A Hilbert space approach to multiple recurrence in ergodic theory

by

Frederik J.C. Beyers

Submitted in partial fulfillment of the requirements for the degree

Magister Scientiae

in the Department of Mathematics and Applied Mathematics

in the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

December 2004

Dankbetuigings

Hiermee my opregte waardering aan Prof. Anton Ströh vir sy ondersteuning en leiding gedurende hierdie studie.

Gedurende my studies het die volgende persone my denke oor dinamiese stelsels help ontwikkel: Richard de Beer, Dr. Rocco Duvenhage, Wha Suck Lee, Rudolph Müller, Prof. Elemer Rosinger en Gusti van Zyl.

Verder, my innige dank aan my vrou Hanni vir haar ondersteuning asook vir haar hulp met die uitleg.

Contents

Chapter 1

Basics of measure-theoretic dynamical systems

1.1 Introduction

1.1.1 Background

The use of Hilbert space theory became an important tool for ergodic theoreticians ever since John von Neumann proved the fundamental Mean Ergodic Theorem in Hilbert space. Recurrence is one of the corner stones in the study of dynamical systems. In this dissertation we will investigate some extended ideas besides those of the basic, well-known results regarding recurrence. We will look at multiple recurrence, and see that a Hilbert space approach will help to prove certain multiple recurrence results as special cases.

Another very important use of Hilbert space theory became evident only relatively recently, when it was realized that non-commutative dynamical systems become accessible to the ergodic theorist through the important Gelfand-Naimark-Segal (GNS) representation of C^*-algebras as Hilbert spaces. Through this construction we are enabled to invoke the rich catalogue of Hilbert space ergodic results to approach the more general, and usually more involved, non-commutative extensions of classical ergodic-theoretical results.

In order to make this text self-contained, we have included in this text the basic, standard, ergodic-theoretical results. Where possible we also noted the instances where these results have a Hilbert space counterpart. Chapters 1 and 2 are devoted to the introduction of the basic ergodic-theoretical results such as an introduction to the idea of measure-theoretic dynamical systems, citing some basic examples, Poincairé's recurrence, the ergodic theorems of Von Neumann and Birkhoff, ergodicity, mixing and weakly mixing. We also state some of the rudimentary results regarding these ideas, and supply many of the basic tools used in proofs of subsequent theorems.

In Chapter 3 we show how a Hilbert space result, i.e. a variant of a result by Van der Corput for uniformly distributed sequences modulo 1, is used to simplify the proofs of some multiple recurrence problems.

First we use it to simplify and clarify the proof of a multiple recurrence result by Furstenberg, and also to extend that result to a more general case, using the same Van der Corput lemma. This may be considered the main result of this thesis, since it supplies an original proof of this result. The Van der Corput lemma helps to simplify many of the tedious terms that are found in Furstenberg's proof.

In Chapter 4 we list and discuss a few important results where classical (commutative) ergodic results were extended to the non-commutative case. As stated before, these extensions are mainly due to the accessibility of Hilbert space theory through the GNS construction. The main result in this section is a result proved by Niculescu, Ströh and Zsidó, which is proved here using a similar Van der Corput lemma as in the commutative case. Although we prove a special case of the theorem by Niculescu, Ströh and Zsidó, the same method (Van der Corput) can be used to prove the generalized result.

1.1.2 Measure-theoretic dynamical systems

The main ideas for the discussion in this section were obtained from [3], p 1187 and [6], p 657.

Ergodic theory can be described as the study of statistical properties of *measure-preserving dynamical systems*.

We say that $(X, \mathcal{B}, \mu, T)$ is a measure-preserving dynamical system, if $(X, \mathcal{B}, \mu)$ is a complete probability space and $T : X \to X$ is a *measure-preserving transformation (m.p.t)* in the sense that T satisfies the following conditions:

(a) T is bijective

(b) $TA, T^{-1}A \in \mathcal{B}$ for all $A \in \mathcal{B}$ and

(c) $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}$.

In some instances the assumption that T is bijective is not always included. It is, however, convenient to assume that T is bijective, but it should be noted that many ergodic results also hold under weaker assumptions.

A fundamental question in statistical mechanics concerns the existence of certain types of time averages. The problem may be formulated as follows: The momentary state of a mechanical system is described by specifying a point in a "phase space" X. When the mechanical system is assumed to be governed by the classical Hamiltonian equations, it is subject to a principle of scientific determinism whereby it is known that an initial state x will, after t seconds have elapsed, have passed into a uniquely determined new state y. Since y is uniquely determined by x and t, a function $T : X \to X$ is defined by the equation $y = T_t(x)$. The flow T_t is assumed to have the property that

$$T_t(T_s(x)) = T_{t+s}(x)$$

for all points x in phase space and for all times s and t. This identity holds for certain mechanical systems and in particular if the Hamiltonian function is independent of time.

Any numerical quantity determined by the momentary state of the mechanical system will be given as a real function f defined on X. If the initial state of the system is specified by the point x in X, the value of the quantity f at a time t will be $f(T_t(x))$. For practical purposes, we are in most cases unable to observe a state directly, but rather an average value of $f(T_t(x))$ i.e

$$\frac{1}{N}\int_0^N f(T_t(x))dt$$

computed over a time interval $0 \leq t \leq N$.

If observations regarding micro-processes are made, like gas in a vessel, we usually find that the quantity N, which is determined by the inertial character of "macroscopic" observational instruments we use, is very large compared to the natural rate of evolution of the mechanical system under consideration. For example, regarding the gas vessel, in each second, the molecules travel thousands of feet and recoil from the wall millions of times. Thus, the time N involved in the experiment is large enough to give a good approximation for the limit

$$\lim_{N\to\infty}\frac{1}{N}\int_0^N f(T_t(x))dt.$$

Thus it is central in ergodic theory to determine whether or not, or under what circumstances the limit above exists.

It should be noted that ergodic theory is not only intended to be used for observation and interpretation at "micro"-level. At macro level, useful averages, or approximate averages can be obtained through scientific study over sufficiently large time intervals, by studying the historical information of a system from the distant past to the present.

Additional to the uses of ergodic theory mentioned above, it also has interesting philosophical implications, such as recurrence, many of which are not understood quite correctly amongst laymen.

Historically, a mechanical system is said to be ergodic if it has the property that the above limit (the time mean) is the constant space mean taken with respect to the ordinary Lebesque measure μ in the phase space X, i.e.

$$\lim_{N\to\infty}\frac{1}{N}\int_0^N f(T_t(x))dt = \frac{\int_X f d\mu}{\mu(X)} = \int_X f d\mu.$$

This tells us that, if the ergodic assumption is accepted, then we can use the averages obtained over sufficiently large time intervals to obtain global information about a state in X. It is hence obvious to note that a mathematical definition of ergodicity will include the idea that the time-orbit of a state x $(T^n x)$ must traverse almost the whole of X eventually, just as the whole space x is traversed during the calculation of the space mean.

For our purposes, it is convenient to focus on the case where the flow T_t is taken over discrete instead of continuous time. This restriction will not impact our study negatively, since we will only be studying the long-term behavior of dynamical systems.

Now, $T_{n+m} = T_n T_m$ and $T_n = T_1^n$, and hence, for a given measure-preserving transformation, the map $(n, x) \mapsto T^n x$ defines a group action of the integers on X, where $n \in \mathbb{Z}$ and $x \in X$. It is indeed of considerable interest to consider more general group or semigroup actions on X of *measure-preserving flows*, that is actions of $\mathbb{R}$ on X, but we will restrict ourselves to the case where the group acting on X is $\mathbb{Z}$. Hence we will be looking at averages of the form

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T_1^n(x)), \quad f \in L_p(X, \mathcal{B}, \mu)$$

where T_1 is a mapping of X into itself, instead of the averages

$$\frac{1}{N}\int_0^N f(T_t(x))dt.$$

Hence we see that in the case where we use discrete time the problem is to determine whether the time mean

$$\lim_{n\to\infty}\sum_{k=0}^{n-1} f(T^k x)$$

exists and is equal to the space mean

$$\int_X f d\mu$$

for almost all $x \in X$ with respect to μ. In other words, the claim is here that the space mean of an observable quantity can be derived almost surely from discrete measurements along the time evolution of a single state $x \in X$.

It has been shown that some stronger assumptions about dynamical systems are also useful. These assumptions are called *strongly mixing* and *weakly mixing*. Both these conditions imply ergodicity. The principal part of this dissertation will involve weakly mixing, and also *multiple weakly mixing*. After the concepts of ergodicity, weakly mixing and strongly mixing have been introduced, we will also be in a position to consider weakly mixing and recurrence of "higher orders" in more detail. We will see that a wide variety of multiple recurrence results can be proved. This adds to the rich store of practical and philosophical implications that ergodic theory provides for the field of dynamical systems.

Besides the question regarding time averages and space averages (which is a quantitative problem) we may investigate the *recurrence* behavior of sets. Poincaré supplied the first recurrence result, and as mentioned earlier, the ideas regarding recurrence provide a rich array of profound philosophical implications for our understanding of dynamical systems. Recurrence and multiple recurrence also provide an important link to the field of combinatorial number theory.

As mentioned in section 1.1.1, many analogues of classical (commutative) ergodic theoretical results exist in a non-commutative setting. This dissertation will be concluded by showing how some recurrence and multiple recurrence results can be extended to more general spaces. Ergodic theory has historically been studied with regard to classical dynamical systems, where the assumption of commutativity of the underlying phase space is plausible. Only fairly recently has it been noted that many ergodic results can be extended to non-commutative spaces such as C^*-algebras. This may have important implications for the field of quantum dynamics.

1.1.3 Examples of measure-preserving transformations, dynamical systems and associated spaces

The following examples were drawn from [11] (examples 1-4,6,7), [21] (examples 5, 8, 11) and [9] (example 10).

Some examples of typical measure spaces we may consider are:

1. Finite-dimensional Euclidian spaces with Borel measurability or Lebesque measure.

2. The unit interval with the same definitions of measurability and measure.

3. The set of all sequences $x = \{x_n\}$ of 0's and 1's, $n \in \mathbb{Z}$; the measurable sets are the elements of the σ-algebra generated by sets of the form $\{x : x_n = 1\}$, and the measure is determined by the condition that its value on the intersection of k generating sets is always $1/2^k$.

4. A simple example of an *invertible measure-preserving transformation* on the real line is defined by $Tx = x + 1$. More generally, in a finite-dimensional Euclidean space, let c be an arbitrary vector and define T by $Tx = x + c$ which is a m.p.t.

