EXTENSIONS OF J. BOURGAIN'S DOUBLE RECURRENCE THEOREM

Ryo Moore

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Mathematics in the College of Arts and Sciences.

Chapel Hill
2016

Approved by:

Idris Assani

Zoltán Buczolich

Jean-Pierre Conze

Richárd Rimányi

Justin Sawon

ABSTRACT

Ryo Moore: Extensions of J. Bourgain's Double Recurrence Theorem
(Under the direction of Idris Assani)

The study of multiple recurrence averages was pioneered by Furstenberg in 1977, when he provided an alternative proof to Szemerédi's theorem using ergodic theory, which states that a set of integer with a positive density contains an arbitrary long arithmetic progression. Since then, many convergence results of multiple recurrence averages have been obtained. Their norm convergence have been studied by Conze and Lesigne (in 1984), Host and Kra (2005), Ziegler (2006), Tao (2008), and the best result was obtained by Walsh (2012).

The results are much scarcer for pointwise convergence. In 1990, Bourgain answered Furstenberg's question by showing that some double recurrence averages converge pointwise. This deep result has not been generalized since then, while some partial results on the pointwise convergence of multiple recurrence averages are obtained by Derrien and Lesigne (1996), Assani (1998, 2005), and recently announced by Huang, Shao, and Ye (2014), and Donoso and Sun (2015). Also, Assani and Buczolich have shown that the pointwise convergence of double recurrence averages need not to hold when both functions are in L^1.

On the other hand, Brunel initiated the study of return times averages in his PhD thesis from 1966, where one concerns ergodic averages with weights that are generated randomly. In 2000, Assani showed that the sequence appearing in the multiple recurrence averages can be a good universal weight for multiple return times averages under some assumptions on the system.

We will show that one can extend Bourgain's double recurrence result in numerous ways. We will first show the Wiener-Wintner extension of the double recurrence theorem, which was a study initiated by Duncan in his doctoral dissertation from 2001. Furthermore, we will show a polynomial Wiener-Wintner result for the double recurrence theorem, extending the work of Lesigne (1990, 1993) and Frantzikinakis (2006). Secondly, from the angle of the Wiener-Wintner extension, we will show that the sequence appearing in the double recurrence averages can be a good universal weight for

some nonconventional ergodic averages in norm, ultimately extending the work of Host and Kra, and Ziegler.

ACKNOWLEDGEMENTS

First and foremost, this dissertation would be far from complete without my sincerest gratitude to my advisor, Professor Idris Assani. His constant support, patience, and numerous helpful advices have helped me achieve this far, and encouraged me to pursue mathematics beyond the graduate school. I highly appreciate the amount of effort that he has put to expose me to the field of ergodic theory and dynamical systems. Words cannot describe how fortunate I was to study mathematics under his guidance.

I would also like to thank Professors Zoltán Buczolich, Jean-Pierre Conze, Richárd Rimányi, and Justin Sawon for their commitment to be part of my dissertation committee. I am grateful of the questions that they raised during the oral exam, and their flexibility when scheduling the exam as well as the defense, especially Profs. Buczolich and Conze for testing the technologies required to hold these events virtually.

Numerous people that I have encountered during my graduate studies deserve my gratitudes, and I apologize in advance in case if I miss anyone. I am thankful for the organizers and the lecturers at ICTP-NLAGA School in Dynamical Systems and Ergodic Theory that took place in M'bour, Senegal on June 2014 for their fantastic lectures and hospitality. At UNC, I was happy to be with wonderful friends and colleagues, such as Perry Harabin, Mayukh Mukherjee, Cass Sherman, Mike Schuster, Nick Battista, Colin Grudzien, and David Lax, to name a few. I am also grateful for the departmental staff members, especially Laurie Straube and Brian Whitling, for their assistance with the administrative matters. Furthermore, I would like to thank Professor Jingfang Huang for his helpful advice when he was the director of graduate studies.

As always, I would like to thank my family for supporting me. It was always my pleasure to visit them in Eugene, Oregon during the summer and winter vacations.

Finally, I would like to thank Michael Malahe for sharing his LaTeX document class of the dissertation in a way that would meet the guidelines set by the UNC Graduate School.

TABLE OF CONTENTS

CHAPTER 1

Introduction

In this chapter, we will recall basic notions of ergodic theory as well as background and historical developments on relevant topics, such as Wiener-Wintner ergodic theorem, multiple recurrence averages, characteristic factors, good universal weights, and the return times.

1.1 Background on Ergodic Theory

1.1.1 Measure preserving systems

Here we recall the notions of measure-preserving systems, ergodicity, factors, conditional expectations, and the classical ergodic theorems. The proofs of the results stated here can be found in many references on ergodic theory (e.g. [51]).

We denote the quadruple $(X, \mathcal{F}, \mu, T)$ to be a (dynamical) system, where $(X, \mathcal{F}, \mu)$ is a probability measure space (i.e. $\mu(X) = 1$), and $T : X \to X$ is a measurable transformation. We say that a transformation T is measure-preserving if the measure μ is preserved under action of T on X, i.e. for any $A \in \mathcal{F}$, we have $\mu(T^{-1}A) = \mu(A)$. If a system $(X, \mathcal{F}, \mu, T)$ is equipped with a measure-preserving transformation T, we say the system is measure-preserving. We say that a measure-preserving transformation T is ergodic if a set $A \in \mathcal{F}$ is invariant under T, i.e. $T^{-1}A = A$, then A is either (almost) everything or (almost) nothing, i.e. $\mu(A) \in \{0, 1\}$. Similarly, if T is ergodic, then we say $(X, \mathcal{F}, \mu, T)$ is an ergodic system.

If we consider a dynamical system with multiple transformations (say, $T_1, T_2, \dots, T_k$), we list each of them (i.e. $(X, \mathcal{F}, \mu, T_1, T_2, \dots, T_k)$). We say the system $(X, \mathcal{F}, \mu, T_1, T_2, \dots, T_k)$ is measure-preserving if each transformation $T_1, T_2, \dots, T_k$ is measure-preserving. More generally, if Γ is a group acting on $(X, \mathcal{F}, \mu)$ in a measure-preserving way (i.e. if $\gamma \in \Gamma$, $\gamma \cdot x = T_\gamma x$ for some μ-measure-preserving action T_γ), then we denote $(X, \mathcal{F}, \mu, \Gamma)$ to be the measure-preserving system. So for instance, the system $(X, \mathcal{F}, \mu, T)$ can be written as $(X, \mathcal{F}, \mu, \mathbb{Z})$, where we identify the action $n \cdot x = T^n x$ for any $n \in \mathbb{Z}$.

The transformation T may also be viewed as a linear operator on $L^p(\mu)$ for $1 \le p \le \infty$, by identifying $Tf = f \circ T$ for any $f \in L^p(\mu)$. Furthermore, if T is measure-preserving, then T is an unitary operator on $L^2(\mu)$, since for any $f, g \in L^2(\mu)$,

$$\langle Tf, g\rangle = \int_X f(Tx)\overline{g(x)}\, d\mu(x) = \int_{TX} f(x)\overline{g(T^{-1}x)}\, d\mu(x) = \left\langle f, T^{-1}g\right\rangle,$$

so that $T^* = T^{-1}$. Consequentially, if f is an eigenfunction of T, i.e. $Tf = \lambda f$, then λ must lie on a unit circle on $\mathbb{C}$.

Let $(X, \mathcal{F}, \mu, \Gamma)$ and $(Y, \mathcal{G}, \nu, \Gamma)$ be two measure-preserving systems where the group Γ acts on both X and Y. Suppose $\phi : X \to Y$ is a measurable map. We say ϕ is measure-preserving if for any $A \in \mathcal{G}$, $\nu(A) = \mu(\phi^{-1}A)$. We say a measure-preserving map ϕ is a homomorphism between the systems $(X, \mathcal{F}, \mu, \Gamma)$ and $(Y, \mathcal{G}, \nu, \Gamma)$ if for any $g \in \Gamma$, $\phi(g \cdot x) = g \cdot \phi(x)$ for μ-a.e. $x \in X$. If ϕ is invertible, and ϕ^{-1} is a measurable map, we say the homomorphism ϕ is an isomorphism. When such homomorphism exists, we say the system $(Y, \mathcal{G}, \nu, \Gamma)$ is a factor of $(X, \mathcal{F}, \mu, \Gamma)$.

We introduce a few examples of factors. For instance, given a measure-preserving system $(X, \mathcal{F}, \mu, \Gamma)$, a system $(X, \mathcal{I}, \mu, \Gamma)$ is a factor of $(X, \mathcal{F}, \mu, \Gamma)$ provided that $\mathcal{I}$ is sub-σ-algebra of $\mathcal{F}$ that is Γ-invariant, i.e. for any $\gamma \in \Gamma$ and $A \in \mathcal{I}$, $\gamma^{-1}A \in \mathcal{I}$. In this case, the homomorphism ϕ is simply the identity map on X. In particular, if $(X, \mathcal{F}, \mu, T)$ is a measure-preserving system with one transformation T, then $(X, \mathcal{I}(T), \mu, T)$, where $\mathcal{I}(T)$ is the σ-algebra of T-invariant sets, i.e.

$$\mathcal{I}(T) = \left\{A \in \mathcal{F} : T^{-1}A = A\right\},$$

is a factor of $(X, \mathcal{F}, \mu, T)$.

Another important factor that we consider extensively is the Kronecker factor. We first follow the notions explained in [6, §2.2]. Suppose $(X, \mathcal{F}, \mu, T)$ be an ergodic system, and consider $K \subset L^2(\mu)$ to be the closed linear span of the eigenfunctions of T (viewing T as a unitary operator on $L^2(\mu)$). If T is ergodic, we observe the following facts:

1. From the definition of ergodicity, any T-invariant function (i.e. $f \in L^1(\mu)$ for which $Tf = f$) is a constant function.

2. If $f \in E_\lambda := \{f \in K : Tf = \lambda f\}$, then $f \in L^\infty(\mu)$. This can be realized by the fact that $|f|$ is a constant function. To see this, suppose $Tf = \lambda f$ for some $\lambda \in S^1$. Then $T|f| = |Tf| =$

$|\lambda||f| = |f|$, so $|f|$ is a T-invariant function.

3. We note that E_λ is one-dimensional subspace of K. First, if $\lambda = 1$, then E_λ consists of all the T-invariant functions, which are constants, so E_1 is one-dimensional. To see this fact for the case $\lambda \neq 1$, suppose $f, g \in E_\lambda$ for some $\lambda \in S^1 - \{1\}$. Then $T(f\bar{g}) = f\bar{g}$, which implies that $f\bar{g} = c$ for some constant $c \in \mathbb{C}$. If c is a nonzero constant, then $f = c(g/|g|)$, and $g/|g| \in E_\lambda$. In fact, $c = 0$ if and only if $f = 0$ or $g = 0$. To see this, assuming that $c = 0$, we have $X = f^{-1}(\{0\}) \cup g^{-1}(\{0\})$. Since $T^{-1}f^{-1}(\{0\}) = f^{-1}(\{0\})$, so $\mu(f^{-1}(\{0\}))$ is either 0 or 1. Because of this, either f or g is identically equal to 0 for μ-a.e. $x \in X$.

If $(X, \mathcal{F}, \mu, T)$ is separable, then the fact (3) tells us that K has a countable orthogonal basis of eigenfunctions. Furthermore, because product of two eigenfunctions is an eigenfunction, it is an $L^\infty(\mu)$ function. Thus, one can find a sub-σ-algebra $\mathcal{K}$ of $\mathcal{F}$, that is the smallest σ-algebra for which all the functions in K are measurable. We call the system $(X, \mathcal{K}, \mu, T)$ the Kronecker factor.

An alternative characteristic of Kronecker factors can be given as follows (cf. [53, §2]): Given an ergodic system $(X, \mathcal{F}, \mu, T)$, there exists a group rotation system (Z, α) such that Z is a compact abelian (additive) group, an element $\alpha \in Z$ that generates a dense cyclic subgroup in Z, a measure-preserving map $\pi : (X, \mathcal{F}, \mu, T) \to (Z, \mathcal{B}, \nu, R_\alpha)$ (where ν is the Haar measure on Z, $\mathcal{B}$ is the Borel measure on Z, and $R_\alpha(z) = z + \alpha$ that satisfies $\pi(Tx) = R_\alpha(\pi(x))$ for μ-a.e. $x \in X$, and furthermore, f is an eigenfunction of T if and only if $f = c\chi \circ \pi$ for some constant c and a character $\chi : Z \to \mathbb{C}$. This shows that $(Z, \mathcal{B}, \nu, R_\alpha)$ is a factor of $(X, \mathcal{F}, \mu, T)$. Furthermore, since $L^2(Z)$ is spanned by characters on Z, so the lift of this space on $L^2(X)$ is a closed linear span of eigenfunctions. Hence, this factor coincides with the Krocnecker factor that we discussed earlier.

Suppose $(Y, \mathcal{G}, \nu, \Gamma)$ is a factor of $(X, \mathcal{F}, \mu, \Gamma)$, where $\phi : X \to Y$ is a homomorphism. We note that $\mathcal{Z} = \phi^{-1}\mathcal{G} = \left\{A \in \mathcal{F} : A = \phi^{-1}(B) \text{ for some } B \in \mathcal{G}\right\}$ is a sub-σ-algebra of $\mathcal{F}$. One can also see that $(X, \mathcal{Z}, \mu, \Gamma)$ is isomorphic to $(Y, \mathcal{G}, \nu, \Gamma)$. If $f \in L^p(\mathcal{G})$, then we see that $f^\phi = f \circ \phi$ is a function in $L^p(\mathcal{Z})$. Since $L^2(\mathcal{Z})$ is a subspace of $L^2(\mathcal{F})$, we define $P : L^2(\mathcal{F}) \to L^2(\mathcal{Z})$ to be the orthogonal projection. For each factor $\mathcal{G}$, one can define a conditional expectation operator $\mathbb{E}(\cdot|\mathcal{G}) : L^2(\mathcal{F}) \to L^2(\mathcal{Z})$ such that if $f \in L^2(\mu)$, then $\mathbb{E}(f|\mathcal{G}) \in L^2(\mathcal{G})$ such that $\mathbb{E}(f|\mathcal{G})^\phi = Pf$. In an event where $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$, then this notion of conditional expectation agrees with the ones we see in a basic probability course. In fact, the following properties are true.

Proposition 1.1.1 (cf. [51, Proposition 5.4]). *Let $(Y, \mathcal{G}, \nu, \Gamma)$ be a factor of $(X, \mathcal{F}, \mu, \Gamma)$ with a homomorphism ϕ. The conditional expectation operator $\mathbb{E}(\cdot|\mathcal{G}) : L^2(\mathcal{F}) \to L^2(\mathcal{G})$ has the following properties.*

(i) The map $f \mapsto \mathbb{E}(f|\mathcal{G})$ is a linear operator from $L^2(\mathcal{F})$ to $L^2(\mathcal{G})$.

(ii) If $f \geq 0$, then $\mathbb{E}(f|\mathcal{G}) > 0$.

(iii) If $g \in L^2(\mathcal{G})$, then $\mathbb{E}(g^\phi|\mathcal{G}) = g$. In particular, $\mathbb{E}(\mathbb{1}_X|\mathcal{G}) = \mathbb{1}_Y$.

(iv) If $g \in L^\infty(\mathcal{G})$, then $\mathbb{E}(g^\phi f|\mathcal{G}) = g\mathbb{E}(f|\mathcal{G})$.

(v) In particular, $\int f \, d\mu = \int \mathbb{E}(g|\mathcal{G}) \, d\nu$.

We also note that the conditional expectation operator is a contraction.

Theorem 1.1.2 ([51, Theorem 5.6]). *The conditional expectation map, $f \mapsto \mathbb{E}(f|\mathcal{G})$, extends to a map of $L^1(\mathcal{F})$ to $L^1(\mathcal{G})$ satisfying (i) - (v) of Proposition* 1.1.1*, and, in addition, it maps each $L^p(\mathcal{F})$ to $L^p(\mathcal{G})$, $1 \leq p \leq \infty$, with $\|\mathbb{E}(f|\mathcal{G})\|_p \leq \|f\|_p$.*

If the isomorphism is known, we may sometimes write $\mathbb{E}(f|\mathcal{Z})$ instead of $\mathbb{E}(f|\mathcal{G})$. We note that if f is a $\mathcal{Z}$-measurable function, then $\mathbb{E}(f|\mathcal{Z}) = f$. In this case, we sometimes say that f is in $\mathcal{Z}$, or $f \in \mathcal{Z}$. Conversely, if $\mathbb{E}(f|\mathcal{Z}) = 0$, then f is orthogonal to $L^2(\mathcal{Z})$. In such instances, we may say f belongs to the orthogonal complement of $\mathcal{Z}$, or simply write $f \in \mathcal{Z}^\perp$.

Here we recall a few convergence theorem in ergodic theory. The first one is due to von Neumann.

Theorem 1.1.3 (Mean ergodic theorem). *Let $\mathcal{H}$ be a Hilbert space, and U be a unitary operator on $\mathcal{H}$. Then for any $x \in \mathcal{H}$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} U^n x = Px,$$

where P is a projection onto $Ker(I - U) = \{x \in \mathcal{H} : Ux = x\}$.

Given a measure-preserving system $(X, \mathcal{F}, \mu, T)$, we let $\mathcal{H} = L^2(\mu)$, and U be a Koopman operator of T i.e. $Uf = f \circ T$ for any $f \in L^2(\mu)$. Thus, the mean ergodic theorem above tells us that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f \circ T^n = \mathbb{E}(f|\mathcal{I}(T)) \text{ in } L^2(\mu.)$$

The next result deals with the pointwise convergence of the same averages, which is due to G. D. Birkhoff.

Theorem 1.1.4 (Pointwise Ergodic Theorem)**.** *Let $(X, \mathcal{F}, \mu, T)$ be a measure-preserving system. Then for any $f \in L^1(\mu)$, then for μ-a.e. $x \in X$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) = \mathbb{E}(f|\mathcal{I}(T))(x).$$

We observe that if T is ergodic, then $\mathcal{I}(T)$ is trivial, so the conditional expectation on the right becomes the integral $\int f d\mu$.

1.1.2 The spectral theorem and ergodic theory

We will briefly recall some of the spectral properties of dynamical systems. Proofs are omitted for most of the statements that are stated in here, since they are provided in the references, such as [63] and [75].

We denote $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ to be the torus. Suppose σ is a Borel measure on $\mathbb{T}$. We say $\hat{\sigma}(n) = \int_0^1 e(nt) d\sigma(t)$ to be the n-th Fourier-Stieltjes coefficient of the measure σ. One useful fact about Fourier-Stieltjes coefficient of a positive measure is that we can associate them with a positive-definite sequence. We say a numerical sequence (a_n) is positive-definite if for any choice of finite set of complex numbers $\{z_n\}$, one can show that

$$\sum_{n,m} a_{n-m} z_n \bar{z}_m \geq 0.$$

Proposition 1.1.5 (Herglotz, cf. [63, §1.7.6])**.** *A numerical sequence (a_n) is positive-definite if and only if there exists a positive Borel measure σ on $\mathbb{T}$ such that $a_n = \hat{\sigma}(n)$.*

This theorem allows us to establish the following spectral theorem for unitary operators.

Theorem 1.1.6 (The spectral theorem, cf. [63, §1.7.8])**.** *Let $\mathcal{H}$ be a Hilbert space, and U a unitary operator on $\mathcal{H}$. Then for any $f \in \mathcal{H}$, there exists a positive Borel measure σ_f on $\mathbb{T}$ such that $\hat{\sigma}_f(n) = \langle U^{-n} f, f \rangle$.*

To illustrate the application of Proposition 1.1.5, we will provide the proof of this theorem.

Proof. By Proposition 1.1.5, it suffices to show that the sequence $a_n = \langle U^{-n}f, f\rangle$ is positive definite. Since U is unitary, we know that $a_{n-m} = \langle U^{m-n}f, f\rangle = \langle U^m f, U^n f\rangle$. Hence, for any finite set of complex numbers $\{z_n\}$, we have

$$\sum_{n,m} a_{n-m} z_n \bar{z}_m = \left\langle \sum_m \bar{z}_m U^m f\,, \sum_n \bar{z}_n U^n f \right\rangle = \left\| \sum_k \bar{z}_k U^k f \right\|_{\mathcal{H}}^2 \geq 0,$$

where $\|\cdot\|_{\mathcal{H}}$ denotes the norm of the Hilbert space $\mathcal{H}$ induced by the inner product $\langle \cdot, \cdot \rangle$. □

This theorem is useful in the context of ergodic theory. Given a measure-preserving system $(X, \mathcal{F}, \mu, T)$, one defines U_T to be an operator on $L^2(\mu)$ so that for any $f \in L^2(\mu)$, $U_T f = f \circ T$. We call the operator U_T the Koopman operator of T. Clearly, U_T is unitary, so one can use the spectral theorem to find a measure σ_f associated to the transformation T and the function f. We call this measure σ_f the spectral measure of f with respect to T.

We say a complex Borel measure σ on $\mathbb{T}$ is continuous if for any $t \in \mathbb{T}$, $\sigma(\{t\}) = 0$. The following proposition associates a continuous measure and its Fourier coefficients.

Proposition 1.1.7 (cf. [63, §1.7.13])**.** *Let σ be a complex Borel measure on $\mathbb{T}$. Then the following statements are true.*

1. *We have*

$$\sigma(\{t\}) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \hat{\sigma}(n) e(nt).$$

2. *(Wiener's Lemma) We have*

$$\sum_{t\in\mathbb{T}} |\sigma(\{t\})|^2 = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |\hat{\sigma}(n)|^2,$$

so in particular, σ is continuous if and only if

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |\hat{\sigma}(n)|^2 = 0.$$

We conclude our remarks on the spectral theorem by introducing the spectral family of a given Hilbert space $\mathcal{H}$ and a unitary operator U. Let σ_h denote the spectral measure of $h \in \mathcal{H}$ with respect to the operator U. Given $f, g \in \mathcal{H}$, we set $a_n = \langle U^n f, g\rangle$. We note that $\langle U^n g, f\rangle = \bar{a}_{-n}$, and furthermore by the elementary identity (which is a variant of the polarization identity for a complex inner product space), we have

$$\begin{aligned} a_n &= \frac{1}{4}\left(\langle U^n(f+g), f+g\rangle - \langle U^n(f-g), f-g\rangle + i\,\langle U^n(f+ig), f+ig\rangle - i\,\langle U^n(f-ig), f-ig\rangle\right) \\ &= \frac{1}{4}(\hat{\sigma}_{f+g} - \hat{\sigma}_{f-g} + i\hat{\sigma}_{f+ig} - i\hat{\sigma}_{f-ig}). \end{aligned}$$

Thus, we observe that (a_n) are Fourier-Stieltjes coefficients of the complex Borel measure

$$\sigma_{f,g} := \frac{1}{4}(\sigma_{f+g} - \sigma_{f-g} + i\sigma_{f+ig} - i\sigma_{f-ig}). \tag{1.1}$$

We say $(\sigma_{f,g})_{f,g\in\mathcal{H}}$ is the spectral family of the operator U. One useful fact regarding the elements of this spectral family is as follows:

Proposition 1.1.8 ([75, Proposition 2.4]). *For any $f, g \in \mathcal{H}$, $\sigma_{f,g}$ is absolutely continuous with respect to both σ_f and σ_g. More precisely, for any Borel set B of $\mathbb{T}$, we have*

$$|\sigma_{f,g}|(B) \leq \sigma_f(B)^{1/2}\sigma_g(B)^{1/2}.$$

1.1.3 Van der Corput's inequality and its variants

The following inequality, which is credited to van der Corput, it utilized quite frequently in ergodic theory, especially when one wishes to show that a sequence of averages converges to zero.

Lemma 1.1.9 (van der Corput). *If (a_n) is a sequence of complex numbers and if H is an integer between 0 and $N-1$, then*

$$\begin{aligned} \left|\frac{1}{N}\sum_{n=0}^{N-1} a_n\right|^2 \leq{}& \frac{N+H}{N^2(H+1)}\sum_{n=0}^{N-1}|a_n|^2 \\ &+ \frac{2(N+H)}{N^2(H+1)^2}\sum_{h=1}^{H}(H+1-h)Re\left(\sum_{n=0}^{N-h-1} a_n\bar{a}_{n+h}\right). \end{aligned} \tag{1.2}$$

where $Re(z)$ denotes the real part of the complex number z.

One may consult [65] for a proof. One immediate corollary of this lemma, which provides an control over the averages of the sequences $a_n = u_n e(nt)$, for some numerical sequence (u_n), is as follows.

Corollary 1.1.10 (cf. [6, Corollary 2.1]). *Given (u_n) a sequence of complex numbers, and if H is an integer between 0 and $N-1$, then*

$$\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=0}^{N-1}u_n e(nt)\right|^2 \le \frac{2}{H+1}\sum_{n=0}^{N-1}|u_n|^2 + \frac{4}{H+1}\sum_{h=1}^{H}\left|\frac{1}{N}\sum_{n=0}^{N-h-1}u_n u_{n+h}\right|. \tag{1.3}$$

We note that the right hand side of this inequality does not depend on the parameter t. This estimate will be useful when we study uniformity of Wiener-Wintner type averages.

The following inequalities, which will be useful when we evaluate the limit of some averages, can be derived directly from Lemma 1.1.9 and Corollary 1.1.10.

Lemma 1.1.11.

- *There exists an absolute constant C such that for any sequence of complex numbers (a_n) such that $\sup_n |a_n| \le 1$ and any positive integer N, we have*

$$\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N}a_n\right|^2 \le \frac{C}{H} + \frac{C}{(H+1)^2}\sum_{h=1}^{H}(H+1-h)Re\left(\limsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}a_n\overline{a}_{n+h}\right) \tag{1.4}$$

  for any $H\in\mathbb{N}$.

- *There exists an absolute constant C such that for any sequence of complex numbers (a_n) such that $\sup_n |a_n| \le 1$ and any positive integer N, we have*

$$\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=1}^{N}a_n e^{2\pi int}\right|^2 \le \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\left|\frac{1}{N}\sum_{n=1}^{N-h}a_n\overline{a}_{n+h}\right| \tag{1.5}$$

  for $1 \le H \le N$.

- *There exists an absolute constant C such that for any sequence of complex numbers (a_n) such that $\sup_n |a_n| \le 1$ and any positive integer N, we have*

$$\limsup_{N\to\infty}\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=1}^{N}a_n e^{2\pi int}\right|^2 \le \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N}a_n\overline{a}_{n+h}\right| \tag{1.6}$$

for all $H \in \mathbb{N}$.

Proof. To show (1.4), we take the limit supremum (as $N \to \infty$) on both sides of (1.2). Then we obtain

$$\limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} a_n \right|^2 \leq \frac{1}{H} + \frac{2}{(H+1)^2} \sum_{h=1}^{H} (H+1-h) Re \left(\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-h-1} a_n \overline{a}_{n+h} \right).$$

Let u_n be another sequence of complex numbers norms bounded by 1. Then, for fixed h, we have

$$\frac{1}{N} \sum_{n=0}^{N-h-1} u_n = \frac{1}{N} \sum_{n=0}^{N} u_n - \frac{1}{N} \sum_{n=N-h}^{N} u_n.$$

Since $|u_n| \leq 1$, we know that for fixed h,

$$\limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=N-h}^{N} u_n \right| \leq \limsup_{N\to\infty} \frac{h}{N} = 0.$$

Therefore,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-h-1} u_n = \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} u_n \,. \tag{1.7}$$

Now apply (1.7) to $u_n = a_n \overline{a}_{n+h}$, we obtain

$$\limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} a_n \right|^2 \leq \frac{1}{H} + \frac{2}{(H+1)^2} \sum_{h=1}^{H} (H+1-h) Re \left(\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} a_n \overline{a}_{n+h} \right),$$

so set $C > 2$, and the claim holds.

To show (1.5), we utilize Corollary 1.1.10 and the fact that $\sup_n |a_n|^2 \leq 1$ to see that

$$\frac{2}{NH} \sum_{n=1}^{N} |a_n|^2 \leq \frac{2}{H}.$$

Choose $C > 4$, and we obtain the desired inequality.

