots f_k\circ T^{p_k(n)} - \mathbb{E}(f_1|\mathcal{G})\circ T^{p_1(n)}\cdots \mathbb{E}(f_k,\mathcal{G})\circ T^{p_k(n)}\right]$$

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

In 1998, D. Rudolph [16] extended the return times theorem to averages with more than two terms with his proof of the multidimensional return times theorem. His proof answered one of the questions on the return times raised by I. Assani[1] who proved the same result for weakly mixing systems in [1].

Theorem 2 (Multidimensional Return Times Theorem)**.** *Let k be any positive integer. For any dynamical system $(X, \mathcal{F}, T, \mu)$ and any $f \in L^\infty(\mu)$, there exists a set of full measure X_f in X such that if $x \in X_f$ for any other dynamical system $(Y_1, \mathcal{G}_1, S_1, \nu_1)$ and any $g_1 \in L^\infty(\nu_1)$ there exists a set of full*

[1] These questions were brought up during D. Rudolph's visit to UNC-CH in 1991 while he was working on his joinings proof of Bourgain's return times theorem [14].

measure Y_{g_1} in Y_1 such that if $y_1 \in Y_{g_1}$ then ... for any other dynamical system $(Y_{k-1}, \mathcal{G}_{k-1}, S_{k-1}, \nu_{k-1})$ and any $g_{k-1} \in L^\infty(\nu_{k-1})$ there exists a set of full measure $Y_{g_{k-1}}$ in Y_{k-1} such that if $y_{k-1} \in Y_{g_{k-1}}$ for any other dynamical system $(Y_k, \mathcal{G}_k, S_k, \nu_k)$ the average:

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S_1^n y_1) g_2(S_2^n y_2) \cdots g_k(S_k^n y_k)$$

converges ν_k-a.e..

D. Rudolph's proof of the multidimensional return times theorem utilized the method of joinings and fully generic sequences. This led to an elegant proof of the theorem which avoided the study of the factor of the σ-algebra which was characteristic for the averages. So the higher order version of the Kronecker factor $\mathcal{K}$ which had been key to the BFKO [9] proof was not needed in D. Rudolph's argument. This paper seeks to determine what factors serve a role similar to the Kronecker factor $\mathcal{K}$ in this higher dimensional setting.

For our purposes we define the notion of pointwise characteristic factors for the multidimensional return times averages as follows.

Definition 3. *Consider $(X, \mathcal{F}, \mu, T)$ a measure preserving system. The factor $\mathcal{A}$ is **pointwise characteristic for the k-th return times averages** if for each $f \in L^\infty(\mu)$ we can find a set of full measure X_f such that for each $x \in X_f$, for any other dynamical system $(Y_1, \mathcal{G}_1, S_1, \nu_1)$ and any $g_1 \in L^\infty(\nu_1)$, there exists a set of full measure Y_{g_1} such that for each y_1 in Y_{g_1} then ... for any other dynamical system $(Y_{k-1}, \mathcal{G}_{k-1}, S_{k-1}, \nu_{k-1})$ and any $g_{k-1} \in L^\infty(\nu_{k-1})$, there exist a set of full measure $Y_{g_{k-1}}$ in Y_{k-1} such that if $y_{k-1} \in Y_{g_{k-1}}$ for any* *other dynamical system* $(Y_k, \mathcal{G}_k, S_k, \nu_k)$ *for* ν_k*-a.e.* y_k *the average*

$$\frac{1}{N}\sum_{n=1}^{N}\left[f(T^n x) - \mathbb{E}(f|\mathcal{A})(T^n x)\right] g_1(S_1^n y_1) g_2(S_2^n y_2)\cdots g_k(S_k^n y_k)$$

converges to 0.

In looking for potential characteristic factors for the general multidimensional return times averages we consider the factors first used by H. Furstenberg to prove Szemeredi's Theorem [10]. We denote these factors as $\mathcal{A}_k$ using the notation from [3] where these factors were shown to be L^2-characteristic for the averages

$$\frac{1}{N}\sum_{n=1}^{N}\prod_{i=1}^{I} f_i \circ T^{in}.$$

As noted in the introduction to [4], while the norm convergence of averages for L^2-characteristic factors can sometimes lead to pointwise characteristic properties, this is not always guaranteed to be the case. Thus it is of consequence to look at pointwise convergence in addition to investigating factors with respect to the norm convergence.

We will show that these $\mathcal{A}_k$ factors can be defined in an inductive way by seminorms using Lemma 1.3 of [15][2] Using these seminorms we will prove our first main result.

Theorem 3. *The factors* $\mathcal{A}_k$ *are pointwise characteristic for the multidimensional return times averages*

[2]This approach was used in two 2003 unpublished papers of the first author ([3] and what was ultimately combined into the published paper [4]). The first author thanks C. Demeter and N. Frantzikinakis for pointing out to him that the factors he defined with these seminorms were in fact the ones introduced by H. Furstenberg in [10].

The study of the nonconventional Furstenberg averages has seen important progress being made in the last seven years. In [12] and [17] the Host-Kra-Ziegler factors $\mathcal{Z}_k$ were created independently by B. Host, B. Kra and T. Ziegler and were shown to be characteristic in L^2 norm for the Furstenberg averages. Using these factors we prove our second main result.