   The following example of a measure-preserving dynamical system is the well-known Hamiltonian dynamical system.

5. *Hamiltonian system*

   The state at any time t of a physical system consisting of N particles can be specified by the three coordinates of position and the three of momentum of each particle, that is by a point in $\mathbb{R}^{6N}$, which is the *phase space* of the system. More generally (allowing for changes of variables and constraints on the system), let the state of the system be described by a pair of vectors (q, p), where $p = (p_1, \ldots, p_n)$ (the "generalized momentum") and $q = (q_1, \ldots, q_n)$ (the "generalized position") are in $\mathbb{R}^n$, in which case the phase space is $\mathbb{R}^{2n}$. There is given a $(\mathcal{C}^2)$ *Hamiltonian function* $H(q, p)$, which we assume to be independent of time, and which is typically the sum of the kinetic energy $K(p)$ and potential energy $U(q)$ of the system. *Hamilton's equations* are

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, 2, \ldots, n).$$

These equations determine the state $T_t(q,p)$ at any time t if the system has initial state (q,p), by the theorem on the existence and uniqueness of solutions of first-order ordinary differential equations. We obtain in this way a one-parameter flow $\{T_t : -\infty < t < \infty\}$ on the phase space $\mathbb{R}^{2n}$.

The study of Hamiltonian systems, brings ergodic theory in contact with a wide range of dynamical systems which can be modelled by the Hamiltonian equations. Such systems range from micro level such as gas molecules to macro level such as planets and galaxies.

Naturally, we can construct numerous examples of transformations which are not both measure-preserving and bijective.

6. A typical example of a bijective *but not measure-preserving* transformation on the real line is given by $Tx = 2x$; we have that $\mu(T^{-1}E) = \frac{1}{2}\mu(E)$ for every Borel set E (where μ is the Lebesque measure).

   We can define a closely related transformation and show that it is measure-preserving. Let $Tx = 2x$ $(mod\ 1)$, i.e. consider the half-open interval $[0,1)$ and write $Tx = 2x$ when $0 \leq x < 1/2$ and $Tx = 2x - 1$ when $1/2 \leq x < 1$. For example, if

$$E = \left[\frac{2}{8}, \frac{5}{8}\right),$$

   then

$$T^{-1}E = \left[\frac{2}{16}, \frac{5}{16}\right) \cup \left[\frac{1}{2}(\frac{2}{8}+1), \frac{1}{2}(\frac{5}{8}+1)\right),$$

   and we calculate that

$$\mu(T^{-1}E) = \frac{3}{16} + \frac{3}{16} = \frac{3}{8} = \mu(E).$$

   Similarly it follows that

$$\mu(T^{-1}E) = \mu(E)$$

   whenever E is a half-open interval with rational endpoints, and hence it follows that T is measure preserving. Since T is not one-to-one, and since it cannot be made one-to-one by any alteration of a set of measure zero, we have obtained an example of a transformation that is *measure-preserving, but not bijective*. In this instance we succeeded to make the transformation measure-preserving, but at the price of bijectivity.

7. An elementary example of a bijective measure preserving transformation is given by the transformation defined on $\mathbb{R}^2$ by $T(x,y) = \left(2x, \frac{1}{2}y\right)$. The inverse image of the unit square is a rectangle with base $\frac{1}{2}$ and altitude 2. Since, similarly, the inverse image of every rectangle is a rectangle of the same area, it follows that T is a measure-preserving transformation; clearly T is invertible.

8. *Stationary stochastic processes*

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\dots f_{-1}, f_0, f_1, f_2, \dots$ a sequence of measurable functions on Ω. Suppose that the sequence is *stationary,* in that for any $n_1, n_2, \dots n_r$, any Borel subsets $B_1, B_2, \dots B_r$ of $\mathbb{R}$, and any $k \in \mathbb{Z}$,

$$P\{\omega : f_{n_1}(\omega) \in B_1, \dots f_{n_r}(\omega) \in B_r\}$$
$$= P\{\omega : f_{n_1+k}(\omega) \in B_1, \dots, f_{n_r+k}(\omega) \in B_r\}.$$

Such a stationary process corresponds to a measure-preserving system in a standard way.

Let $\mathbb{R}^{\mathbb{Z}} = \{(\dots x_{-1}, x_0, x_1, \dots) : \text{each } x_i \in \mathbb{R}\}$, define

$$\phi : \Omega \to \mathbb{R}^{\mathbb{Z}} \text{ by } (\phi\omega)_n = f_n(\omega)$$

for all $n \in \mathbb{Z}$, and define μ on the Borel subsets of $\mathbb{R}^{\mathbb{Z}}$ by

$$\mu(E) = P(\phi^{-1}E).$$

Extend μ to the completion $\mathcal{B}$ of the Borel field. Let $\sigma : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$ be the *shift transformation* defined by

$$(\sigma x)_n = x_{n+1}.$$

Because of the stationarity of $\{f_n\}$, μ is shift-invariant on cylinder sets and hence on all of $\mathcal{B}$, so that we have constructed a measure-preserving system $(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}, \mu, \sigma)$.

9. *The Baker's transformation*

Let X be the space of the sequences $x = \{x_n\}$, $n = 0, \pm 1, \pm 2, \cdots$, described in (3); let T be the transformation induced by a unit shift on the indices, i.e.

$$Tx = y = \{y_n\}, \text{ where } y_n = x_{n+1}.$$

The mapping is measure-preserving and bijective.

If the elements of X are restricted to sequences $\{x_n\}$ (with $n = 0, 1, 2, \cdots$ i.e. X is the unilateral sequence space), the same equation defines a measure-preserving but non-invertible (two-to-one) transformation.

There is a mapping S from the unilateral sequence space to the unit interval; S sends the sequence $\{x_n\}$ of 0's and 1's onto the binary expansion $.x_1 x_2 \cdots$. The transformation S is measure-preserving and can be made one-to-one by suitably defining S. Since the set of sequences whose image is rational and the set of rational numbers are both countably infinite, we can suitably define S on these sequences, whereby we obtain an invertible measure-preserving transformation from the sequences onto the unit interval. The existence of such a transformation shows that the measure-theoretic structures of the two spaces are isomorphic. The isomorphism (i.e. the transformation S) carries the unilateral shift T onto an invertible measure-preserving transformation T' on the interval; T' is defined by $T' = STS^{-1}$. An examination of the definitions of S and T shows that T' is: $T'x = 2x \ (mod\ 1)$ (as seen in (6)) almost everywhere.

There is a natural correspondence between the bilateral sequence space and the Cartesian product of the unilateral sequence space with itself; the correspondence sends

$$\{\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots\} \text{ onto } (\{x_0, x_1, x_2, \cdots\}, \{x_{-1}, x_{-2}, \cdots\}).$$

This correspondence is an invertible measure-preserving transformation and therefore, a measure-theoretic isomorphism. If we denote this isomorphism by P and if we denote by Q the Cartesian product $S \times S$ (so that $Q(x,y) = (Sx, Sy)$ whenever x and y are unilateral sequences), then the composite transformation QP is an isomorphism from the bilateral sequence space onto the unit square. This isomorphism carries the bilateral shift onto an invertible measure-preserving transformation T'' on the square. An examination of the definition shows that T'' is close to the transformation in (8), given by $T''(x,y) = \left(2x, \frac{1}{2}y\right)$ when $0 \leq x < \frac{1}{2}$ and $T''(x,y) = \left(2x, \frac{1}{2}(y+1)\right)$ when $\frac{1}{2} \leq x < 1$. (These equations, valid almost everywhere, must be taken modulo 1).

The transformation T'' can be described geometrically as follows. Transform the unit square by the linear transformation that sends (x,y) onto $\left(2x, \frac{1}{2}y\right)$, getting a rectangle whose bottom edge is $[0,2)$ and whose left edge is $[0,1/2)$; cut of the right half of this rectangle (with bottom edge $[1,2)$) and move it, by translation, to the top half of the unit square.

These actions remind one of the action of kneading dough, and hence the transformation T'' is often called the baker's transformation.

10. *Bernoulli shifts*

A Bernoulli system is the dynamical system that corresponds to a stochastic process of infinitely many independent, identically distributed (i.i.d.), Bernoulli trials. More precisely, a Bernoulli system consists of a space Ω which is the space of all sequences $\{\omega_n\}_{n\in\mathbb{Z}}$ with values in a finite set, say, $\Gamma = \{0, 1, \ldots r\}$. We assign to each ω_n a weight p_n such that all $p_n > 0$ and $\sum_{n=0}^{r} p_n = 1$. A σ-algebra of sets $\mathcal{B}$ is obtained in Ω by letting $\mathcal{B}$ be the smallest σ-algebra for which $\omega \to \omega_n$ is measurable:

$$\mu\{\omega : \omega_{i_1} = j_1, \omega_{i_2} = j_2, \ldots \omega_{i_k} = j_k\} = p_{j_1} p_{j_2}, \ldots p_{j_k}.$$

Hence the measure-preserving transformation T in this system is the *shift* $T\{\omega_n\} = \{\omega_{n+1}\}$.

11. *Automorphisms of compact groups*

Let G be a compact group and $T : G \to G$ a continuous automorphism. The uniqueness of normalized Haar measure implies that it is T-invariant.

1.2 Recurrence

In order to discuss the asymptotic properties of a measure-preserving transformation T, i.e. the properties of the sequence T^n, the powers of T must make sense; for this reason we shall restrict our attention to transformations from a set X into itself [11].

The earliest and simplest asymptotic questions were raised by Poincairé (*Calcul des probabilités*, 1912), concerning recurrence.

Definition: recurrence

If T is a measurable transformation on X and $B \in \mathcal{B}$, then a point $x \in B$ is said to be *recurrent with respect to* B if there is a $k \geq 1$ for which $T^k x \in B$.

1.2.1 Theorem (Poincaré Recurrence Theorem)

Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ a measure-preserving transformation. For each $B \in \mathcal{B}$, almost every point of B is recurrent with respect to B.

Proof:

Let F be the set of all those points of B which are not recurrent with respect to B; then

$$F = B - \bigcup_{k=1}^{\infty} T^{-k}B = B \cap T^{-1}(X - B) \cap T^{-2}(X - B) \cap \ldots \ .$$

If $x \in F$, then $T^n x \notin F$ for each $n \geq 1$. Thus $F \cap T^{-n}F = \varnothing$ for $n \geq 1$, and hence $T^{-k}F \cap T^{-(n+k)}F = \varnothing$ for each $n \geq 1$ and each $k \geq 0$. Then the sets $F, T^{-1}F, T^{-2}F, \ldots$ are pairwise disjoint and each has measure $\mu(F)$.