To show (1.6), we apply limit supremum (as $N \to \infty$) to both sides of (1.5), which gives us

$$\limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} a_n e^{2\pi i n t} \right|^2 \leq \frac{C}{H} + \frac{C}{H} \sum_{h=1}^{H} \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n-1}^{N} a_n \overline{a}_{n+h} \right|.$$

We apply (1.7) to $u_n = a_n\overline{a}_{n+h}$, and we obtain the desired inequality. $\square$

1.2 Wiener-Wintner Ergodic Theorem

The study of Wiener-Wintner averages originally appeared in the work of N. Wiener and A. Wintner from 1941, which is strengthens Birkhoff's pointwise ergodic theorem (Theorem 1.1.4) in the following way:

Theorem 1.2.1 (Wiener-Wintner Ergodic Theorem, [82]). *Let $(X, \mathcal{F}, \mu, T)$ be a measure-preserving system, and $f \in L^\infty(\mu)$. Then there exists a set of full measure X_f such that for any $x \in X_f$, the limit*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)e(nt)$$

exists for all $t \in \mathbb{R}$ (where $e(\alpha) := e^{2\pi t\alpha}$).

We recall that the case $t = 0$ gives the pointwise ergodic theorem. The novelty of this claim is that the set of full measure X_f does not depend on the real number $t \in \mathbb{R}$; if one wishes to show the averages converge almost everywhere for a particular value of t, one can simply apply the pointwise ergodic theorem on the product space $X \times \mathbb{T}$ with the transformation $T \times R_t$, where R_t is the rotation on $\mathbb{T}$ by t. While there was an error in the original proof of Wiener and Wintner, the statement is true, as numerous correct proofs were provided later, including the one by H. Furstenberg [49] in 1960.

In [31], Bourgain announced the following uniform Wiener-Wintner result.

Theorem 1.2.2 (Uniform Wiener-Wintner Theorem). *Let $(X, \mathcal{F}, \mu, T)$ an ergodic system, and $f \in L^\infty(\mu)$. Let $\mathcal{K}$ be the Kronecker factor of T. Then the following statements are equivalent.*

1. *The function f belongs to the orthogonal complement of $\mathcal{K}$.*

2. *For μ-a.e. $x \in X$, we have*

$$\lim_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)e(nt) \right| = 0.$$

One may consult [6, Chapter 2] for multiple proofs the Wiener-Wintner theorem, and the uniform Wiener-Wintner theorem. One of the key arguments in establishing this uniformity result is the characterization of a function in the orthogonal complement of the Kronecker factor and its spectral measure.

Proposition 1.2.3 (Characterization of $\mathcal{K}^{\perp}$, cf. [6, Proposition 2.2])**.** *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system. A function f belongs to $\mathcal{K}^{\perp}$ if and only if its spectral measure σ_f is continuous.*

One notices from the uniformity theorem that $\mathcal{K}$ is a (pointwise) characteristic factor for the Wiener-Wintner averages, which means that we can characterize the limit of the averages by projecting the function onto the factor $\mathcal{K}$, i.e.

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(T^n x)e(nt) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}(f|\mathcal{K})(T^n x)e(nt).$$

So if one would finish the proof of the Wiener-Wintner Ergodic Theorem for an ergodic system $(X, \mathcal{F}, \mu, T)$, one can decompose $f = f_1 + f_2$, where $f_1 = \mathbb{E}(f|\mathcal{K})$, and $f_2 = f - f_1$. This means that we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(T^n x)e(nt) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_1(T^n x)e(nt) + \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_2(T^n x)e(nt).$$

Since $f_2 \in \mathcal{K}^{\perp}$, the uniform Wiener-Wintner theorem (Theorem 1.2.2) asserts that the second limit of the averages on the right hand side of this equation is 0. Thus, it remains to show that the limit of the first averages exist. But since $f_1 \in \mathcal{K}$, and since $\mathcal{K}$ is the closed linear span of eigenfunctions of T in $L^2(\mu)$, one can prove the existence of the limit by assuming that f_1 is an eigenfunction of T.

The way of showing the convergence of ergodic averages by decomposing a function into a sum of two functions, where one function is the orthogonal projection to a factor, and the other to its orthogonal complement, is sometimes called the method of characteristic factor decomposition. This method was pioneered by Furstenberg [50], and made more explicit in the work of Furstenberg and Weiss [53]. The goal of this method is to show that (1) if the functions belongs to the factor, use the structure of the factor to show the convergence, and (2) if the functions belongs to the orthogonal complement of the factor, show that the averages converge to zero by applying certain estimates on the averages. More on this will be discussed in the next section.

The Wiener-Wintner theorem has been extended in various directions, and we will discuss some of them. Recently, G. Hong and M. Sun has announced an operator-algebraic extension of the

Wiener-Wintner result to noncommutative trace preserving dynamical system [57], which provides a multi-parameter version of Bellow and Losert's Wiener-Wintner type ergodic theorem [26].

1.3 Multiple recurrence, Host-Kra-Ziegler factors, and Gowers-Host-Kra seminorms

In 1990, Bourgain utilized Theorem 1.2.2 to establish the double recurrence theorem, which is stated as follows:

Theorem 1.3.1 (Bourgain's Double Recurrence Theorem, [31]). *Let $(X, \mathcal{F}, \mu, T)$ an ergodic system, and $f_1, f_2 \in L^\infty(\mu)$. Then for any distinct nonzero integers $a, b \in \mathbb{Z}$, the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x)f_2(T^{bn}x) \tag{1.8}$$

converge for μ-a.e. $x \in X$ as $N \to \infty$.

In fact, the averages above converge if $f_1 \in L^p(\mu)$ and $f_2 \in L^q(\mu)$ for any $p, q \in [1, \infty]$, provided that $1/p + 1/q \leq 1$ (here we treat $1/\infty = 0$). We note that the convergence does not need to hold when $p = q = 1$, as it was shown by Assani and Buczolich [14, Theorem 3].

The averages seen in (1.8) is an example of multiple recurrence averages, which are averages of the form

$$\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{k} f_i(T_i^n x), \tag{1.9}$$

where x is an element of a measure space $(X, \mathcal{F}, \mu)$, $f_1, \dots, f_k \in L^\infty(\mu)$, and $T_1, \dots, T_k$ are measure-preserving transformations on X. These averages are also commonly referred to as multiple ergodic averages or nonconventional ergodic averages. In 1977, H. Furstenberg provided an ergodic theoretic proof of Szemerédi's theorem, which states that a subset of $\mathbb{Z}$ with positive upper density has an arbitrarily long arithmetic progression, by studying structures behind multiple recurrence averages [50]. More precisely, he has shown that there is a correspondence, often referred to as Furstenberg's correspondence principle, between the structure of arithmetic progressions and a measure preserving system $(X, \mathcal{F}, \mu, T)$, and Szemerédi's theorem can be proven by showing that for any measurable set B with positive measure and for any positive integer k, one has

$$\mu(B \cap T^{-n}B \cap T^{-2n}B \cap \cdots \cap T^{-(k-1)n}B) > 0$$

for any nonzero integer n (cf. [50, Theorem 1.4]). In fact, Furstenberg showed something stronger in a sense that one has

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mu(B\cap T^{-n}B\cap\cdots T^{-(k-1)n}B) > 0 \tag{1.10}$$

for any positive integer k and any set B with a positive measure. This result was later generalized by Furstenberg and Katznelson [52, Theorem A] for the case when one has a measure-preserving system with commuting transformations $(X, \mathcal{F}, \mu, T_1, T_2, \ldots, T_k)$, and $B \in \mathcal{F}$ with a positive measure, and k a positive integer, one has

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mu(T_1^{-n}B\cap T_2^{-n}B\cap\cdots T_k^{-n}B) > 0, \tag{1.11}$$

and they used this result to prove the multidimensional version of Szemerédi's theorem. Later, (1.11) was generalized by Bergelson and Leibman to the case of polynomial actions [28, Theorem A], and furthermore by Leibman to the case where the group of transformations generate a nilpotent group [66, Theorem NM'].

Since Furstenberg's study on multiple recurrence, the averages on (1.9) have been studied extensively. The problem regarding $L^2(\mu)$-norm convergence of these averages has been settled. In 1984, Conze and Lesigne showed this for the case for $k = 2$, and also for $k = 3$ if each T_i is a power of a single measure-preserving transformation [35]. Zhang later showed that the norm convergence holds for the case $k = 3$ while assuming that each T_i and $T_i \circ T_j^{-1}$ were ergodic, for $i \neq j$ in 1996 [83], and under the same assumptions, Frantzikinakis and Kra showed the convergence for any $k \geq 1$ in 2004 [48][1]. In 2005, Host and Kra showed that the averages in (1.9) converge in norm for the case when $T_i = T^i$ where T is a measure-preserving transformation, for each $i = 1, 2, \ldots, k$ [59]; this result was also obtained by Ziegler independently [84]. In 2008, Tao showed that (1.9) converge in norm when the transformations commute, without assuming any ergodicity on the transformations [79]. Shortly after, alternative proofs of Tao's result were provided by Austin [24], Host [58], and Towsner

[1]One of the key ingredients of the proof of this result was to show that when T_1 and T_2 are two commuting ergodic transformations and $T_1 \circ T_2^{-1}$ is ergodic, then their Host-Kra-Ziegler factors $\mathcal{Z}(T_1)$ and $\mathcal{Z}(T_2)$ are identical, and one can use an appropriate Gowers-Host-Kra seminorm to obtain an upper bound of the averages [48, Propositions 3.1 and 3.2]. These results were obtained by Assani independently, and communicated to the authors.

[80]. In 2012, Walsh showed that those averages converge in norm provided that $T_1, \ldots, T_k$ generate a nilpotent group, and in a view of the counterexamples provided by Bergelson and Leibman in 2002 [29] when the transformations generate a solvable group, Walsh's result is considered to be the complete result for the norm convergence of multiple recurrence averages. An alternative proof to Walsh's result that uses coupling was later given by Austin in 2013 [25], assuming the result regarding finite complexity of a system with nilpotent group action that was shown by Walsh [81, Theorem 4.2]. Also, some extensions of Walsh's result were provided by Zorin-Kranich [86] (for amenable group actions onto a nilpotent group) and by Mesón and Vericat [74] (for the spherical averages of Markov group actions onto a nilpotent group)—we remark both of these work used elements of Walsh's work, regarding the Hahn-Banach theorem as well as the notion of the complexity of the system.

For the a.e. convergence avenue of the multiple recurrence averages, however, the results are much scarcer. In fact, Bourgain's double recurrence theorem (Theorem 1.3.1) remains to be the best result in this direction. However, there are some significant partial results when one assumes more about the dynamical system and/or the transformations. For instance, Assani has shown that if the system $(X, \mathcal{F}, \mu, T)$ is a weakly-mixing space for which the restriction of T on its Pinsker algebra (i.e. the maximal sub-σ-algebra of $\mathcal{F}$ for which T has zero entropy), then the multiple recurrence averages (1.9) for the case $T_i = T$ for each i converge for μ-a.e. $x \in X$ [2, Theorem 2]. Furthermore, when $k = 2$, Assani has shown that the averages in (1.9) converge a.e. under certain assumptions [8, Proposition 5, Theorem 6], answering some of the questions raised by Furstenberg. Other pointwise convergence results were obtained by Derrien and Lesigne [39], where they have shown for the case where they have $T_i = T$, where T is an exact automorphism or K-automorphism, and each exponent in is replaced by a integer-valued polynomial $q_i(n)$ with rational coefficients. Recently, Huang, Shao, and Ye announced that (1.9) converge for the case each $T_i = T$, and $(X, \mathcal{F}, \mu, T)$ is a distal system [62, Theorem C], using some of the matters discussed in Assani's attempts on the pointwise convergence of some averages in the form of (1.9) [11, 13]. Furthermore, Donoso and Sun have recently announced a pointwise convergence result for the case $k = 2$, and T_1 and T_2 commuting, provided that $(X, \mathcal{F}, \mu, T_1, T_2)$ is a distal system [41].

It is worth mentioning that much simpler proof of the double recurrence theorem is available if we assume the system to be a Wiener-Wintner dynamical system, which is a collection of dynamical systems introduced by I. Assani that contains many types of dynamical systems, such as K-systems, systems with discrete spectrum, and some skew products (see [6, Theorem 6.3] for this simple proof, and see [5, 7] for more information on Wiener-Wintner dynamical systems).

In the works of Conze-Lesigne, Host-Kra, and Ziegler, (norm) characteristic factors were identified in order to show the convergences. An advantage of this strategy is that one can use nice algebraic structure of such factor to show convergence for the case where each function is measurable with respect to this factor. Once this is accomplished, the remaining task is to show that the averages converge to zero when one of the function belongs to the orthogonal complement of such factor. Here, we will focus on the factors that appeared in the work of Host-Kra and Ziegler, which we shall refer to as Host-Kra-Ziegler factors. We remark that these characteristic factors were for mean convergence, as the pointwise characteristic factors may not coincide with the mean characteristic factors (these differences of notions were mentioned explicitly by Assani in [10]).

1.3.1 Averages along cubes

While the results regarding pointwise convergence of the multiple recurrence averages are scarce, the story is quite different for the pointwise convergence of averages along cubes. Given a positive integer k, we let $V_k^* = V_k - \{\mathbf{0}\}$, where we recall $V_k = \{0,1\}^k$ and $\mathbf{0} = (0,0,\ldots,0) \in V_k$. Note that V_k^* has $2^k - 1$ elements. Suppose a probability measure space $(X, \mathcal{F}, \mu)$ is equipped with $2^k - 1$ bounded functions $\{f_\epsilon\}_{\epsilon \in V_k^*} \subset L^\infty(\mu)$, and $2^k - 1$ measure-preserving transformations $\{T_\epsilon\}_{\epsilon \in V_k^*}$. Suppose $\mathbf{n} = (n_1, n_2, \ldots, n_k) \in \mathbb{Z}^k$, and $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_k) \in V_k$. We define $\mathbf{n} \cdot \epsilon := \sum_{i=1}^k n_i \epsilon_i$ to be the usual dot product. Finally, we define the k-term averages along cubes to be the averages of the form

$$\frac{1}{N^k} \sum_{\mathbf{n} \in [1,N]^k} \prod_{\epsilon \in V_k^*} f_\epsilon(T_\epsilon^{\mathbf{n}\cdot\epsilon} x). \tag{1.12}$$

For instance, when $k = 2$, the averages can be written as

$$\frac{1}{N^2} \sum_{n,m=1}^{N} f_1(T_1^n x) f_2(T_2^m x) f_3(T_3^{n+m} x),$$

and when $k = 3$, the averages can be written as

$$\frac{1}{N^3} \sum_{n,m,p=1}^{N} f_1(T_1^n x) f_2(T_2^m x) f_3(T_3^{n+m} x) f_4(T_4^p x) f_5(T_5^{n+p} x) f_6(T_6^{m+p} x) f_7(T_7^{n+m+p} x).$$

The averages of these forms arise naturally in the study of multiple ergodic averages and arithmetic progressions. Furthermore, these averages can be viewed as discrete versions of the continuous averages introduced by T. Gowers [54].

For the case $T_\epsilon = T$ is a single measure-preserving transformation on $(X, \mathcal{F}, \mu)$, the $L^2(\mu)$-norm convergence of the averages in (1.12) was shown by Host and Kra for $k = 3$ in [61, Theorem 3], and later for any $k \in \mathbb{N}$ in [59, Theorem 1.2]. For pointwise convergence, Assani has shown in 2003 that the averages in (1.12) converge for any $k \in \mathbb{N}$ when each T_ϵ commute [10]. In 2007, he has also shown that the averages converge for six bounded functions when the transformations $T_1, T_2, \ldots, T_6$ do not necessarily commute [9]. Later, by using Assani's matter, Chu and Frantzikinakis have shown that the averages in (1.12) converge for any $k \in \mathbb{N}$ for noncommuting transformations.

For $k = 2$ and $T_1 = T_2 = T_3$, the cubic averages are known to converge pointwise for the case $f_i \in L^{p_i}(\mu)$ for $i = 1, 2, 3$, when $1/p_1 + 1/p_2, +1/p_3 \leq 5/2$. This result was shown by Demeter, Tao, and Thiele [38]. In 2014, Donoso and Sun have shown that the averages

$$\frac{1}{N^2} \sum_{n,m=1}^{N} f_1(S^n x) f_2(T^m x) f_3(S^n T^m x)$$

converge for μ-a.e. $x \in X$, where S and T are measure-preserving transformations on $(X, \mathcal{F}, \mu)$ that commutes with each other [40].

1.3.2 Gowers-Host-Kra seminorms

Following the work of Host and Kra [59], we define a class of seminorms on a set of bounded and measurable functions. We recall that given a measure space $(X, \mathcal{F}, \mu)$, the map $|||\cdot||| : L^\infty(\mu) \to [0, \infty)$ is a seminorm if it satisfies

1. If $a \in \mathbb{C}$ and $f \in L^\infty(\mu)$, then $|||af||| = |a| |||f|||$, and

2. If $f, g \in L^\infty(\mu)$, then $|||f + g||| \leq |||f||| + |||g|||$.

In other words, a seminorm satisfies all of the properties of norms, except that it is possible for $|||f||| = 0$ even for a nonzero function $f \in L^\infty(\mu)$.

Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system. For each nonnegative integer, we denote $X^{[k]} = X^{2^k}$, and $\mathcal{F}^{[k]} = \mathcal{F}^{2^k}$. The coordinate on $X^{[k]}$ will be denoted in terms of a k-dimensional vector of 0's and 1's, i.e. an element of $V_k := \{0, 1\}^k$, and in particular, we denote $\mathbf{0} = (0, 0, \ldots, 0) \in V_k$.

By following [59, §3], we construct measure $\mu^{[k]}$ on $X^{[k]}$ inductively. First, $\mu^{[0]} = \mu$. Next, we construct $\mu^{[1]}$ from relative products (cf. [51, Definition 5.7]) of $\mu^{[0]}$ as follows: Given $f_0, f_1 \in L^\infty(\mu)$, we define

$$\int_{X^{[1]}} f_0 \otimes f_1(x_0, x_1) d\mu^{[1]}(x_0, x_1) := \int_X \mathbb{E}(f_0|\mathcal{I}(T))(x)\mathbb{E}(f_1|\mathcal{I}(T))(x) d\mu(x).$$

Note that, in this special case where T is ergodic, $\mathcal{I}(T)$ is a trivial σ-algebra, so $\mathbb{E}(f_i|\mathcal{I}(T)) = \int f_i d\mu$ for both $i = 0, 1$.

Now to define $\mu^{[2]}$ from $\mu^{[1]}$, we consider four functions $f_{00}, f_{01}, f_{10}, f_{11} \in L^\infty(\mu)$. Then we define

$$\int_{X^{[2]}} \bigotimes_{\epsilon \in V_2} f_\epsilon d\mu^{[2]} = \int_{X^{[1]}} \mathbb{E}\left(\bigotimes_{\eta \in V_1} f_{\eta 0} | \mathcal{I}(T \times T)\right) \mathbb{E}\left(\bigotimes_{\eta \in V_1} f_{\eta 1} | \mathcal{I}(T \times T)\right) d\mu^{[1]},$$

and similarly, for any positive integer k, we define $\mu^{[k]}$ to be a measure on $X^{[k]}$ such that

$$\int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} f_\epsilon d\mu^{[k]} = \int_{X^{[k-1]}} \mathbb{E}\left(\bigotimes_{\eta \in V_{k-1}} f_{\eta 0} | \mathcal{I}^{[k-1]}\right) \mathbb{E}\left(\bigotimes_{\eta \in V_{k-1}} f_{\eta 1} | \mathcal{I}^{[k-1]}\right) d\mu^{[k-1]},$$

where where $\mathcal{I}^{[k-1]}$ denotes the collection of sets in $X^{[k]}$ that are invariant under the transformation $T \times T \times \cdots \times T$ (2^k times).

Using these measures, we construct the following seminorms on $L^\infty(\mu)$ as follows: For any $f \in L^\infty(\mu)$, we define

$$|||f|||_k = \left(\int_{X^{[k]}} \prod_{\epsilon \in V_k} f(x_\epsilon)\, d\mu^{[k]}\right)^{1/2^k} = \left(\int_{X^{[k-1]}} \left(\mathbb{E}\left(\bigotimes_{\epsilon \in V_{k-1}} f | \mathcal{I}^{[k-1]}\right)\right)^2 d\mu^{[k-1]}\right)^{1/2^k}. \quad (1.13)$$

Since the value inside the integral of the third expression of (1.13) is nonnegative, $|||f|||_k$ is nonnegative for each positive integer k and $f \in L^\infty(\mu)$. In fact, one can understand from [59, Lemma 3.9] that $|||\cdot|||_k$ is indeed a seminorm for each k, and furthermore, $|||f|||_k \leq |||f|||_{k+1}$ for every $f \in L^\infty(X)$.

If one works with a complex-valued function, one makes the following modification to the seminorms above. First if $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_k) \in V_k$, we define $|\epsilon| := \sum_{i=1}^k \epsilon_i$. We also let $C : \mathbb{C} \to \mathbb{C}$ to be the complex conjugate map, i.e. $Cz = \bar{z}$. In particular, $C^m z = z$ if m is even, and $C^m z = \bar{z}$ if

m is odd. So if $f \in L^\infty(X, \mathbb{C})$, then we define

$$\left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_k = \left(\int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} C^{|\epsilon|} f \, d\mu^{[k]} \right)^{1/2^k} . \tag{1.14}$$

Another way to characterize these seminorms is by using the ergodic averages, which is much more convenient for our purposes. When $k = 1$,the fact that T is ergodic tells us that

$$\left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_1^2 = \int \mathbb{E}(f|\mathcal{I}(T))^2 d\mu = \left(\int f d\mu \right)^2 .$$

Next, for $k = 2$, we use the fact that the conditional expectation is operator is an orthogonal projection to see that

$$\left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_2^4 = \int \mathbb{E}(f \otimes f|\mathcal{I}^{[1]})^2 d\mu^{[1]} = \int f \otimes f(x_0, x_1) \cdot \mathbb{E}(f \otimes f|\mathcal{I}^{[1]})(x_0, x_1) d\mu^{[1]}.$$

By the pointwise ergodic theorem as well as the dominated convergence theorem, we have

$$\left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_2^4 = \lim_{H \to \infty} \frac{1}{H} \sum_{h=1}^{H} \int (f \cdot f \circ T^h)(x_0)(f \cdot f \circ T^h)(x_1) d\mu^{[1]} = \lim_{H \to \infty} \frac{1}{H} \sum_{h=1}^{H} \left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \cdot f \circ T^h \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_1^2 .$$

We note that one can express $\left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_4^2$ in terms of the limit of the Cesàro averages of $\left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \cdot f \circ T^h \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_1^2$. Indeed, we can express these seminorms recursively for any positive integer k. To demonstrate this, we denote $\otimes^{[k]} f = f \otimes f \otimes \cdots \otimes f$ (2^k times), and we compute

$$\left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_k^{2^k} = \int \mathbb{E}(\otimes^{[k-1]} f|\mathcal{I}^{[k-1]})^2 \, d\mu^{[k-1]} = \int \otimes^{[k-1]} f \cdot \mathbb{E}(\otimes^{[k-1]} f|\mathcal{I}^{[k-1]}) \, d\mu^{[k-1]} ,$$

and again, by the pointwise ergodic theorem, we obtain

$$\left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_k^{2^k} = \lim_{H \to \infty} \frac{1}{H} \sum_{h=1}^{H} \left\vert\kern-0.25ex\left\vert\kern-0.25ex\left\vert f \cdot f \circ T^h \right\vert\kern-0.25ex\right\vert\kern-0.25ex\right\vert_{k-1}^{2^{k-1}} . \tag{1.15}$$

We remark here that if X is a finite cyclic group (i.e. $X = \mathbb{Z}/N\mathbb{Z}$ for some positive integer N), and if T is a transformation for which $T(a) = a + 1$, then these seminorms are the same ones that were used in the work of W. T. Gowers [54]. Henceforth, we refer to these seminorms as Gowers-Host-Kra seminorms (or "GHK seminorms" for short).

If f is a complex-valued function, one can derive from (1.14) that

$$|||f|||_k^{2^k} = \lim_{H\to\infty} \frac{1}{H}\sum_{h=1}^{H} \left|\left|\left|\bar{f}\cdot f\circ T^h\right|\right|\right|_{k-1}^{2^{k-1}}. \tag{1.16}$$

1.3.3 Host-Kra-Ziegler factors

Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system. In [59, §4], Host and Kra constructed factors of the system that are characterized by the Gowers-Host-Kra seminorms, which are also norm characteristic factors for the multiple recurrence averages for the case one has a single ergodic transformation (i.e. (1.9) for the case when $T_i = T$). In [84], T. Ziegler independently obtained characteristic factors for the same averages, and while the method of constructing them was different (in particular, the seminorms were not used), these factors correspond to the ones obtained by Host and Kra. Here, we refer to these characteristic factors as Host-Kra-Ziegler factors.

In [59, §4], Host and Kra have shown that there exists a factor of $(Z_k, \mathcal{Z}_k, \mu_k, T)$ such that given $f \in L^\infty(\mu)$, one has

$$\mathbb{E}(f|\mathcal{Z}_k) = 0 \text{ if and only if } |||f|||_{k+1} = 0 \text{ (cf. [59, Lemma 4.3])}.$$

The factor $\mathcal{Z}_{k-1}$ turns out to be a norm characteristic factors for the averages of the form

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x) f_2(T^{2n}x)\cdots f_k(T^{kn}x) \tag{1.17}$$

for any $f_1, f_2, \ldots, f_k \in L^\infty(\mu)$. In fact, they have shown that if $\|f_i\|_{L^\infty(\mu)} \le 1$, then

$$\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{k} f_i\circ T^{in}\right\|_{L^2(\mu)} \le \min_{1\le l\le k}(l\cdot |||f_l|||_k) \text{ (cf. [59, Theorem 12.1])}.$$

1.3.4 Nilsystems, nilsequences, and the structure theorem

We recall the notions of nilsystems that appear in numerous results of ergodic theory briefly. Let G be a k-step nilpotent Lie group for some positive integer k, and Γ be its co-compact subgroup.

We say the homogeneous space $X = G/\Gamma$ is a k-step nilmanifold. Given $g \in G$, we let T_g to be the left group action on X by g (i.e. for any $x \in X$, $T_g x = g \cdot x$. Since X is a compact manifold, there exists a unique probability Haar measure μ. We call the system $(X, \mathcal{B}, \mu, T_g)$ a k-step nilsystem, where $\mathcal{B}$ is the Borel σ-algebra.

One of the remarkable features of the Host-Kra-Ziegler factors is that they are inverse limits of nilsystems of the same degree. More precisely,

Theorem 1.3.2 (The structure theorem, [59, Theorem 10.1]). *The k-th Host-Kra-Ziegler factor $(Z_k, \mathcal{Z}_k)$, for $k \geq 1$, can be expressed as an inverse limit of a sequence of k-step nilsystems.*

Combined with the fact that $\mathcal{Z}_k$ is a norm characteristic factor of the averages (1.17), the structure theorem and a density argument will show that it is sufficient to show that the averages converge on a k-step nilsystem. The study of ergodic averages on nilmanifolds have been done extensively by, for instance, E. Lesigne [72] and A. Leibman (e.g. [67, 68, 69, 70]). The following result is of our interest is due to Leibman.

Theorem 1.3.3 ([68, Theorem A]). *Let $Y = G/\Gamma$ be a nilmanifold. If $g : \mathbb{Z} \to G$ such that $g(n) = \prod_{i=1}^{k} a_i^{p_i(n)}$ for some $a_1, \ldots, a_k \in G$ and integer-valued polynomials $p_1(n), \ldots, p_k(n)$, then for any continuous function F on Y, and for any $y \in Y$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F(g(n)y) \text{ exists.}$$

We remark that the limit exists for all $y \in Y$, rather than for almost every $y \in Y$. To apply Leibman's result to show the convergence of (1.17), we first assume that $X = G/\Gamma$ is a k-step nilsystem, and $f_1, f_2, \ldots, f_k$ are continuous functions on X. We set $F := f_1 \otimes f_2 \otimes \cdots \otimes f_k$ to be a continuous function on $Y = X^k = (G/\Gamma)^k$, which is a k-step nilmanifold, and then we define $y = (x, x, \ldots, x)$, and $g(n) = \prod_{i=1}^{k} a_i^{in}$, where $a_1 = T \times \mathrm{Id} \times \cdots \times \mathrm{Id}, a_2 = \mathrm{Id} \times T \times \cdots \times \mathrm{Id}, \ldots, a_k = \mathrm{Id} \times \mathrm{Id} \times \ldots \mathrm{Id} \times T$ are all elements of G^k. Now we apply Theorem 1.3.3, and the norm convergence of (1.17) holds after standard approximation arguments (first so that the convergence holds for L^∞ functions on a nilsystem, and then to the original space given using the structure theorem).