Theorem 4. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving system. The Host-Kra-Ziegler factors $\mathcal{Z}_k$ are pointwise characteristic for the multidimensional return times averages.*

As the $\mathcal{Z}_k$ factors are smaller than the factors $\mathcal{A}_k$, and thus $\mathcal{A}_k^{\perp} \subseteq \mathcal{Z}_k^{\perp}$, Theorem 3 is a consequence of Theorem 4. But our Theorem 3 with the use of the seminorm defining the factors $\mathcal{A}_k$ gives a different set of information. More precisely, using the factors $\mathcal{A}_k$ we obtain pointwise uniform bounds of the multidimensional return times averages. With the $\mathcal{Z}_k$ factors we do not have such pointwise estimates. The uniform upper bounds are derived after integration combined with a $\limsup$ argument.

2 The $\mathcal{A}_k$ factors are pointwise characteristic for the multidimensional return times averages

Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system on a probability measure space.

Definition 4. *The factors $\mathcal{A}_k$ are defined in the following inductive way.*

- *The factor $\mathcal{A}_0$ is equal to the trivial σ-algebra $\{X, \emptyset\}$*

- *For $k \geq 0$ the factor $\mathcal{A}_{k+1}$ is characterized by the following. A function $f \in \mathcal{A}_{k+1}^{\perp}$ if and only if*

$$N_{k+1}(f)^4 := \lim_H \frac{1}{H} \sum_{h=1}^{H} \left\| \mathbb{E}(f \cdot f \circ T^h | \mathcal{A}_k) \right\|_2^2 = 0$$

Note that the factor $\mathcal{A}_1$ is the Kronecker factor of our ergodic operator T because

$$N_1(f)^4 = \lim_H \frac{1}{H} \sum_{h=1}^{H} \left\| \mathbb{E}(f \cdot f \circ T^h | \mathcal{A}_0) \right\|_2^2 = \lim_H \frac{1}{H} \sum_{h=1}^{H} \left| \int f \cdot f \circ T^h d\mu \right|^2.$$

The next lemma shows that these seminorms are well-defined and characterize factors of T. It combines some results in the unpublished preprint [3].

Lemma 1. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic dynamical system on a probability measure space, For $k \geq 2$ for each function $f \in L^\infty(\mu)$ the quantities $N_k(f)$ are well defined. Furthermore, they characterize factors of T which are successive maximal isometric extensions.*

Proof. Let us consider a general factor $\mathcal{A}$ of T and $\mathbb{E}(\cdot, \mathcal{A})$ the projection onto this factor. The relatively independent joining of $T \times T$ over the factor $\mathcal{A}$ is the measure $\mu_{\mathcal{A}}$ defined for f, g bounded functions as

$$\int f \times g d\mu_{\mathcal{A}} := \int \mathbb{E}(f, \mathcal{A}) \mathbb{E}(g, \mathcal{A}) d\mu.$$

By Birkhoff's ergodic theorem applied to $T \times T$ and the invariant measure $\mu_{\mathcal{A}}$ we have

$$\begin{aligned} \lim_H \frac{1}{H} \sum_{h=1}^{H} \left\| \mathbb{E}(f \cdot f \circ T^h, \mathcal{A}) \right\|_2^2 &= \lim_H \frac{1}{H} \sum_{h=1}^{H} \int (f \cdot f \circ T^h)(x)(f \cdot f \circ T^h)(y) d\mu_{\mathcal{A}} \\ &= \left\| \mathbb{E}(f \times f, \mathcal{I}_{\mathcal{A}}) \right\|_{L^2(\mu_{\mathcal{A}})}^2. \end{aligned}$$

where $\mathcal{I}_{\mathcal{A}}$ is the $T\times T$-$\mu_{\mathcal{A}}$ invariant σ-algebra. If we denote by $N(f)$ the quantity

$$N(f)^4 = \lim_H \frac{1}{H}\sum_{h=1}^{H} \left\|\mathbb{E}(f\cdot f\circ T^h, \mathcal{A})\right\|_2^2$$

then Lemma 1.3 in [15] tells us that $N(f)=0$ if and only if $\mathbb{E}(f,\mathcal{K}\mathcal{A})=0$ where $\mathcal{K}\mathcal{A}$ is the maximal isometric extension of $\mathcal{A}$.

Using these observations one can characterize the successive maximal isometric extensions. The trivial σ-algebra is $\mathcal{A}_0$. Then we define $\mathcal{A}_1 = \mathcal{K}\mathcal{A}_0$, $\mathcal{A}_2 = \mathcal{K}\mathcal{A}_1$ and more generally $\mathcal{A}_{k+1} = \mathcal{K}\mathcal{A}_k$. The seminorms characterizing these factors are well defined as $N_k(f)$ where

$$N_k(f)^4 = \lim_H \frac{1}{H}\sum_{h=1}^{H} \left\|\mathbb{E}(f\cdot f\circ T^h, \mathcal{A}_k)\right\|_2^2 .$$

□

In order to simplify the inductive parts of our argument, we first clarify the techniques that we will use in a series of small lemmas. This next lemma relies on an application of the spectral theorem which allows us to alternate between Wiener-Wintner and return times averages in our inductive argument.