Since $\mu(X) = 1 < \infty$, $\mu(F)$ =0. ■

The recurrence theorem above implies a stronger version of itself. Not only is it true that for almost every $x \in B$ at least one term of the sequence $T^n x$ belongs to B; in fact:

1.2.2 Corollary ([11], p 10)

For almost every $x \in B$, there are infinitely many values of n such that $T^n x \in B$.

Proof:

If F_n is the set of points of B that never return to B under the action of T^n, then, by the recurrence theorem, $\mu(F_n) = 0$. If

$$x \in B - (F_1 \cup F_2 \cup \ldots),$$

then $T^n x \in B$ for some positive n, since $x \in B - F_1$. Similarly, since $x \in B - F_n$, it follows that

$$T^{kn} x \in B$$

for some positive k. The stronger version of the recurrence theorem then follows by an inductive repetition of this twice-repeated argument. ■

The conclusion of the recurrence theorem can be formulated in terms of the characteristic function χ_B as follows: for almost every $x \in B$, the series $\sum \chi_B(T^n x)$ diverges. This conclusion can be generalized:

1.2.3 COROLLARY ([11], p 12)

If f is an arbitrary non-negative measurable function, then for almost every $x \in \{x : f(x) > 0\}$, the series

$$\sum f(T^n x)$$

diverges.

Proof:

Consider, for every positive integer k, the set B_k where $f(x) > \frac{1}{k}$. The recurrence theorem implies that for almost every $x \in B_k$ the point $T^k x$ will return to B_k infinitely often; the desired result follows by forming the union of the B_k's. ■

1.3 The fundamental ergodic theorems

1.3.1 The Mean Ergodic Theorem

The main ideas for the discussion that follows, including the introduction and discussion of the unitary operator U and the subsequent example, was drawn from [11].

Poincairé's Recurrence Theorem states that almost every point of each measurable set B returns to B infinitely often under a transformation T, and under appropriate conditions.

The following question may be asked: how long do the recurring points spend in B? This problem can be formulated more precisely as follows: given a point x (in B or not), and given a positive integer n, form the ratio of the number of these points that belong to B to the total number (i.e. to n), and evaluate the limit of these ratios as n tends to infinity, if this limit exist in a meaningful sense.

Hence we should consider the ratio

$$\frac{1}{n}\sum_{k=0}^{n-1} \chi_B(T^k x).$$

This average is called the mean sojourn of x and we are therefore concerned with the problem of convergence of the mean sojourn.

The unitary operator U.

We will not restrict ourselves to characteristic functions. If f is any arbitrary function on X, then another function g on X may be defined by $g(x) = f(Tx)$. If we write $g = Uf$, then U is a mapping that operates on functions. The mapping U has some important properties.

1. The most obvious property of U is its *linearity*, i.e.

$$U(af + bg)(x) = (af + bg)(Tx) = (af)(Tx) + (bg)(Tx)$$
$$= af(Tx) + bg(Tx) = aUf(x) + bUg(x)$$

for any complex-valued functions f and g on X, complex scalars a and b and any $x \in X$.

2. If T is measure-preserving, then U also has the property that it sends $L_1(X, \mathcal{B}, \mu)$ into itself, and moreover, is an isometry on $L_1(X, \mathcal{B}, \mu)$. This implies that if $f \in L_1$, then

$$Uf \in L_1 \text{ and } \|f\|_1 = \|Uf\|_1.$$

To show that U is an isometry, we follow a standard approximation tool. If χ_B is the characteristic function of the set B of finite measure, then $U\chi_B$ is the characteristic function of $T^{-1}B$. Also, $\|\chi_B\|_1 = \mu(B)$. From this and from the linearity of U it follows that U is norm-preserving on finite linear combinations of such characteristic functions, i.e. on simple functions. If f is a non-negative function, then f is the pointwise limit of an increasing sequence f_n of simple functions. Since Uf_n is also an increasing sequence of non-negative functions, it follows from the theorem on integration of monotone sequences that

$$\lim_{n\to\infty} \|Uf_n\|_1 = \|Uf\|_1$$

as well as

$$\lim_{n\to\infty} \|f_n\|_1 = \|f\|_1.$$

This proves the result for non-negative functions. The general case follows from the fact that the norm of every f in L_1 is the same as the norm of $|f|$. (Note that it was not necessary to assume that $\mu(X) < \infty$).

3. The fact that U is an isometry on L_1 implies that U is an isometry on L_2. To see this, note that $\|f\|_2 = \sqrt{\|f^2\|_1}$. If T is a bijective measure-preserving transformation, then U is a bijective isometry, with $U^{-1}f(x) = f(T^{-1}x)$. An invertible isometry on a Hilbert space is a unitary operator ([16] Theorem 3.10-6(f)). This $U : L^2 \to L^2$ is called *the unitary operator induced by* T.

4. An auxiliary fact about isometries may be noted at this stage: if U is an isometry, then $Uf = f$ if and only if $U^*f = f$. To see this, note that $Uf = f \Longrightarrow U^*Uf = U^*f \Longrightarrow f = U^*f$, if we recall for an isometry U, $U^*U = 1$. Conversely, if $U^*f = f$, then

$$\|Uf - f\|^2 = \langle Uf - f, Uf - f\rangle = \|Uf\|^2 - \langle f, Uf\rangle - \langle Uf, f\rangle + \|f\|^2.$$

Since $\langle f, Uf\rangle = \langle U^*f, f\rangle = \|f\|^2$, and $\langle Uf, f\rangle = \langle f, U^*f\rangle = \|f\|^2$.

Hence

$$\|Uf - f\|^2 = 0 \text{ and } Uf = f.$$

Considering the properties of the unitary operator U, we find that one of the basic asymptotic problems of ergodic theory reduces to the limiting behavior of averages

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k$$

where U is an isometry on a Hilbert space.

Example

If the Hilbert space under consideration is one-dimensional, the Mean Ergodic Theorem is quite simple, but still interesting. In this case, the isometry is determined by a complex number u such that $|u| = 1$. Now consider the average

$$\frac{1}{n}\sum_{k=0}^{n-1} u^k.$$

If $u = 1$, then each average is equal to 1. If $u \neq 1$, then

$$\left|\frac{1}{n}\sum_{k=0}^{n-1} u^k\right| = \left|\frac{1 - u^n}{n(1-u)}\right| \leq \frac{2}{n|1-u|} \longrightarrow 0,$$

as $n \to \infty$. Hence the averages converge to 0.

We see that the averages converge to a function p, which can be seen to be a projection on the space of all elements f such that $uf = f$.

If we consider the finite-dimensional case, every isometry is given by a unitary matrix, which, without loss of generality, may be assumed to be a diagonal matrix. Since the diagonal entries of such a matrix U are complex numbers with absolute value 1, it follows that the averages converge to a diagonal matrix with diagonal entries 0's and 1's. The limit matrix, say P, is also a projection in this case, i.e. the projection on the space of all vectors f such that $Uf = f$.

We will now proceed to state and prove the Mean Ergodic Theorem in Hilbert space. We will also state and prove Von Neumann's Mean Ergodic Theorem in L_2.

1.3.1.1 Mean Ergodic Theorem in Hilbert space. ([3], p 1192)

Let H be a Hilbert space, $U : H \mapsto H$ a unitary operator and let $M = \{f \in H : Uf = f\}$. If $P : H \mapsto M$ is the projection of H onto M, then

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k f - Pf \right\| \longrightarrow 0.$$

Proof:

M is a closed linear (vector) subspace of H.

Linearity of M: If $f, g \in M$, and α some constant, then $\alpha f + g = \alpha Uf + Ug = U(\alpha f) + Ug = U(\alpha f + g)$.

M *is closed*: Suppose $f \in \overline{M}$, then there is a sequence $(f_n) \subset M$, $f_n \to f$. Hence the sequence (Uf_n) converges to f. Also, since

$$\|Uf_n - Uf\| = \|U(f_n - f)\| = \|f_n - f\| \to 0, Uf_n \to Uf.$$

Thus $f = Uf$. Hence $\overline{M} \subset M$, so M is closed.

Let $N := \{Uf - f | f \in H\}$. Now N is also a closed linear subspace of H.

Linearity of N: If $f, g \in N$, and α some constant, then

$$\alpha f + g = \alpha(Uh - h) + Ur - r$$

for some $h, r \in H$. We now obtain the following:

$$\alpha(Uh - h) + Ur - r = U(\alpha h) - \alpha h + Ur - r = U(\alpha h + r) - (\alpha h + r),$$

which is an element of N.

N *is closed*: Suppose $f \in \overline{N}$, then there is a sequence $(f_n) \subset M$, $f_n \to f$. Hence there is a sequence

$$(g_n) \subset H \text{ (with } f_n = Ug_n - g_n), Ug_n - g_n \to f.$$

It follows that the sequences (Ug_n) and (g_n) are both convergent in H. Let $Ug_n \to h \in H$ and $g_n \to g \in H$.

Now,

$$\|Ug_n - Ug\| = \|U(g_n - g)\| = \|g_n - g\| \to 0, \; Ug_n \to Ug.$$

Hence $h = Ug$. Therefore $Ug_n - g_n \to Ug - g$, so $f = Ug - g$. Hence $\overline{N} \subset M$, so N is closed.

We now proceed to show that $H = M \oplus N$.