We conclude this section with a summary of nilsequences. Let G/Γ be a k-step nilmanifold. We say a sequence (a_n) is a basic k-step nilsequence if there exists a continuous function f on G/Γ, a group element $\tau \in G$ and $x \in G/\Gamma$, one has $a_n = f(\tau^n x)$. We say a sequence (b_n) is a k-step nilsequence if it is a uniform limit of basic k-step nilsequence. Nilsequences have been studied and utilized in various contexts of ergodic theory (e.g. [27, 60]), and the interests in these sequences also appear in number theory, as one can see in the work of Green, Tao, and Ziegler (e.g. [55, 56]).

1.4 Weighted averages and return times

1.4.1 Good universal weights

In some literatures (e.g. [6, Definitions 3.1-3.3]), the sequence (a_n) is called a *good universal weight for the pointwise ergodic theorem* if for any probability measure preserving system $(Y, \mathcal{G}, \nu, S)$ and any $g \in L^\infty(\nu)$, the averages

$$\frac{1}{N}\sum_{n=0}^{N-1} a_n g(S^n y) \tag{1.18}$$

converge for ν-a.e. $y \in Y$. Similarly, the sequence (a_n) is called a *good universal weight for the mean ergodic theorem* if the averages in (1.18) converge in $L^2(\nu)$.

In this paper, we will extend these classical notions of good universal weights to discuss the case where the sequence $(g \circ S^n)_n$ in (1.18) is replaced by other sequences of bounded and measurable functions $(X_n)_n$.

Definition 1.4.1. *Let $(\Omega, \mathcal{S}, \mathbb{P})$ be a probability measure space. We say $(X_n)_n$ is a process on $(\Omega, \mathcal{S}, \mathbb{P})$ if for all nonnegative integers $n \geq 0$, X_n is a bounded and measurable function on $(\Omega, \mathcal{S}, \mathbb{P})$.*

For instance, a sequence of bounded and measurable functions $(X_n)_n = (g \circ S^n)_n$ for any $g \in L^\infty(\nu)$ on any probability measure-preserving system $(Y, \mathcal{G}, \nu, S)$ is a process. Another process $(X_n)_n$ of our interest is the sequence that appears in the multiple ergodic averages (1.9), i.e. for any measure-preserving system $(Y, \mathcal{G}, \nu, S_1, S_2, \ldots, S_k)$, one may have

$$X_n^k(y) = g_1(S_1^n y) g_2(S_2^n y) \cdots g_k(S_k^n y),$$

for any positive integer k, where $g_1, g_2, \ldots, g_k \in L^\infty(\nu)$ on any measure-preserving system $(Y, \mathcal{G}, \nu, S)$.

Definition 1.4.2. *We denote by*

$$M_1 = \left\{ (a_n) : \sup_N \frac{1}{N} \sum_{n=1}^{N} |a_n| < \infty \right\}.$$

We denote Π to be a collection of probability measure spaces, and $\mathfrak{X}(\Pi)$ be a collection of processes on a probability measure space $(\Omega, \mathcal{S}, \mathbb{P}) \in \Pi$.

- *We say a sequence $(a_n) \in M_1$ is a good universal weight for $\mathfrak{X}(\Pi)$ (a.e.) pointwise, if for any probability space $(\Omega, \mathcal{S}, \mathbb{P}) \in \Pi$, and any process $(X_n) \in \mathfrak{X}(\Pi)$ on Ω, the averages*

$$\frac{1}{N}\sum_{n=1}^{N} a_n X_n(\omega)$$

converge for $\mathbb{P}$-a.e. $\omega \in \Omega$.

- *We say a sequence $(a_n) \in M_1$ is a good universal weight for $\mathfrak{X}(\Pi)$ in norm, if for any probability space $(\Omega, \mathcal{S}, \mathbb{P}) \in \Pi$ and any process $(X_n) \in \mathfrak{X}(\Pi)$ on Ω, the averages*

$$\frac{1}{N}\sum_{n=1}^{N} a_n X_n(\omega)$$

converge in $L^2(\mathbb{P})$.

For example, if (a_n) is a good universal weight for the pointwise ergodic theorem, one can say that (a_n) is a good universal weight for $\mathfrak{X}(\Pi)$ pointwise, where Π is a collection of all the measure-preserving system $(Y, \mathcal{G}, \nu, S)$, and $\mathfrak{X}(\Pi)$ is a collection of sequences of the form $(g \circ S^n)$, where $(Y, \mathcal{G}, \nu, S) \in \Pi$, and $g \in L^\infty(\nu)$.

1.4.2 History of the return times theorem

The studies of the return times theorem have shown that we can randomly generate good universal weights. The basic principle of the return times theorem that has been initially studied by A. Brunel in his Ph.D. thesis in 1966 [33] is as follows: Given a process $X_n(\omega)$ converging in average (in norm or pointwise) and the characteristic function of a measurable set with positive measure, $\mathbb{1}_A$, do we still have the convergence of the averages along the subsequence given by the return times of $T^n x$ to the set A? In other words, is the sequence $(\mathbb{1}_A(T^n x))_n$ a good universal weight (in norm or pointwise) for the averages of $\mathbb{1}_A(T^n x)X_n(\omega)$? In 1969, A. Brunel and M. Keane answered this question positively for a particular class of dynamical systems for both pointwise and norm convergence [34]. Krengel's book highlights some of the generalization of their work [64].

One of the important results in ergodic theory is the proof of return times theorem by J. Bourgain

[30], which was later simplified by J. Bourgain himself, along with H. Furstenberg, Y. Katznelson, and D. Ornstein (this simplified proof is sometimes referred to as the "BFKO" argument, where the first letter of of the last name of each author is taken) [32]. This result strengthens Birkhoff's pointwise ergodic theorem and generalizes the above-mentioned results on return times.

Theorem 1.4.3 (Bourgain's Return Times Theorem)**.** *Let $(X, \mathcal{F}, \mu, T)$ be a probability measure-preserving system and $f \in L^\infty(\mu)$. Then there exists a set $X_f \subset X$ of full measure such that for any other probability measure-preserving system $(Y, \mathcal{G}, \nu, S)$ and any $g \in L^\infty(\nu)$,*

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

converges ν-almost everywhere for all $x \in X_f$.

An alternative proof of this result was obtained by Rudolph, which uses the notions of joinings [77]. It can be seen in Rudolph's work that the return times theorem holds if $f \in L^p(\mu)$, $g \in L^q(\nu)$, where $1 \leq p, q \leq \infty$, and $1/p + 1/q \leq 1$. But we note that I. Assani, Z. Buczolich, and R. D. Mauldin have shown that the convergence need not hold when $p = q = 1$ [15, Corollary 3], answering Assani's question on the break of duality for the return times theorem that was raised in 1991 [23, §3]. On the other hand, C. Demeter, M. Lacey, T. Tao, and C. Thiele have shown that the return times theorem holds when $1 < p \leq \infty$ and $q \geq 2$ [37, Theorem 3.1], which also answers Assani's question on the break of duality.

While the set of full measure X_f depends on the function f and the transformation T, it is independent of every other ergodic system. We notice that the statement of the return times theorem is in some way similar to that of the Wiener-Wintner theorem, in a sense that one can obtain a set of full measure independent of uncountably many averages. In fact, by an application of the spectral theorem (Theorem 1.1.6), one has the following norm convergence result.

Theorem 1.4.4 (cf. [6, Theorem 3.1])**.** *Suppose $(a_n) \in M_1$ (where M_1 is as in Definition 1.4.2). The sequence is a good universal weight for the mean ergodic theorem if and only if the averages*

$$\frac{1}{N}\sum_{n=1}^{N} a_n e(nt)$$

converge for every $t \in \mathbb{R}$.

1.4.3 Extensions of the return times theorem

Much of the background, historical development, and current status of the return times can be found in the survey paper prepared by I. Assani and K. Presser [23]. Here, we will focus on discussing some of the developments on the return times theorem regarding mixing of multiple recurrence and multi-term return times problems. Some new results that appeared since the emergence of the survey paper are mentioned as well.

Since the result of Bourgain emerged, the return times theorem has been extended in multiple direction. One way is to find a new universal weight in which the return-times averages converge. For instance, Assani shows in [6, Proposition 5.3] that if $(X, \mathcal{F}, \mu, T)$ is a weakly-mixing, standard uniquely ergodic system with Lebesgue spectrum, and $f \in \mathcal{C}(X)$, then $(f(T^n x))$ is a good universal weight for the pointwise ergodic theorem for all $x \in X$. Recently, P. Zorin-Kranich announced the extension of Bourgain's return times theorem by showing that the double recurrence sequence is a good universal weight for the pointwise ergodic theorem for μ-a.e. $x \in X$ [87].

The return times theorem has also been extended to averages with more than two terms. One example of such is the multiterm return times theorem that was obtained by D. Rudolph in 1998 [78], which answers one of the questions raised by Assani in 1991. Rudolph's proof utilized the method of joinings and fully generic sequences, while the method of factor decomposition was absent, which was one of the key tools in the BFKO argument of the return times theorem. Later, Assani and Presser identified pointwise characteristic factors for the multiterm return times theorem [21, 22]. Furthermore, P. Zorin-Kranich provided a different proof of the multiterm return times theorem based on these factor structures, and showed that multiterm return times averages can be extended to Wiener-Wintner type averages with nilsequences [85]. Also, T. Eisner showed the convergence of Wiener-Wintner type averages for multiterm return times theorem with linear sequences [43].

In another direction, the return times theorem has been extended by mixing weights from the a.e. multiple recurrence and the multiterm return times theorem. This idea was introduced by Assani in 1998, in which he proved the following:

Theorem 1.4.5 ([3, Theorem 3])**.** *Let $(X, \mathcal{F}, \mu, T)$ be a weakly mixing dynamical system such that for all positive integers H, for all $f_1, f_2, \dots f_H \in L^\infty(\mu)$, for all $(b_1, b_2, \dots, b_H) \in \mathbb{Z}^H$ where b_i*

distinct and not equal to zero, the sequence

$$\frac{1}{N}\sum_{n=1}^{N}\left(\prod_{i=1}^{H} f_i(T^{b_i n}x)\right) \text{ converges a.e. to } \prod_{i=1}^{H}\int f_i d\mu\,.$$

Then there exists a set of full measure X' for any other weakly mixing 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} in Y_1 such that if $y_1 \in Y_{g_1}$, then . . . for any other weakly mixing 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 $y_{k-1} \in Y_{g_{k-1}}$, then for any other weakly mixing system $(Y_k, \mathcal{G}_k, S_k, \nu_k)$, the sequence

$$\xi_n(x, y_1, y_2, \ldots, y_k) = \left(\prod_{i=1}^{H} f_i(T^{b_i n}x)\right)\left(\prod_{j=1}^{k} g_j(S_j^n y_j)\right)$$

is a good universal weight for the pointwise ergodic theorem for ν_k-a.e. $y_k \in Y_k$.

For instance, if $(X, \mathcal{F}, \mu, T)$ is a weakly mixing system for which the restriction of T to its Pinsker algebra has singular spectrum, then the hypothesis of the theorem above holds. This result was proven by Assani in 1998 [2].

In terms of Definition 1.4.2, Theorem 1.4.5 says that for $k = 1$, there exists a set of full measure $Y_{g_1} \subset Y_1$ such that for all $y_1 \in Y_1$, the sequence $(\prod_{i=1}^{H} f_i(T^{b_i n}x))_n$ is a good universal weight for μ-a.e. $x \in X$ for the process $X_n(z) = X_n[y_1, g_1, S_1](z) = g_1(S_1^n y_1)h(R^n z)$ pointwise, for any measure-preserving system $(Z, \mathcal{Z}, \eta, R)$ and a function $h \in L^\infty(\eta)$.

In 2009, B. Host and B. Kra showed that given an ergodic system $(X, \mathcal{F}, \mu, T)$ and $f \in L^\infty(\mu)$, the sequence $(f(T^n x))$ is a good universal weight for μ-a.e. $x \in X$ for the convergence in L^2-norm of the Furstenberg averages, i.e. they have shown that there exists a set of full measure $X' \subset X$ such that for any $x \in X'$ and any other measure-preserving system $(Y, \mathcal{G}, \nu, S)$ with functions $g_1, \ldots, g_k \in L^\infty(\nu)$, the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x)\prod_{i=1}^{k} g_i \circ S^{in}, \tag{1.19}$$

converge in $L^2(\nu)$ [60, Theorem 2.25]. In particular, if $f = \mathbb{1}_A$ for some measurable set $A \in \mathcal{F}$, then they have shown that the averages of the sequence $(\prod_{i=1}^{k} g_i(S^{in}y))_n$ along the subsequence of the return times of $T^n x$ to the set A converge in $L^2(\nu)$-norm. In the language of Definition 1.4.2, for μ-a.e. $x \in X$, the sequence $(f(T^n x))$ is a good universal weight for $\mathfrak{X}(\Pi)$ in norm, where Π is a collection of all the measure-preserving systems, and

$$\mathfrak{X}(\Pi) = \left\{ X_n = \prod_{i=1}^{k} g_i \circ S^{in} : (Y, \mathcal{G}, \nu, S) \in \Pi, g \in L^\infty(\nu) \right\}. \tag{1.20}$$

This result extends their earlier work in [59], where they proved the result for $f = \mathbb{1}_X$. To show this result, they used the machinery of nilsequences; they showed that if a bounded sequence $(a_n)_n \in \ell^\infty$ has a property that the Cesaro averages of $a_n b_n$ converge for any k-step nilsequence $(b_n)_n$, then $(a_n)_n$ is a good universal weight for k-term multiple recurrent averages in the L^2-norm. Then the convergence of the averages in (1.19) follows from the fact that there exists a set of full measure X' so that for any $x \in X'$ and any nilsequence $(b_n)_n$, the Cesaro averages of $f(T^n x) b_n$ converge; this is referred to as the generalized Wiener-Wintner theorem [60, Theorem 2.22]. Later, in the work of T. Eisner and P. Zorin-Kranich, the generalized Wiener-Wintner theorem was extended to any measure-preserving system (not necessarily ergodic) with uniform counterpart, and used this to extend the result to a case with polynomial actions [45].

1.5 Conventions

Unless stated otherwise, the following conventions are assumed throughout this dissertation.

- If functions $f \in L^\infty(\mu)$ is given for some measure space $(X, \mathcal{F}, \mu)$, we assume the function is real-valued, and bounded above by 1.

- In certain computations, C denotes a universal constant. This constant may change from one line to another.

- When we say $A \lesssim B$, this implies that there exists a universal constant C such that $A \leq CB$. If we say $A \lesssim_{D_1, D_2, \ldots, D_j} B$, then there exists a constant C that only depends on the quantities $D_1, D_2, \ldots, D_j$ such that $A \leq CB$.

CHAPTER 2

Wiener-Wintner double recurrence I: Linear case

In this chapter[1], as well as in Chapter 3, we will discuss the extension of the double recurrence theorem to Wiener-Wintner type averages.

2.1 Background

In 2001, D. Duncan (a former PhD student of I. Assani) worked on extending Bourgain's double recurrence theorem (cf. Theorem 1.3.1) in his PhD thesis [42], and proved the following result.

Theorem 2.1.1 ([42]). *Let $(X, \mathcal{F}, \mu, T)$ be a standard ergodic dynamical system (i.e. X is a compact metrizable space, $\mathcal{F}$ is a Borelian sigma-algebra, μ is a probability Borel measure, and T is a self-homeomorphism). Suppose f_1 and f_2 belong to $L^2(X)$. Let $\mathcal{CL}$ be the maximal isometric extension (see, for example, [42, §1.3] or [53, §5] for a definition) of the Kronecker factor of T. Let*

$$W_N(f_1, f_2, x, t) = \frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) e(nt)$$

1. *(Double Uniform Wiener-Wintner Theorem) If either f_1 or f_2 belongs to $\mathcal{CL}^\perp$, then there exists a set of full measure X_{f_1,f_2} such that for all $x \in X_{f_1,f_2}$,*

$$\limsup_{N\to\infty} \sup_{t\in\mathbb{R}} |W_N(f_1, f_2, x, t)| = 0$$

2. *(General Convergence) If $f_1, f_2 \in \mathcal{CL}$, then $W_N(f_1, f_2, x, t)$ converges for μ-a.e. $x \in X$ for all $t \in \mathbb{R}$, provided that the cocycle associated with $\mathcal{CL}$ is affine.*

Theorem 2.1.1 was proved in several stages. For (1), one first identifies the pointwise limit of the

[1] The material presented in this chapter originally appeared in [16], in the Ergodic Theory & Dynamical Systems. The original citation is as follows: I. Assani, D. Duncan, and R. Moore. Pointwise characteristic factors for Wiener-Wintner double recurrence theorem. Ergod. Th. and Dynam. Sys., 2015. Available on CJO 2015 doi:10.1017/etds.2014.99.

double recurrence averages as an integral with respect to a particular Borel measure (disintegration). Then one uses Wiener's lemma on the continuity of spectral measures and van der Corput's inequality to show that the double recurrence average converges to 0. For the second part, one first shows that the total ergodicity of T asserts that $\mathcal{CL}$ for every integer power of T are the same, which allows one to assume that both functions lie in the same factor of $L^2(X, \mu)$. Furthermore, the assumption that the measurable cocycle associated with $\mathcal{CL}$ is affine allows one to use the homomorphism property to simplify the computations.

A little was known about characteristic factors back then, especially for pointwise convergence. Originally in [42], the factor $\mathcal{CL}$ is referred to as "Conze-Lesigne" factor, as they first appeared in series of work by J.-P. Conze and E. Lesigne (see, for example, [35, 36] for details), and named so by D. Rudolph [76]. But with the works of B. Host and B. Kra in [59, 61], the definition of the Conze-Lesigne factor has been updated when the Host-Kra-Ziegler factors emerged in 2005. It is noted that the updated Conze-Lesigne factor $\mathcal{Z}_2$, which is the second Host-Kra-Ziegler factor, is smaller than $\mathcal{CL}$, so more work is needed to prove the uniform double recurrence Wiener-Wintner theorem for the case either $f_1, f_2 \in \mathcal{Z}_2^\perp$ since $\mathcal{CL}^\perp \subset \mathcal{Z}_2^\perp$.

2.2 Main results

The following double recurrence Wiener-Wintner result extends Theorem 2.1.1.

Theorem 2.2.1. *Let $(X, \mathcal{F}, \mu, T)$ be a standard ergodic dynamical system, and $f_1, f_2 \in L^\infty(X)$. Let*

$$W_N(f_1, f_2, x, t) = \frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x)f_2(T^{bn}x)e(nt).$$

1. *(Double Uniform Wiener-Wintner Theorem) If either f_1 or f_2 belongs to $\mathcal{Z}_2^\perp$, then there exists a set of full measure X_{f_1,f_2} such that for all $x \in X_{f_1,f_2}$,*

$$\limsup_{N\to\infty}\sup_{t\in\mathbb{R}} |W_N(f_1, f_2, x, t)| = 0$$

2. *(General Convergence) If $f_1, f_2 \in \mathcal{Z}_2$, then for μ-a.e. $x \in X$, $W_N(f_1, f_2, x, t)$ converges for all $t \in \mathbb{R}$.*

The preceding theorem is the extension of Duncan's result in the following way:

- We assume that the transformation T is ergodic, rather than totally ergodic.
- We show that $\mathcal{Z}_2$ is a characteristic factor for this Wiener-Wintner average, i.e. we will prove the uniform double Wiener-Wintner result for the case either $f_1 \in \mathcal{Z}_2^{\perp}$ or $f_2 \in \mathcal{Z}_2^{\perp}$ rather than $\mathcal{CL}^{\perp}$.
- We show that the convergence holds in general for case $f_1, f_2 \in \mathcal{Z}_2$ (i.e. we drop the assumption regarding cocycles).

While we will show that $\mathcal{Z}_2$ is a pointwise characteristic factor of the double recurrence Wiener-Winter averages, we will also consider other factors. In 2012, the I. Assani and K. Presser published an update [22] of their earlier unpublished work [21] on characteristic factors and the multiterm return times theorem. The following seminorms were discussed in their work:

Definition 2.2.2. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic dynamical system on a probability measure space. We define factors $\mathcal{A}_k$ in the following inductive way.*

- *The factor $\mathcal{A}_0$ is the trivial σ-algebra $\{X, \emptyset\}$.*
- *The factor $\mathcal{A}_1$ is the Kronecker factor of T. We denote $N_1(f) = \|\mathbb{E}(f|\mathcal{A}_1)\|_2$.*
- *For $k \geq 1$, the factor $\mathcal{A}_{k+1}$ is characterized by the following: A function $f \in \mathcal{A}_{k+1}^{\perp}$ if and only if*

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

It was proven that the quantities $N_k(f)$ are well-defined in [4], and they characterize factors $\mathcal{A}_k$ of T which are successive maximal isometric extensions. These successive factors turned out to be the k-step distal factors introduced by H. Furstenberg in [50].

In [22], it was shown that given an ergodic system $(X, \mathcal{F}, \mu, T)$ and $f \in L^\infty(\mu)$, there exists a set of full measure X_f such that for any $x \in X_f$ and for any measure-preserving system $(Y, \mathcal{G}, \nu, S)$ and $g \in L^\infty(\nu)$ such that $\|g\|_{L^\infty(\nu)} \leq 1$, the average

$$\limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(S^n y) e^{2\pi int} \right| \leq C N_3(f)^2$$

converges for ν-a.e. $y \in Y$ for some absolute constant C independent of f_1, f_2, S, and y.

It is known that $\mathcal{Z}_k \subset \mathcal{A}_k$ for each k (in fact, $\mathcal{Z}_0$ equals $\mathcal{A}_0$ and $\mathcal{Z}_1$ equals $\mathcal{A}_1$, but $\mathcal{Z}_2 \subsetneq \mathcal{A}_2$), so $\mathcal{Z}_k^\perp \supset \mathcal{A}_k^\perp$. In [22], it was proven that $\mathcal{Z}_k$ and $\mathcal{A}_k$ are both pointwise characteristic for the k-term return times averages.

By Theorem 2.2.1, we now know that $\mathcal{A}_2$ is also a characteristic factor for the double recurrence Wiener-Wintner averages. However, the uniformity part of this result can be obtained independently of that particular theorem. In particular, we obtain the following pointwise estimate with the N_3-norm, which was not the case with the Gowers-Host-Kra seminorm (cf. [22, Remark (1), p. 359]).

Theorem 2.2.3. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system, and $f_1, f_2 \in L^\infty(\mu)$ for which $\|f_i\|_{L^\infty(\mu)} \leq 1$ for both $i = 1, 2$. Then we have*

$$\limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^n x) f_2(T^{2n} x) e(nt) \right| \leq \min_{i=1,2} [N_2(f_i)]^2$$

for μ-a.e. $x \in X$.

2.3 Proof of (1) Theorem 2.2.1

The proof of the Wiener-Wintner double recurrence theorem can be obtained in multiple steps. We first assume that either f_1 or f_2 belongs to the orthogonal complement of $\mathcal{Z}_2$, and show that the averages vanish uniformly on t. Secondly, if both f_1 and f_2 are measurable with respect to $\mathcal{Z}_2$, we will use the structures of nilsystems to show the convergence.

Our goal in this section is to find an upper estimate of the expression

$$\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an} x) f_2(T^{bn} x) e(nt) \right|^2 d\mu(x).$$

We will first check that the function $x \mapsto \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(nt) \right|$ is indeed measurable. In a view of extending this result to the polynomial case (cf. Chapter 3), we will prove something stronger, by showing that the function $x \mapsto \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)) \right|$ is measurable, where $\mathbb{R}_k[\xi]$ is the set of all the degree-k polynomials with real coefficients.

Lemma 2.3.1. *For each positive integers N and k, the map*

$$x \in X \mapsto F_{N,k}(x) = \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)) \right|$$

is measurable.

Proof. If we denote $p(n) = \sum_{j=0}^{k} c_j n^j$, then

$$\begin{aligned}
&\sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)) \right| \\
&= \sup_{(c_0,c_1,c_2,\ldots,c_k)\in\mathbb{R}^{k+1}} \left| \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) e\left(\sum_{j=0}^{k} c_j n^j\right) \right| \\
&= \sup_{(c_0,c_1,c_2,\ldots,c_k)\in\mathbb{Q}^{k+1}} \left| \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) e\left(\sum_{j=0}^{k} c_j n^j\right) \right|,
\end{aligned}$$

where the last equality follows from the fact that $\mathbb{Q}^{k+1}$ is dense in $\mathbb{R}^{k+1}$, and the map

$$(c_0, c_1, c_2, \ldots, c_k) \mapsto e\left(\sum_{j=0}^{k} c_j n^j\right)$$

is a continuous one from $\mathbb{R}^{k+1}$ to $\mathbb{T}$ for each $n \in \mathbb{Z}$. Since $\mathbb{Q}^{k+1}$ is countable, it follows that the map $x \mapsto F_{N,k}(x)$ is measurable for each k and N. □

One of the difficulties of the proof, which stems from the fact that we are no longer assuming that T is totally ergodic, is that the transformations T^a, T^b, and T^{b-a} are not ergodic (see [16, §4] for the case when $b - a = 1$, which simplifies the matter significantly). This means that the limit of the ergodic averages along these transformations will not be a constant, but rather a conditional expectation onto the invariant set, i.e. for any $f \in L^\infty(\mu)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^{(b-a)n}x) = \mathbb{E}(f|\mathcal{I}(T^{b-a}))(x)$$

for μ-a.e. $x \in X$. To overcome this difficulty, the following integral kernel will be useful. This kernel appeared in the proof work of Furstenberg and Weiss [53, Proof of Theorem 2.1].

Lemma 2.3.2. *Let T be an ergodic map, and s be a positive integer. Then there exist a positive integer l, a disjoint partition of T^s-invariant sets $A_1, \ldots, A_l$ such that every T^s-invariant function f*

can be expressed as an integral with respect to the kernel

$$K(x, y) = l \sum_{k=1}^{l} \mathbb{1}_{A_i}(x)\mathbb{1}_{A_i}(y). \tag{2.1}$$

Proof. If T^s is ergodic, we are done, since f is a constant. If not, suppose A is a T^s-invariant subset of X such that $0 < \mu(A) < 1$. Define a function

$$f_A := \mathbb{1}_A + \mathbb{1}_{T^{-1}A} + \mathbb{1}_{T^{-2}A} + \cdots + \mathbb{1}_{T^{-(s-1)}A}.$$

Observe that f_A is T-invariant, and since T is ergodic, f_A must be a constant. Therefore,

$$\mathbb{1}_A + \mathbb{1}_{T^{-1}A} + \mathbb{1}_{T^{-2}A} + \cdots + \mathbb{1}_{T^{-(s-1)}A} = \int f_A d\mu = s\mu(A).$$

Note that $f_A \neq 0$, since $\mu(A) \neq 0$. Similarly, $f_A \neq s$, since $\mu(A) \neq 1$. If $f_A = 1$, then for μ-a.e. $x \in X$, $\mathbb{1}_A + \mathbb{1}_{T^{-1}A} + \mathbb{1}_{T^{-2}A} + \cdots + \mathbb{1}_{T^{-(s-1)}A} = 1$, which implies that $\mu(T^{-i}A \cap T^{-j}A) = 0$ for any $0 \leq i < j \leq s-1$. Hence, $A, T^{-1}A, \ldots, T^{-(s-1)}A$ are disjoint, and furthermore, $\mu(X) = \sum_{k=0}^{s-1} \mu(T^{-k}A) = 1$, so $A, T^{-1}A, \ldots, T^{-(s-1)}A$ is a partition of X.