Lemma 2. *Let $\{a_n\}$ be a sequence of complex numbers. If*

$$\sup_N \frac{1}{N}\sum_{n=1}^{N} |a_n|^2 < \infty \text{ and } \sup_\epsilon \left|\frac{1}{N}\sum_{1}^{N} a_n e^{2\pi i n\epsilon}\right| \to 0,$$

then

$$\frac{1}{N}\sum_{1}^{N} a_n g(S^n y) \to 0$$

in $L^2(\nu)$ for all measure-preserving systems $(Y,\mathcal{G},S,\nu)$.

Proof. This follows immediately from the proof of Theorem 3.1 in [2]. □

Next, we will use the following lemma which is an easy consequence of the Van der Corput Lemma [13]. It will help us simplify the Wiener-Wintner averages which will appear in the inductive argument.

Lemma 3. *Let $\{a_n\}$ be a bounded sequence of complex numbers. Then*

$$\sup_{\epsilon}\left|\frac{1}{N}\sum_{n=1}^{N} a_n e^{2\pi i n\epsilon}\right|^2 \leq C\left(\frac{1}{H}+\frac{1}{H}\sum_{h=1}^{H}\left|\frac{1}{N}\sum_{n=1}^{N-h} a_n\overline{a_{n+h}}\right|\right)$$

for some constant C and $1 \leq H \leq N$.

We will prove our main result, Theorem 3, in the course of proving the following more detailed statement.

Theorem 5. *Let k be any positive integer. For each $f \in L^\infty(\mu)$ we can find a set of full measure X_f such that for each $x \in X_f$, for any other dynamical system $(Y_1, \mathcal{G}_1, S_1, \nu_1)$ and any $g_1 \in L^\infty(\nu_1)$ with $\|g_1\|_\infty \leq 1$, there exists a set of full measure Y_{g_1} such that for each y_1 in Y_{g_1} then ... for any other dynamical system $(Y_{k-1}, \mathcal{G}_{k-1}, S_{k-1}, \nu_{k-1})$ and any $g_{k-1} \in L^\infty(\nu_{k-1})$ with $\|g_k\|_\infty \leq 1$ there exist a set of full measure $Y_{g_{k-1}}$ in Y_{k-1} such that if $y_{k-1} \in Y_{g_{k-1}}$ for any other dynamical system $(Y_k, \mathcal{G}_k, S_k, \nu_k)$ for ν_k-a.e. y_k the average*

$$\frac{1}{N}\sum_{n=1}^{N}\left[f(T^n x) - \mathbb{E}(f|\mathcal{A}_k)(T^n x)\right] g_1(S_1^n y_1) g_2(S_2^n y_2)\cdots g_k(S_k^n y_k)$$

converges to 0. Thus for $f \in \mathcal{A}_k^\perp$ the average

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S_1^n y_1) g_2(S_2^n y_2)\cdots g_k(S_k^n y_k)$$

converges to 0 ν_k-a.e..

Proof. The basis step for the induction was done in the BFKO [9] proof of Bourgain's Return Times Theorem. Here it was shown that $\mathcal{A}_1 = \mathcal{K}$ was pointwise characteristic for averages of the type

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S_1^n y_1)$$

By using bases e_j and γ_j of eigenfunctions for T and S_1 of modulus 1, respectively, one can even evaluate the limit, assuming for the moment that S_1 is ergodic. The limit is equal to

$$\sum_{s \in E_{T,S}} \left(\int f \cdot \overline{e_s} d\mu \right) e_s(x) \left(\int g_1 \cdot \overline{\gamma_s} d\nu \right) \gamma_s(y)$$

where $E_{T,S}$ is the subset of common eigenvalues for T and S. By using the Cauchy-Schwartz inequality one gets the upper bound

$$\left(\sum_{j=0}^{\infty} \left| \int f \cdot \overline{e_j} d\mu \right|^2 \right)^{1/2} \left(\sum_{k=0}^{\infty} \left| \int g_1 \cdot \overline{\gamma_k} d\nu \right|^2 \right)^{1/2}$$

which is equal to

$$\|\mathbb{E}(f|\mathcal{A}_1)\|_2 \|\mathbb{E}(g|\mathcal{K}_{S_1})\|_2,$$

where $\mathcal{K}_{S_1}$ is the Kronecker factor for S_1. This last term is itself less than $\|\mathbb{E}(f|\mathcal{A}_1)\|_2$ because $\|g\|_\infty \leq 1$. We can remove the ergodic assumption on S_1 by using the ergodic decomposition. Thus we have reached the inequality

$$\limsup_N \left| \frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S^n y) \right|^2 \leq \|\mathbb{E}(f|\mathcal{A}_1)\|_2^2 \qquad (1)$$

which shows clearly that $\mathcal{A}_1$ is pointwise characteristic for the return times average.

Assume that for any $f \in L^\infty(\mu)$ and $1 \leq j < k$ we can find sets $X_{f,j}$ of full measure such that if $x \in X_{f,j}$, then for any other dynamical system

$(Y_1, \mathcal{G}_1, S_1, \nu_1)$ and any $g_1 \in L^\infty(\nu_1)$ with $\|g_1\|_\infty \leq 1$, there exists a set of full measure Y_{g_1} such that for each y_1 in Y_{g_1} then … for any other dynamical system $(Y_{j-1}, \mathcal{G}_{j-1}, S_{j-1}, \nu_{j-1})$ and any $g_{j-1} \in L^\infty(\nu_{j-1})$ with $\|g_{j-1}\|_\infty \leq 1$ there exist a set of full measure $Y_{g_{j-1}}$ in Y_{j-1} such that if $y_{j-1} \in Y_{g_{j-1}}$ for any other dynamical system $(Y_j, \mathcal{G}_j, S_j, \nu_j)$ and any $g_j \in L^\infty(\nu_j)$ with $\|g_j\|_\infty \leq 1$ for ν_j-a.e. y_j we have

$$\frac{1}{N}\sum_{n=1}^{N}\left[f(T^n x) - \mathbb{E}(f|\mathcal{A}_j)(T^n x)\right] g_1(S_1^n y_1)\cdots g_j(S_j^n y_j)$$

converges to 0.