If $g \perp N$ then $\langle Uf, g\rangle = \langle f, g\rangle$ for all $f \in H$ since

$$\langle Uf, g\rangle - \langle f, g\rangle = \langle\, Uf - f, g\rangle = 0.$$

Hence

$$\langle f, U^{-1}g - g\rangle = \langle f, U^{-1}g\rangle - \langle\, f, g\rangle = \langle Uf, g\rangle - \langle f, g\rangle = 0$$

for all $f \in H$. Therefore $U^{-1}g = g$ so $g = Ug$ and $N^{\perp} \subset M$. If $Uf = f$ then

$$\langle f, Ug - g\rangle = \langle f, Ug\rangle - \langle\, f, g\rangle = \langle Uf, Ug\rangle - \langle\, f, g\rangle = \langle f, U^{-1}Ug\rangle - \langle\, f, g\rangle = 0$$

for all $g \in H$. Hence we also have $M \subset N^{\perp}$, so $M = N^{\perp}$. Thus $H = N^{\perp} \oplus N = M \oplus N$. This implies that if $g := f - Pf$ (for any $f \in H$), then $f = Pf + (f - Pf) = Pf + g$ and since $Pf \in M$, $g \in N$ (since $H = M \oplus N$). Note also that any unitary operator U is an isometry, i.e

$$\|Uf\|^2 = \langle Uf, Uf\rangle = \langle f, U^{-1}Uf\rangle = \langle f, f\rangle = \|f\|^2.$$

Therefore, there is an $h \in H$ such that

$$\begin{aligned}
\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k g\right\| &= \left\|\frac{1}{n}\sum_{k=0}^{n-1}(U^k(Uh - h))\right\| \\
&= \left\|\frac{1}{n}\sum_{k=0}^{n-1}(U^{k+1}h - U^k h)\right\| \\
&= \left\|\frac{1}{n}(U^n h - h)\right\| \\
&\leqslant 1/n[\|U^n h\| + \|h\|] \\
&= 2/n\,\|h\| \longrightarrow 0 \quad as\ n \to \infty.
\end{aligned}$$

Now, since

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k g = \frac{1}{n}\sum_{k=0}^{n-1} U^k(f - Pf) = \frac{1}{n}\sum_{k=0}^{n-1}(U^k f - Pf) = \frac{1}{n}\sum_{k=0}^{n-1} U^k f - Pf,$$

it follows immediately that

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k f - Pf\right\| = \left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k g\right\| \longrightarrow 0$$

as $n \to \infty$. ■

From this result follows the Mean Ergodic Theorem of Von Neumann. Recall our definition of the term *measure preserving transformation (m.p.t)* . Let $(X, \mathcal{B}, \mu)$ be a probability space such that $\mathcal{B}$ contains all subsets of sets of measure 0. Let $T : X \mapsto X$ be a bijective $(a.e.)$ map such that T and T^{-1} are both measurable and $T^{-1}\mathcal{B} = T\mathcal{B} = \mathcal{B}$. Assume further that $\mu(T^{-1}E) = \mu(E)$ for all $E \in \mathcal{B}$. Then T is called a *measure preserving transformation (m.p.t.).*

1.3.1.2 Corollary

Let $(X, \mathcal{B}, \mu)$ *be a probability space,* $T : X \mapsto X$ *a measure-preserving transformation, and let* $f \in L^2(X, \mathcal{B}, \mu)$*. Then there is a function* $\overline{f} \in L^2(X, \mathcal{B}, \mu)$ *for which*

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k - \overline{f} \right\|_{L^2(X)} \longrightarrow 0$$

as $n \to \infty$.

Proof:

Let $U : L^2 \to L^2$ be the *unitary operator induced by* T, as discussed above.

Now since $Uf = f \circ T$, $U^k f = f \circ T^k$. Let $H = L^2(X, \mathcal{B}, \mu)$ and let the set M and the projection P be as in Theorem 1.3.1.1. Let $\overline{f} := Pf$ for $f \in H$. Then it follows immediately from Theorem 1.3.1.1 that

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k - \overline{f} \right\|_2 = \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k f - Pf \right\| \longrightarrow 0$$

as $n \to \infty$. ■

As promised earlier, we can now state a recurrence result for *sets*, called Khintchine's theorem. We need a preliminary notion: i.e. *relative density*. We give the formulation supplied in [21]. A set $E \subset \mathbb{Z}$ is called *relatively dense* if there is a positive integer k such that

$$E \cap \{j, j+1, \ldots, j+k-1\} \neq \varnothing \quad \text{for each} \quad j \in \mathbb{Z}.$$

We also say that, in this case, the set E has *bounded gaps*.

First, we will state and prove a generalized Hilbert space version of Khintchine's theorem, and give the conventional theorem thereafter as a direct corollary.

1.3.1.3 Theorem ([7], p 435)

Let H*,* U *and* P *be as in the Mean Ergodic Theorem above. Consider any* $x, y \in H$ *and* $\varepsilon > 0$*. Then the set*

$$E = \left\{ k \in \mathbb{N} : \left| \langle x, U^k y \rangle \right| > |\langle x, Py \rangle| - \varepsilon \right\}$$

is relatively dense in $\mathbb{N}$*.*

Proof:

Given any $\varepsilon > 0$. By the Mean Ergodic Theorem (Theorem 1.3.1.1) there exists an $n \in \mathbb{N}$ such that

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k y - Py \right\| < \frac{\varepsilon}{\|x\| + 1}.$$

Since $UPy = Py$ and $\|U\| \leq 1$, it follows for any $j \in \mathbb{N}$ that

$$\begin{aligned} \left\| \frac{1}{n} \sum_{k=j}^{j+n-1} U^k y - Py \right\| &= \left\| U^j \left(\frac{1}{n} \sum_{k=0}^{n-1} U^k y - Py \right) \right\| \\ &\leq \|U^j\| \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k y - Py \right\| \\ &\leq \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k y - Py \right\| \\ &< \frac{\varepsilon}{\|x\| + 1} \end{aligned}$$

therefore

$$\left| \left\langle x, \frac{1}{n} \sum_{k=j}^{j+n-1} U^k y - Py \right\rangle \right| \leq \|x\| \left\| \frac{1}{n} \sum_{k=j}^{j+n-1} U^k y - Py \right\| < \varepsilon.$$

Hence

$$\left| \frac{1}{n} \sum_{k=j}^{j+n-1} \langle x, U^k y \rangle - \langle x, Py \rangle \right| \leq \left| \left| \frac{1}{n} \sum_{k=j}^{j+n-1} \langle x, U^k y \rangle \right| - |\langle x, Py \rangle| \right| < \varepsilon,$$

and then we obtain that

$$|\langle x, Py \rangle| - \varepsilon < \left| \frac{1}{n} \sum_{k=j}^{j+n-1} \langle x, U^k y \rangle \right| \leq \frac{1}{n} \sum_{k=j}^{j+n-1} \left| \langle x, U^k y \rangle \right|$$

and so $\left| \langle x, U^k y \rangle \right| > |\langle x, Py \rangle| - \varepsilon$ for some $k \in \{j, j+1, ..., j+n-1\}$, i.e. E is relatively dense in $\mathbb{N}$. ■

We see from Khintchine's theorem that for every $k \in E$, the set B contains a set $B \cap T^{-k} B$ of measure larger than $\mu(B)^2 - \varepsilon$ which is mapped back into B by T^k.

1.3.1.4. Khintchine's theorem

For any $B \in \mathcal{B}$ and any $\varepsilon > 0$, the set

$$E = \{k \in \mathbb{N} : \mu(B \cap T^{-k} B) \geq \mu(B)^2 - \varepsilon\}$$

is relatively dense.

Proof:

Let $H = L^2(X, \mathcal{B}, \mu)$ and let $y = x = \chi_B \in L^2(X, \mathcal{B}, \mu)$ in Theorem 1.3.1.3. ■

Hence Khintchine's theorem provides a sharper, quantitative indication of the strength of Poincairé recurrence.

1.3.2 The Pointwise Ergodic Theorem

Another fundamental ergodic-theoretical result is Birkhoff's ergodic theorem, which is also known as the Individual (or Pointwise) Ergodic Theorem.

This theorem is proved using a standard preliminary result, called the Maximal Ergodic Theorem, which will show that the average

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

is convergent almost everywhere. In Theorems 1.3.2.1 - 1.3.2.3 we will follow the formulation given in [21], p 27.

The notation f^* will be used to indicate the following maximum of the above averages, i.e.

$$f^*(x) = \sup_{n\geq 1} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x).$$

1.3.2.1 Maximal Ergodic Theorem

If $f \in L^1(X, \mathcal{B}, \mu)$, *then*

$$\int_{\{f^*>0\}} f d\mu \geqslant 0.$$

Then also

$$\int_{\{f^*\geqslant 0\}} f d\mu \geqslant 0,$$

since for each $\varepsilon > 0$ *we have*

$$0 \leqslant \int_{\{(f+\varepsilon)^*>0\}} (f+\varepsilon) d\mu \leqslant \int_{\{f^*>-\varepsilon\}} f d\mu + \varepsilon,$$

and we can let ε *decrease to* 0.

Proof:

The set $\{x \in X : f^*(x) > 0\}$ is the disjoint union of the sets

$B_1 = \{x : f(x) > 0\}$
$B_2 = \{x : f(x) \leqslant 0, f(x) + f(Tx) > 0\}$
$B_3 = \{x : f(x) \leqslant 0, f(x) + f(Tx) < 0, f(x) + f(Tx) + f(T^2x) > 0\}$
$\vdots$
$B_n = \{x : f(x) \leqslant 0, ..., f(x) + ... + f(T^{n-2}x) \leqslant 0, f(x) + ... + f(T^{n-1}x) > 0\}$
$\vdots$

We will show that

$$\int_{B_1\cup\ldots\cup B_n} f d\mu \geqslant 0$$

for all $n = 1, 2, \ldots$. Since

$$\{x : f^*(x) > 0\} = \bigcup_{n=1}^{\infty} B_n,$$

and

$$\lim_{n\to\infty} \chi_{B_1\cup\ldots\cup B_n} f$$

it will follow by the Dominated Convergence Theorem applied to $\chi_{B_1\cup\ldots\cup B_n} f$ that

$$\begin{aligned}\int_{\{f^*>0\}} f d\mu &= \lim_{n\to\infty} \int_{\{f^*>0\}} \chi_{B_1\cup\ldots\cup B_n} f d\mu \\ &= \lim_{n\to\infty} \int_{B_1\cup\ldots\cup B_n} f d\mu \\ &= \geq 0.\end{aligned}$$

Fix an $n = 1, 2, \ldots$. We break $B_1 \cup \ldots \cup B_n$ into a union of disjoint pieces, over each of which the integral of f is non-negative.

We now make three observations:

1. $T^k B_n \subset B_1 \cup \ldots \cup B_{n-k}$ for $k = 1, 2, \ldots, n-1$.

   This is so because if $x \in B_n$, then

   $$f(x) + f(Tx) + \ldots + f(T^{k-1}x) \leq 0,$$

   while

   $$f(x) + f(Tx) + \ldots + f(T^{k-1}x) + f(T^k x) + \ldots + f(T^{n-1}x) > 0,$$

   and it follows that

   $$f(T^k x) + \ldots + f(T^{n-1}x) > 0,$$

   i.e.,

   $$f(T^k x) + f(T(T^k x)) + \ldots + f(T^{n-k-1}(T^k x)) > 0.$$

   $\therefore T^k x \in B_1 \cup \ldots \cup B_{n-k}$ by the definition of the B_i's.