Now we show that A (and similarly, $T^{-1}A, \ldots, T^{-(s-1)}A$) is an atom (of a collection of T^s-invariant sets). If $B \subset A$ and B is T^s-invariant, then

$$f_B = \mathbb{1}_B + \mathbb{1}_{T^{-1}B} + \mathbb{1}_{T^{-2}B} + \cdots + \mathbb{1}_{T^{-(s-1)}B} = s\mu(B) \leq s\mu(A) = 1.$$

The above holds only when $\mu(B) = 0$ or $\mu(B) = 1/s = \mu(A)$, which implies that $B = A$ μ-a.e. For $k > 0$, we note that if $B \subset T^{-k}A$ is T^s-invariant, then $T^kB \subset A$ is also T^s-invariant, so if $\mu(B) \neq 0$, then $\mu(B) = \mu(T^kB) = \mu(A) = \mu(T^{-k}A)$, which proves that $T^{-k}A$ is also an atom for $k > 0$.

If f is a T^s-invariant function, then we claim that

$$f = \sum_{k=0}^{s-1} \left(\frac{\int_{T^{-k}A} f d\mu}{\mu(T^{-k}A)} \right) \mathbb{1}_{T^{-k}A} = s \sum_{k=0}^{s-1} \left(\int_{T^{-k}A} f d\mu \right) \mathbb{1}_{T^{-k}A}. \tag{2.2}$$

First, we note that $\mathcal{S}$, the σ-algebra generated by the sets $A, T^{-1}A, \ldots, T^{-(s-1)}A$, is a collection of finite union of sets $A, T^{-1}A, \ldots, T^{-(s-1)}A$. We also know that f is $\mathcal{S}$-measurable, since for any

$\lambda \in \mathbb{R}$,

$$\{f > \lambda\} = \bigcup_{k=0}^{s-1} \left(\{f > \lambda\} \cap T^{-k}A \right),$$

and we note that $\{f > \lambda\} \cap T^{-k}A$ is T^s-invariant. Since $T^{-k}A$ is an atom for each k, we know $\{f > \lambda\} \cap T^{-k}A$ equals either $T^{-k}A$ or the empty set. This implies that $\{f > \lambda\} \in \mathcal{S}$.

Since we know that f is $\mathcal{S}$-measurable, we note that f can be expressed as the expression above (a fact regarding conditional expectation). This proves (2.2), and if we denote $T^{-k}A = A_k$, then we have

$$f \circ T^s(x) = f(x) = \int s \sum_{k=0}^{s-1} \mathbb{1}_{A_k}(y) \mathbb{1}_{A_k}(x) f(y) d\mu(y) = \int f(y) K(x,y) d\mu(y),$$

which proves the lemma for the case $f_A = 1$.

Now, suppose $f_A = k$ for $2 \leq k \leq s-1$. Let $B = T^{-l_1}A \cap T^{-l_2}A \cap \cdots T^{-l_k}A$, where $0 \leq l_1 < l_2 < \cdots < l_k \leq s-1$, and $\mu(B) > 0$ (we know such B exists since $f_A = k$). Define

$$f_B = \mathbb{1}_B + \mathbb{1}_{T^{-1}B} + \cdots + \mathbb{1}_{T^{-(s-1)}B}$$

Note that f_B is T-invariant, so it must be a constant function that equals to $s\mu(B)$. Since $\mu(B) > 0$, we know that $f_B > 0$.

Also, note that each $T^{-j}B$ is disjoint for $0 \leq j \leq s-1$. Assume it is not. Then for some $0 \leq i < j \leq s-1$, there exists $x \in T^{-i}B \cap T^{-j}B$ such that $f_A(x) > k$, which is a contradiction. Therefore, we must have $f_B \leq 1$, and we can conclude that $f_B = 1$. By letting $A_i = T^{-i}B$, we have proved the lemma. □

We will use the preceding lemma to extend the uniform Wiener-Wintner theorem (Theorem 1.2.2) for the case where we have a power of an ergodic transformation.

Proposition 2.3.3 (Uniform Wiener-Wintner Theorem for a power of ergodic transformation)**.** *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system. Suppose $f \in \mathcal{Z}_1^{\perp}$. Then there exists a set of full measure X_f such that for any $x \in X_f$ and for any integer a, we have*

$$\limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f(T^{an}x) e^{2\pi int} \right| = 0$$

Proof. To show that the uniform convergence holds, we apply the inequality (1.6) for $a_n = f(T^{an}x)$ pointwise, and use the pointwise ergodic theorem and the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}\limsup_{N\to\infty}\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=1}^{N} f(T^{an}x)e^{2\pi int}\right|^2 &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\left|\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}(f\cdot f\circ T^{ah})(T^n x)\right| \\ &= \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\left|\mathbb{E}(f\cdot f\circ T^{ah}|\mathcal{I}(T^a))(x)\right| \\ &\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}\left|\mathbb{E}(f\cdot f\circ T^{ah}|\mathcal{I}(T^a))(x)\right|^2\right)^{1/2}. \qquad (2.3)\end{aligned}$$

The proof is complete if we can show that the last line of (2.3) converges to 0 for μ-a.e. $x \in X$. To show this, we use the integral kernel from Lemma 2.3.2 so that for some positive integer l, we have

$$\mathbb{E}(f\cdot f\circ T^{ah}|\mathcal{I}(T^a))(x) = l\int \mathbb{1}_A(y)f(y)f(T^{ah}y)d\mu(y)$$

where A is one of the sets of the partition of X given in Lemma 2.3.2 such that $x \in A$. Set $g(y) = \mathbb{1}_A(y)f(y)$. Then we notice that

$$l^{-1}\mathbb{E}(f\cdot f\circ T^{ah}|\mathcal{I}(T^a))(x) = \int g(y)f(T^{ah}y)d\mu(y) = \hat{\sigma}_{f,g,T^a}(h),$$

where σ_{f,g,T^a} is the complex Borel measure on $\mathbb{T}$ defined in (1.1) for the functions f and g with respect to the transformation T^a. By Proposition 1.1.8, σ_{f,g,T^a} is absolutely continuous with respect to σ_{f,T^a}. We claim that σ_{f,T^a} is a continuous measure. Since $f \in \mathcal{Z}_1^{\perp}$ and $\sigma_{f,T}$ is a continuous measure, Wiener's lemma (Proposition 1.1.7) tells us that $\lim_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}|\hat{\sigma}_{f,T}(h)|^2 = 0$. Since $\left(\frac{1}{|a|H}\sum_{h=1}^{|a|H}|\hat{\sigma}_{f,T}(h)|^2\right)_H$ is a subsequence of $\left(\frac{1}{H}\sum_{h=1}^{H}|\hat{\sigma}_{f,T}(h)|^2\right)_H$, we have

$$\lim_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}|\hat{\sigma}_{f,T^a}(h)|^2 = \lim_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}|\hat{\sigma}_{f,T}(ah)|^2 \leq \lim_{H\to\infty}|a|\left(\frac{1}{|a|H}\sum_{h=1}^{|a|H}|\hat{\sigma}_{f,T}(h)|^2\right) = 0,$$

and again, by Wiener's lemma, σ_{f,T^a} is a continuous measure. Hence, σ_{f,g,T^a} is continuous, so we have

$$0 = \lim_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} |\hat{\sigma}_{f,g,T^a}(h)|^2 = l^{-2} \cdot \lim_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} |\mathbb{E}(f \cdot f \circ T^{ah} | \mathcal{I}(T^a))(x)|^2.$$

□

We will use the Gowers-Host-Kra seminorms in our arguments. While these seminorms (and the associated characteristic factors)are defined for ergodic transformations, they are still useful tools when we have a power of ergodic transformation, as one can see in the following lemma.

Lemma 2.3.4. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic dynamical system, and a a nonzero integer. Then for any positive integer k, we have*

$$\limsup_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \left|\left|\left| f \cdot f \circ T^{ah} \right|\right|\right|_k^{2^k} \leq |a| |||f|||_{k+1}^{2^{k+1}}.$$

Proof. We note that

$$\frac{1}{H} \sum_{h=1}^{H} \left|\left|\left| f \cdot f \circ T^{ah} \right|\right|\right|_k^{2^k} \leq \frac{1}{H} \sum_{h=1}^{|a|H} \left|\left|\left| f \cdot f \circ T^{h} \right|\right|\right|_k^{2^k} = |a| \left(\frac{1}{|a|H} \sum_{h=1}^{|a|H} \left|\left|\left| f \cdot f \circ T^{h} \right|\right|\right|_k^{2^k} \right).$$

The sequence $\left(\frac{1}{|a|H} \sum_{h=1}^{|a|H} \left|\left|\left| f \cdot f \circ T^{h} \right|\right|\right|_k^{2^k} \right)_H$ is a subsequence of $\left(\frac{1}{H} \sum_{h=1}^{H} \left|\left|\left| f \cdot f \circ T^{h} \right|\right|\right|_k^{2^k} \right)_H$, which converges to $|||f|||_{k+1}^{2^{k+1}}$. By taking the limit supremum on both sides of the inequality above, we get

$$\limsup_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \left|\left|\left| f \cdot f \circ T^{ah} \right|\right|\right|_k^{2^k} \leq |a| \left(\lim_{H\to\infty} \frac{1}{|a|H} \sum_{h=1}^{|a|H} \left|\left|\left| f \cdot f \circ T^{h} \right|\right|\right|_k^{2^k} \right) = |a| |||f|||_{k+1}^{2^{k+1}}.$$

□

By utilizing Lemma 2.3.4, we will make a series of statements that will allow us to overcome the difficulty of the omission of the total ergodicity assumption. First, the following inequality applies for any measure-preserving system.

Lemma 2.3.5. *Suppose $(Y, \mathcal{Y}, \nu, U)$ is a measure preserving system, and $f \in L^\infty(\nu)$. Then there* *exists a universal constant C such that for any positive integer H, one has*

$$\int \limsup_{N\to\infty} \left| \frac{1}{N}\sum_{n=1}^{N} f(U^n y)\right|^2 d\nu \le C\left(\frac{1}{H} + \left(\frac{1}{H}\sum_{h=1}^{H}\left|\int f\cdot f\circ U^h d\nu\right|^2\right)^{1/2}\right). \tag{2.4}$$

Proof. We denote $F_h(x) = f(x) f\circ U^h(x)$. We apply the inequality (1.4) by setting $a_n = f(U^n y)$ pointwise and the pointwise ergodic theorem to obtain

$$\begin{aligned}\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N} f(U^n y)\right|^2 &\le \frac{C}{H} + \frac{C}{(H+1)^2}\sum_{h=1}^{H}(H+1-h)\left(\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N}F_h(U^n y)\right) \quad (2.5)\\ &= \frac{C}{H} + \frac{C}{(H+1)^2}\sum_{h=1}^{H}(H+1-h)\mathbb{E}(f\cdot f\circ U^h|\mathcal{I}(U))(y).\end{aligned}$$

Note that $\int \mathbb{E}(f\cdot f\circ U^h|\mathcal{I}(U))\, d\nu = \int f\cdot f\circ U^h\, d\nu$. So if we take the integral on both sides of the inequality (2.5), we would obtain the desired estimate (2.4) after applying the Cauchy-Schwarz inequality. □

The previous lemma allows one to prove the following inequality, which is an integral of the limit superior of the square of the magnitude of the averages $\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}y)$ over the product measure-preserving system $(X^2, \mathcal{F}^2, \mu\otimes\mu, T^a\times T^b)$. While it may seem slightly peculiar to focus on the product space and these averages considering that we are working with the double recurrence averages (i.e. $\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x)$), we remark here that this is where the strength of the integral kernel seen in Lemma 2.3.2 can be observed: The double recurrence theorem (Theorem 1.3.1) and the dominated convergence theorem allows one to switch the limit and the integral, and then one applies the pointwise ergodic theorem to obtain the conditional expectation, which can be turned into a function that depends on both the first variable x and the second variable y.

Lemma 2.3.6. *Suppose $(X, \mathcal{F}, \mu, T)$ is an ergodic dynamical system, and f_1 and f_2 be bounded and measurable functions for which $\|f_i\| \le 1$ for both $i = 1, 2$. Then for any nonzero distinct integers a_1 and a_2, there exists a universal constant C such that*

$$\int \limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N} f_1(T^{a_1 n}x) f_2(T^{a_2 n}y)\right|^2 d\mu\otimes\mu(x,y) \le C\min_{i=1,2}|a_i|\||f_i\||_2^2.$$

Proof. We denote $F_{1,h}(x) = f_1(x) f_1 \circ T^{a_1 h}(x)$, and $F_{2,h}(x) = f_2(x) f_2 \circ T^{a_2 h}(x)$. If $U = T^{a_1} \times T^{a_2}$, then $(X^2, \mathcal{F} \otimes \mathcal{F}, \mu \otimes \mu, U)$ is a measure preserving system. Hence, we can apply (2.4) in Lemma 2.3.5 to obtain, for any $H \in \mathbb{N}$,

$$\begin{aligned}
&\int \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} f_1 \otimes f_2(U^n(x,y)) \right|^2 d\mu \otimes \mu(x,y) \\
&\leq \frac{C}{H} + C \left(\frac{1}{H} \sum_{h=1}^{H} \left| \int F_{1,h} \otimes F_{2,h}(x,y)\, d\mu \otimes \mu(x,y) \right|^2 \right)^{1/2} \\
&\leq \frac{C}{H} + C \left(\frac{1}{H} \min_{i=1,2} \sum_{h=1}^{H} \left| \int f_i \cdot f_i \circ T^{a_i h} d\mu \right|^2 \right)^{1/2}.
\end{aligned}$$

As we let $H \to \infty$, we obtain the desired inequality by Lemma 2.3.4. □

Next, we will extend the previous lemma to Wiener-Wintner type averages.

Lemma 2.3.7. *Suppose $(X, \mathcal{F}, \mu, T)$ is an ergodic system, and $f_1, f_2 \in L^\infty(\mu)$ for which $\|f_i\|_\infty \leq 1$ for both $i = 1, 2$. Then for any distinct nonzero integers a_1 and a_2, there exists a universal constant C for which*

$$\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{a_1 n}x) f_2(T^{a_2 n}y) e(nt) \right|^2 d\mu \otimes \mu(x,y) \leq C \min_{i=1,2} |a_i|^{1/2} |||f_i|||_3^2.$$

Proof. We denote $F_{1,h}(x) = f_1(x) f_1 \circ T^{a_1 h}(x)$, and $F_{2,h}(x) = f_2(x) f_2 \circ T^{a_2 h}(x)$. By applying the inequality (1.6) for $a_n = f_1(T^{a_1 n}x) f_2(T^{a_2 n}y)$ pointwise, we obtain

$$\begin{aligned}
&\limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{a_1 n}x) f_2(T^{a_2 n}y) e(nt) \right|^2 \\
&\leq \frac{C}{H} + \frac{C}{H} \sum_{h=1}^{H} \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}y) \right|.
\end{aligned}$$

Again, if we set $U = T^{a_1} \times T^{a_2}$, then $(X^2, \mathcal{F} \otimes \mathcal{F}, \mu \otimes \mu, U)$ is a measure preserving system. Hence, Birkhoff's pointwise ergodic theorem asserts that the average $\frac{1}{N} \sum_{n=1}^{N} F_{1,h} \otimes F_{2,h}(U^n(x,y))$ converges

$\mu \otimes \mu$-a.e. By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{a_1 n}x) f_2(T^{a_2 n}y) e(nt) \right|^2 d\mu\otimes\mu(x,y) \\
&\le \frac{C}{H} + \frac{C}{H} \sum_{h=1}^{H} \int \lim_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}y) \right| d\mu\otimes\mu(x,y) \\
&\le \frac{C}{H} + C \left(\frac{1}{H} \sum_{h=1}^{H} \int \lim_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}y) \right|^2 d\mu\otimes\mu(x,y) \right)^{1/2}.
\end{aligned}$$

By Lemma 2.3.6, we know that

$$\int \lim_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,h}(T^{an}x) F_{2,h}(T^{bn}y) \right|^2 d\mu\otimes\mu(x,y) \le C \min_{i=1,2} |a_i| \vert\vert\vert F_{i,h} \vert\vert\vert_2^2.$$

Hence,

$$\begin{aligned}
&\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}y) e(nt) \right|^2 d\mu \\
&\le \frac{C}{H} + C \min_{i=1,2} \left(\frac{|a_i|}{H} \sum_{h=1}^{H} \left\vert\left\vert\left\vert f_i \cdot f_i \circ T^{a_i h} \right\vert\right\vert\right\vert_2^2 \right)^{1/2} \le \frac{C}{H} + C \min_{i=1,2} \left(\frac{|a_i|^2}{H} \sum_{h=1}^{H} \left\vert\left\vert\left\vert f_i \cdot f_i \circ T^{a_i h} \right\vert\right\vert\right\vert_2^4 \right)^{1/4}.
\end{aligned}$$

Let $H \to \infty$, and apply Lemma 2.3.4 to obtain the desired result. □

We present one more inequality before we present the proof of (1) of Theorem 2.2.1. This inequality, which is similar to the ones appeared in the work of Assani [10, Lemma 5]. In our proof of the main result, we apply (the variants of) van der Corput's lemma twice, which lead us to have averages involving multiple integer parameters. The following estimate will be useful when controlling those averages.

Lemma 2.3.8. *Let a_n, b_n, and c_n, $n \in \mathbb{N}$ be three complex-valued sequences, norm of each bounded above by 1. Then for each positive integer N,*

$$\left| \frac{1}{H(H+1)^2} \sum_{h,k=0}^{H-1} (H+1-k) a_h \cdot b_k \cdot c_{h+k} \right|^2 \le \sup_t \left| \frac{1}{H} \sum_{k'=1}^{2(H-1)} c_{k'} e^{2\pi i k' t} \right|^2 \tag{2.6}$$

Proof. This proof is a small modification of the proof provided in [10, Lemma 5]. By the Cauchy-

Schwarz inequality, we have

$$\left|\frac{1}{H(H+1)^2}\sum_{h,k=0}^{H-1}(H+1-k)a_h b_k c_{h+k}\right|^2 \leq \|a\|_\infty^2\left(\frac{1}{H}\sum_{h=0}^{H-1}\left|\frac{1}{(H+1)^2}\sum_{k=0}^{H-1}(H+1-k)b_k c_{h+k}\right|^2\right).$$

Set $B_k = b_k\dfrac{(H+1-k)}{H+1}$, and the inequality above becomes

$$\begin{aligned}&\left|\frac{1}{H(H+1)^2}\sum_{h,k=0}^{H-1}(H+1-k)a_h b_k c_{h+k}\right|^2\\ &\leq \|a\|_\infty^2\left(\frac{1}{H}\sum_{h=0}^{H-1}\left|\frac{1}{H+1}\sum_{k=0}^{H-1}B_k c_{h+k}\right|^2\right)\\ &\leq \|a\|_\infty^2\frac{1}{H}\sum_{h=0}^{H-1}\left|\int\left(\sum_{k=0}^{H-1}B_k e^{-2\pi ikt}\right)\left(\frac{1}{H+1}\sum_{k=0}^{2(H-1)}c_k e^{2\pi ikt}\right)e^{-2\pi iht}dt\right|^2,\end{aligned}$$

where $\|a_n\|_\infty = \|a_n\|_{\ell^\infty}$. We apply Parseval's inequality to the integral above to obtain

$$\begin{aligned}&\left|\frac{1}{H(H+1)^2}\sum_{h,k=0}^{H-1}(H+1-k)a_h b_k c_{h+k}\right|^2\\ &\leq \|a\|_\infty^2\frac{1}{H}\sum_{h=0}^{H-1}\int\left|\sum_{k=0}^{H-1}B_k e^{-2\pi ikt}\right|^2\left|\frac{1}{H+1}\sum_{k'=0}^{2(H-1)}c_{k'}e^{2\pi ik't}\right|^2 dt\\ &\leq \|a\|_\infty^2\sup_{t\in\mathbb{R}}\left|\frac{1}{H+1}\sum_{k=0}^{2(H-1)}c_k e^{2\pi ikt}\right|^2\frac{1}{H}\sum_{h=0}^{H-1}\int\left|\sum_{k=0}^{H-1}B_k e^{-2\pi int}\right|^2 dt\\ &\leq \|a\|_\infty^2\sup_{t\in\mathbb{R}}\left|\frac{1}{H+1}\sum_{k=0}^{2(H-1)}c_k e^{2\pi ikt}\right|^2\frac{1}{H+1}\sum_{h=0}^{H-1}|B_k|^2.\end{aligned}$$

Since $|B_k| < 1$, we know that $\dfrac{1}{H+1}\displaystyle\sum_{h=0}^{H-1}|B_k|^2 \leq 1$. Thus, (2.6) holds. □

Proof of (1) of Theorem 2.2.1. We denote $F_{1,h}(x) = f_1(x)f_1 \circ T^{a_1h}(x)$, and $F_{2,h}(x) = f_2(x)f_2 \circ T^{a_2h}(x)$, where $a_1 = a$ and $a_2 = b$. We apply the inequality (1.6) by setting $a_n = f_1(T^{a_1n}x)f_2(T^{a_2n}x)$, we obtain the following for all $H \in \mathbb{N}$:

$$\begin{aligned}
&\limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} f_1(T^{a_1 n}x) f_2(T^{a_2 n}x) e(nt) \right|^2 \\
&\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H} \limsup_{N\to\infty} \left| \frac{1}{N}\sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}x) \right| \\
&\leq \frac{C}{H} + C\left(\frac{1}{H}\sum_{h=1}^{H} \limsup_{N\to\infty} \left| \frac{1}{N}\sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}x) \right|^2 \right)^{1/2},
\end{aligned}$$

where the second inequality is the consequence of the Cauchy-Schwarz inequality. Note that we can apply the inequality (1.4) on the average $\left| \frac{1}{N}\sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}x) \right|^2$ to obtain the following bound for $0 < K < N$:

$$\begin{aligned}
&\limsup_{N\to\infty} \left| \frac{1}{N}\sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}x) \right|^2 \\
&\leq \frac{C}{K} + \frac{C}{(K+1)^2}\sum_{k=1}^{K} (K+1-k) \\
&\times \left(\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \left(F_{1,h} \cdot F_{1,h} \circ T^{a_1 k} \right)(T^{a_1 n}x) \left(F_{2,h} \cdot F_{2,h} \circ T^{a_2 k} \right)(T^{a_2 n}x) \right).
\end{aligned}$$

Note that the averages

$$\frac{1}{N}\sum_{n=1}^{N} \left(F_{1,h} \cdot F_{1,h} \circ T^{a_1 k} \right)(T^{a_1 n}x) \cdot \left(F_{2,h} \cdot F_{2,h} \circ T^{a_2 k} \right)(T^{a_2 n}x)$$

converge for μ-a.e. $x \in X$ as $N \to \infty$ by the double recurrence theorem (Theorem 1.3.1). Therefore,

$$\begin{aligned}
&\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} f_1(T^{a_1 n}x) f_2(T^{a_2 n}x) e(nt) \right|^2 d\mu(x) \\
&\leq \frac{C}{H} + C\int \left(\frac{1}{H}\sum_{h=1}^{H} \limsup_{N\to\infty} \left| \frac{1}{N}\sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}x) \right|^2 \right)^{1/2} d\mu(x) \\
&\leq \frac{C}{H} + C\left(\frac{1}{H}\sum_{h=1}^{H} \int \limsup_{N\to\infty} \left| \frac{1}{N}\sum_{n=1}^{N} F_{1,h}(T^{a_1 n}x) F_{2,h}(T^{a_2 n}x) \right|^2 d\mu(x) \right)^{1/2} \quad \text{(by Hölder's inequality)}
\end{aligned}$$

$$\leq \frac{C}{H} + C\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{C}{K} + \frac{C}{(K+1)^2}\sum_{k=1}^{K}(K+1-k)\right.\right.$$
$$\left.\left.\left(\lim_{N\to\infty}\int \frac{1}{N}\sum_{n=1}^{N}\left(F_{1,h}\cdot F_{1,h}\circ T^{a_1k}\right)(T^{a_1n}x)\left(F_{2,h}\cdot F_{2,h}\circ T^{a_2k}\right)(T^{a_2n}x)\,d\mu(x)\right)\right)\right)^{1/2}.$$

Since $T^{a_2-a_1}$ is a measure preserving transformation, we can apply the mean ergodic theorem to obtain

$$\lim_{N\to\infty}\int \frac{1}{N}\sum_{n=1}^{N}\left(F_{1,h}\cdot F_{1,h}\circ T^{a_1k}\right)(T^{a_1n}x)\left(F_{2,h}\cdot F_{2,h}\circ T^{a_2k}\right)(T^{a_2n}x)\,d\mu(x)$$
$$= \lim_{N\to\infty}\int \left(F_{1,h}\cdot F_{1,h}\circ T^{a_1k}\right)(x)\left(\frac{1}{N}\sum_{n=1}^{N}\left(F_{2,h}\cdot F_{2,h}\circ T^{a_2k}\right)(T^{(a_2-a_1)n}x)\right)d\mu(x)$$
$$= \int \left(F_{1,h}\cdot F_{1,h}\circ T^{a_1k}\right)(x)\mathbb{E}(F_{2,h}\cdot F_{2,h}\circ T^{a_2k}|\mathcal{I}(T^{a_2-a_1}))(x)d\mu(x).$$

By Lemma 2.3.2, there exists a positive integer l_{b-a} and partition $A_1, \ldots, A_{l_{b-a}}$ of X such that

$$\mathbb{E}(F_{2,h}\cdot F_{2,h}\circ T^{a_2k}|\mathcal{I}(T^{a_2-a_1}))(x) = \int \left(F_{2,h}\cdot F_{2,h}\circ T^{a_2k}\right)(y)K_{a_2-a_1}(x,y)d\mu(y),$$

where $K_{a_2-a_1}(x,y) = l_{a_2-a_1}\sum_{i=1}^{l_{a_2-a_1}} \mathbb{1}_{A_i}(x)\mathbb{1}_{A_i}(y)$. Using this kernel expression and writing $U = T^{a_1}\times T^{a_2}$ as a measure-preserving transformation on the product space $(X^2, \mathcal{F}^2, \mu\otimes\mu)$, we obtain

$$\int \left(F_{1,h}(x)\cdot F_{1,h}\circ T^{a_1k}\right)(x)\left(F_{2,h}\cdot F_{2,h}\circ T^{a_2k}\right)(y)K_{b-a}(x,y)\,d\mu\otimes\mu(x,y)$$
$$= \int f_1\otimes f_2(x,y)K_{b-a}(x,y)$$
$$\times\left[f_1\otimes f_2(U^h(x,y))\,f_1\otimes f_2(U^k(x,y))\,f_1\otimes f_2(U^{h+k}(x,y))\right]d\mu\otimes\mu(x,y),$$

Let $H = K$. Note that, on the system $(X^2, \mu\otimes\mu, U)$, the inequality (2.6) tells us that we have

$$\left|\frac{1}{H(H+1)^2}\sum_{h,k=0}^{H-1}(H+1-k)f_1\otimes f_2(U^h(x,y))\,f_1\otimes f_2(U^k(x,y))\,f_1\otimes f_2(U^{h+k}(x,y))\right|^2$$
$$\leq \sup_{t\in\mathbb{R}}\left|\frac{1}{H}\sum_{h=1}^{2(H-1)} f_1\otimes f_2(U^h(x,y))e(ht)\right|^2. \tag{2.7}$$

By Lemma 2.3.7, we know that

$$\begin{aligned}
&\int \limsup_{H\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{H} \sum_{h=1}^{H} f_1 \otimes f_2(U^h(x,y)) e(ht) \right|^2 \, d\mu\otimes\mu(x,y) \\
&= \int \limsup_{H\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{H} \sum_{h=1}^{H} f_1(T^{a_1 h}x) f_2(T^{a_2 h}y) e(ht) \right|^2 \, d\mu\otimes\mu(x,y) \\
&\le C \min_{i=1,2} |a_i|^{1/2} |||f_i|||_3^2.
\end{aligned}$$

Hence, by letting $H \to \infty$ and applying the Hölder's inequality, we obtain

$$\begin{aligned}
&\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e^{2\pi i n t} \right|^2 \, d\mu(x) \\
&\le \Bigg(\int f_1 \otimes f_2(x,y) K(x,y) \\
&\times \limsup_{H\to\infty} \frac{1}{H(H+1)^2} \sum_{h,k=0}^{H-1} (H+1-k) \\
&\times \Big(f_1\otimes f_2(U^h(x,y))\, f_1\otimes f_2(U^k(x,y))\, f_1\otimes f_2(U^{h+k}(x,y)) \Big) \, d\mu\otimes\mu(x,y) \Bigg)^{1/2} \\
&\le C \Bigg(\int \limsup_{H\to\infty} \Bigg| \frac{1}{H(H+1)^2} \sum_{h,k=0}^{H-1} (H+1-k) \\
&\times \Big(f_1\otimes f_2(U^h(x,y))\, f_1\otimes f_2(U^k(x,y))\, f_1\otimes f_2(U^{h+k}(x,y)) \Big|^2 \, d\mu\otimes\mu(x,y) \Bigg)^{1/4} \\
&\le C \left(\int \limsup_{H\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{H} \sum_{h=1}^{H} f_1 \otimes f_2(U^h(x,y)) e(ht) \right|^2 \, d\mu\otimes\mu(x,y) \right)^{1/4} \\
&\le C \min_{i=1} |a_i|^{1/8} |||f_i|||_3^{1/2}.
\end{aligned}$$

Since either f_1 or f_2 belongs to $\mathcal{Z}_2^{\perp}$, we know that either $|||f_i|||_3 = 0$. This completes the proof. □

As a corollary of this proof, we have the following estimate, which will be useful later:

$$\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) e(nt) \right|^2 \, d\mu(x) \lesssim_{a,b} \min_{i=1,2} |||f_i|||_3^{1/2}. \tag{2.8}$$

2.4 Proof of (2) of Theorem 2.2.1

Here we prove the convergence of double recurrence Wiener-Wintner averages for the case where $f_1, f_2 \in \mathcal{Z}_2$. There are multiple ways of proving this part of the theorem. One way is to notice that the sequence $(f_1(T^{an}x)f_2(T^{bn}x)e(nt))$ is a 2-step nilsequence when f_1 and f_2 are both measurable with respect to $\mathcal{Z}_2$. Alternatively, one can view $(\mathbb{T}, m, R_t)$ is a one-step nilsystem, where m is a normed Lebesgue probability measure, and R_t is a rotation by t ($R_t(\alpha) = t + \alpha$), and use Leibman's convergence result (cf. Theorem 1.3.3); this method will be used later when we extend Theorem 2.2.1 to polynomials (cf. Lemma 3.2.6). In this section, we will present an elementary proof that relies on the structure of the second Host-Kra-Ziegler factor $\mathcal{Z}_2$ in terms of the extension of $\mathcal{Z}_1$, which is the Kronecker factor of T. Having done this, we apply Leibman's convergence result.

Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system. Recall that X is called a k-step nilsystem if X is a homogeneous space of a k-step nilpotent Lie group G (such a manifold is called a nilmanifold). Let Λ be a discrete cocompact subgroup of G such that $X = G/\Lambda$. The outline of the proof of the following theorem, which is the special case of the structure theorem (Theorem 1.3.2), when $\mathcal{Z}_2$ was referred to as the Conze-Lesigne system, is given in [61].

Theorem 2.4.1 ([61, Theorem 18])**.** *If X is a Conze-Lesigne system, then it is the inverse limit of a sequence of 2-step nilsystems.*

In the outline of the proof, X is reduced to the case where X is a group extension of the Kronecker factor Z_1 and torus U, with cocycle $\rho : Z_1 \to U$. A group G is defined to be a family of transformations of $X = Z_1 \times U$, where U is a finite dimensional torus and Z_1 is the Kronecker factor of X that has the structure of compact abelian Lie group. If $g \in G$, $(z, u) \in X$, then $g \cdot (z, u) = (sz, uf(z))$ where $s \in Z_1$ and $f : Z_1 \to U$ satisfy the Conze-Lesigne equation $\rho(sz)\rho(z)^{-1} = f(Rz)f(z)^{-1}c$ for some constant $c \in U$. It can be easily verified that G is a 2-step nilpotent group, and T corresponds to $(\beta, \rho) \in G$, where $\beta \in Z_1$ such that if $\pi_1 : Z_2 \to Z_1$ is a factor map, then $\pi_1(Tx) = \beta\pi_1(x)$. Furthermore, if G is given a topology of convergence in probability, then we know that G is a Lie group.

The outline of the proof given in [61] concludes by stating that G acts on X transitively, and X can be identified with the nilmanifold G/Λ, where Λ is a stabilizer group of a point $x_0 \in X$ (hence it is a discrete cocompact subgroup of G). Furthermore, μ is a Haar measure on X, and T

is a translation by the element $(\beta, \rho) \in G$. Hence, T acts on X by translation. We will use this fact to prove the convergence of the double recurrence Wiener-Wintner average for the case when $f_1, f_2 \in \mathcal{Z}_2$.

Proof of (2) of Theorem 2.2.1. In this proof, we will consider two cases: The case when t is rational, and the case when t is irrational.

Case I: When t is rational. Fix $t \in \mathbb{Q}$. Let S_t be a rotation on $\mathbb{T}$ by $e^{2\pi i t}$. Let $(X \times \mathbb{T}, \mu \otimes m, U)$ be a measure preserving system, where m is the Lebesgue measure on $\mathbb{T}$, and $U = T \otimes S_t$. Define $F_1(x, y) = f_1(x)e^{2\pi i \alpha_1 y}$, and $F_2(x, y) = f_2(x)e^{2\pi i \alpha_2 y}$, where $\alpha_1, \alpha_2 \in \mathbb{R}$ such that $\alpha_1 a + \alpha_2 b = 1$. Then

$$\frac{1}{N}\sum_{n=1}^{N} F_1(U^{an}(x, y))F_2(U^{bn}(x, y)) = \frac{e^{2\pi i y}}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e^{2\pi i n t} \tag{2.9}$$

Note that the averages on the left hand side of (2.9) converges $\mu \otimes m$-a.e. as $N \to \infty$ by the double recurrence theorem (Theorem 1.3.1). So there exists a set of full measure $V_t \subset X \times \mathbb{T}$ such that the average in (2.9) converges for all $(x, y) \in V_t$. If $V = \bigcap_{t\in\mathbb{Q}} V_t$, then V is a set of full measure such that the average on (2.9) converges for all $(x, y) \in V$ for all $t \in \mathbb{Q}$. This implies that the claim holds for μ-a.e. $x \in X$ when $t \in \mathbb{Q}$.

Case II: When t is irrational. First, we let $X = Z_2$, the Conze-Lesigne system. Let $\beta \in Z_1$ is an element such that for any $(z, u) \in Z_1 \times U = X$, $T(z, u) = (\beta z, u\rho(z))$. In other words, T acts on Z_1 as a rotation by β (here, we let Z_1 be a multiplicative abelian group). Then note that $B = \langle\beta\rangle$, the cyclic subgroup generated by β, is dense in the Kronecker factor Z_1. Define a character $\phi_t : B \to \mathbb{T}$ such that $\phi_t(\beta) = e^{2\pi i t}$. Such group homomorphism exists since t is irrational, and $\left\langle e^{2\pi i t}\right\rangle$ generates a dense cyclic subgroup in $\mathbb{T}$.

We claim that there exists a multiplicative character $\bar{\phi}_t : Z_1 \to \mathbb{T}$ such that $\bar{\phi}_t|_B = \phi_t$. Since B is dense in Z_1, for any $z \in Z_1$, there exists a sequence $(\beta^{n_k})_k$ such that $\lim_{k\to\infty} \beta^{n_k} = z$. So we define

$$\bar{\phi}_t(z) = \lim_{k\to\infty} \phi_t(\beta)^{n_k}.$$

We must show that this limit converges, which would show that $\overline{\phi}_t$ is well-defined by the continuity of ϕ. Note that $\mathbb{T}$ is compact, so there exists a converging subsequence $(\phi_t(\beta)^{n_{k_l}}) \in \mathbb{T}$ such that $\lim_{l\to\infty} \phi_t(\beta)^{n_{k_l}} = \gamma$ for some $\gamma \in \mathbb{T}$. We will show that $\lim_{k\to\infty} \phi_t(\beta)^{n_k} = \gamma$. Assuming on the contrary, suppose that there exists a subsequence $(\phi_t(\beta)^{n_{k_m}})_m$ such that $|\phi_t(\beta)^{n_{k_m}} - \gamma| > \epsilon$ for all $m \in \mathbb{N}$. This implies that, for sufficiently large l, we have $|\phi_t(\beta)^{n_{k_m}} - \phi_t(\beta)^{n_{k_l}}| > \epsilon/2$. This however contradicts the continuity of ϕ_t, since if d_{Z_1} is the metric on Z_1, then $d_{Z_1}(\beta^{n_{k_l}}, \beta^{n_{k_m}}) \to 0$ as $l, m \to \infty$, because both $\beta^{n_{k_l}}$ and $\beta^{n_{k_m}}$ converges to the same limit z. This proves that $\bar{\phi}_t$ is well-defined for all $z \in Z_1$. The fact that $\bar{\phi}_t$ is a multiplicative character is obvious from the way $\bar{\phi}_t$ is defined in terms of ϕ_t.

We define a continuous function $f_t := \bar{\phi}_t \circ \pi_1$, where $\pi_1 : Z_2 \to Z_1$ is the factor map. We note that

$$f_t(T^n x) = \bar{\phi}_t(\pi_1(T^n x)) = \bar{\phi}_t(\pi_1(x)\beta^n) = f_t(x)\phi_t(\beta)^n = f_t(x)e^{2\pi i n t}.$$

Therefore,

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) f_t(T^n x) = \frac{f_t(x)}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e^{2\pi i n t}.$$

To show the convergence of this average, let $F(x_1, x_2, x_3) = f_1(x_1)f_2(x_2)f_t(x_3)$ be a function on $X^3 = G^3/\Lambda^3$. Let $T_1 = T \times \mathrm{Id} \times \mathrm{Id}$, $T_2 = \mathrm{Id} \times T \times \mathrm{Id}$, and $T_3 = \mathrm{Id} \times \mathrm{Id} \times T$. Note that an action of T_1 on X^3 corresponds to $g_1 = ((\beta, \rho), e, e) \in G^3$ (where e is the identity element of G), and similarly, T_2 corresponds to $g_2 = (e, (\beta, \rho), e) \in G^3$, and T_3 corresponds to $g_3 = (e, e, (\beta, \rho)) \in G^3$. Thus,

$$g(n) = g_1^{an} g_2^{bn} g_3^{n}$$

is a polynomial sequence. Furthermore, if $\vec{x} = (x, x, x) \in X^3$, then

$$\frac{1}{N}\sum_{n=1}^{N} F(g(n)\vec{x}) = \frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) f_t(T^n x)$$

converges by Theorem 1.3.3.

It remains to show that the averages converges for any $f_1, f_2 \in L^\infty(\mu) \cap \mathcal{Z}_2$. We actually postpone this for now, since these approximation arguments are given later for more general case, where we have $e(p(n))$, where p is a polynomial with real coefficient, instead of $e(nt)$ (cf. Lemma 3.2.5). □

2.5 Proof of Theorem 2.2.3: The pointwise estimate

In this section, we will show that we can obtain a pointwise estimate to the Wiener-Wintner double recurrence averages using the seminorm of $\mathcal{A}_2$. This means that we can bound the Wiener-

Wintner double recurrence averages using the seminorm $N_2(\cdot)$ without taking the integral of the norm of the averages. This was not the case when we used the Host-Kra seminorm $\|\!|\cdot|\!\|_3$, where we obtained the norm bound

$$\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} f_1(T^n x) f_2(T^{2n}x) e(nt) \right|^2 d\mu \le C \|\!| f_1 |\!\|_3^{1/2}.$$

We recall that $(X, \mathcal{F}, \mu, T)$ is an ergodic system, $f_1, f_2 \in L^\infty(\mu)$ real-valued functions such that $\|f_i\|_\infty \le 1$ for both $i = 1, 2$. For the sake of simplicity, we will prove this result only for the case $a = 1$ and $b = 2$.

First, we will show that the Kronecker factor $\mathcal{A}_1$ is a pointwise characteristic factor for the double recurrence theorem.

Lemma 2.5.1. *Suppose $F_1, F_2 \in L^\infty(X)$ such that $\|F_1\|_\infty, \|F_2\|_\infty \le 1$. If $F_1 \in \mathcal{A}_1^\perp$, then for μ-a.e. $x \in X$,*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} F_1(T^n x) F_2(T^{2n}x) = 0.$$

Thus, $\mathcal{A}_1$ is a pointwise characteristic factor for the double recurrence averages, i.e.

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} F_1(T^n x) F_2(T^{2n}x) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}(F_1|\mathcal{A}_1)(T^n x)\mathbb{E}(F_2|\mathcal{A}_1)(T^{2n}x). \tag{2.10}$$

Proof. Since $\left|\frac{1}{N}\sum_{n=1}^{N} F_1(T^n x) F_2(T^{2n}x)\right|$ is non-negative, we can prove this lemma by showing

$$\int \limsup_{N\to\infty} \left| \frac{1}{N}\sum_{n=1}^{N} F_1(T^n x) F_2(T^{2n}x) \right| d\mu(x) = \lim_{N\to\infty} \int \left| \frac{1}{N}\sum_{n=1}^{N} F_1(T^n x) F_2(T^{2n}x) \right| d\mu(x) = 0$$

where the first equality holds by Bourgain's double recurrence theorem and Lebesgue's dominated convergence theorem. Note that the Cauchy-Schwarz inequality asserts that

$$\int \left| \frac{1}{N}\sum_{n=1}^{N} F_1(T^n x) F_2(T^{2n}x) \right| d\mu(x) \le \left(\int \left| \frac{1}{N}\sum_{n=1}^{N} F_1(T^n x) F_2(T^{2n}x) \right|^2 d\mu(x) \right)^{1/2}.$$

We will proceed by applying the reverse Fatou's lemma and the inequality (1.4) to the sequence

$F_1(T^n x)F_2(T^{2n}x)$ pointwise. We have

$$\begin{aligned}
&\limsup_{N\to\infty}\int\left|\frac{1}{N}\sum_{n=1}^{N}F_1(T^n x)F_2(T^{2n}x)\right|^2 d\mu(x)\\
&\leq \int\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N}F_1(T^n x)F_2(T^{2n}x)\right|^2 d\mu(x)\\
&\leq \frac{C}{H}+\frac{C}{(H+1)^2}\sum_{h=1}^{H}(H+1-h)\int\limsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}(F_1\cdot F_1\circ T^h)(T^n x)(F_2\cdot F_2\circ T^{2h})(T^{2n}x)d\mu\\
&\leq \frac{C}{H}+\frac{C}{H}\sum_{h=1}^{H}\left|\int\limsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}(F_1\cdot F_1\circ T^h)(T^n x)(F_2\cdot F_2\circ T^{2h})(T^{2n}x)d\mu\right|.
\end{aligned}$$

Note that the limit inside the integral exists by the double recurrence theorem. Hence, the dominated convergence theorem tells us that

$$\begin{aligned}
&\limsup_{N\to\infty}\int\left|\frac{1}{N}\sum_{n=1}^{N}F_1(T^n x)F_2(T^{2n}x)\right|^2 d\mu(x)\\
&\leq \frac{C}{H}+\frac{C}{H}\sum_{h=1}^{H}\left|\lim_{N\to\infty}\int\frac{1}{N}\sum_{n=1}^{N}(F_1\cdot F_1\circ T^h)(T^n x)(F_2\cdot F_2\circ T^{2h})(T^{2n}x)d\mu\right|\\
&= \frac{C}{H}+\frac{C}{H}\sum_{h=1}^{H}\left|\lim_{N\to\infty}\int(F_1\cdot F_1\circ T^h)(x)\frac{1}{N}\sum_{n=1}^{N}(F_2\cdot F_2\circ T^{2h})(T^n x)d\mu\right|
\end{aligned}$$

Then we use the mean ergodic theorem and the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}
\limsup_{N\to\infty}\int\left|\frac{1}{N}\sum_{n=1}^{N}F_1(T^n x)F_2(T^{2n}x)\right|^2 d\mu(x) &\leq \frac{C}{H}+\frac{C}{H}\sum_{h=1}^{H}\|F_2\|_\infty^2\left|\int(F_1\cdot F_1\circ T^h)(x)d\mu\right|\\
&\leq \frac{C}{H}+C\left(\frac{1}{H}\sum_{h=1}^{H}\left|\int(F_1\cdot F_1\circ T^h)(x)d\mu\right|^2\right)^{1/2}\\
&= \frac{C}{H}+C\left(\frac{1}{H}\sum_{h=1}^{H}|\hat{\sigma}_{F_1}(h)|^2\right)^{1/2},
\end{aligned}$$

where σ_{F_1} is the spectral measure of F_1 with respect to the transformation T. Now we let $H\to\infty$ to obtain

$$\limsup_{N\to\infty}\int\frac{1}{N}\sum_{n=1}^{N}\left|F_1(T^n x)F_2(T^{2n}x)\right|^2 d\mu(x)\leq C\left(\lim_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}|\hat{\sigma}_{F_1}(h)|^2\right)^{1/2},$$

and because $F_1 \in \mathcal{A}_1^{\perp}$, the spectral measure σ_{F_1} is continuous, so the Wiener's lemma implies the limit of the right hand side of the inequality above equals 0. $\square$

Proof of Theorem 2.2.3. We first apply the inequality (1.6) to the sequence $a_n = f_1(T^n x) f_2(T^{2n} x)$ pointwise to obtain

$$\begin{aligned}
&\limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^n x) f_2(T^{2n} x) e^{2\pi i n t} \right|^2 \\
&\leq \frac{C}{H} + \frac{C}{H} \sum_{h=1}^{H} \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} (f_1 \cdot f_1 \circ T^h)(T^n x)(f_2 \cdot f_2 \circ T^{2h})(T^{2n} x) \right|.
\end{aligned}$$

Our main task is to show that

$$\limsup_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \lim_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} \left(f_1 \cdot f_1 \circ T^h\right)(T^n x) \left(f_2 \cdot f_2 \circ T^{2h}\right)(T^{2n} x) \right| \leq [N_2(f_1)]^2 \tag{2.11}$$

for μ-a.e. $x \in X$. By Lemma 2.2.3, it would be sufficient to show that

$$\limsup_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \lim_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\left(f_1 \cdot f_1 \circ T^h | \mathcal{A}_1\right)(T^n x)\, \mathbb{E}\left(f_2 \cdot f_2 \circ T^{2h} | \mathcal{A}_1\right)(T^{2n} x) \right| \leq [N_2(f_1)]^2.$$

So set $F_{1,h} = f_1 \cdot f_1 \circ T^h$, and $F_{2,h} = f_2 \cdot f_2 \circ T^{2h}$. Denote

$$P_N(F_{1,h}, F_{2,h}) = \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}(F_{1,h}|\mathcal{A}_1) \circ T^n \cdot \mathbb{E}(F_{2,h}|\mathcal{A}_1) \circ T^{2n}.$$

Let $\{e_j\}$ be an eigenbasis of $\mathcal{A}_1$, where λ_j is the corresponding eigenvalue of e_j. Then we would have

$$\mathbb{E}(F_{1,h}|\mathcal{A}_1) \circ T^n = \sum_{j=0}^{\infty} \left(\int F_{1,h} \bar{e}_j \, d\mu \right) \lambda_j^n e_j \text{ and } \mathbb{E}(F_{2,h}|\mathcal{A}_1) \circ T^{2n} = \sum_{l=0}^{\infty} \left(\int F_{2,h} \bar{e}_l \, d\mu \right) \lambda_l^{2n} e_l$$

in the L^2-norm. Hence,

$$\lim_{N\to\infty} P_N(F_{1,h}, F_{2,h}) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} \left(\int F_{1,h} \overline{e_j} \, d\mu \right) \left(\int F_{2,h} \overline{e_l} \, d\mu \right) \lambda_j^n \lambda_l^{2n} e_j e_l$$

in the L^2-norm. Note that for each j and l,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \lambda_j^n \lambda_l^{2n} = \begin{cases} 1 & \text{if } \lambda_j = \overline{\lambda}_l^2, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, if we denote $R = \left\{(j, l_j) \in \mathbb{N}^2 : \lambda_j = \overline{\lambda}_{l_j}^2\right\}$, then

$$\lim_{N\to\infty} P_N(F_{1,h}, F_{2,h}) = \sum_{(j,l_j)\in R} \left(\int F_1 \overline{e_j}\, d\mu\right)\left(\int F_{2,h}\overline{e_{l_j}}\, d\mu\right) e_j e_{l_j}$$

in the L^2-norm. Note that the sequence

$$B_J = \left(\sum_{(j,l_j)\in R, j\le J} \left(\int F_{1,h}\overline{e_j}\, d\mu\right)\left(\int F_{2,h}\overline{e_{l_j}}\, d\mu\right) e_j e_{l_j}\right)_J$$

converges to $\lim_{N\to\infty} P_N(F_{1,h}, F_{2,h})$ in the L^2-norm as $J \to \infty$. Therefore, there exists a subsequence $(B_{J_k})_k$ that converges to $\lim_{N\to\infty} P_N(F_{1,h}, F_{2,h})(x)$ for μ-a.e. $x \in X$. Thus,

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}(F_{1,h}|\mathcal{A}_1)(T^n x)\mathbb{E}(F_{2,h}|\mathcal{A}_1)(T^{2n}x) \\
&= \lim_{k\to\infty} \sum_{(j,l_j)\in R, j\le J_k} \left(\int F_{1,h}\overline{e_j}\, d\mu\right)\left(\int F_{2,h}\overline{e_{l_j}}\, d\mu\right) e_j(x)e_{l_j}(x) \\
&\le \lim_{k\to\infty} \left(\sum_{(j,l_j)\in R, j\le J_k} \left|\int F_{1,h}\overline{e_j}\, d\mu\right|^2\right)^{1/2} \left(\sum_{(j,l_j)\in R, j\le J_k} \left|\int F_{2,h}\overline{e_{l_j}}\, d\mu\right|^2\right)^{1/2} \\
&\text{(by the Cauchy-Schwarz Inequality)} \\
&\le \left(\sum_{j=1}^{\infty} \left|\int F_{1,h}\, \overline{e_j}\, d\mu\right|^2\right)^{1/2} \left(\sum_{l=1}^{\infty} \left|\int F_{2,h}\overline{e_l}\, d\mu\right|^2\right)^{1/2} \\
&= \|\mathbb{E}(F_{1,h}|\mathcal{A}_1)\|_2 \|\mathbb{E}(F_{2,h}|\mathcal{A}_1)\|_2 \le \min_{i=1,2} \|\mathbb{E}(F_{i,h}|\mathcal{A}_1)\|_2,
\end{aligned}$$

where the last inequality follows from the fact that $\|f_2\|_\infty \le 1$. Therefore,

$$\limsup_{H\to\infty} \frac{1}{H}\sum_{h=1}^{H} \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \left(f_1 \cdot f_1 \circ T^h\right)(T^n x)\left(f_2 \cdot f_2 \circ T^{2h}\right)(T^{2n}x)$$

$$
\begin{aligned}
&\leq \min_{i=1,2} \limsup_{H\to\infty} \frac{1}{H}\sum_{h=1}^{H} \|\mathbb{E}(f_i \cdot f_i \circ T^h | \mathcal{A}_1)\|_2 \\
&\leq \min_{i=1,2} \left(\limsup_{H\to\infty} \frac{1}{H}\sum_{h=1}^{H} \|\mathbb{E}(f_i \cdot f_i \circ T^h | \mathcal{A}_1)\|_2^2 \right)^{1/2} = \min_{i=1,2} [N_2(f_i)]^2,
\end{aligned}
$$

where the second inequality holds by the Cauchy-Schwarz inequality. □

CHAPTER 3

Wiener-Wintner double recurrence II: Polynomial Case

In this chapter[1], we will extend the result presented in Chapter 2.

3.1 Introduction

The Wiener-Wintner averages have also been generalized to the context of having a polynomial as the exponent of the complex number, i.e. study of the averages of the form

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

where p is a real-valued polynomial. The convergence result of these types of averages were first obtained by E. Lesigne in 1990 [71]. Furthermore, if p is a degree-k polynomial, Lesigne has shown that if f belongs to the orthogonal complement of the k-th Abramov factor (i.e. span of k quasi-eigenfunctions) and T is totally ergodic (i.e. T^a is ergodic for any $a \in \mathbb{Z} - \{0\}$), then the averages above converge to zero in 1993 [73]. With the same assumption, Lesigne's work was extended by N. Frantzikinakis to a uniformity result in 2006, but one cannot remove the totally ergodicity assumption on T to obtain the similar uniformity result [46, Theorems 1.3 and 1.4].

On the other hand, B. Host and B. Kra extended Lesigne's work to a nilsequence Wiener-Wintner result in 2008 [60, Theorem 2.22]. In other words, they have shown that given an ergodic system $(X, \mathcal{F}, \mu, T)$ and a function $f \in L^\infty(\mu)$, there exists a set of full measure $X_f \subset X$ such that for any $x \in X_f$, for any nilsequence (a_n), the averages

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

[1]The material presented in this chapter originally appeared in [18], an article to appear on Journal d'Analyse Mathematique. The original citation is as follows: I. Assani and R. Moore. Extension of Wiener-Wintner double recurrence theorem to polynomials. To appear in Journal d'Analyse Mathematique, arXiv:1409.0463, 2015.

This result was later extended by T. Eisner and P. Zorin-Kranich in 2013, who showed that if we assume the function f is orthogonal to the k-th Host-Kra-Ziegler factor $\mathcal{Z}_k$, then the uniformity result for a k-step nilsequence (in particular, $a_n = e(p(n))$ may be obtained when T is ergodic [45, Theorem 1.2].

Recently, T. Eisner and B. Krause obtained a uniform Wiener-Wintner results for averages with weights involving Hardy functions and for "twisted" polynomial ergodic averages [44].

3.2 The main result

Throughout this section, we denote $\mathbb{R}_k[\xi]$ to be the set of all the k-th degree polynomials with real coefficients, whereas $\mathbb{R}[\xi]$ denotes the set of all polynomials with real coefficients.

Theorem 3.2.1. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic dynamical system, a and b be distinct nonzero integers, and $f_1, f_2 \in L^\infty(\mu)$ for which $\|f_i\|_\infty \leq 1$ for both $i = 1, 2$. Let*

$$W_N(f_1, f_2, x, p) = \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)),$$

where p is a polynomial with real coefficients. Then the following statements are true for any positive integer k:

1. *If either f_1 or f_2 belongs to $\mathcal{Z}_{k+1}^\perp$, then there exists a set of full measure $X_{f_1,f_2}^{1,k}$ such that for all $x \in X_{f_1,f_2}^{1,k}$,*
$$\limsup_{N\to\infty} \sup_{p \in \mathbb{R}_k[\xi]} |W_N(f_1, f_2, x, p)| = 0.$$

2. *If $f_1, f_2 \in \mathcal{Z}_{k+1}$, then there exists a set of full measure $X_{f_1,f_2}^{2,k}$ such that for all $x \in X_{f_1,f_2}^{2,k}$, the averages $W_N(f_1, f_2, x, p)$ converge for all $p \in \mathbb{R}_k[\xi]$.*

3. *For any character $\phi : \mathbb{T} \to \mathbb{C}$, there exists a set of full measure X_{f_1,f_2} such that for any $x \in X_{f_1,f_2}$, the averages*
$$\frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) \phi(p(n))$$
converge for any real polynomial $p \in \mathbb{R}[\xi]$.

Consequentially, there exists a set of full measure X_{f_1,f_2} such that for any $x \in X_{f_1,f_2}$ and $p \in \mathbb{R}[\xi]$, the averages $W_N(f_1, f_2, x, p)$ converge.

Certainly, one can obtain the set of full measure X_{f_1,f_2} mentioned in the theorem by taking the intersection of the countable intersections of $X^{1,k}_{f_1,f_2}$ and $X^{2,k}_{f_1,f_2}$, i.e.

$$X_{f_1,f_2} = \bigcap_{i=1}^{2}\bigcap_{k=1}^{\infty} X^{i,k}_{f_1,f_2}.$$

Prior to this theorem, the author and I. Assani obtained a weaker version of the aforementioned theorem. More specifically, we obtained convergence on a set of full measure that is independent of a real parameter t as opposed to all the polynomials with real coefficients.