Lemma 4. *Let f be an element of $f \in L^\infty$ and let g_i, S_i and y_i be as defined in the preceding paragraph. If*

$$B_N = \sup_\epsilon \left|\frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S_1^n y_1)\cdots g_{k-1}(S_{k-1}^n y_{k-1}) e^{2\pi i n\epsilon}\right|^2$$

then

$$\limsup_N B_N \leq C N_k(f)^2$$

for some absolute constant C.

Proof. By Lemma 3, there exists a constant C such that for $1 \leq H \leq N$

$$\begin{aligned} B_N \quad \leq \quad & C\Bigg(\frac{1}{H} + \frac{1}{H}\sum_{h=1}^{H}\Bigg|\frac{1}{N}\sum_{n=1}^{N-h}(f\cdot\overline{f\circ T^h})(T^n x) \\ & \cdot(g_1\cdot\overline{g_1\circ S_1^h})(S_1^n y_1)\cdots(g_{k-1}\cdot\overline{g_{k-1}\circ S_{k-1}^h})(S_{k-1}^n y_{k-1})\Bigg|\Bigg). \end{aligned}$$

From our inductive hypothesis, we know that for each h there is a set of full measure $X_{f\cdot\overline{f\circ T^h}}$ on which

$$\left|\frac{1}{N}\sum_{n=1}^{N-h}\Big[(f\cdot\overline{f\circ T^h})(T^n x)-(\mathbb{E}(f\cdot\overline{f\circ T^h},\mathcal{A}_{k-1})(T^n x)\Big]\right.$$
$$\left.\cdot(g_1\cdot\overline{g_1\circ S_1^h})(S_1^n y_1)\cdots(g_{k-1}\cdot\overline{g_{k-1}\circ S_{k-1}^h})(S_{k-1}^n y_{k-1})\right| \quad\to\quad 0.$$

Therefore, the intersection of these sets $X_{f\cdot\overline{f\circ T^h}}$ over h gives a set of full measure $\widehat{X_f}$ on which

$$\begin{aligned}\limsup_N B_N \quad\leq\quad & \limsup_N C\Bigg(\frac{1}{H}+\frac{1}{H}\sum_{h=1}^{H}\Bigg|\frac{1}{N}\sum_{n=1}^{N-h}(\mathbb{E}(f\cdot\overline{f\circ T^h},\mathcal{A}_{k-1})(T^n x)\\ &\cdot(g_1\cdot\overline{g_1\circ S_1^h})(S_1^n y_1)\cdots(g_{k-1}\cdot\overline{g_{k-1}\circ S_{k-1}^h})(S_{k-1}^n y_{k-1})\Bigg|\Bigg)\end{aligned}$$

for all H. The Cauchy-Schwartz inequality gives us

$$\begin{aligned}\limsup_N B_N \quad\leq\quad & \limsup_N C\Bigg(\frac{1}{H}+\frac{1}{H}\sum_{h=1}^{H}\Bigg(\frac{1}{N}\sum_{n=1}^{N-h}\Big|(\mathbb{E}(f\cdot\overline{f\circ T^h},\mathcal{A}_{k-1})(T^n x)\Big|^2\\ &\cdot\Big|(g_1\cdot\overline{g_1\circ S_1^h})(S_1^n y_1)\Big|^2\cdots\Big|(g_{k-1}\cdot\overline{g_{k-1}\circ S_{k-1}^h})(S_{k-1}^n y_{k-1})\Big|^2\Bigg)^{\frac{1}{2}}\Bigg)\\ \leq\quad & \limsup_N C\Bigg(\frac{1}{H}+\frac{\|g_1\|_\infty^2\ldots\|g_{k-1}\|_\infty^2}{H}\cdot\\ &\sum_{h=1}^{H}\Bigg(\frac{1}{N}\sum_{n=1}^{N-h}\Big|(\mathbb{E}(f\cdot\overline{f\circ T^h},\mathcal{A}_{k-1})(T^n x)\Big|^2\Bigg)^{\frac{1}{2}}\Bigg)\end{aligned}$$

By Birkhoff's Pointwise Ergodic Theorem we know that there is a set of full measure X_{k-1} on which for each h the average over n in the above inequality converges to

$$\int\Big|(\mathbb{E}(f\cdot\overline{f\circ T^h},\mathcal{A}_{k-1})\Big|^2\,d\mu=\Big\|\mathbb{E}(f\cdot\overline{f\circ T^h},\mathcal{A}_{k-1})\Big\|_2^2.$$