2. The sets $B_n, TB_n, \ldots, T^{n-1}B_n$ are pairwise disjoint. Indeed, if $T^i B_n \cap T^j B_n \neq \varnothing$ for some $i < j$, then $B_n \cap T^{j-i} B_n \neq \varnothing$, which contradicts the fact that the sets $B_n, n = 1, 2, \ldots$ are pairwise disjoint.

3. If we let

$$\begin{aligned} B'_n &= B_n, C_n = B_n \cup TB_n \cup \ldots \cup T^{n-1}B_n, \\ B'_{n-1} &= B_{n-1} \setminus C_n, C_{n-1} = B'_{n-1} \cup TB'_{n-1} \cup \ldots \cup T^{n-2}B'_{n-1}, \\ &\vdots \\ B'_1 &= B_1 \setminus (C_2 \cup \ldots \cup C_n), C_1 = B'_1, \end{aligned}$$

we find the following: Each $B'_k \subset B_k$, hence the B'_k's are pairwise disjoint. Also, since $B_k, TB_k, \ldots, T^{k-1}B_k$ are pairwise disjoint, $B'_k, TB'_k, \ldots, T^{k-1}B'_k$ are also pairwise disjoint within each C_k. Since the sets B_k are mutually disjoint, and the sets $B'_k \subset B_k$, the sets C_k are mutually disjoint. Finally we also have that $B_1 \cup \ldots B_n = C_1 \cup \ldots C_n$.

Finally we conclude that

$$\begin{aligned} \int_{B_1 \cup \ldots \cup B_n} f d\mu = \sum_{k=1}^{n} \int_{C_k} f d\mu &= \sum_{k=1}^{n} \int_{B'_k \cup TB'_k \cup \ldots \cup T^{k-1}B'_k} f d\mu \\ &= \sum_{k=1}^{n} \sum_{j=0}^{k-1} \int_{T^j B'_k} f d\mu \\ &= \sum_{k=1}^{n} \sum_{j=0}^{k-1} \int_{B'_k} fT^j d\mu \\ &= \sum_{k=1}^{n} \int_{B'_k} (f + fT + \ldots + fT^{k-1}) d\mu \\ &\geq 0, \end{aligned}$$

since $f + fT + \ldots + fT^{k-1} > 0$, by definition, on $B_k \supset B'_k$. ■

1.3.2.2 Corollary

For each $\alpha \in \mathbb{R}$,

$$\int_{f^* > \alpha} f d\mu \geq \alpha \mu\{f^* > \alpha\}.$$

Proof:

Let $g = f - \alpha$, hence $\{f^* > \alpha\} = \{g^* > 0\}$, so that

$$0 \leq \int_{g^* > 0} g d\mu = \int_{f^* > \alpha} (f - \alpha) d\mu,$$

and hence

$$\int_{f^* > \alpha} f d\mu \geq \alpha \mu\{f^* > \alpha\}.$$

■

We now state and prove the Pointwise Ergodic Theorem of G.D. Birkhoff.

1.3.2.3 Theorem

Let $(X, \mathcal{B}, \mu)$ *be a probability space,* $T : X \to X$ *a m.p.t. and* $f \in L^1(X, \mathcal{B}, \mu)$*. Then*

(1) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \bar{f}$ *exists a.e.;*

(2) $\bar{f}(Tx) = \bar{f}(x)$ *a.e.;*

(3) $\bar{f} \in L^1$*, and in fact* $\|\bar{f}\|_1 \leq \|f\|_1$*;*

(4) $\frac{1}{n} \sum_{k=0}^{n-1} fT^k \to \bar{f}$ *in* L^1*;*

(5) *if* $A \in \mathcal{B}$ *with* $T^{-1}A = A$*, then* $\int_A f d\mu = \int_A \bar{f} d\mu$.

Proof of the ergodic theorem:

1. For each $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$, let

$$E_{\alpha,\beta} = \left\{ x \in X : \liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) < \alpha < \beta < \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) \right\}.$$

If we have that $\mu(E_{\alpha,\beta}) = 0$ for each α, β as defined above, then the union over all rational α, β will also have measure 0, and hence the limit exists a.e. In this case we will have that the set in X for which the limit in (1) does not exists, has measure zero.

Now $E_{\alpha,\beta}$ is an invariant subset of $\{f^* > \beta\}$, and then by considering T restricted to $E_{\alpha,\beta}$, we see that

$$\int_{E_{\alpha,\beta}} f d\mu \geq \beta\mu(E_{\alpha,\beta})$$

by the Maximal Ergodic Theorem.

Next we consider $-f$. Since if $x \in E_{\alpha,\beta}$ there is an $n \geq 1$ with

$$\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) < \alpha,$$

we see that

$$E_{\alpha,\beta} \subset \{(-f)^* > -\alpha\}.$$

Then, by the Maximal Ergodic Theorem,

$$\int_{E_{\alpha,\beta}} -f d\mu \geq -\alpha\mu(E_{\alpha,\beta}),$$

i.e.

$$\int_{E_{\alpha,\beta}} f d\mu \leq \alpha\mu(E_{\alpha,\beta}).$$

Therefore

$$\beta\mu(E_{\alpha,\beta}) \leq \int_{E_{\alpha,\beta}} f d\mu \leq \alpha\mu(E_{\alpha,\beta}),$$

and hence $\mu(E_{\alpha,\beta}) = 0$, since $\alpha < \beta$.

2.
$$\begin{aligned}\bar{f}T &= \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(Tx)) &&= \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^{k+1}(x)) \\ &= \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(Tx)) &&= \bar{f} \text{ a.e.}\end{aligned}$$

3. Since
$$\left|\frac{1}{n}\sum_{k=0}^{n-1} fT^k\right| \leq \frac{1}{n}\sum_{k=0}^{n-1} |f|T^k,$$
we have
$$\left|\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} fT^k\right| \leq \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} |f|T^k,$$
hence, $|\bar{f}| \leq |f|^-$ with
$$|f|^- = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} |f|T^k$$
and thus
$$\int |\bar{f}|d\mu \leq \int |f|^- d\mu \leq \liminf_{n\to\infty}\int \frac{1}{n}\sum_{k=0}^{n-1}|f(T^k x)|d\mu = \int |f|d\mu < \infty$$
using Fatou's Lemma.

4. For bounded functions, the L^1 convergence would follow from the Bounded Convergence Theorem. The general case can then be proved by approximating by bounded functions (which are dense in L^1) and using (3). Since we may write $f = f^+ - f^-$ we can assume that $f \geq 0$. If g is bounded and $0 \leq g \leq f$, then
$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} fT^k - \bar{f}\right\|_1 \leq \left\|\frac{1}{n}\sum_{k=0}^{n-1}(fT^k - gT^k)\right\|_1 + \left\|\frac{1}{n}\sum_{k=0}^{n-1} gT^k - \bar{g}\right\|_1 + \|\bar{g} + \bar{f}\|_1.$$

By (3), the third term is less than or equal to $\|g - f\|_1$, which can be made arbitrarily small by appropriate choice of g. Similarly, the first term is also less than or equal to $\|f - g\|_1$. But once g is fixed, the second term approaches 0 as $n \to \infty$, by the Bounded Convergence Theorem. This proves (4).

5.
$$\left|\int_A f d\mu - \sum_A \bar{f} d\mu\right| = \left|\int_A \left(\frac{1}{n}\sum_{k=0}^{n-1} fT^k - \bar{f}\right) d\mu\right|$$
$$\leq \int_A \left|\frac{1}{n}\sum_{k=0}^{n-1} fT^k - \bar{f}\right| d\mu = \left\|\frac{1}{n}\sum_{k=0}^{n-1} fT^k - \bar{f}\right\|_{L^1(A)} \to 0,$$

by (4). ■

Chapter 2

Ergodicity, Mixing, Weakly Mixing

2.1 Ergodicity

As discussed in [11], p 25, if T is a measure-preserving transformation on X and if X is the union of two disjoint measurable sets E and F of positive measure, each of which is invariant under T, then the study of any property of T on X reduces to the separate studies of the corresponding properties of T on E and T on F. In such a situation T may be called decomposable. The most significant transformations are the indecomposable ones - usually referred to as *metrically transitive or ergodic*.

According to Halmos [11], p 25, "Ergodicity is one of the precise formulations of the natural requirement that a transformation do a good job of stirring up the points of the space it acts on."

We will define ergodicity in the following way: First, a set $B \in \mathcal{B}$ is called *invariant* if $\mu(T^{-1}B \triangle B) = 0$. This means that $T^{-1}B = B$ a.e. The transformation T (or, more properly, the system $(X, \mathcal{B}, \mu, T)$) is called *ergodic* or *metrically transitive* if every invariant set has measure 0 or 1.

Examples ([11], p 26)

1. The translation T defined by $Tx = x + 1$ on the space of integers is ergodic; the translation T defined by $Tx = x+2$ is not, since the set of even numbers is invariant. The translation T defined by $Tx = x + 1$ on the real line is not ergodic (an example of a non-trivial invariant set is $\bigcup_{n\in\mathbb{N}}(n, n + \frac{1}{2})$).

2. If X is the circle group $\{z \in \mathbb{C} : |z| = 1\}$, if $c \in X$, and T is defined by $Tx = cx$, then T is ergodic for some values for c and not ergodic for others. If c is a root of unity, i.e. $c^n = 1$ for some $n \in \mathbb{N}$, then T is not ergodic. To see this, note that the function $f(x) = x^n$ is invariant under T since $Tf(x) = f(cx) = c^n x^n = x^n$, and T is non-constant. If c is not a root of unity, then T is ergodic. To see this, note that the functions f_n defined by $f_n(x) = x^n, n \in \mathbb{Z}$ form a complete orthonormal set in L_2. It follows that if f is in L_2, then $f = \sum_n a_n f_n$, where the series converge in

the mean. Define the functional operator U as before, i.e. $Uf(x) = f(Tx)$; since $Uf_n = c^n f_n$, it follows that $Uf = \sum_n a_n c^n f_n$. If f is invariant, then $a_n = a_n c^n$ for all n, and hence $a_n = 0$ whenever $n \neq 0$. This shows that every invariant function in L_2 is a constant, hence that T is ergodic.