Theorem 3.2.2 ([17]). *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic dynamical system, a and b be distinct nonzero integers, and $f_1, f_2 \in L^\infty(\mu)$ for which $\|f_i\|_\infty \le 1$ for both $i = 1, 2$. Let*

$$W_N(f_1, f_2, x, p, t) = \frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(tp(n)),$$

where p is a polynomial with real coefficients, and t a real number. Then the following statements are true for any positive integer k:

1. *If either f_1 or f_2 belongs to $\mathcal{Z}_{k+1}^{\perp}$, then for any $p \in \mathbb{R}_k[\xi]$, there exists a set of full measure $X^{1,k}_{f_1,f_2,p}$ such that for all $x \in X^{1,k}_{f_1,f_2,p}$,*

$$\limsup_{N\to\infty}\sup_{t\in\mathbb{R}} |W_N(f_1, f_2, x, p, t)| = 0.$$

2. *If $f_1, f_2 \in \mathcal{Z}_{k+1}$, then for any $p \in \mathbb{R}_k[\xi]$, there exists a set of full measure $X^{2,k}_{f_1,f_2,p}$ such that for all $x \in X^{2,k}_{f_1,f_2,p}$, the averages $W_N(f_1, f_2, x, p, t)$ converge for all $t \in \mathbb{R}$.*

Consequentially, for any polynomial $p \in \mathbb{R}[\xi]$, there exists a set of full measure $X_{f_1,f_2,p}$ such that for any $x \in X_{f_1,f_2,p}$ and $t \in \mathbb{R}$, the averages $W_N(f_1, f_2, x, p, t)$ converge.

While this result follows immediately from Theorem 3.2.1, one can obtain an independent proof by using Anzai's skew product [1] (in addition to Theorem 2.2.1 and the Host-Kra-Ziegler factors). In fact, this weaker theorem provided an angle to prove more general case, which is Theorem 3.2.1. Furthermore, Theorem 3.2.2 is enough to show that the sequence $(f_1(T^{an}x)f_2(T^{bn}x))_n$ is a good universal weight for $\mathfrak{X}(\Pi)$ in norm, where Π is a collection of the measure-preserving systems, and

$\mathfrak{X}(\Pi)$ is collection of all the processes of the form $X_n = g \circ S^{p(n)}$, where $(Y, \mathcal{G}, \nu, S) \in \Pi$, $g \in L^\infty(\nu)$, and p an integer-value polynomial. In other words,

Corollary 3.2.3. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic theorem, and a and b be distinct nonzero integers. Then there exists a set of full measure X_{f_1,f_2} such that for any $x \in X_{f_1,f_2}$ and any integer-valued polynomial p, and any other measure-preserving system $(Y, \mathcal{G}, \nu, S)$ and $g \in L^\infty(\nu)$, the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) g \circ S^{p(n)}$$

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

Of course, if S^t is well-defined for any $t \in \mathbb{R}$, the previous corollary holds for any real-valued polynomial $p \in \mathbb{R}[\xi]$.

3.2.1 Proof of (1) of Theorem 3.2.1

In this section, we prove (1) of Theorem 3.2.1. To illustrate the proof better, we will first proceed for the case where we have a second degree polynomial.

Proof for the case $k = 2$ Suppose that either f_1 or f_2 belongs to the orthogonal complement of $\mathcal{Z}_3$. Then we apply van der Corput's lemma and the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}
&\limsup_{N\to\infty} \sup_{\alpha,\beta\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(\alpha n^2 + \beta n) \right|^2 \\
&\leq \frac{2}{H} + \frac{4}{(H+1)^2}\sum_{h=1}^{H}(H+1-h) \\
&\cdot \limsup_{N\to\infty} \sup_{\alpha,\beta\in\mathbb{R}} Re \left(\frac{1}{N}\sum_{n=1}^{N} (f_1 \cdot f_1 \circ T^{ah})(T^{an}x)(f_2 \cdot f_2 \circ T^{bh})(T^{bn}x) e(-(\alpha h^2 + 2\alpha h n + \beta h)) \right) \\
&\leq \frac{2}{H} + 4 \left(\frac{1}{H+1}\sum_{h=1}^{H} \right. \\
&\left. \limsup_{N\to\infty} \sup_{\alpha\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} (f_1 \cdot f_1 \circ T^{ah})(T^{an}x)(f_2 \cdot f_2 \circ T^{bh})(T^{bn}x) e(-2\alpha h n) \right|^2 \right)^{1/2}.
\end{aligned}$$

We integrate both sides of the inequality above (which can be done by Lemma 2.3.1), and by Hölder's inequality, we obtain

$$\begin{aligned}
&\int \limsup_{N\to\infty} \sup_{\alpha,\beta\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(\alpha n^2+\beta n)\right|^2 d\mu(x) \\
&\le \frac{2}{H} + 4\Bigg(\frac{1}{H+1}\sum_{h=1}^{H} \\
&\int \limsup_{N\to\infty} \sup_{\alpha\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} (f_1\cdot f_1\circ T^{ah})(T^{an}x)(f_2\cdot f_2\circ T^{bh})(T^{bn}x)e(-2\alpha hn)\right|^2 d\mu(x)\Bigg)^{1/2}.
\end{aligned} \tag{3.1}$$

Note that the inside of the integral on the right hand side of (3.1) is a double recurrence Wiener-Wintner average (by setting $t = -2\alpha h$) for each h. By (2.8), we have

$$\begin{aligned}
&\int \limsup_{N\to\infty} \sup_{\alpha\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} (f_1\cdot f_1\circ T^{ah})(T^{an}x)(f_2\cdot f_2\circ T^{bh})(T^{bn}x)e(-2\alpha hn)\right|^2 d\mu(x) \\
&\lesssim_{a,b} \min\left\{ \left|\!\left|\!\left| f_1\cdot f_1\circ T^{ah}\right|\!\right|\!\right|_3^{1/2}, \left|\!\left|\!\left| f_2\cdot f_2\circ T^{bh}\right|\!\right|\!\right|_3^{1/2} \right\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\int \limsup_{N\to\infty} \sup_{\alpha,\beta\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(\alpha n^2+\beta n)\right|^2 d\mu(x) \\
&\lesssim_{a,b} \frac{1}{H} + \min\left\{ \left(\frac{1}{H}\sum_{h=1}^{H} \left|\!\left|\!\left| f_1\cdot f_1\circ T^{ah}\right|\!\right|\!\right|_3^{1/2}\right)^{1/2}, \left(\frac{1}{H}\sum_{h=1}^{H} \left|\!\left|\!\left| f_2\cdot f_2\circ T^{bh}\right|\!\right|\!\right|_3^{1/2}\right)^{1/2} \right\} \\
&\lesssim_{a,b} \frac{1}{H} + \min\left\{ \left(\frac{1}{H}\sum_{h=1}^{H} \left|\!\left|\!\left| f_1\cdot f_1\circ T^{ah}\right|\!\right|\!\right|_3^{8}\right)^{1/32}, \left(\frac{1}{H}\sum_{h=1}^{H} \left|\!\left|\!\left| f_2\cdot f_2\circ T^{bh}\right|\!\right|\!\right|_3^{8}\right)^{1/32} \right\},
\end{aligned}$$

and by letting $H \to \infty$, we obtain

$$\int \limsup_{N\to\infty} \sup_{\alpha,\beta\in\mathbb{R}} \left| \frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(\alpha n^2+\beta n)\right|^2 d\mu(x) \lesssim_{a,b} \min\left\{ |\!|\!| f_1|\!|\!|_4^{1/2}, |\!|\!| f_2|\!|\!|_4^{1/2}\right\}. \tag{3.2}$$

Since either f_1 or f_2 belongs to $\mathcal{Z}_3^{\perp}$, either $|\!|\!| f_1|\!|\!|_4$ or $|\!|\!| f_2|\!|\!|_4$ equals 0. This completes the proof for the case where $p(n) = \alpha n^2 + \beta n$.

The proof for any positive integer k One of the key inequalities (besides van der Corput's lemma) used for the case where $k = 2$ is (3.2), where we controlled the integral of the lim sup of the averages by an appropriate Gowers-Host-Kra seminorm. We generalize this inequality for polynomials with higher degree with induction (on k) and van der Corput's inequality.

Lemma 3.2.4. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system, and $f_1, f_2 \in L^\infty$. Then*

$$\int \limsup_{N\to\infty} \sup_{p\in\mathbb{R}_k[\xi]} \left|\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(p(n))\right|^2 d\mu(x) \lesssim_{a,b,k} \min\left\{|||f_1|||_{k+2}^{1/2}, |||f_2|||_{k+2}^{1/2}\right\}. \quad (3.3)$$

Proof of Lemma 3.2.4. We proceed by induction on k. The base case $k = 1$ is clear from the estimate (2.8). Now suppose the claim holds for $k = 1, 2, \ldots, l$. Let $p(n)$ be a polynomial with degree $l+1$. If $q_h(n) = p(n+h) - p(n)$, then $q_h(n)$ is a polynomial of degree less than or equal to l for all h, viewing n as the variable. By van der Corput's lemma and the Cauchy-Schwarz inequality, we know that

$$\begin{aligned}
&\limsup_{N\to\infty} \sup_{p\in\mathbb{R}_{l+1}[t]} \left|\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(p(n))\right|^2 \\
&\leq \frac{2}{H+1} + \frac{4}{H+1}\sum_{h=1}^{H} \limsup_{N\to\infty} \sup_{q_h\in\mathbb{R}_l[t]} \left|\frac{1}{N}\sum_{n=1}^{N-h-1} (f_1\cdot f_1\circ T^{ah})(T^{an}x)(f_2\cdot f_2\circ T^{bh})(T^{bn}x)e(q_h(n))\right| \\
&\leq \frac{2}{H+1} + 4\left(\frac{1}{H+1}\sum_{h=1}^{H}\right. \\
&\left.\limsup_{N\to\infty} \sup_{q_h\in\mathbb{R}_l[t]} \left|\frac{1}{N}\sum_{n=1}^{N-h-1} (f_1\cdot f_1\circ T^{ah})(T^{an}x)(f_2\cdot f_2\circ T^{bh})(T^{bn}x)e(q_h(n))\right|^2\right)^{1/2}.
\end{aligned}$$

By integrating both sides (which is possible by Lemma 2.3.1) and applying Hölder's inequality, we have

$$\begin{aligned}
&\int \limsup_{N\to\infty} \sup_{p\in\mathbb{R}_{l+1}[t]} \left|\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(p(n))\right|^2 d\mu(x) \\
&\leq \frac{2}{H+1} + 4\left(\frac{1}{H+1}\sum_{h=1}^{H}\right. \\
&\left.\int \limsup_{N\to\infty} \sup_{q_h\in\mathbb{R}_l[t]} \left|\frac{1}{N}\sum_{n=1}^{N-h-1} (f_1\cdot f_1\circ T^{ah})(T^{an}x)(f_2\cdot f_2\circ T^{bh})(T^{bn}x)e(q_h(n))\right|^2 d\mu\right)^{1/2}.
\end{aligned}$$

For any $1 \leq h \leq H$, the inductive hypothesis tells us that

$$\int \limsup_{N\to\infty} \sup_{q_h\in\mathbb{R}_l[t]} \left| \frac{1}{N} \sum_{n=1}^{N-h-1} (f_1 \cdot f_1 \circ T^{ah})(T^{an}x)(f_2 \cdot f_2 \circ T^{bh})(T^{bn}x)e(q_h(n)) \right|^2 d\mu$$
$$\lesssim_{a,b,l} \min\left\{ \left|\!\left|\!\left| f_1 \cdot f_1 \circ T^{ah} \right|\!\right|\!\right|_{l+2}^{1/2}, \left|\!\left|\!\left| f_2 \cdot f_2 \circ T^{bh} \right|\!\right|\!\right|_{l+2}^{1/2} \right\}.$$

Therefore,

$$\int \limsup_{N\to\infty} \sup_{p\in\mathbb{R}_{l+1}[t]} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)) \right|^2 d\mu(x)$$
$$\lesssim_{a,b,l} \frac{1}{H} + \min\left\{ \left(\frac{1}{H} \sum_{h=1}^{H} \left|\!\left|\!\left| f_1 \cdot f_1 \circ T^{ah} \right|\!\right|\!\right|_{l+2}^{1/2} \right)^{1/2}, \left(\frac{1}{H} \sum_{h=1}^{H} \left|\!\left|\!\left| f_2 \cdot f_2 \circ T^{bh} \right|\!\right|\!\right|_{l+2}^{1/2} \right)^{1/2} \right\}$$
$$\lesssim_{a,b,l} \frac{1}{H} + \min\left\{ \left(\frac{1}{H} \sum_{h=1}^{H} \left|\!\left|\!\left| f_1 \cdot f_1 \circ T^{ah} \right|\!\right|\!\right|_{l+2}^{2^{l+2}} \right)^{2^{-(l+4)}}, \left(\frac{1}{H} \sum_{h=1}^{H} \left|\!\left|\!\left| f_2 \cdot f_2 \circ T^{bh} \right|\!\right|\!\right|_{l+2}^{2^{l+2}} \right)^{2^{-(l+4)}} \right\},$$

and if we let $H \to \infty$, we obtain

$$\int \limsup_{N\to\infty} \sup_{p\in\mathbb{R}_{l+1}[t]} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)) \right|^2 d\mu(x) \lesssim_{a,b,l} \min\left\{ |\!|\!| f_1 |\!|\!|_{l+3}^{1/2}, |\!|\!| f_2 |\!|\!|_{l+3}^{1/2} \right\}.$$

□

Proof of (1) of Theorem 3.2.1. By our assumption, either f_1 or f_2 belongs to $\mathcal{Z}_{k+1}^{\perp}$, which implies that either $|\!|\!| f_1 |\!|\!|_{k+2}$ or $|\!|\!| f_2 |\!|\!|_{k+2}$ equals 0, hence the right hand side of the inequality (3.3) equals 0. Thus, there exists a set of full measure X_{f_1,f_2} such that for any $x \in X_{f_1,f_2}$ and $p \in \mathbb{R}_k[\xi]$, we have

$$\limsup_{N\to\infty} \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)) \right| = 0.$$

□

3.2.2 Proofs of (2) and (3) of Theorem 3.2.1

In this section, we first prove (2) of Theorem 3.2.1. Then we use this, together with (1), to prove (3) of the same theorem.

First, we prove the following approximation lemma that allows us to reduce our proof to the case where f_1 and f_2 are both continuous functions on an ergodic nilsystem. The following inequality will be useful when dominating the averages in norm: Given a measure-preserving system $(X, \mathcal{F}, \mu, T)$ and $F \in L^\alpha(\mu)$ for $\alpha \in (1, \infty)$, we have

$$\left\| \sup_N \frac{1}{N} \sum_{n=1}^{N} F(T^n x) \right\|_\alpha \leq \frac{\alpha}{\alpha - 1} \|F\|_\alpha . \tag{3.4}$$

This inequality can be obtained by using the maximal ergodic theorem (see, for example, [6, Theorem 1.8] for a proof).

Lemma 3.2.5. *Let $(X, \mathcal{F}, \mu, T)$ be a measure-preserving system, and a and b be distinct integers. Let $f_1, f_2 \in L^\infty(\mu)$. Suppose that there exist two sequences of functions $(f_1^i)_i$ and $(f_2^i)_i$ in $L^\infty(\mu)$ such that $\|f_1^i\|_{L^\infty(\mu)} < M$ for some constant $M > 0$ for any $i \in \mathbb{N}$, $f_j^i \to f_j$ in $L^2(\mu)$-norm as $i \to \infty$ for each $j = 1, 2$, and for each i, there exists a set of full measure X_i such that for any $x \in X_i$ and any $p \in \mathbb{R}_k[\xi]$ for each $k \in \mathbb{N}$, the averages*

$$\frac{1}{N} \sum_{n=1}^{N} f_1^i(T^{an}x) f_2^i(T^{bn}x) e(p(n))$$

converge. Then there exists a set of full measure $X_\infty \subset X$ such that for any $x \in X_\infty$ and any $p \in \mathbb{R}_k[\xi]$, the averages

$$\frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(p(n))$$

converge for each $k \in \mathbb{N}$.

Proof. For each $j = 1, 2$, we can write $f_j = (f_j - f_j^i) + f_j^i$ for each i, so we can rewrite the averages as follows:

$$\begin{aligned}
&W_N(f_1, f_2, x, p) = \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)) \\
&= \frac{1}{N} \sum_{n=1}^{N} (f_1 - f_1^i)(T^{an}x) f_2(T^{bn}x) e(p(n)) + \frac{1}{N} \sum_{n=1}^{N} f_1^i(T^{an}x)(f_2 - f_2^i)(T^{bn}x) e(p(n)) \qquad (3.5) \\
&+ \frac{1}{N} \sum_{n=1}^{N} f_1^i(T^{an}x) f_2^i(T^{bn}x) e(p(n)).
\end{aligned}$$

Ultimately, we would like to show that there exists a set of full measure $X_\infty \subset X$ such that for any

$x \in X_\infty$,

$$\mathcal{L}_R(W_N(f_1, f_2, x, p)) \tag{3.6}$$
$$= \sup_{p \in \mathbb{R}_k[\xi]} \left(\limsup_{N\to\infty} Re(W_N(f_1, f_2, x, p)) - \liminf_{N\to\infty} Re(W_N(f_1, f_2, x, p)) \right) = 0,$$

and

$$\mathcal{L}_I(W_N(f_1, f_2, x, p)) \tag{3.7}$$
$$= \sup_{p \in \mathbb{R}_k[\xi]} \left(\limsup_{N\to\infty} Im(W_N(f_1, f_2, x, p)) - \liminf_{N\to\infty} Im(W_N(f_1, f_2, x, p)) \right) = 0.$$

To show (3.6), we first note that the third term on the right hand side of (3.5) vanishes for μ-a.e. $x \in X$ for each i after applying $\mathcal{L}_R$ since we know that the averages converge for all $x \in \bigcap_{i=1}^\infty X_i$, which is a set of full measure, and for any $p \in \mathbb{R}_k[\xi]$ and $i \in \mathbb{N}$. To show the remaining terms vanish, we apply Hölder's inequality as well as the inequality (3.4). For instance, for the first term of (3.5), we have that

$$\sup_{p \in \mathbb{R}_k[\xi]} \left| \frac{1}{N} \sum_{n=1}^N (f_1 - f_1^i)(T^{an}x) f_2(T^{bn}x) e(p(n)) \right| \leq \frac{1}{N} \sum_{n=1}^N \left| (f_1 - f_1^i)(T^{an}x) f_2(T^{bn}x) \right|$$
$$\leq \|f_2\|_{L^\infty(\mu)} \frac{1}{N} \sum_{n=1}^N \left| f_1 - f_1^i \right| (T^{an}x).$$

If we take supremum over N on both sides, we would have

$$\sup_{N\geq 1} \sup_{p \in \mathbb{R}_k[\xi]} \left| \frac{1}{N} \sum_{n=1}^N (f_1 - f_1^i)(T^{an}x) f_2(T^{bn}x) e(p(n)) \right| \leq \|f_2\|_{L^\infty(\mu)} \left(\sup_{N\geq 1} \frac{1}{N} \sum_{n=1}^N \left| f_1 - f_1^i \right| (T^{an}x) \right),$$

so we integrate both sides (which is possible by Lemma 2.3.1) and apply Hölder's inequality for the right hand side to obtain

$$\int \sup_{N\geq 1} \sup_{p \in \mathbb{R}_k[\xi]} \left| \frac{1}{N} \sum_{n=1}^N (f_1 - f_1^i)(T^{an}x) f_2(T^{bn}x) e(p(n)) \right| d\mu(x)$$
$$\leq \|f_2\|_{L^\infty(\mu)} \left\| \sup_{N\geq 1} \frac{1}{N} \sum_{n=1}^N \left| f_1 - f_1^i \right| (T^{an}x) \right\|_{L^2(\mu)}.$$

We apply the inequality (3.4) for the case where $\alpha = 2$ to the $L^2(\mu)$-norm on the right hand side to obtain

$$\int \sup_{N\geq 1} \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N}\sum_{n=1}^{N} (f_1 - f_1^i)(T^{an}x) f_2(T^{bn}x) e(p(n)) \right| d\mu(x) \leq 2\left\| f_2 \right\|_{L^\infty(\mu)} \left\| f_1 - f_1^i \right\|_{L^2(\mu)} . \quad (3.8)$$

By the similar argument as in the first term of (3.5) (and recalling that $\left\| f_1^i \right\|_{L^\infty(\mu)} \leq M$ for all i), we can also obtain an estimate for the second term:

$$\int \sup_{N\geq 1} \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N}\sum_{n=1}^{N} f_1^i(T^{an}x)(f_2 - f_2^i)(T^{bn}x) e(p(n)) \right| d\mu(x) \leq 2M \| f_2 - f_2^i \|_{L^2(\mu)} . \quad (3.9)$$

We are now ready to verify (3.6). We note that

$$\begin{aligned} 0 &\leq \sup_{p\in\mathbb{R}_k[\xi]} \left(\limsup_{N\to\infty} Re\left(W_N(f_1,f_2,x,p)\right) - \liminf_{N\to\infty} Re\left(W_N(f_1,f_2,x,p)\right) \right) \\ &\leq 2 \liminf_{i\to\infty} \left(\sup_{N\geq 1} \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N}\sum_{n=1}^{N} (f_1 - f_1^i)(T^{an}x) f_2(T^{bn}x) e(p(n)) \right| \right. \\ &\left. + \sup_{N\geq 1} \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N}\sum_{n=1}^{N} f_1^i(T^{an}x)(f_2 - f_2^i)(T^{bn}x) e(p(n)) \right| \right) . \end{aligned}$$

According to the inequalities (3.8) and (3.9), the integral of each average in the right-hand side of the inequality above is bounded by a constant multiple of either $\|f_1 - f_1^i\|_{L^2(\mu)}$ or $\|f_2 - f_2^i\|_{L^2(\mu)}$. These norms converge to 0 as $i \to \infty$. Using those inequalities together with Fatou's lemma, we obtain

$$\begin{aligned} &\int \sup_{p\in\mathbb{R}_k[\xi]} \left(\limsup_{N\to\infty} Re\left(W_N(f_1,f_2,x,p)\right) - \liminf_{N\to\infty} Re\left(W_N(f_1,f_2,x,p)\right) \right) d\mu \quad (3.10) \\ &\leq 2 \liminf_{i\to\infty} \left(\int \sup_{N\geq 1} \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N}\sum_{n=1}^{N} (f_1 - f_1^i)(T^{an}x) f_2(T^{bn}x) e(p(n)) \right| d\mu(x) \right. \\ &\left. + \int \sup_{N\geq 1} \sup_{p\in\mathbb{R}_k[\xi]} \left| \frac{1}{N}\sum_{n=1}^{N} f_1^i(T^{an}x)(f_2 - f_2^i)(T^{bn}x) e(p(n)) \right| d\mu(x) \right) = 0. \end{aligned}$$

Since inside the integral of (3.10) is nonnegative, (3.6) is established for μ-a.e. $x \in X$. Therefore, there exists a set of full measure $X_R \subset \bigcap_{i=1}^{\infty} X_i$ such that for any $x \in X_R$ and any polynomial $p \in \mathbb{R}_k[\xi]$, the real part of the sequence $(W_N(f_1, f_2, x, p))_N$ converge for any $k \in \mathbb{N}$.

Similarly, we can show that (3.7) holds for μ-a.e. $x \in \bigcap_{i=1}^{\infty} X_i$, so there exists a set of full measure $X_I \subset \bigcap_{i=1}^{\infty} X_i$ such that for any $x \in X_I$ and any polynomial $p \in \mathbb{R}_k[\xi]$, the imaginary part of the sequence $(W_N(f_1, f_2, x, p))_N$ converge for any $k \in \mathbb{N}$. So if we set $X_\infty = X_R \cap X_I$, we obtain the desired set of full measure. □

To prove (2) of Theorem 3.2.1, we first prove this for the case where $(X, \mathcal{F}, \mu, T)$ is an ergodic nilsystem, and f_1 and f_2 are both continuous functions on X; under these assumptions, the averages converge *for all* $x \in X$. The key ingredient of this proof is Leibman's pointwise convergence theorem of polynomial actions on a nilsystem (Theorem 1.3.3), which is used to prove the following lemma.

Lemma 3.2.6. *Let $(X, \mathcal{F}, \mu, T)$ be a $(k+1)$-step ergodic nilsystem. Suppose f_1 and f_2 are continuous functions on X, and $a, b \in \mathbb{Z}$ such that $a \neq b$. Then for any $x \in X$ and $p \in \mathbb{R}_k[\xi]$, the averages*

$$\frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) e(p(n)) \tag{3.11}$$

converge as $N \to \infty$.

Proof of Lemma 3.2.6. Let t be any real number. Suppose first that we fix a polynomial $q \in \mathbb{R}[\xi]$. Suppose also that X is a $(k+1)$-step nilsystem. Since we know that $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is a one-step nilmanifold, the system $(\mathbb{T}, \mathcal{B}, m, R_t)$ is a nilsystem, where $\mathcal{B}$ is a Borel σ-algebra of $\mathbb{T}$, m is the usual Borel probability measure, and R_t is a rotation by t i.e. for any $\alpha \in (0, 1]$, we have $R_t(e(\alpha)) = e(\alpha + t)$. Thus, $(X^2 \times \mathbb{T}, \mathcal{F}^2 \otimes \mathcal{B}, \mu^2 \otimes m)$ is a $k+1$-step nilmanifold. Suppose $F : X^2 \times \mathbb{T} \to \mathbb{C}$ for which

$$F(x_1, x_2, e(\alpha)) = f_1(x_1) f_2(x_2) e(\alpha).$$

Then F is continuous on $X^2 \times \mathbb{T}$. Hence, for any $\alpha \in [0, 1)$, we see that

$$\frac{1}{N} \sum_{n=1}^{N} F(T^{an}x_1, T^{bn}x_2, R_t^{q(n)} e(\alpha)) = \frac{e(\alpha)}{N} \sum_{n=1}^{N} f_1(T^{an}x_1) f_2(T^{bn}x_2) e(q(n)t).$$

Note that the left-hand side of the equation above converges by Theorem 1.3.3 for all $(x_1, x_2, e(\alpha)) \in$

$X^2 \times \mathbb{T}^2$, so the averages in (3.11) converge for *all* $(x_1, x_2, e(\alpha)) \in X^2 \times \mathbb{T}$ for any $t \in \mathbb{R}$. In particular, it converges when $x_1 = x_2 = x$ and $\alpha = 0$, so we have shown that for any $x \in X$, the averages

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(tq(n)) \tag{3.12}$$

converge as $N \to \infty$.

Now we fix $t \in (0, 1)$. Then for any k-th degree polynomial p, there exists another k-th degree polynomial q such that $p = tq$ (e.g. if $p(n) = \sum_{l=0}^{k} c_l n^l$, then we can set $q(n) = \sum_{l=0}^{k} c_l' n^l$, where $c_l' = c_l/t$). Thus, we have

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(p(n)) = \frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(tq(n)),$$

and we know that the averages in the right hand side is (3.12), so the averages on the left hand side converge for all $x \in X$ as $N \to \infty$. Hence, we have shown that the claim holds for the case where f_1 and f_2 are both continuous functions on an ergodic nilsystem. □

Using the preceding lemma as well as Lemma 3.2.5, we are ready to prove (2) of Theorem 3.2.1.