Therefore on the set of full measure $X_f = \widehat{X_f} \bigcap X_{k-1}$

$$\begin{aligned}\limsup_N B_N &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H} \left\|\mathbb{E}(f\cdot \overline{f\circ T^h}, \mathcal{A}_{k-1})\right\|_2 \\ &\leq C \lim_H \left(\frac{1}{H} + \left(\frac{1}{H}\sum_{h=1}^{H}\left\|\mathbb{E}(f\cdot \overline{f\circ T^h}, \mathcal{A}_{k-1})\right\|_2^2\right)^{\frac{1}{2}}\right) \\ &= C\cdot N_k(f)^2.\end{aligned}$$

□

As functions f in $\mathcal{A}_k^\perp$ are characterized by the property that $N_k(f)^4 = 0$, Lemma 4 implies that when f is an element of $L^\infty(\mu)\bigcap \mathcal{A}_k^\perp$ we have

$$\limsup_N B_N = 0$$

on the set of full measure $X_f = \widehat{X_f}\bigcap X_{k-1}$. Therefore

$$\sup_\epsilon \left|\frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S_1^n y_1)\cdots g_{k-1}(S_{k-1}^n y_{k-1}) e^{2\pi i n\epsilon}\right|$$

converges to 0 μ-a.e.. Hence by an application of Lemma 2, we know that for any other dynamical system $(Y_k, \mathcal{G}_k, S_k, \nu_k)$ and any $g \in L^\infty(\nu_k)$

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S_1^n y_1)\cdots g_k(S_k^n y_k) \tag{2}$$

converges to 0 in $L^2(\nu_k)$. As pointwise convergence of the average in Equation (2) follows from Theorem 2, we have

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S_1^n y_1)\cdots g_k(S_k^n y_k)$$

converges to 0 ν_k-a.e., when f is in $L^\infty(\mu)\bigcap \mathcal{A}_k^\perp$. Therefore for all $f \in L^\infty(\mu)$ we have

$$\frac{1}{N}\sum_{n=1}^{N} \left[f(T^n x) - \mathbb{E}(f, \mathcal{A}_k)(T^n x)\right] g_1(S_1^n y_1) g_2(S_2^n y_2)\cdots g_k(S_k^n y_k)$$

converges to 0 ν_k-a.e.. Thus, we have shown that the factors $\mathcal{A}_k$ are pointwise characteristic for the multiple term return times averages. □

3 The $\mathcal{Z}_k$ factors are pointwise characteristic for the multidimensional return times averages

As noted above, the factors $\mathcal{Z}_k$ are smaller than the $\mathcal{A}_k$ factors and thus their orthogonal complements $\mathcal{Z}_k^{\perp}$ are bigger. Therefore Theorem 4, which we are proving in this section, is an extension of Theorem 3. We will prove this fact directly from the properties of the factors $\mathcal{Z}_k$. As shown in [12] the Host-Kra-Ziegler factors, $\mathcal{Z}_k$, can be characterized by the following seminorms.

Definition 5.

- *The factors $\mathcal{A}_0$ and $\mathcal{Z}_0$ are equal to the trivial σ-algebra.*
- *The factors $\mathcal{A}_1$ and $\mathcal{Z}_1$ are also identical. They can be characterized by the seminorms $|||f|||_2$ or $N_2(f)$ where*

$$|||f|||_2^4 = \lim_H \frac{1}{H} \sum_{h=1}^{H} \left| \int f \cdot f \circ T^h d\mu \right|^2 = N_2(f)^4.$$

- *The difference starts with the factors $\mathcal{A}_2$ and $\mathcal{Z}_2$. The factor $\mathcal{Z}_2$ is the Conze-Lesigne factor, $\mathcal{CL}$. Functions in this factor are characterized by the seminorm $|||\cdot|||_3$ such that*

$$|||f|||_3^8 = \lim_H \frac{1}{H} \sum_{h=1}^{H} |||f \cdot f \circ T^h|||_2^4.$$

  A function $f \in CL^{\perp}$ if and only $|||f|||_3 = 0$.

- *More generally B. Host and B. Kra showed in [12] that for each positive integer k we have*

$$|||f|||_{k+1}^{2^{k+1}} = \lim_H \frac{1}{H}\sum_{h=1}^{H} |||f\cdot f\circ T^h|||_k^{2^k},$$

with the condition that $f \in \mathcal{Z}_{k-1}$ if and only if $|||f|||_k = 0$.

Our induction argument comes from reducing the return times averages by looking at an associated Wiener-Wintner type average using the following lemma.

Lemma 5. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system and $f \in L^\infty(\mu)$. Then for all positive integers H we have*

$$\limsup_N \sup_t \left|\frac{1}{N}\sum_{n=1}^{N} f(T^n x)e^{2\pi i n t}\right|^2 \leq C\left(\frac{1}{H} + \frac{1}{H}\sum_{h=1}^{H}\left|\int f\cdot f\circ T^h d\mu\right|\right)$$

In particular we have for μ-a.e. x

$$\limsup_N \sup_t \left|\frac{1}{N}\sum_{n=1}^{N} f(T^n x)e^{2\pi i n t}\right|^2 \leq C|||f|||_2^2.$$

Proof. This is Lemma 2 from the paper [4]. □

Using this result, we can deduce the following lemma concerning the integral of the lim sup of our averages.