We will look at a few characterizations of ergodicity.

We say that a function f on X is *invariant* if $f \circ T = f$ a.e.

2.1.1 Lemma ([21], p 42)

$(X, \mathcal{B}, \mu, T)$ is ergodic if and only if every invariant measurable function on X is constant a.e.

Proof:

Suppose every invariant measurable function is constant a.e., and let $E \in \mathcal{B}$ be an invariant set. Then χ_E is constant a.e., so χ_E is either 0 a.e. or 1 a.e. Thus $\mu(E)$ is 0 or 1.

Conversely, suppose $(X, \mathcal{B}, \mu, T)$ is ergodic and f is an invariant measurable function. Then for each $r \in \mathbb{R}, E_r = \{x \in X : f(x) > r\}$ is measurable and invariant, hence has measure 0 or 1. But if f is not constant a.e., there exists an $r \in \mathbb{R}$ such that $0 < \mu(E_r) < 1$. Therefore f must be constant a.e. ■

2.1.2 Lemma ([21], p 43 and [11], p 34)

$(X, \mathcal{B}, \mu, T)$ is ergodic if and only if 1 is a simple eigenvalue of the transformation U induced on $L^2(X, \mathcal{B}, \mu)$ (complex) by T.

Moreover, if $(X, \mathcal{B}, \mu, T)$ is ergodic, then every eigenvalue of U is simple and the set of all eigenvalues of U is a subgroup of the circle group $\mathbb{K} = \{z \in \mathbb{C} : |z| = 1\}$.

Proof:

Since the space has finite measure, every constant function f is in L_2. Since $Uf = f$, the number 1 is always an eigenvalue of U. Since the set of all constant functions is a one-dimensional subspace of L_2, and since T is ergodic if and only if the only invariant functions in L_2 are the constants, the first assertion is proved.

Since U is unitary, every eigenvalue of U has absolute value 1. This is so since if $Uf = cf$, then $f = cU^*f$, so $\langle Uf, f\rangle = c\langle f, f\rangle$ and $\langle Uf, f\rangle = \langle f, U^*f\rangle = 1/\bar{c}\langle f, f\rangle$. Thus $c\bar{c} = 1$. Hence, if f is an eigenfunction with eigenvalue c, i.e. $f(Tx) = cf(x)$ a.e., then $|f|$ is invariant since $|Uf| = |cf| = |f|$. If both f and g are eigenfunctions with eigenvalues c, then f/g is an invariant function since $U(f/g) = (cf)/(cg) = f/g$. (Note that since $|g|$ is a nonzero constant, f/g makes sense). This proves the simplicity of each eigenvalue. If, finally, b and c are eigenvalues of U, with corresponding eigenfunctions f and g, then f/g is an eigenfunction of U with eigenvalue b/c. This shows that the eigenvalues of U form a group. ■

2.1.3 Theorem ([21], p 44)

$(X, \mathcal{B}, \mu, T)$ is ergodic if and only if for each $f \in L^1(X, \mathcal{B}, \mu)$ the time mean of f equals the space mean of f a.e.:

$$\bar{f}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \int_X f d\mu \quad \text{a.e.}$$

Proof:

Let $(X, \mathcal{B}, \mu, T)$ be ergodic. From Theorem 1.3.2.3 (2), $\bar{f}$ is invariant, and hence it is constant a.e. Also from the same theorem (5) we have

$$\int_X f d\mu = \int_X \bar{f} d\mu = \bar{f}(x) \quad \text{a.e.}$$

Conversely, suppose that for each $f \in L^1(X, \mathcal{B}, \mu)$, we have that

$$\bar{f}(x) = \int_X f d\mu \quad \text{a.e.}$$

This means that $\bar{f}$ is constant a.e. Let f be any invariant function in $L^1(X, \mathcal{B}, \mu)$. Then

$$\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = f(x) \quad \text{a.e.},$$

for all $n \in \mathbb{N}$, and so $f = \bar{f}$ a.e. Thus f is constant a.e., and Lemma 2.1.1 implies that $(X, \mathcal{B}, \mu, T)$ is ergodic. ■

2.1.4 Theorem ([21], p 45)

$(X, \mathcal{B}, \mu, T)$ is ergodic if and only if for each $f, g \in L^2(X, \mathcal{B}, \mu)$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \langle U^k f, g\rangle = \langle f, 1\rangle \langle \overline{g, 1}\rangle.$$

Proof:

If $(X, \mathcal{B}, \mu, T)$ is ergodic, then

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \langle U^k f, g\rangle &= \langle \bar{f}, g\rangle = \left\langle \int_X f d\mu, g \right\rangle \\ &= \langle \langle f, 1\rangle, g\rangle = \langle f, 1\rangle \langle 1, g\rangle = \langle f, 1\rangle \langle \overline{g, 1}\rangle. \end{aligned}$$

Conversely, given $f \in L^2(X, \mathcal{B}, \mu) \subset L^1(X, \mathcal{B}, \mu)$, suppose $\langle \bar{f}, g\rangle = \langle f, 1\rangle\overline{\langle g, 1\rangle}$ for all $g \in L^2(X, \mathcal{B}, \mu)$. Then we must have $\bar{f} = \langle f, 1\rangle = \int_X f d\mu$ a.e., so T is ergodic by the preceding theorem. ■

2.1.5 Corollary

T is ergodic if and only if

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu(T^{-k}A \cap B) \to \mu(A)\mu(B) \quad \text{for all} \quad A, B \in \mathcal{B}.$$

Proof:

In Theorem 2.1.4 let $f = \chi_A$ and $g = \chi_B$, and the result follows immediately. ■

The *mean sojourn time* of x in E is defined to be

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \chi_E(T^k x) = \bar{\chi}_E(x).$$

2.1.6 Lemma

T is ergodic if and only if the mean sojourn time in a measurable set equals the measure of the set for almost all points of X, i.e. if and only if

$$\bar{\chi}_E = \int_X \chi_E d\mu = \mu(E).$$

Proof:

If T is ergodic, then

$$\bar{\chi}_E = \int_X \chi_E d\mu = \mu(E)$$

a.e.

Conversely, if $\bar{\chi}_E = \chi_E = \mu(E)$ a.e. then, for an invariant set $E \in \mathcal{B}$, $\mu(E)$ must be 0 or 1 (the possible values of χ_E). Thus T is ergodic. ■

2.2 Strong mixing

The concept of ergodicity can be strengthened by introducing two more classes of transformations which satisfy even more stringent quantitative recurrence conditions, i.e. strong and weakly mixing.

Halmos used an example to supply a physical illustration of the concepts of a typical dynamical system as well as ideas of ergodicity, strong- and weakly mixing. The example runs as follows: Consider the physical system consisting of a cocktail shaker containing ice and gin into which a few drops of vermouth have been introduced, and suppose that the system is acted upon by the flow induced by a conscientious application of a swizzle-stick.

In this context, if ergodicity is expressed by saying that on the average a "measurable subset" A has 10 per cent vermouth, and if strong mixing is expressed by saying that after a while A will have 10 per cent vermouth in it, then weakly mixing can be expressed by saying that after a while A will have 10 per cent vermouth in it, with the exception of a few rare instants during which it may be either too strong or too sweet.

A m.p.t. $T : X \to X$ is strongly mixing if and only if

$$|\mu(T^{-k}A \cap B) - \mu(A)\mu(B)| \to 0 \quad \text{for all} \quad A, B \in \mathcal{B}.$$

Thus T is strongly mixing if and only if

$$\mu(T^{-k}A \cap B) \to \mu(A)\mu(B),$$

i.e.,

$$\frac{\mu(T^{-k}A \cap B)}{\mu(B)} \to \mu(A) \quad \text{for all} \quad A, B \in \mathcal{B}, \quad \text{with} \quad \mu(B) > 0.$$

This says that eventually T distributes A fairly evenly throughout the space X: for large k, the proportion of $T^{-k}A$ that lies in B, namely $\mu(T^{-k}A \cap B)/\mu(B)$, is approximately the same as the relative size of A in X, namely $\mu(A)$. ([21], p 57)

We may mention a few results regarding strong mixing at this stage.

2.2.1 Lemma

T is strongly mixing if and only if

$$\langle U^n f, g\rangle \to \langle f, 1\rangle\overline{\langle g, 1\rangle} \quad \textit{for all} \quad f, g \in L^2.$$

Proof:

Let $A, B \in \mathcal{B}$ and χ_A and χ_B be characteristic functions of A and B.

Then

$$\langle U^n\chi_A, \chi_B\rangle = \mu(T^{-n}A \cap B)$$

and

$$\langle \chi_A, 1\rangle\langle 1, \chi_B\rangle = \mu(A)\mu(B).$$

Now apply a double approximation process:

For each fixed χ_B, the result

$$\langle U^n\chi_A, \chi_B\rangle \longrightarrow \langle \chi_A, 1\rangle\langle 1, \chi_B\rangle$$

holds for all simple functions χ_A, and therefore, by L_2-approximation, for all functions $f \in L_2$.

Second, for each fixed χ_A, we can argue similarly about χ_B and all functions $g \in L_2$ obtained by L_2-approximation. ■

Next we show that the statement regarding recurrence obtained from the Khintchine theorem can be strengthened under the assumption of strongly mixing:

2.2.2 Lemma ([21], p 58)

T is strongly mixing if and only if for each $A \in \mathcal{B}$

$$\lim_{n\to\infty} \mu(T^{-n}A \cap A) = \mu(A)^2.$$

Proof:

We will show that if $\lim_{n\to\infty} \mu(T^{-n}A\cap A) = \mu(A)^2$ for all $A \in \mathcal{B}$, then $(X, \mathcal{B}, \mu, T)$ is strongly mixing. The converse follows immediately from the definition of strongly mixing with $A = B$.

Let $A \in \mathcal{B}$ be fixed and let M be the closed linear subspace of $L^2(X, \mathcal{B}, \mu)$ generated by (i) the constant functions and (ii) the set $\{U^k\chi_A : k \in \mathbb{Z}\}$.