Proof of (2) *of Theorem* 3.2.1. By the structure theorem [59, Theorem 10.1], the $k+1$-th Host-Kra-Ziegler factor is the inverse limit of a sequence of $k+1$-step ergodic nilsystems that are factor of $(X, \mathcal{F}, \mu, T)$. This implies that if $f_1 \in \mathcal{Z}_{k+1}$, then for any $k+1$-step ergodic nilsystem $(N, \mathcal{N}, \mu, T)$ that is a factor of $(X, \mathcal{F}, \mu, T)$, we have $\|\mathbb{E}(f_1|\mathcal{N})\|_{L^\infty(\mu)} \le \|f_1\|_{L^\infty(\mu)}$ by Theorem 1.1.2. Hence, by Lemma 3.2.5, it suffices to show that the statement of the theorem holds for the case where f_1 and f_2 are bounded and measurable with respect to a $k+1$-step ergodic nilsystem $(N, \mathcal{N}, \mu, T)$ that is a factor of $(X, \mathcal{F}, \mu, T)$. But since we know that if g_1 and g_2 are continuous functions on N, then the averages

$$\frac{1}{N}\sum_{n=1}^{N} g_1(T^{an}x)g_2(T^{bn}x)e(p(n))$$

[2] In Leibman's paper, the polynomial sequences are defined for polynomials with integer coefficients. However, the cited theorems are proven for the case where one has a polynomial with real coefficients, as it is mentioned in [68, §3.13], provided that this action makes sense. Since the element of the group corresponding to the rotation in $\mathbb{T}$ belongs to the identity component of the group, a real polynomial exponential makes sense.

converge for all $x \in N$ and $p \in \mathbb{R}_k[\xi]$ by Lemma 3.2.6. Furthermore, by density, there exist sequences of continuous functions $(\tilde{g}_1^i)_i$ and $(\tilde{g}_2^i)_i$ on N such that $\tilde{g}_j^i \to f_j$ in $L^2(\mu)$ as $i \to \infty$ for each $j = 1, 2$. We can construct another sequence of continuous functions $(g_1^i)_i$ such that

$$g_1^i(x) = \begin{cases} \min\left(\tilde{g}_1^i(x), \|f_1\|_{L^\infty(\mu)}\right) & \text{if } \tilde{g}_1^i(x) \geq 0, \\ \max\left(\tilde{g}_1^i(x), -\|f_1\|_{L^\infty(\mu)}\right) & \text{if } \tilde{g}_1^i(x) < 0, \end{cases}$$

so that $g_1^i \to f_1$ in $L^2(\mu)$ as $i \to \infty$, and $\left\|g_1^i\right\|_{L^\infty(\mu)} < \|f_1\|_{L^\infty(\mu)}$ for each $i \in \mathbb{N}$. Thus, we can apply Lemma 3.2.5 again (for the sequences $(g_1^i)_i$ and $(\tilde{g}_2^i)_i$) to show that the statement of the theorem holds for the case where f_1 and f_2 are bounded and measurable functions on N. □

Now we are ready to prove (3) of Theorem 3.2.1 using (1) and (2).

Proof of (3) *of Theorem* 3.2.1. Since any continuous function ϕ on $\mathbb{T}$ can be approximated by a linear combination of complex trigonometric functions, it suffices to prove this claim by showing that the averages

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e(p(n)) \tag{3.13}$$

converge off a single null-set independent of p. First we find a single set of full measure independent of $p \in \mathbb{R}_k[\xi]$ for which the averages converge. For each $j = 1, 2$, we write $f_j = \mathbb{E}(f_j|\mathcal{Z}_{k+1}) + f_j^\perp$, where $f_j^\perp \in \mathcal{Z}_{k+1}^\perp$. By (1) of Theorem 3.2.1, we merely need to show that the averages

$$\frac{1}{N}\sum_{n=1}^{N} \mathbb{E}(f_1|\mathcal{Z}_{k+1})(T^{an}x)\mathbb{E}(f_2|\mathcal{Z}_{k+1})(T^{bn}x)e(p(n)) \tag{3.14}$$

converge on a set of full measure independent of $p \in \mathbb{R}_k[\xi]$. By (2) of Theorem 3.2.1, we know that there exists a set of full measure $X_{f_1,f_2,k}$ such that for any $x \in X_{f_1,f_2,k}$, the averages in (3.14) converges for all $p \in \mathbb{R}_k[\xi]$. Thus, if we set

$$X_{f_1,f_2} = \bigcap_{k=1}^{\infty} X_{f_1,f_2,k},$$

then X_{f_1,f_2} is a set of full measure independent for which the averages in (3.13) converge for all $x \in X_{f_1,f_2}$ and for all polynomials p with real coefficients. □

CHAPTER 4

Weighted multiple ergodic averages: Powers of a single transformation

In this chapter[1], we will show that given a measure-preserving system $(X, \mathcal{F}, \mu, T)$, bounded functions $f_1, f_2 \in L^\infty(\mu)$, and distinct nonzero integers a and b, the sequence averaged in the double recurrence theorem, $(f_1(T^{an}x)f_2(T^{bn}x))$, is a good universal weight for certain multiple recurrent averages in norm. The key ideas behind this chapter is that one can obtain the norm convergence of these averages using the Wiener-Wintner theorem.

4.1 The statement of the main result

In this chapter, we will prove the following:

Theorem 4.1.1. *Let $(X, \mathcal{F}, \mu, T)$ be a probability measure-preserving system, with functions $f_1, f_2 \in L^\infty(\mu)$. Then for μ-a.e. $x \in X$, the sequence $u_n = f_1(T^{an}x)f_2(T^{bn}x)$ is a good universal weight for a k-term Furstenberg averages in norm for any positive integer k. More precisely, there exists a set of full measure $X_{f_1,f_2} \subset X$ such that for any $x \in X_{f_1,f_2}$, $a, b \in \mathbb{Z}$ and any positive integer $k \geq 1$, and any other probability measure-preserving system $(Y, \mathcal{G}, \nu, S)$ with $g_1, \dots, g_k \in L^\infty(\nu)$, the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)\prod_{i=1}^{k} g_i \circ S^{in}$$

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

In particular, if $f_1 = \mathbb{1}_A$ and $f_2 = \mathbb{1}_B$ for some measurable sets $A, B \in \mathcal{F}$ with positive measures, then we see that the averages of the sequence $(\prod_{i=1}^{k} g_i(S^{in}y))_n$ along the subsequence of the return times of $T^{an}x$ to the set A and $T^{bn}x$ to the set B converge in $L^2(\nu)$-norm. This theorem mixes

[1] The material presented in this chapter originally appeared in [19], an article in Ergodic Theory & Dynamical Systems. The original citation is as follows: I. Assani and R. Moore. A good universal weight for nonconventional ergodic averages in norm. Ergod. Th. and Dynam. Sys. Available on CJO 2015 doi:10.1017/ etds.2015.76. Also available on arXiv:1503.08863. Also available on arXiv:1503.08863., 2015.

the weights from the a.e. double recurrent convergence result and the norm convergence of the multiple recurrent theorem. In terms of Definition 1.4.2, we show that for μ-a.e. $x \in X$, the sequence $(f_1(T^{an}x)f_2(T^{bn}x))_n$ is a good universal weight for the process $(X_n)_n$ of the form in (1.20) in norm.

Note that this theorem generalizes the result obtained by B. Host and B. Kra, since if $a = 1$ and $f_2 = \mathbb{1}_X$, then the averages in the theorem become the averages seen in (1.19). We remark that in the work of Host and Kra, it was shown that the sequence $f(T^n x)$ is a good universal weight for this process by noting that the sequence $f(T^n x)$ satisfies the generalized Wiener-Wintner averages (see [60, Theorem 2.22, Theorem 2.24]). Later, it was shown by Assani [12], and independently by Zorin-Kranich [88], that the sequence $(f_1(T^{an}x)f_2(T^{bn}x))$ satisfies the nilsequence Wiener-Wintner result as well.

We observe that the case $k = 1$ of the main result follows immediately from Theorem 2.2.1 and Theorem 1.4.4. In fact, we will show that this is the key step required to establish the base case of our inductive argument in the proof.

In the proof, we will assume that the systems $(X, \mathcal{F}, \mu, T)$ and $(Y, \mathcal{G}, \nu, S)$ are ergodic, since we can apply the ergodic decomposition to show that the result holds for general measure-preserving systems. To prove the main result for $k \geq 2$, we will first decompose the functions f_1 and f_2 into the appropriate Host-Kra-Ziegler factor of $(X, \mathcal{F}, \mu, T)$, and treat the cases when either f_1 or f_2 belongs to the orthogonal complement of this factor, or the case when both of them belong to the factor. For the first case, we will proceed by induction on the number of transformations $g_1, g_2, \ldots, g_k$; to do so, we will show that the $L^2(\nu)$-norm limit of the averages can be controlled by the limit of the double recurrence Wiener-Wintner averages. For the second case, we will decompose the functions $g_1, \ldots, g_k$ into an appropriate characteristic factor, and treat the sub-cases when either one of $g_1, \ldots, g_k$ belongs to the orthogonal complement of this factor, and when all of them belong to the factor separately. For the first sub-case, we will control the norm limit of the averages with the Gowers-Host-Kra seminorm that characterizes this factor, and uses this to show that the norm averages converge to 0. For the second sub-case, we will use the structure of nilmanifolds and Leibman's convergence result (Theorem 1.3.3) to prove the claim. The following theorem summarizes the steps required to prove Theorem 4.1.1.

Theorem 4.1.2. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system, and $f_1, f_2 \in L^\infty(\mu)$ such that $\|f_i\|_{L^\infty(\nu)} \leq 1$*

for both $i = 1, 2$. Fix a positive integer $k \geq 1$. Then the following statements are true.

(a) *Suppose either $f_1, f_2 \in \mathcal{Z}_{k+1}(T)^{\perp}$. Then there exists a set of full measure $\tilde{X}_k \subset X$ such that for any $x \in \tilde{X}$, any other measure-preserving system $(Y, \mathcal{G}, \nu, S)$, and functions $g_1, g_2, \ldots, g_k \in L^{\infty}(\nu)$ where $\|g_j\|_{L^{\infty}(\nu)} \leq 1$ for each $1 \leq j \leq k$, the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) \prod_{j=1}^{k} g_j \circ S^{jn} \tag{4.1}$$

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

(b) *For any $f_1, f_2 \in L^{\infty}(\mu)$, there exists a set of full measure $\hat{X}_k$ such that for any $x \in \hat{X}$ and any other ergodic system $(Y, \mathcal{G}, \nu, S)$, and functions $g_1, \ldots, g_k \in L^{\infty}(\nu)$ with one of them belonging to $\mathcal{Z}_k(S)^{\perp}$, the averages in (4.1) converge to 0 in $L^2(\nu)$.*

(c) *Suppose both $f_1, f_2 \in \mathcal{Z}_{k+1}(T)$. Then there exists a set of full measure $X_k' \in X$ such that for any $x \in X'$, and for any other ergodic system $(Y, \mathcal{G}, \nu, S)$ and functions $g_1, \ldots, g_k \in L^{\infty}(\nu) \cap \mathcal{Z}_k(S)$, the averages in (4.1) converge in $L^2(\nu)$.*

Proof that Theorem 4.1.2 implies Theorem 4.1.1. Fix a positive integer $k \geq 1$. Let $f_i' = f_i - \mathbb{E}(f_i|\mathcal{Z}_{k+1})$ for $i = 1, 2$. We rewrite the averages in (4.1) as follows:

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) \prod_{j=1}^{k} g_j \circ S^{jn} \qquad (4.2)\\
&= \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}(f_1|\mathcal{Z}_{k+1})(T^{an}x)\mathbb{E}(f_2|\mathcal{Z}_{k+1})(T^{bn}) \prod_{j=1}^{k} g_j \circ S^{jn}\\
&+ \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}(f_1|\mathcal{Z}_{k+1})(T^{an}x) f_2'(T^{bn}x) \prod_{j=1}^{k} g_j \circ S^{jn}\\
&+ \frac{1}{N}\sum_{n=1}^{N} f_1'(T^{an}x)\mathbb{E}(f_2|\mathcal{Z}_{k+1})(T^{bn}x) \prod_{j=1}^{k} g_j \circ S^{jn}\\
&+ \frac{1}{N}\sum_{n=1}^{N} f_1'(T^{an}x) f_2'(T^{bn}x) \prod_{j=1}^{k} g_j \circ S^{jn}.
\end{aligned}$$

We know that, by Theorem 4.1.2(a), there exists a universal set of full measure $\tilde{X}_k$ such that for all $x \in \tilde{X}_k$, the last three averages of the right hand side of (4.2) converge to 0 in $L^2(\nu)$. And by

Theorem 4.1.2(b-c), the first averages also converge in $L^2(\nu)$ for all $x \in \hat{X}_k \cap X'_k$. So if we set

$$X_{f_1,f_2,k} = \tilde{X}_k \cap \hat{X}_k \cap X'_k,$$

then $X_{f_1,f_2,k}$ is a set full measure that only depends on f_1, f_2, the transformation T, and the positive integer k, since it is a finite intersection of the sets of full measure, each only depending on the functions f_1, f_2, and the transformation T. Thus, for any $x \in X_{f_1,f_2,k}$, $a, b \in \mathbb{Z}$, and any other ergodic system $(Y, \mathcal{G}, \nu, S)$ with functions $g_1, \ldots, g_k \in L^\infty(\nu)$, the averages in (4.1) converge in $L^2(\nu)$. This implies that the set

$$X_{f_1,f_2} = \bigcap_{k=1}^{\infty} X_{f_1,f_2,k}$$

is a set of full measure that only depends on the functions f_1, f_2, and the transformation T, and this is indeed the desired universal set for Theorem 4.1.1. □

4.2 The case where either f_1 or f_2 belongs to $\mathcal{Z}_{k+1}(T)^\perp$ (Proof of (a) of Theorem 4.1.2)

For pedagogical purposes, we will first prove the statement for the case $k = 2$. We first identify the set of full measure for which the averages in (4.1) converges to 0; the fact that this is indeed a set of full measure can be shown by using Fatou's lemma and (2.8).

The key observation of the proof is the fact that S is a measure-preserving transformation allows us to bound the $L^2(\nu)$-norm of the averages by the double recurrence Wiener-Wintner averages; to do so, we apply van der Corput's lemma, Hölder's inequality, and the spectral theorem. This allows us to show that the averages in (4.1) indeed converge to 0 when $k = 2$ for this set of full measure.

Then we will proceed for the case $k = 3$ to demonstrate that the claim can be proven inductively for the case $k > 2$. Again we start by identifying the set of full measure. To show that the averages converge to 0 on this set, we rely on the result obtained for the case $k = 2$.

Before we prove this part of the theorem, we will prove this for the case where $k = 2, 3$ to demonstrate the inductive step for simple cases. For the case $k = 2$, we would like to show that

$$\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) g_1 \circ S^n g_2 \circ S^{2n} \right\|_{L^2(\nu)}^2 = 0. \tag{4.3}$$

Consider a set

$$\tilde{X}_2 = \Bigg\{ x \in X : \liminf_{H\to\infty} \Bigg(\frac{1}{H} \sum_{h=1}^{H} \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N-h} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x) e^{2\pi i n t} \right|^2 \Bigg)^{1/2} = 0 \Bigg\}. \tag{4.4}$$

First we show that $\tilde{X}_2$ is a set of full measure. To do so, we apply Fatou's lemma and the inequality (2.8) to obtain

$$\begin{aligned}
&\int \liminf_{H\to\infty} \left(\frac{1}{H} \sum_{h=1}^{H} \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N-h} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x) e^{2\pi i n t} \right|^2 \right)^{1/2} d\mu \\
&\le \liminf_{H\to\infty} \left(\frac{1}{H} \sum_{h=1}^{H} \int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N-h} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x) e^{2\pi i n t} \right|^2 d\mu \right)^{1/2} \\
&\lesssim_{a,b} \min_{i=1,2} \liminf_{H\to\infty} \left(\frac{1}{H} \sum_{h=1}^{H} \left|\left|\left| f_i \cdot f_i \circ T^h \right|\right|\right|_3^{1/2} \right)^{1/2} \le \min_{i=1,2} \left(\liminf_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \left|\left|\left| f_i \cdot f_i \circ T^h \right|\right|\right|_3^{8} \right)^{1/32} \\
&= \min_{i=1,2} |||f_i|||_4^{1/2}.
\end{aligned}$$

Since either f_1 or f_2 belongs to $\mathcal{Z}_3(T)^{\perp}$, either $|||f_1|||_4$ or $|||f_2|||_4$ equals zero. This shows that $\mu(\tilde{X}_2) = 1$.

Now we claim (4.3) holds for all $x \in \tilde{X}_2$. In fact, we show that for any $1 \le H < N$, we have

$$\begin{aligned}
&\left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) g_1 \circ S^n g_2 \circ S^{2n} \right\|_{L^2(\nu)}^2 \\
&\lesssim_{a,b} \frac{1}{H} + \left(\frac{1}{H} \sum_{h=1}^{H} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N-h} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x) e^{2\pi i n t} \right|^2 \right)^{1/2}.
\end{aligned} \tag{4.5}$$

To do so, we proceed with van der Corput's lemma; using the fact that S is a measure preserving transformation, we obtain

$$\left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) g_1 \circ S^n g_2 \circ S^{2n} \right\|_{L^2(\nu)}^2$$

$$\leq \frac{2}{H} + \frac{4}{H}\sum_{h=1}^{H}\left|\int \frac{1}{N}\sum_{n=1}^{N-h} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x)\right.$$

$$\left. \times (g_1 \cdot g_1 \circ S^h) \circ S^n (g_2 \cdot g_2 \circ S^{2h}) \circ S^{2n} d\nu \right|$$

$$= \frac{2}{H} + \frac{4}{H}\sum_{h=1}^{H}\left|\int \frac{1}{N}\sum_{n=1}^{N-h} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x)(g_1 \cdot g_1 \circ S^h)(g_2 \cdot g_2 \circ S^{2h}) \circ S^n d\nu\right|$$

$$\leq \frac{2}{H} + \frac{4}{H}\sum_{h=1}^{H}\left\|\frac{1}{N}\sum_{n=1}^{N} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x)(g_2 \cdot g_2 \circ S^{2h}) \circ S^n\right\|_{L^2(\nu)}$$

(by Hölder's inequality)

$$\leq \frac{2}{H} + \left(\frac{16}{H}\sum_{h=1}^{H}\left\|\frac{1}{N}\sum_{n=1}^{N} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x)(g_2 \cdot g_2 \circ S^{2h}) \circ S^n\right\|_{L^2(\nu)}^2\right)^{1/2},$$

where the last inequality follows from the Cauchy-Schwarz inequality. We apply the spectral theorem to the square of the $L^2(\nu)$-norm in the last line to obtain

$$\left\|\frac{1}{N}\sum_{n=1}^{N} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x)(g_2 \cdot g_2 \circ S^{2h})\right\|_{L^2(\nu)}^2$$

$$= \int \left|\frac{1}{N}\sum_{n=1}^{N} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x) e^{2\pi int}\right|^2 d\sigma_{g_2 \cdot g_2 \circ S^{2h}}(t)$$

$$\leq \sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=1}^{N} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x) e^{2\pi int}\right|^2,$$

which tells us that (4.5) holds. Thus, if $x \in \tilde{X}^2$, and we let $N \to \infty$ (and consequently $H \to \infty$) in (4.5), we obtain

$$\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) g_1 \circ S^n g_2 \circ S^{2n}\right\|_{L^2(\nu)}^2$$

$$\lesssim_{a,b} \left(\liminf_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}\limsup_{N\to\infty}\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=1}^{N-h} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh}(T^{bn}x) e^{2\pi int}\right|^2\right)^{1/2} = 0.$$

This proves the case for $k = 2$. Now we show that the holds for the case $k = 3$ using the fact that the convergence to 0 holds for $k=2$. We let $F_{1,h_1} = f_1 \cdot f_1 \circ T^{ah_1}$ and $F_{2,h_1} = f_2 \cdot f_2 \circ T^{bh_1}$. Then we set

$$\tilde{X}_3 = \Bigg\{ x \in X : \liminf_{H_1\to\infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,h_1} \cdot F_{1,h_1} \circ T^{ah_2}(T^{an}x) F_{2,h_1} \cdot F_{2,h_1} \circ T^{bh_2}(T^{bn}x) e^{2\pi i n t} \right|^2 \Bigg)^{1/4} = 0 \Bigg\}.$$

We first show that $\tilde{X}_3$ is a set of full measure. To see that, we apply Fatou's lemma twice to interchange the integral and the $\liminf$'s, Hölder's inequality, the inequality (2.8), and the Cauchy-Schwarz inequality multiple times to obtain

$$\begin{aligned}
&\int \liminf_{H_1\to\infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,h_1} \cdot F_{1,h_1} \circ T^{ah_2}(T^{an}x) F_{2,h_1} \cdot F_{2,h_1} \circ T^{bh_2}(T^{bn}x) e^{2\pi i n t} \right|^2 d\mu(x) \Bigg)^{1/4} \\
&\le \liminf_{H_1\to\infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,h_1} \cdot F_{1,h_1} \circ T^{ah_2}(T^{an}x) F_{2,h_1} \cdot F_{2,h_1} \circ T^{bh_2}(T^{bn}x) e^{2\pi i n t} \right|^2 d\mu(x) \Bigg)^{1/4} \\
&\lesssim_{a_1,a_2} \liminf_{H_1\to\infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \min_{i=1,2} \left|\left|\left| F_{i,h_1} \cdot F_{i,h_1} \circ T^{a_i h_2} \right|\right|\right|_3^{1/2} \Bigg)^{1/4} \qquad (\text{where } a_1 = a,\ a_2 = b) \\
&\le \liminf_{H_1\to\infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \Bigg(\liminf_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \min_{i=1,2} \left|\left|\left| F_{i,h_1} \cdot F_{i,h_1} \circ T^{a_i h_2} \right|\right|\right|_3^{8} \Bigg)^{1/16} \Bigg)^{1/4} \\
&\lesssim_{a_1,a_2} \liminf_{H_1\to\infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \min_{i=1,2} \left|\left|\left| f_i \cdot f_i \circ T^{a_i h_1} \right|\right|\right|_4 \Bigg)^{1/4} \\
&\le \liminf_{H_1\to\infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \min_{i=1,2} \left|\left|\left| f_i \cdot f_i \circ T^{a_i h_1} \right|\right|\right|_4^{16} \Bigg)^{1/64} \lesssim_{a_1,a_2} \min_{i=1,2} |||f_i|||_5^{1/2},
\end{aligned}$$

and since either f_1 or f_2 belongs to $\mathcal{Z}_4(T)^\perp$, either $|||f_1|||_5$ or $|||f_2|||_5$ equals zero. Hence, we know that $\tilde{X}_3$ is a set of full measure.

Now we will show that the averages converge to 0 when $x \in \tilde{X}_3$. To do so, we wish to show that

$$\left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) \prod_{j=1}^{3} g_j \circ S^{jn} \right\|_{L^2(\nu)}^2$$
$$\lesssim_{a,b} \frac{1}{H_1} + \left(\frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \left(\frac{2}{H_2} + \right.\right.$$
$$\left.\left.\left(\frac{16}{H_2} \sum_{h_2=0}^{H_2-1} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,h_1} \cdot F_{1,h_1} \circ T^{ah_2}(T^{an}x) F_{2,h_1} \cdot F_{2,h_1} \circ T^{bh_2}(T^{bn}x) e^{2\pi int} \right|^2 \right)\right)\right)^{1/4}$$

Indeed, we apply van der Corput's lemma and the Cauchy-Schwarz inequality to show that

$$\left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) \prod_{j=1}^{3} g_j \circ S^{jn} \right\|_{L^2(\nu)}^2$$
$$\leq \frac{2}{H_1} + \frac{4}{H_1} \sum_{h=1}^{H} \left| \int \frac{1}{N} \sum_{n=1}^{N-h_1} f_1 \cdot f_1 \circ T^{ah_1}(T^{an}x) f_2 \cdot f_2 \circ T^{bh_1}(T^{bn}x) \prod_{j=1}^{3} (g_j \cdot g_j \circ S^{jh_1}) \circ S^{jn} d\nu \right|$$
$$= \frac{2}{H_1} + \frac{4}{H_1} \sum_{h_1=1}^{H_1} \left| \int \frac{1}{N} \sum_{n=1}^{N-h_1} f_1 \cdot f_1 \circ T^{ah}(T^{an}x) f_2 \cdot f_2 \circ T^{bh_1}(T^{bn}x) \prod_{j=1}^{3} (g_j \cdot g_j \circ S^{jh_1}) \circ S^{(j-1)n} d\nu \right|$$
$$\leq \frac{2}{H_1} + \left(\frac{16}{H_1} \sum_{h_1=1}^{H_1} \right.$$
$$\left. \left\| \frac{1}{N} \sum_{n=1}^{N-h_1} f_1 \cdot f_1 \circ T^{ah_1}(T^{an}x) f_2 \cdot f_2 \circ T^{bh_1}(T^{bn}x) \prod_{j=2}^{3} (g_j \cdot g_j \circ S^{jh_1}) \circ S^{(j-1)n} \right\|_{L^2(\nu)}^2 \right)^{1/2}$$

We can now apply the inequality (4.5) and the Cauchy-Schwarz inequality to obtain

$$\left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{an}x) f_2(T^{bn}x) \prod_{j=1}^{3} g_j \circ S^{jn} \right\|_{L^2(\nu)}^2$$
$$\lesssim_{a,b} \frac{1}{H_1} + \left(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \left(\frac{1}{H_2} + \right.\right.$$
$$\left.\left.\left(\frac{1}{H_2} \sum_{h_2=1}^{H_2} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N-h_1-h_2} F_{1,h_1} \cdot F_{1,h_1} \circ T^{ah_2}(T^{an}x) F_{2,h_1} \cdot F_{2,h_1} \circ T^{bh_2}(T^{bn}x) e^{2\pi int} \right|^2 \right)\right)\right)^{1/4}$$

Therefore, we have shown that the averages converge to 0 in $L^2(\nu)$ when $x \in \tilde{X}_3$.

One of the key observations in showing that the case $k = 2$ implies the case $k = 3$ was the use of the inequality (4.5). We will show that this can be done for $k \geq 4$. For the following lemma, we will use the following notations for our convenience: We shall denote $a_1 = a$ and $a_2 = b$. Let $\vec{h}(l) = (h_1, h_2, \ldots, h_l) \in \mathbb{N}^l$. With this notation, we define the following functions recursively:

$$
\begin{aligned}
&F_{1,\vec{h}(1)} = f_1 \cdot f_1 \circ T^{a_1 h_1}, && F_{2,\vec{h}(1)} = f_2 \cdot f_2 \circ T^{a_2 h_1},\\
&F_{1,\vec{h}(2)} = F_{1,\vec{h}(1)} \cdot F_{1,\vec{h}(1)} \circ T^{a_1 h_2}, && F_{2,\vec{h}(2)} = F_{2,\vec{h}(1)} \cdot F_{2,\vec{h}(1)} \circ T^{a_2 h_2},\\
&\cdots, && \cdots,\\
&F_{1,\vec{h}(k-1)} = F_{1,\vec{h}(k-2)} \cdot F_{1,\vec{h}(k-2)} \circ T^{a_1 h_{k-1}}, && F_{2,\vec{h}(k-1)} = F_{2,\vec{h}(k-2)} \cdot F_{2,\vec{h}(k-2)} \circ T^{a_2 h_{k-1}}.
\end{aligned}
$$

Lemma 4.2.1. *Let everything as in (a) of Theorem* 4.1.2. *Then for each positive integer $k \geq 2$, we have*

$$
\begin{aligned}
&\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{a_1 n}x) f_2(T^{a_2 n}x) \prod_{i=1}^{k} g_i \circ S^{in} \right\|_{L^2(\nu)}^2 \qquad (4.6)\\
&\lesssim_{a_1,a_2} \liminf_{H_1\to\infty} \left(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \cdots \right.\\
&\left. \liminf_{H_{k-1}\to\infty} \frac{1}{H_{k-1}} \sum_{h_{k-1}=1}^{H_{k-1}} \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,\vec{h}(1)}(T^{a_1 n}x) F_{2,\vec{h}(1)}(T^{a_2 n}x) e^{2\pi i n t} \right|^2 \right)^{2^{-(k-1)}}
\end{aligned}
$$

Proof. We will show this by using induction. The base case $k = 2$ has been treated by the estimate (4.5) after we let $N \to \infty$ and $H \to \infty$. Now suppose the estimate holds when we have $k - 1$ terms. By applying the van der Corput's lemma and the Cauchy-Schwarz inequality, the left hand side of the estimate (4.6) is bounded above by the universal constant depending on a_1 and a_2 times

$$
\begin{aligned}
&\liminf_{H_1\to\infty} \left(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \right.\\
&\left. \limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} F_{1,\vec{h}(k-1)}(T^{a_1 n}x) F_{2,\vec{h}(k-1)}(T^{a_2 n}x) \prod_{i=2}^{k} (g_i \cdot g_i \circ S^{i h_1}) \circ S^{(i-1)n} \right\|_{L^2(\nu)}^2 \right)^{1/2},
\end{aligned}
$$

and we can apply the inductive hypothesis on this lim sup of the square of the L^2-norm and the

Cauchy-Schwarz inequality to obtain the desired estimate. □

Proof of Theorem 4.1.2(a). The set $\tilde{X}_1$ can be obtained from the double recurrence Wiener-Wintner result [16] by applying the spectral theorem. For $k \geq 2$, we consider a set

$$\tilde{X}_k = \Bigg\{ x \in X : \liminf_{H_1 \to \infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2 \to \infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \cdots$$
$$\liminf_{H_k \to \infty} \frac{1}{H_{k-1}} \sum_{h_k=1}^{H_k} \limsup_{N \to \infty} \sup_{t \in \mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,\vec{h}(k-1)}(T^{a_1 n}x) F_{2,\vec{h}(k-1)}(T^{a_2 n}x) e^{2\pi i n t} \right|^2 \Bigg)^{2^{-(k-1)}} = 0 \Bigg\}.$$

We will show that this set is the desired set of full measure. To show that $\mu(\tilde{X}_k) = 1$, we will show that the integral

$$\int \liminf_{H_1 \to \infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2 \to \infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \cdots \tag{4.7}$$
$$\liminf_{H_{k-1} \to \infty} \frac{1}{H_{k-1}} \sum_{h_{k-1}=1}^{H_{k-1}} \limsup_{N \to \infty} \sup_{t \in \mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,\vec{h}(k-1)}(T^{a_1 n}x) F_{2,\vec{h}(k-1)}(T^{a_2 n}x) e^{2\pi i n t} \right|^2 \Bigg)^{2^{-(k-1)}} d\mu = 0,$$

which would show that the averages inside the integral equals zero for μ-a.e. $x \in X$ since the averages are nonnegative. To do so, we apply Fatou's lemma and Hölder's inequality to show that the integral above is bounded above by

$$\liminf_{H_1 \to \infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2 \to \infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \cdots$$
$$\liminf_{H_{k-1} \to \infty} \frac{1}{H_{k-1}} \sum_{h_{k-1}=1}^{H_{k-1}} \int \limsup_{N \to \infty} \sup_{t \in \mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,\vec{h}(k-1)}(T^{a_1 n}x) F_{2,\vec{h}(k-1)}(T^{a_2 n}x) e^{2\pi i n t} \right|^2 d\mu \Bigg)^{2^{-(k-1)}}.$$

Note that the integral above is bounded above by $\min_{i=1,2} \left|\left|\left| F_{i,\vec{h}(k-1)} \right|\right|\right|_3^{1/2}$ by the estimate (2.8). By applying $\liminf_{H_j} \to \infty$ for each $i = 1, 2, \ldots, k-1$, we conclude that the integral on the left hand side of (4.7) is bounded above by the minimum of $|||f_1|||_{k+2}^{1/2}$ or $|||f_2|||_{k+2}^{1/2}$. Since either f_1 or f_2 belongs to $\mathcal{Z}_{k+1}(T)^\perp$, we know that either $|||f_1|||_{k+2} = 0$ or $|||f_2|||_{k+2} = 0$. Thus, (4.7) holds, which implies that $\tilde{X}_k$ is indeed a set of full measure.