Lemma 6. *Given $(X, \mathcal{B}, \mu, T)$ an ergodic measure preserving system on a probability measure space and $f \in L^\infty$ then we can find a set of full measure X_f such that for every $x \in X_f$ for each measure preserving dynamical system $(Y, \mathcal{G}, \nu, S)$ and each $g \in L^\infty(\nu)$ we have*

$$\int \limsup_N F_N(y) d\nu := \int \limsup_N \left|\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n y)\right|^2 d\nu \leq C|||f|||_2^2\|g\|_2^2.$$

Proof. By the BFKO [9] proof of the Return Times Theorem we have pointwise convergence of the above averages, therefore the lim sup on the left hand side of the above expression becomes a limit. Therefore, we have

$$\begin{aligned}\int \limsup_N F_N(y)d\nu &= \lim_N \int F_N(y)d\nu \\ &= \lim_N \int \left| \frac{1}{N}\sum_{n=1}^{N} f(T^n x)e^{2\pi int}\right|^2 d\sigma_g(t)\end{aligned}$$

where σ_g is the spectral measure associated to g with respect to the dynamical system $(Y, \mathcal{G}, \nu, S)$. Thus

$$\int \limsup_N F_N(y)d\nu \leq \limsup_N \sup_t \left| \frac{1}{N}\sum_{n=1}^{N} f(T^n x)e^{2\pi int}\right|^2 \|g\|_2^2.$$

By Lemma 5 we derive the inequality

$$\int \limsup_N F_N(y)d\nu \leq C|||f|||_2^2\|g\|_2^2.$$

□

From Lemma 6 the iteration process follows. For instance, we can use this lemma to prove the following Wiener-Wintner return times result which refines the one obtained in [5].

Lemma 7. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving system on a probability measure space and $f \in L^\infty(\mu)$. Then for μ-a.e. $x \in X$ for every measure preserving system $(Y, \mathcal{G}, \nu, S)$ and each $g \in L^\infty(\nu)$ and ν-a.e. y we have*

$$\int \limsup_N \sup_t \left| \frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n y)e^{2\pi int}\right|^2 d\nu \leq C|||f|||_3^2\|g\|_\infty^2. \quad (3)$$

In particular, for $f \in \mathcal{CL}^\perp$ (or equivalently $|||f|||_3 = 0$) we have for ν-a.e. y

$$\limsup_N \sup_t \left| \frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n y)e^{2\pi int}\right| = 0. \quad (4)$$

Proof. By the Van Der Corput Lemma [13] we have

$$\begin{aligned}
\int \limsup_N \sup_t \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) e^{2\pi i n t} \right|^2 d\nu \quad &\le \\
C\Bigg(\frac{1}{H} + \frac{1}{H} \sum_{h=1}^{H} \int \limsup_N \Big| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) f(T^{n+h} x) \cdot \\
g(S^n y) g(S^{n+h} y) \Big| d\nu \Bigg) \quad &\le \\
C\Bigg(\frac{1}{H} + \frac{1}{H} \sum_{h=1}^{H} \Bigg(\int \limsup_N \Big| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) f(T^{n+h} x) \cdot \\
g(S^n y) g(S^{n+h} y) \Big|^2 d\nu \Bigg)^{1/2} \Bigg) \qquad (5)
\end{aligned}$$

By using Lemma 5 the expression in (5) is

$$\begin{aligned}
&\le \quad C \frac{1}{H} \sum_{h=1}^{H} |||f \cdot f \circ T^h|||_2 \|g \cdot g \circ S^h|||_2 \\
&\le \quad C \frac{1}{H} \sum_{h=1}^{H} |||f \cdot f \circ T^h|||_2 \|g\|_\infty^2
\end{aligned}$$

on a set of full measure depending only on f. It is, in fact, the intersection of the sets of full measure obtained by the BFKO [9] proof of the Return Times Theorem for each function $f \cdot f \circ T^h$. As

$$\lim_H \frac{1}{H} \sum_{h=1}^{H} |||f \cdot f \circ T^h|||_2^4 = |||f|||_3^8$$

this proves (3) of Lemma 7. Equation (4) follows directly from the characterization of the $\mathcal{CL}$ factor. □

The induction assumption giving the result on the pointwise characteristic factors for the $\mathcal{Z}_k$ factors can now be made. To end it at the $\mathcal{CL} = \mathcal{Z}_2$ level we prove the next lemma.

Lemma 8. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving system on a probability measure space and $f \in L^\infty$ Then for μ-a.e. $x \in X$ for every measure preserving system $(Y, \mathcal{G}, \nu, S)$ and each $g \in L^\infty(\nu)$, for ν-a.e. y, for every measure preserving system $\Gamma_1 = (Z, \mathcal{F}, \rho, V)$ and each $\phi \in L^\infty(\rho)$ we have*

$$\int \sup_{\Gamma_1,\phi} \limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) \phi(V^n z) \right|^2 d\rho(z) d\nu(y) \leq C |||f|||_3^2 \|g\|_\infty^2.$$

Proof. By Theorem 2 for averages with three terms we have

$$\begin{aligned} \int \limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) \phi(V^n z) \right|^2 d\rho(z) \quad &= \\ \lim_N \int \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) \phi(V^n z) \right|^2 d\rho(z). \end{aligned} \tag{6}$$

Using the spectral measure as before we have that the expression in (6) is equal to

$$\begin{aligned} \lim_N \int \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) e^{2\pi i n t} \right|^2 d\sigma_\phi \quad &\leq \\ \limsup_N \sup_t \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) e^{2\pi i n t} \right|^2 \|\phi\|_\infty^2 \quad &\leq \\ \limsup_N \sup_t \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) e^{2\pi i n t} \right|^2 \end{aligned}$$

as $\|\phi\|_\infty| \leq 1$. Therefore by Lemma 7 we have

$$\int \left(\sup_{\Gamma_1,\phi} \int \limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) \phi(V^n z) \right|^2 d\rho(z) \right) d\nu(y) \leq C |||f|||_3^2 \|g\|_\infty^2.$$

□

We now have the tools necessary to prove our second main result, Theorem 4, that the $\mathcal{Z}_k$ averages are pointwise characteristic for the multidimensional return times averages.