Since

$$\begin{aligned}
\lim_{n\to\infty} \langle U^n\chi_A, 1\rangle &= \lim_{n\to\infty} \mu(A) = \langle \chi_A, 1\rangle\langle 1, 1\rangle \quad \text{and} \\
\lim_{n\to\infty} \langle U^n\chi_A, U^k\chi_A\rangle &= \lim_{n\to\infty} \langle U^{n-k}\chi_A, \chi_A\rangle = \lim_{n\to\infty} \mu(T^{-n+k}A \cap A) \\
&= \mu(A)^2 = \langle \chi_A, 1\rangle\langle U^k\chi_A, 1\rangle,
\end{aligned}$$

we obtain that

$$\lim_{n\to\infty} \langle U^n\chi_A, f\rangle = \langle \chi_A, 1\rangle\overline{\langle f, 1\rangle} \quad \text{for all} \quad f \in M.$$

For any given $f \in L^2(X, \mathcal{B}, \mu)$, we write $f = f_1 + f_2$, where $f_1 \in M$ and $f_2 \in M^\perp$. Then

$$\begin{aligned}
\lim_{n\to\infty} \langle U^n\chi_A, f\rangle &= \lim_{n\to\infty} \langle U^n\chi_A, f_1\rangle + \lim_{n\to\infty} \langle U^n\chi_A, f_2\rangle \\
&= \langle \chi_A, 1\rangle\overline{\langle f_1, 1\rangle} = \langle \chi_A, 1\rangle\overline{\langle f, 1\rangle},
\end{aligned}$$

since $(f_2, 1) = 0$. Therefore, if we let $f = \chi_B$, $B \in \mathcal{B}$ we see that

$$\lim_{n\to\infty} \mu(T^{-n}A \cap B) = \lim_{n\to\infty} \langle U^n \chi_A, \chi_B \rangle = \langle \chi_A, 1 \rangle \langle \chi_B, 1 \rangle = \mu(A)\mu(B);$$

hence $(X, \mathcal{B}, \mu, T)$ is strongly mixing. ∎

2.3 Weakly mixing equivalent statements and other results

A m.p.t. $T : X \to X$ on a probability space $(X, \mathcal{B}, \mu)$ is said to be *weakly mixing* in case for every pair of measurable sets A and B,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(T^{-k}A \cap B) - \mu(A)\mu(B)| = 0.$$

Example: Bernoulli shifts

As introduced in Example 10, we now briefly show that Bernoulli shifts are weakly mixing systems. Since the cylinder sets generate $\mathcal{B}$, we will restrict our attention to this class of sets. Thus, let A, B be arbitrary cylinder sets:

$$A = \{\omega : \omega_{i_1} = j_1, \omega_{i_2} = j_2, \ldots, \omega_{i_k} = j_k\}$$

and

$$B = \{\omega : \omega_{l_1} = s_1, \omega_{l_2} = s_2, \ldots, \omega_{l_h} = s_h\}$$

for some $k, h \in \mathbb{N}$ and as discussed in Example 10.

Recall that the measure-preserving transformation T in this system is given by the shift $T\{\omega_n\} = \{\omega_{n+1}\}$. Consider the terms $\mu(A \cap T^{-n}B)$, $n = 1, 2, \ldots$. For some $n_0 \in \mathbb{N}$ we will have that the defining coordinates for A and $T^{-n}B$ are disjoint for all $n \geq n_0$. Hence for such n,

$$A \cap T^{-n}B = \{\omega : \omega_{i_1} = j_1, \omega_{i_2} = j_2, \ldots, \omega_{i_k} = j_k, \omega_{l_1} = s_1, \omega_{l_2} = s_2, \ldots, \omega_{l_h} = s_h\}$$

such that

$$\mu(A \cap T^{-n}B) = p_{j_1} p_{j_2} \cdots p_{j_k} p_{s_1} p_{s_2} \cdots p_{s_h} = \mu(A)\mu(B).$$

Therefore

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(A \cap T^{-k}B) - \mu(A)\mu(B)| = 0,$$

which means that the system is weakly mixing. In fact, this system is also strongly mixing.

Multiple weakly mixing

Most of our multiple recurrence results will be stated in terms of multiple weakly mixing, which can be defined as follows:

The system $(X, \mathcal{B}, \mu, T)$ is said to be *multiply weakly mixing* (weakly mixing of order k) if for sets $A_0, A_1, \ldots, A_k$ in $\mathcal{B}$, we have that

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

We can extend this idea even further and say that the system $(X, \mathcal{B}, \mu, T)$ is *weakly mixing of all orders* (or weakly mixing of order $(m_1, m_2, \cdots, m_k)$) if for any sets $A_0, A_1, \ldots, A_k$ in $\mathcal{B}$, and for any choice of non-negative integers $m_1, m_2, \cdots, m_k$ such that

$$1 \le m_1 < m_2 < \cdots < m_k,$$

we have that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\mu(A_0 \cap T^{-m_1 n}A_1 \cap T^{-m_2 n}A_2 \cap \ldots \cap T^{-m_k n}A_k) - \mu(A_0)\mu(A_1)\ldots\mu(A_k)\right)^2 = 0.$$

In this section we will supply a number of results (Propositions 2.3.1 to 2.3.8) which will help to explain most of the implications in this theorem.

Some useful results

At this stage it may be opportune to state and prove some elementary, but useful results which are used in the proof of Theorem 2.3.9 as well as in several other proofs in the remainder of this text.

2.3.1 PROPOSITION

If a_n *is a sequence of complex numbers such that*

$$\lim_{n\to\infty} a_n = a,$$

then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\, a_k - a| = 0.$$

Proof:

Given any $\varepsilon > 0$. There exists an $N \in \mathbb{N}$ such that $|a_n - a| < \varepsilon$ for all $n \geq N$.

Consider the average

$$\frac{1}{n}\sum_{k=0}^{n-1}|\ a_k - a|.$$

For $n > N$ we have that

$$\frac{1}{n}\sum_{k=0}^{n-1}|\ a_k - a| = \frac{1}{n}\sum_{k=0}^{N-1}|\ a_k - a| + \frac{1}{n}\sum_{k=N}^{n-1}|\ a_k - a|.$$

It is clear that the finite sum $\frac{1}{n}\sum_{k=0}^{N-1}|\ a_k - a| \to 0$ if $n \to \infty$.

Furthermore,

$$\frac{1}{n}\sum_{k=N}^{n-1}|\ a_k - a| < \frac{1}{n}(n-1-N)\varepsilon < \varepsilon.$$

Hence, since ε is arbitrary,

$$\frac{1}{n}\sum_{k=N}^{n-1}|\ a_k - a| \longrightarrow 0$$

and the proposition is proved. ■

2.3.2 Proposition

If a_n *is a sequence of complex numbers such that*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}|\ a_k - a| = 0$$

then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}a_k = a.$$

Proof:

We have that

$$0 \le \left|\frac{1}{n}\sum_{k=0}^{n-1}(a_k - a)\right| \le \frac{1}{n}\sum_{k=0}^{n-1}|\ a_k - a|\ ,$$

and the Proposition follows immediately. ■

Remark: From the previous two Propositions we see that strong mixing implies weakly mixing and weakly mixing implies ergodicity.

Zero density sets and convergence in density

1. A subset $E \subset \mathbb{N}$ of positive integers is said to have *density zero* if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_E(k) = 0.$$

A few basic properties of zero density sets can readily be verified:

(a) If $B \subset \mathbb{N}$ has density zero and $A \subset B$, then A has density zero. To see this note that $\chi_A(k) \leq \chi_B(k)$ for all $k \in \mathbb{N}$. Hence

$$0 \leq \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(k) \leq \frac{1}{n} \sum_{k=0}^{n-1} \chi_B(k)$$

and so

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(k) = 0.$$

(b) If $A \subset \mathbb{N}$ and $B \subset \mathbb{N}$ both have density zero then $A \cup B$ has density zero. We have that $\chi_{A\cup B}(k) \leq \chi_A(k) + \chi_B(k)$ for all $k \in \mathbb{N}$. Therefore

$$\frac{1}{n} \sum_{k=0}^{n-1} \chi_{A\cup B}(k) \leq \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(k) + \frac{1}{n} \sum_{k=0}^{n-1} \chi_B(k),$$

and it follows immediately that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_{A\cup B}(k) = 0.$$

(c) If either a set $A \subset \mathbb{N}$ or $B \subset \mathbb{N}$ has density zero, then $A \cap B$ has density zero. Suppose that A has density zero, then since $\chi_{A\cap B} \leq \chi_A$, it follows similarly than above that $A \cap B$ has density zero.

2. A sequence $\{a_n\}_{n\geq 1}$ in a topological space X is said to converge in density to an element $a \in X$ if there exists a subset $E \subset \mathbb{N}$ of density zero such that

$$\lim_{n\to\infty, n\notin E} a_n = a.$$

We will also write

$$D - \lim_{n\to\infty} a_n = a.$$

2.3.3 LEMMA (KOOPMAN-VON NEUMANN) ([21], p 65 and [20], p 47)

Let $\{a_k\}_{k\geqslant 1}$ *be a sequence in* $[0,\infty)$. *Then the following statements are equivalent:*

1. *There exists a subset* $E \subset \mathbb{N}$ *of density zero such that*
$$\lim_{n\to\infty, n\notin E} a_n = 0.$$

2. $$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} a_k = 0.$$

3. *The set* $E_\varepsilon := \{k \in \mathbb{N} : \; a_k > \varepsilon\}$ *has density zero for all* $\varepsilon > 0$.

Proof:

(1) $\Longrightarrow$ (2): We have that
$$\frac{1}{n}\sum_{k=0}^{n-1} a_k = \frac{1}{n}\sum_{k\leq n-1, k\in E} a_k + \frac{1}{n}\sum_{k\leq n-1, k\notin E} a_k.$$

The first term is close to zero for large n, since E has density 0 and the sequence $\{a_k\}$ is bounded. To see this, let $c = \sup_{k\in\mathbb{N}}(a_k)$. Then
$$0 \leq \frac{1}{n}\sum_{k\leq n-1, k\in E} a_k = \frac{1}{n}\sum_{k=0}^{n-1}\chi_E a_k \leq \frac{c}{n}\sum_{k=0}^{n-1}\chi_E \to 0$$
as $n \to \infty$.

The second term is also close to zero for large n, since we have that
$$\lim_{n\to\infty} f(k)\chi_{E^c}(k) = \lim_{n\to\infty, n\notin E} f(n) = 0$$
and ordinary convergence implies convergence in the mean by Proposition 2.3.1.