Now we need to show that if $x \in \tilde{X}_k$, then the averages in (4.1) converge to 0 in $L^2(\nu)$. But this follows immediately from Lemma 4.2.1, since if $x \in \tilde{X}_k$, the right hand side of (4.6) is 0. $\square$

4.3 When both f_1 and f_2 are in $\mathcal{Z}_{k+1}(T)$ (Proof of (b) and (c) of Theorem 4.1.2)

4.3.1 When one of the functions $g_1, g_2, \dots g_k$ belongs to $\mathcal{Z}_k(S)^\perp$

We first consider the case where either one of the functions $g_1, \dots, g_k$ belongs to $\mathcal{Z}_k(S)^\perp$. In fact, we will show that the averages can be controlled by a seminorm on $L^\infty(\nu)$ with respect to the transformation S. We remark here that B. Host and B. Kra have obtained an estimate sharper than the one we provide, using the tools of nilsequences (cf. [60, Corollary 7.3]). However, the less sharp estimate that we provide here is sufficient to prove our claim. We will also achieve this estimate without the machinery of nilsequences.

We prove this for the case $(Y, \mathcal{G}, \nu, S)$ is an ergodic system, and the general case holds by applying an ergodic decomposition on (Y, S).

Proposition 4.3.1 (See also: [60, Corollary 7.3])**.** *Let $(Y, \mathcal{G}, \nu, S)$ is an ergodic system, $(a_n)_n \in \ell^\infty$ such that $|a_n| \leq 1$ for each n, and functions $g_1, \dots, g_k \in L^\infty(\nu)$ such that $\|g_i\|_{L^\infty(\nu)} \leq 1$ for each $i = 1, 2, \dots, k$. Then there exists a nonnegative constant C_k (that depends only on k) for which*

$$\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \prod_{i=1}^{k} g_i \circ S^{in} \right\|_{L^2(\nu)}^2 \leq C_k \min_{1\leq i\leq k} i \cdot |||g_i|||_{k+1}^2. \tag{4.8}$$

Proof. We proceed by induction on k. For the case $k = 1$, we apply van der Corput's lemma and the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}
&\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n g_1 \circ S^n \right\|_{L^2(\nu)}^2 \\
&\leq \liminf_{H\to\infty} \left(\frac{2}{H} + \frac{4}{H} \sum_{h=0}^{H-1} \limsup_{N\to\infty} \left| \left(\frac{1}{N} \sum_{n=0}^{N-h} a_n \overline{a}_{n+h} \right) \int g_1 \cdot g_1 \circ S^h d\nu \right| \right) \\
&\leq \liminf_{H\to\infty} \left(\frac{2}{H} + 4 \left(\frac{1}{H} \sum_{h=0}^{H-1} \left| \int g_1 \cdot g_1 \circ S^h d\nu \right|^2 \right)^{1/2} \right) = 4|||g_1|||_2^2,
\end{aligned}$$

which proves the base case after setting $C_1 = 4$.

Now suppose the statement holds for $k = l - 1$; i.e. we assume that for any $(b_n)_n \in \ell^\infty$ and

$G_1, \ldots, G_{l-1} \in L^\infty(\nu)$, we have

$$\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} b_n \prod_{i=1}^{l-1} G_i \circ S^{in} \right\|_{L^2(\nu)}^2 \leq C_{l-1} \min_{1\leq i\leq l} i \cdot |||G_i|||_l^2 \tag{4.9}$$

To prove this for the case $k = l$, we again apply van der Corput's lemma and the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}
&\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \prod_{i=1}^{l} g_i \circ S^{in} \right\|_{L^2(\nu)}^2 \\
&\leq \liminf_{H\to\infty} \frac{4}{H} \sum_{h=0}^{H-1} \limsup_{N\to\infty} \left| \int \frac{1}{N} \sum_{n=0}^{N-1} a_n \overline{a}_{n+h} \prod_{i=1}^{l} \left(g_i \cdot g_i \circ S^{ih} \right) \circ S^{(i-1)n} d\nu \right| \\
&\leq \liminf_{H\to\infty} 4 \left(\frac{1}{H} \sum_{h=0}^{H-1} \limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \overline{a}_{n+h} \prod_{i=2}^{l} \left(g_i \cdot g_i \circ S^{ih} \right) \circ S^{(i-1)n} \right\|_{L^2(\nu)}^2 \right)^{1/2} .
\end{aligned}$$

By setting $b_n = a_n \overline{a}_{n+h}$ and $G_i = g_{i-1} \cdot g_{i-1} \circ S^{(i-1)h}$ for each h, we can apply the inequality (4.9) to show that

$$\begin{aligned}
\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \prod_{i=1}^{l} g_i \circ S^{in} \right\|_{L^2(\nu)}^2 &\leq 4C_{l-1}^{1/2} \min_{2\leq i\leq l} \liminf_{H\to\infty} \left(\frac{1}{H} \sum_{h=0}^{H-1} \left|\left|\left| g_i \cdot g_i \circ S^{ih} \right|\right|\right|_l^2 \right)^{1/2} \\
&\leq 4C_{l-1}^{1/2} \min_{2\leq i\leq l} i \cdot \liminf_{H\to\infty} \left(\frac{1}{H} \sum_{h=0}^{H} \left|\left|\left| g_i \cdot g_i \circ S^{ih} \right|\right|\right|_l^{2^l} \right)^{2^{-l}} \\
&= 4C_{l-1}^{1/2} \min_{2\leq i\leq l} i \cdot |||g_i|||_{l+1}^2 .
\end{aligned}$$

To keep $|||g_1|||_{l+1}^2$, we compute

$$\begin{aligned}
&\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \prod_{i=1}^{l} g_i \circ S^{in} \right\|_{L^2(\nu)}^2 \\
&\leq \liminf_{H\to\infty} \frac{4}{H} \sum_{h=0}^{H-1} \limsup_{N\to\infty} \left| \int \frac{1}{N} \sum_{n=0}^{N-1} a_n \overline{a}_{n+h} \prod_{1\leq i\leq l, i\neq j} \left(g_i \cdot g_i \circ S^{ih} \right) \circ S^{(i-j)n} d\nu \right| \\
&\leq 4C_{l-1}^{1/2} \min_{1\leq i\leq l, i\neq j} i \cdot |||g_i|||_{l+1}^2
\end{aligned}$$

for a fixed j. When these estimates are combined, we have

$$\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \prod_{i=1}^{l} g_i \circ S^{in} \right\|_{L^2(\nu)}^2 \leq 4C_{l-1}^{1/2} \min_{1\leq i\leq l} i \cdot |||g_i|||_{l+1}^2 , \tag{4.10}$$

which completes the proof after we set $C_l = 4C_{l-1}^{1/2}$. □

With the estimate (4.8), Theorem 4.1.2(b) can be proven immediately.

Proof of (b) of Theorem 4.1.2. Set $a_n = f_1(T^{an}x)f_2(T^{bn}x)$ in Proposition 4.3.1. Since $f_1, f_2 \in L^\infty$, there exists a set of full measure $\hat{X}_k$ for which the sequence $a_n \in \ell^\infty$. Because one of the functions $g_1, \ldots, g_k$ belongs to $\mathcal{Z}_k(S)^\perp$, we must have $|||g_i|||_k = 0$ for one of them. Hence, we know from (4.8) that the averages must converge to 0 in $L^2(\nu)$. □

4.3.2 When all of the functions $g_1, \ldots, g_k$ belong to $\mathcal{Z}_k(S)$

Here we use Leibman's pointwise convergence result on nilmanifold (i.e. Theorem 1.3.3) to show that the averages converges if all the functions belong to the appropriate Host-Kra-Ziegler factors.

Proof of (c) of Theorem 4.1.2. With appropriate factors maps, we assume $(X, \mathcal{F}, \mu, T)$ and $(Y, \mathcal{G}, \nu, S)$ to be nilsystems, i.e. $X = G_1/\Gamma_1$, and $Y = G_2/\Gamma_2$, where G_1 is a $(k+1)$-step nilpotent Lie group, G_2 is a k-step nilpotent Lie group , and Γ_1 and Γ_2 are discrete co-compact subgroups of G_1 and G_2, respectively. In this proof, we will assume that $f_1, f_2 \in \mathcal{C}(X)$, and $g_1, \ldots, g_k \in \mathcal{C}(Y)$. By taking the product of X^2 and Y^k, we would have another nilmanifold:

$$X^2 \times Y^k = (G_1/\Gamma_1)^2 \times (G_2/\Gamma_2)^k \cong (G_1^2 \times G_2^k)/(\Gamma_1^2 \times \Gamma_2^k).$$

Let $\tau \in G_1$ such that the action of τ on an element of X is determined to be $\tau \cdot x = Tx$. Similarly, we define $\sigma \in G_2$ so that $\sigma \cdot y = Sy$. We define a polynomial sequence p on $X^2 \times Y^k$ as follows:

$$p(n) = (\tau^{an}, \tau^{bn}, \sigma^n, \sigma^{2n}, \ldots, \sigma^{kn}).$$

Clearly, $p(n) \in G_1^2 \times G_2^k$ for all $n \in \mathbb{Z}$, and it acts on $X^2 \times Y^k$ in a way that

$$p(n) \cdot (x_1, x_2, y_1, \ldots, y_k) = (T^{an}x_1, T^{bn}x_2, S^n y_1, \ldots, S^{kn} y_k).$$

Define a continuous function $F \in \mathcal{C}(X^2 \times Y^k)$ such that

$$F(x_1, x_2, y_1, \ldots, y_k) = f_1(x_1) f_2(x_2) \prod_{j=1}^{k} g_j(y_j).$$

Theorem 1.3.3 tells us that the averages

$$\frac{1}{N}\sum_{n=1}^{N} F(p(n)\cdot(x_1, x_2, y_1, \ldots, y_k)) = \frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x_1) f_2(T^{bn}x_2) \prod_{j=1}^{k} g_j(S^{jn}y_j)$$

converge for all $(x_1, x_2, y_1, \ldots, y_k) \in X^2 \times Y^k$. So in particular, if the averages were taken a point $(x, x, y, \ldots, y) \in X^2 \times Y^k$ for any $x \in X$ and $y \in Y$, the desired convergence result holds.

By a standard approximation argument, we can extend this result for the case $f_1, f_2 \in L^\infty(\mu) \cap \mathcal{Z}_{k+1}(T)$ and $g_1, \ldots, g_k \in L^\infty(\nu) \cap \mathcal{Z}_k(S)$. In this process, we neglect a null-set for which the averages may not converge, which allows us to obtain a set of full measure $X'_k \subset X$ that satisfies (c) of Theorem 4.1.2. □

Remark: I. Assani and the author have extended Theorem 4.1.1 to the case where we have multiple commuting transformations [20]. In particular, we have shown the following:

Theorem 4.3.2 ([20, Theorem 1.3]). *Let $(X, \mathcal{F}, \mu, T)$ be a measure-preserving system, and suppose $f_1, f_2 \in L^\infty(\mu)$. Then there exists a set of full measure X_{f_1,f_2} such that for any $x \in X_{f_1,f_2}$, for any $a, b \in \mathbb{Z}$ and any positive integer $k \geq 1$, for any other measure-preserving system with k commuting transformations $(Y, \mathcal{G}, \nu, S_1, S_2, \ldots S_k)$ for any $k \in \mathbb{N}$, and for any $g_1, g_2, \ldots g_k \in L^\infty(\nu)$, the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) \prod_{i=1}^{k} g_i \circ S_i^n \text{ converge in } L^2(\nu).$$

This is the first weighted ergodic averages result that involve multiple commuting transformations, while extending both of the results of Bourgain [31] and Tao [79]. To prove this result, we used the angle of the Wiener-Wintner theorem, as well as the magic systems that were introduced by Host [58].

Later, Frantzikinakis and Host posted another result in this direction on arXiv [47].

REFERENCES

[1] H. Anzai. Ergodic skew product transformations on the torus. *Osaka Mathematical Journal*, 3(1):83–99, 1951.

[2] I. Assani. Multiple recurrence and almost sure convergence for weakly mixing dynamical systems. *Israel J. Math.*, 103:111–124, 1998.

[3] I. Assani. Multiple return times theorems for weakly mixing systems. *Ann. Inst. Henri Poincaré, Probabilités et Statistiques*, 36:153–165, 2000.

[4] I. Assani. Characteristic factors for some nonconventional ergodic averages. Preprint, 2003.

[5] I. Assani. Wiener-Wintner dynamical systems. *Ergodic Theory Dynam. Systems*, 23(6):1637–1654, 2003.

[6] I. Assani. *Wiener Wintner Ergodic Theorems.* World Science Pub Co Inc, May 2003. ISBN: 981-02-4439-8.

[7] I. Assani. Spectral characterization of Wiener-Wintner dynamical systems. *Ergodic Theory Dynam. Systems*, 24(2):347–365, 2004.

[8] I. Assani. Pointwise convergence of nonconventional averages. *Colloq. Math.*, 102(2):245–262, 2005.

[9] I. Assani. Averages along cubes for not necessarily commuting m.p.t. In *Ergodic theory and related fields*, volume 430 of *Contemp. Math.*, pages 1–19. Amer. Math. Soc., Providence, RI, 2007.

[10] I. Assani. Pointwise convergence of ergodic averages along cubes. *J. Analyse Math.*, 110:241–269, 2010. Preprint announced in 2003, arXiv:math/0305388.

[11] I. Assani. A.e. multiple recurrence for weakly mixing commuting actions. Preprint, arXiv:1312.5270v1, 2013.

[12] I. Assani. Pointwise double recurrence and nilsequences. Preprint. Available from arXiv:1504.05732, 2015.

[13] I. Assani. Pointwise recurrence for commuting measure preserving transformations. Preprint, arXiv:1312.5270v2, 2015.

[14] I. Assani and Z. Buczolich. The (L^1, L^1) bilinear Hardy-Littlewood function and Furstenberg average. *Rev. Mat. Iberoam.*, 26(3):861–890, 2010.

[15] I. Assani, Z. Buczolich, and R. D. Mauldin. An L^1 counting problem in ergodic theory. *J. Anal. Math.*, 95:221–241, 2005.

[16] I. Assani, D. Duncan, and R. Moore. Pointwise characteristic factors for Wiener-Wintner double recurrence theorem. *Ergod. Th. and Dynam. Sys.*, 2015. Available on CJO 2015 doi:10.1017/etds.2014.99.

[17] I. Assani and R. Moore. Extension of Wiener-Wintner double recurrence theorem to polynomials. Preliminary version, arXiv:1408.3064, 2014.

[18] I. Assani and R. Moore. Extension of Wiener-Wintner double recurrence theorem to polynomials. To appear in *Journal d'Analyse Mathematique*, arXiv:1409.0463, 2015.

[19] I. Assani and R. Moore. A good universal weight for nonconventional ergodic averages in norm. *Ergod. Th. and Dynam. Sys.* Available on CJO 2015 doi:10.1017/ etds.2015.76. Also available on arXiv:1503.08863., 2015.

[20] I. Assani and R. Moore. A good universal weight for multiple recurrence averages with commuting transformations in norm. Preprint, available on arXiv:1506.05370, *submitted*, June 2015.

[21] I. Assani and K. Presser. Pointwise charateristic factors for multiple term return times theorem. Preprint, 2003.

[22] I. Assani and K. Presser. Pointwise characteristic factors for the multiterm return times theorem. *Ergod. Th. and Dynam Sys.*, 32:341–360, 2012.

[23] I. Assani and K. Presser. Survey of the return times theorem. In *Ergodic Theory and Dynamical Systems: Proceedings of the Ergodic Theory Workshops at University of North Carolina at Chapel Hill, 2011-2012*, De Gruyter Proc. Math., pages 19–58. De Gruyter, Berlin, 2014.

[24] T. Austin. On the norm convergence of non-conventional ergodic averages. *Ergodic Theory Dynam. Systems*, 30(2):321–338, 2010.

[25] T. Austin. A proof of Walsh's convergence theorem using couplings. To appear, Int. Math. Res. Not. IMRN. Preprint available on arXiv:1310.3219, 2013.

[26] A. Bellow and V. Losert. The weighted pointwise ergodic theorem and the individual ergodic theorem along subsequences. *Trans. Amer. Math. Soc.*, 288(1):307–345, 1985.

[27] V. Bergelson, B. Host, and B. Kra with appendix by I Ruzsa. Multiple recurrence and nilsequences. *Invent. Math.*, 160:261–303, 2005.

[28] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.*, 9(3):725–753, 1996.

[29] V. Bergelson and A. Leibman. A nilpotent Roth theorem. *Invent. Math.*, 147(2):429–470, 2002.

[30] J. Bourgain. Return time sequences of dynamical systems. Unpublished preprint, 1988.

[31] J. Bourgain. Double recurrence and almost sure convergence. *J. reine angew. Math.*, 404:140–161, 1990.

[32] J. Bourgain, H. Furstenberg, Y. Katznelson, and D. Ornstein. Appendix on return-time sequences. *Publ. Math. Inst. Hautes Études Sci*, 69:42–45, 1989.

[33] A. Brunel. *Sur quelques problèmes de la thèorie ergodique ponctuelle.* PhD thesis, 1966.

[34] A. Brunel and M. Keane. Ergodic theorems for operator sequences. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 12:231–240, 1969.

[35] J.-P. Conze and E. Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France*, 112:143–175, 1984.

[36] J.-P. Conze and E. Lesigne. Sur un théorème ergodique pour des mesures diagonaless. *Publications de l'Institut de Recherche de Mathématiques de Rennes, Probabilitiés*, 1987-1:1–31, 1988.

[37] C. Demeter, M. T. Lacey, T. Tao, and C. Thiele. Breaking the duality in the return times theorem. *Duke Math J.*, 143:281–355, 2008.

[38] Ciprian Demeter, Terence Tao, and Christoph Thiele. Maximal multilinear operators. *Trans. Amer. Math. Soc.*, 360(9):4989–5042, 2008.

[39] J.-M. Derrien and E. Lesigne. Un théorème ergodique polynomial ponctuel pour les endomorphismes exacts et les K-systèmes. *Ann. Inst. H. Poincaré Probab. Statist.*, 32(6):765–778, 1996.

[40] S. Donoso and W. Sun. A pointwise cubic average for two commuting transformations. Preprint, arXiv:1410.4887, to appear Israel J. Math, 2014.

[41] S. Donoso and W. Sun. Pointwise multiple averages for systems with two commuting transformations. Preprint, arXiv:1509.09310, 2015.

[42] D. Duncan. *A Wiener-Wintner Double Recurrence Theorem*. PhD thesis, The University of North Carolina at Chapel Hill, 2001. Advisor: I. Assani.

[43] T. Eisner. Linear sequences on weighted ergodic theorems. *Abstr. Appl. Anal.*, 2013. Art ID. 815726.

[44] T. Eisner and B. Krause. (Uniform) convergence of twisted ergodic averages. To appear in Ergod. Th. and Dynam. Sys. Preprint available from arXiv:1407.4736, 2014.

[45] T. Eisner and P. Zorin-Kranich. Uniformity in the Wiener-Wintner theorem for nilsequences. *Discrete and Continuous Dynamical Systems*, 33(8):3497–3516, 2013.

[46] N. Frantzikinakis. Uniformity in the polynomial Wiener-Wintner theorem. *Ergod. Th. and Dynam. Sys.*, 26(4):1061–1071, 2006.

[47] N. Frantzikinakis and B. Host. Weighted multiple ergodic averages and correlation sequences. Preprint, arXiv:1602.00927, 2016.

[48] N. Frantzikinakis and B. Kra. Convergence of multiple ergodic averages for some commuting transformations. *Ergodic Theory Dynam. Systems*, 25:799–809, 2005.

[49] H. Furstenberg. *Stationary processes and prediction theory.* Annals of Mathematics Studies, No. 44. Princeton University Press, Princeton, N.J., 1960.

[50] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math*, 31:204–256, 1977.

[51] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton University Press, Princeton, N.J., 1981. M. B. Porter Lectures.

[52] H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.*, 34:275–291, 1978.

[53] H. Furstenberg and B. Weiss. A mean ergodic theorem for $(1/N)\sum_{n=1}^{N} f(T^n x)g(T^{n^2} x)$. In *Convergence in Ergodic Theory and Probability*, Ohio State University Mathematical Research Institute Publications. De Gruyter, Berlin, 1996.

[54] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11:465–588, 2001.

[55] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers U^4-norm. *Glasg. Math. J.*, 53(1):1–50, 2011.

[56] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Ann. of Math. (2)*, 176(2):1231–1372, 2012.

[57] G. Hong and M. Sun. Noncommutative multi-parameter Wiener-Wintner type ergodic theorem. Preprint, arXiv:1602.00927, 2016.

[58] B. Host. Ergodic seminorms for commuting transformations and applications. *Studia Math.*, 195(1):31–49, 2009.

[59] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.*, 161:387–488, 2005.

[60] B. Host and B. Kra. Uniformity seminorms on ℓ^∞ and applications. *J. Anal. Math*, 108:219–276, 2009.

[61] B. Host and B. Kra. Average along cubes. In *Modern Dynamical Systems and Applications*, pages 369 – 432. Cambridge University Press, Cambridge, August 2004.

[62] W. Huang, S. Shao, and X. Ye. Pointwise convergence of multiple ergodic averages and strictly ergodic models. Preprint, arXiv1406.5930, 2014.

[63] Y. Katznelson. *An introduction to harmonic analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2004.

[64] U. Krengel. *Ergodic Theorems*. De Gruyter studies in Mathematics, 6. Walter de Gruyter, 1985.

[65] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences*. John Wiley and Sons, 1974.

[66] A. Leibman. Multiple recurrence theorem for measure preserving actions of a nilpotent group. *Geom. Funct. Anal.*, 8(5):853–931, 1998.

[67] A. Leibman. Pointwise convergence of ergodic averages for polynomial actions of $\mathbb{Z}^d$ by translations on a nilmanifold. *Ergodic Theory Dynam. Systems*, 25(1):215–225, 2005.

[68] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergodic Theory Dynam. Systems*, 25(1):201–213, 2005.

[69] A. Leibman. Rational sub-nilmanifolds of a compact nilmanifold. *Ergodic Theory Dynam. Systems*, 26(3):787–798, 2006.

[70] A. Leibman. Orbits on a nilmanifold under the action of a polynomial sequence of translations. *Ergodic Theory Dynam. Systems*, 27(4):1239–1252, 2007.

[71] E. Lesigne. Un théorème de disjonction de systèmes dynamiques et une généralisation du théorème ergodique de Wiener-Wintner. *Ergod. Th. and Dynam. Sys.*, 10:513–521, 1990.

[72] E. Lesigne. Sur une nil-variété, les parties minimales associées à une translation sont uniquement ergodiques. *Ergodic Theory Dynam. Systems*, 11(2):379–391, 1991.

[73] E. Lesigne. Spectre quasi-discret et théorème ergodique de Wiener-Wintner pur les polynômes. *Ergod. Th. and Dynam. Sys.*, 13:767–784, 1993.

[74] A. M. Mesón and F. Vericat. On the convergence of spherical multiergodic averages for Markov action groups. *Appl. Sci.*, 17:62–74, 2015.

[75] M. Queffelec. *Substitution Dynamical Systems–Spectral Analysis*. Springer, 2nd edition, 2010.

[76] D. Rudolph. Eigenfunctions of $T \times S$ and the Conze-Lesigne algebra. In *Ergodic Theory and Harmonic Analysis*. Cambridge University Press, Cambridge, 1993.

[77] D. Rudolph. A joinings proof of Bourgain's return time theorem. *Ergod. Th. and Dynam Sys.*, 14:197–203, 1994.

[78] D. Rudolph. Fully generic sequences and a multiple-term return-times theorem. *Invent. Math.*, 131(1):199–228, 1998.

[79] T. Tao. Norm convergence for multiple ergodic averages for commuting transformations. *Ergod. Th. and Dynam Sys.*, 28:657–688, 2008.

[80] H. Towsner. Convergence of diagonal ergodic averages. *Ergod. Th. and Dynam. Sys.*, 29(4):1309–1326, 2009.

[81] M. Walsh. Norm convergence of nilpotent ergodic averages. *Ann. of Math. (2)*, 175(3):1667–1688, 2012.

[82] Norbert Wiener and Aurel Wintner. Harmonic analysis and ergodic theory. *Amer. J. Math.*, 63:415–426, 1941.

[83] Q. Zhang. On convergence of the averages $\frac{1}{N}\sum_{n=1}^{N} f_1(R^n x) f_2(S^n x) f_3(T^n x)$. *Monatsh. Math.*, 122(3):275–300, 1996.

[84] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.*, 20(1):53–97, 2006.

[85] P. Zorin-Kranich. Cube spaces and the multiple term return times theorem. *Ergod. Th. and Dynam Sys.*, 34(5):1747–1760, 2014.

[86] P. Zorin-Kranich. Norm convergence of multiple ergodic averages on amenable group. In: J. Analyse Math. (2014). To appear. arXiv:1111.7292, 2014.

[87] P. Zorin-Kranich. A double return times theorem. Preprint. Available from arXiv:1506.05748, 2015.

[88] P. Zorin-Kranich. A nilsequence Wiener-Wintner theorem for bilinear ergodic averages. Preprint. Available from arXiv:1504.04647, 2015.