Proof. It remains to finish the induction argument which we have started in the above lemmas. Assume that for each $i \geq 2$, and for μ-a.e. x, and each g with $\|g\|_\infty \leq 1$ we have

$$\int \left(\sup_{\Gamma_1,\phi_1} \cdots \int \left(\sup_{\Gamma_{i-1},\phi_{i-1}} \int \limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) \cdot \right.\right.\right.$$
$$\left.\left.\left. g(S^n y)\phi_1(V_1^n z_1) \cdots \phi_{i-1}(V_{i-1}^n z_{i-1}) \right|^2 d\rho(z_{i-1}) \right) \cdots d\rho(z_1) \right) d\nu(y) \quad \leq \quad C|||f|||_{i+1}^2$$

We would like to show that we have

$$\int \left(\sup_{\Gamma_1,\phi_1} \cdots \int \left(\sup_{\Gamma_i,\phi_i} \int \limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) \cdot \right.\right.\right.$$
$$\left.\left.\left. g(S^n y)\phi_1(V_1^n z_1) \cdots \phi_i(V_i^n z_i) \right|^2 d\rho(z_i) \right) \cdots d\rho(z_1) \right) d\nu(y) \quad \leq \quad C|||f|||_{i+2}^2$$

As previously, we use the Van Der Corput Lemma [13] to estimate that

$$\limsup_N \sup_t \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) \phi_1(V_1^n z_1) \cdots \phi_{i-1}(V_{i-1}^n z_{i-1}) e^{2\pi i n t} \right|^2 \quad <$$
$$C \left(\frac{1}{H} + \frac{1}{H} \sum_{h=1}^{H} \limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} (f \cdot f \circ T^h)(T^n x) \cdot \right.\right.$$
$$\left.\left. (g \cdot g \circ S^h)(S^n y)(\phi_1 \cdot \phi \circ V_1^h)(V_1^n z_1) \cdots (\phi_{i-1} \cdot \phi_{i-1} \circ V_{i-1}^h)(V_{i-1}^n z_{i-1}) \right| \right)$$

Integrating and using the induction assumption we get

$$\int \left(\sup_{\Gamma_1,\phi_1} \cdots \int \left(\sup_{\Gamma_{i-1},\phi_{i-1}} \int \limsup_N \sup_t \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) \cdot \right.\right.\right.$$
$$\left.\left.\left. g(S^n y)\phi_1(V_1^n z_1) \cdots \phi_{i-1}(V_{i-1}^n z_{i-1}) e^{2\pi n t} \right|^2 d\rho(z_{i-1}) \right) \cdots d\rho(z_1) \right) d\nu(y) \quad \leq$$
$$C \left(\frac{1}{H} + \left(\frac{1}{H} \sum_{h=1}^{H} |||f \cdot f \circ T^h|||_{i-1}^2 \right)^{1/2} \right)$$

After taking the limit with respect to H we have

$$C \left(\frac{1}{H} + \left(\frac{1}{H} \sum_{h=1}^{H} |||f \cdot f \circ T^h|||_{i-1}^2 \right)^{1/2} \right) \leq C|||f|||_{i+1}^2.$$

Finally by applying Theorem 2 and applying Lemma 8 we have

$$\int\Big(\sup_{\Gamma_1,\phi_1}\cdots\int\Big(\sup_{\Gamma_i,\phi_i}\int\limsup_N\Big|\frac{1}{N}\sum_{n=1}^N f(T^nx)\cdot$$

$$g(S^ny)\phi_1(V_1^nz_1)\cdots\phi_i(V_i^nz_i)\Big|^2 d\rho(z_i)\Big)d\rho(z_{i-1})\Big)\cdots d\rho(z_1)\Big)d\nu(y) \quad =$$

$$\int\Big(\sup_{\Gamma_1,\phi_1}\cdots\int\Big(\sup_{\Gamma_i,\phi_i}\lim_N\int\Big|\frac{1}{N}\sum_{n=1}^N f(T^nx)\cdot$$

$$g(S^ny)\phi_1(V_1^nz_1)\cdots\phi_i(V_i^nz_i)\Big|^2 d\rho(z_i)\Big)d\rho(z_{i-1})\Big)\cdots d\rho(z_1)\Big)d\nu(y) \quad =$$

$$\int\Big(\sup_{\Gamma_1,\phi_1}\cdots\int\Big(\sup_{\Gamma_i,\phi_i}\lim_N\int\Big|\frac{1}{N}\sum_{n=1}^N f(T^nx)\cdot$$

$$g(S^ny)\phi_1(V_1^nz_1)\cdots\phi_{i-1}(V_{i-1}^nz_{i-1})e^{2\pi int}\Big|^2 d\sigma_{\phi_i}\Big)d\rho(z_{i-1})\Big)\cdots d\rho(z_1)\Big)d\nu(y) \quad \leq$$

$$\int\Big(\sup_{\Gamma_1,\phi_1}\cdots\int\Big(\sup_{\Gamma_i,\phi_i}\limsup_N\sup_t\Big|\frac{1}{N}\sum_{n=1}^N f(T^nx)\cdot$$

$$g(S^ny)\phi_1(V_1^nz_1)\cdots\phi_{i-1}(V_{i-1}^nz_{i-1})e^{2\pi int}\Big|^2\Big)d\rho(z_{i-1})\Big)\cdots d\rho(z_1)\Big)d\nu(y) \quad \leq$$

$$C|||f|||_{i+2}^2$$

This proves the induction step and proves the theorem. □

4 Remarks

1. If one uses instead the maximal isometric extensions and the factors $\mathcal{A}_k$ we get at the BFKO stage a pointwise upper bound for T and S ergodic. Namely we saw in Equation (1) (for $\|g\|_\infty \leq 1$)

$$\limsup_N\left|\frac{1}{N}\sum_{n=1}^N f(T^nx)g(S^ny)\right|^2 \leq C\|E(f|\mathcal{K}_T)\|_2^2.$$

From this by applying Lemma 3 as in Lemma 4 we can prove that the average of two terms is estimated by

$$\limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) \right|^2 \le C \left[\frac{1}{H} + \left(\frac{1}{H} \sum_{h=1}^{H} \| \mathbb{E}(f \cdot f \circ T^h | \mathcal{A}_1) \|_2^2 \right)^{1/2} \right].$$

By taking the limit with H we get the better estimate

$$\limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) \right|^2 \le C N_2(f)^2.$$

This argument can be generalized to the average of multiple terms resulting in the inequality

$$\limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g_1(S_1^n y_1) \cdots g_k(S_k^n y_k) \right|^2 \le C N_{k+1}(f)^2.$$

2. It is not clear to us if for $k \ge 2$, one can replace in these inequalities the N_k seminorms with those defining the $\mathcal{Z}_k$ factors.

3. If one uses instead the maximal isometric extensions and the factors $\mathcal{A}_k$ we get at the BFKO stage a pointwise upper bound for T and S ergodic . Namely we have

$$\limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) \right| \le C \| E(f | \mathcal{K}_T) \|_2 \| E(g | \mathcal{K}_S) \|_2.$$

It is not clear if one can have a similar upper bound for the seminorms $|||f|||_2$.

4. The authors of this paper are writing a survey of the Return Times Theorem [6] which will include more details of the historical developments of Theorem 1 and 2 and related questions such as the ones noted above.

References

[1] I. Assani. "Multiterm return time theorem for weakly mixing systems" *Annales de L'Institut Henri Poincare*,vol 36,2,153-165,(2000).

[2] I. Assani. *Wiener Wintner Ergodic Theorems.* World Scientific Publishing Co., New Jersey, London, Singapore, Hong Kong, 2003.

[3] I. Assani. "Characteristic factors for some non conventional ergodic averages." *Preprint*, 2003.

[4] I. Assani. "Pointwise Convergence of Ergodic Averages along Cubes." *J. Analyse Math.* 110 (2010), 241-269.

[5] I. Assani, E. Lesigne and D. Rudolph. "Wiener Wintner return-times ergodic theorem." *Isr. J. Math.* 92 (1995) 375-395.

[6] I. Assani and K. Presser. "Survey of the Return Times Theorem." *In Progress*, 2010.

[7] G. D. Birkhoff. "Proof of the ergodic theorem." *Proc. Nat. Acad. Sci.* 17 (1931), 656-660.

[8] J. Bourgain. "Return time sequences of dynamical systems." *Unpublished Preprint*, 1988.

[9] J. Bourgain, H. Furstenberg, Y. Katznelson, and D. Ornstein. Appendix to J. Bourgain: "Pointwise ergodic theorems for arithmetic sets." *Inst. Hautes Études Sci. Publ. Math.* 69 (1989), 5-45.

[10] H. Furstenberg. "Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions." *J. Analyse Math.* 31 (1977), 204 - 256.

[11] H. Furstenberg and B. Weiss. "A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{2n}x)$." *Convergence in Ergodic Theory and Probability* 5 (1996), 193-227.

[12] B. Host and B. Kra. "Nonconventional ergodic averages and nilmanifolds." *Ann. of Math. (2)* 161 (2005), 397-488.

[13] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience, John Wiley & Sons, New York, 1974.

[14] D. Rudolph. "A joinings proof of Bourgain's return time theorem." *Ergodic Theory Dynam. Systems*14 no. 1 (1994), 197 - 203.

[15] D. Rudolph. "Eigenfunctions of $T \times S$ and the Conze-Lesigne algebra." Ergodic theory and its connections with harmonic analysis. *London Math. Soc. Lecture Note Ser.* 205 (1995), 369-432.

[16] D. Rudolph. "Fully generic sequences and a multiple-term return times theorem." *Invent. Math.* 131 no. 1 (1998), 199-228.

[17] T. Ziegler. "Universal characteristic factors and Furstenberg averages." *J. Amer. Math. Soc.*, 20, (2007), 53-97.