(2) $\Longrightarrow$ (3): Suppose that
$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} a_k = 0.$$

For each $m = 1, 2, \ldots$ let
$$E_m = \left\{k \in \mathbb{N} : f(k) > \frac{1}{m}\right\}.$$

We clearly have that for each $m \in \mathbb{N}$, $(1/m)\chi_{E_m} < a_k$, i.e. $\chi_{E_m} < ma_k$ for all $k \in \mathbb{N}$. Also, $E_1 \subset E_2 \subset \ldots$, and each E_m has density zero. This is so since for each $m \in \mathbb{N}$ we have that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_{E_m} \leq m \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k = 0.$$

Now given any $\varepsilon > 0$, we can find an $m \in \mathbb{N}$ such that $\frac{1}{m} < \varepsilon$. Then we have that

$$E_\varepsilon = \{k \in \mathbb{N} : \ a_k > \varepsilon\} \subset \left\{k \in \mathbb{N} : f(k) > \frac{1}{m}\right\} = E_m.$$

Since E_m has density zero, E_ε also has density zero.

(3) $\Longrightarrow$ (1): For each $m = 1, 2, \ldots$ let

$$E_m = \left\{k \in \mathbb{N} : f(k) > \frac{1}{m}\right\}.$$

Then $E_1 \subset E_2 \subset \ldots$, and each E_m has density zero, by assumption. Therefore, for each $m = 1, 2, \ldots$ we may choose $i_m > 0$ such that $i_1 < i_2 < \ldots$ and

$$\frac{1}{n} \sum_{k=0}^{n-1} \chi_{E_m}(k) < \frac{1}{m} \quad \text{for} \quad n \geq i_{m-1} (m = 2, 3, \ldots).$$

Now let $i_0 = 0$ and define

$$E = \bigcup_{m=1}^{\infty} E_m \cap (i_{m-1}, i_m);$$

we will show that E has density zero and

$$\lim_{n\to\infty, n\notin E} f(n) = 0.$$

Let us consider $f(k)$ for $k \notin E$. In the interval $(0, i_1]$ we have removed those values of k for which $f(k) > 1$; in $(i_1, i_2]$ we have removed those values of k for which $f(k) > \frac{1}{2}$; and in $(i_{m-1}, i_m]$ we have removed those values of k for which $f(k) > 1/m$. Thus clearly

$$\lim_{n\to\infty, n\notin E} f(n) = 0.$$

However, we have only removed a set of values of k of density zero. For if $i_{m-1} < n \leq i_m$, then

$$\begin{aligned}
\frac{1}{n} \sum_{k=0}^{n-1} \chi_E(k) &= \frac{1}{n} \sum_{k=0}^{i_{m-1}-1} \chi_E(k) + \frac{1}{n} \sum_{k=i_{m-1}}^{n-1} \chi_E(k) \\
&\leq \frac{1}{n} \sum_{k=0}^{i_{m-1}-1} \chi_{E_{m-1}}(k) + \frac{1}{n} \sum_{k=0}^{n-1} \chi_{E_m}(k) \\
&\leq \frac{1}{i_{m-1}} \sum_{k=0}^{i_{m-1}-1} \chi_{E_{m-1}}(k) + \frac{1}{n} \sum_{k=0}^{n-1} \chi_{E_m}(k) \\
&< \frac{1}{m-1} + \frac{1}{m}. \quad \blacksquare
\end{aligned}$$

In this proof we can isolate an interesting fact for density zero sequences. In the implication (3) $\Longrightarrow$ (1) we constructed a zero density set E from a sequence of zero density sets $\{E_i\}$ such that $E_i \backslash E$ is finite for every i. This type of construction will prove to be very useful as will be shown in Chapter 4, Theorem 4.4.1.1.

From Theorem 2.3.3 we immediately get

2.3.4 Corollary

If a_n is a bounded sequence of complex numbers, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\ a_k - a| = 0$$

if and only if there exists a set $K \subset \mathbb{N}$ of density zero such that

$$D - \lim_{n\to\infty} a_n = a.$$

2.3.5 Proposition

If a_n is a sequence of complex numbers with

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k = \alpha, \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k^2 = \alpha^2,$$

then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (a_k - \alpha)^2 = 0.$$

Proof:

The proof is obtained by direct computation and applying basic convergence theorems, i.e.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (a_k - \alpha)^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (a_k^2 - 2\alpha a_k + \alpha^2) = 0.$$

■

2.3.6 Proposition

If a_n is a sequence of real numbers, then

$$D - \lim_{n\to\infty} a_n^2 = 0 \Longleftrightarrow D - \lim_{n\to\infty} |a_n| = 0.$$

Proof: Let $\varepsilon > 0$.

Suppose that $D - \lim_{n\to\infty} a_n^2 = 0$. Then there is an $N \in \mathbb{N}$ such that

$$\begin{aligned} &|a_n^2| < \varepsilon^2 \text{ for all } n \geq N \text{ except on a set } E \text{ of zero density} \\ \Leftrightarrow\ &|a_n| < \varepsilon \text{ for all } n \geq N \text{ except on } E. \end{aligned}$$

which yields the result since ε is arbitrary. ■

2.3.7 Corollary

If a_n is a sequence of real numbers, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k^2 = 0 \Longleftrightarrow \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |a_k| = 0.$$

Proof: This result follows immediately from Corollary 2.3.4 and Proposition 2.3.6. ■

We now see from Corollary 2.3.7 that if we work in a real space, then weakly mixing implies that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (\mu(T^k A \cap B) - \mu(A)\mu(B))^2 = 0$$

(and vice versa) and this is in turn a special case of the equation

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left(\int fU^n g\ d\mu - \int f\ d\mu \int g\ d\mu \right)^2 = 0,$$

where $f, g \in L^2(X, \mathcal{B}, \mu)$. The special case is obtained when $f = \chi_A$ and $g = \chi_B$.

In the main theorems of this dissertation in Chapter 3, we will use this latter characterization of weakly mixing when it is assumed that the function space under consideration is real.

2.3.8 Corollary

If a_n is a sequence of real numbers with

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k^2 = 0,$$

then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k = 0.$$

Proof:

This follows immediately from Corollary 2.3.7 and Proposition 2.3.2. ■

The next theorem is basic to many theorems regarding weakly mixing systems. In Chapter 3 some of these properties of weakly mixing systems will be needed.

2.3.9 THEOREM ([21], p 65)

Let $T : X \to X$ be a m.p.t. on a probability space $(X, \mathcal{B}, \mu)$. Then the following are equivalent:

(1) T is weakly mixing.

(2) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle U_T^k f, g\rangle - \langle f, 1\rangle\overline{\langle g, 1\rangle}| = 0$ for all $f, g \in L^2(X, \mathcal{B}, \mu)$.

(3) Given $A, B \in \mathcal{B}$, there is a set $J \subset \mathbb{Z}^+$ of density 0 such that

$$\lim_{n\to\infty, n\notin J} \mu(T^n A \cap B) = \mu(A)\mu(B),$$

i.e.

$$D - \lim_{n\to\infty} \mu(T^n A \cap B) = \mu(A)\mu(B).$$

(4) $T \times T$ is weakly mixing.

(5) $T \times S$ is ergodic on $X \times Y$ for each ergodic $(Y, \mathcal{C}, v, S)$.

(6) $T \times T$ is ergodic.

Proof:

(1) $\Longrightarrow$ (2): This is true when f and g are characteristic functions of measurable sets. The general statement follows by forming linear combinations and approximating.

(2) $\Longrightarrow$ (1): Again this can be seen by letting f and g be two characteristic functions of measurable sets.

(1) $\Longleftrightarrow$ (3): This follows directly from the Koopman-von Neumann Lemma 2.3.3, with

$$a_n = |\mu(T^n A \cap B) - \mu(A)\mu(B)|.$$

(3) $\Longrightarrow$ (4): Let $A, B, C, D \in \mathcal{B}$. By (3), there are sets $J_1, J_2 \subset \mathbb{N}$, each of density zero, such that

$$\lim_{n\to\infty, n\notin J_1} |\mu(T^n A \cap C) - \mu(A)\mu(C)| = 0, \text{ and}$$

$$\lim_{n\to\infty, n\notin J_2} |\mu(T^n B \cap D) - \mu(B)\mu(D)| = 0.$$

Let $J = J_1 \cup J_2$. Then J has density zero, as seen earlier, and

$$\begin{aligned}
& \lim_{n\to\infty, n\notin J} |\mu \times \mu(T \times T)^n((A \times B) \cap (C \times D)) \\
& -\mu \times \mu(A \times B)\mu \times \mu(C \times D)| \\
= & \lim_{n\to\infty, n\notin J} |\mu(T^n A \cap C)\mu(T^n B \cap D) - \mu(A)\mu(B)\mu(C)\mu(D)| \\
\leq & \lim_{n\to\infty, n\notin J} [\mu(T^n A \cap C)|\mu(T^n B \cap D) - \mu(B)\mu(D)| \\
& +\mu(B)\mu(D)|\mu(T^n A \cap C) - \mu(A)\mu(C)|] = 0,
\end{aligned}$$

which yields the result by the equivalence (3) $\Leftrightarrow$ (1) applied to $T \times T$.

(4) $\Longrightarrow$ (5): If $T \times T$ is weakly mixing, it is straightforward to show that T is also weakly mixing. Let $(Y, \mathcal{C}, v, S)$ be ergodic. To show that $T \times S$ is ergodic on $X \times Y$, we prove that if $A, B \in \mathcal{B}$ and $C, D \in \mathcal{C}$, then

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu \times v[(T \times S)^k(A \times C) \cap (B \times D)] \to \mu(A)\mu(B)v(C)v(D).$$

This will be sufficient by Corollary 2.1.5.

We have that

$$\begin{aligned}
& \frac{1}{n}\sum_{k=0}^{n-1} \mu \times v[(T \times S)^k(A \times C) \cap (B \times D)] \\
& = \frac{1}{n}\sum_{k=0}^{n-1} \mu(T^k A \cap B)v(S^k C \cap D) \\
& = \frac{1}{n}\sum_{k=0}^{n-1} \{\mu(A)\mu(B)v(S^k C \cap D) + [\mu(T^k A \cap B) - \mu(A)\mu(B)]v(S^k C \cap D)\}.
\end{aligned}$$

By ergodicity of S, we have that

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu(A)\mu(B)v(S^k C \cap D) \longrightarrow \mu(A)\mu(B)v(C)v(D).$$

as $n \to \infty$.

Also,

$$\left|\frac{1}{n}\sum_{k=0}^{n-1}[\mu(T^k A \cap B) - \mu(A)\mu(B)]v(S^k C \cap D)\right| \leq \frac{1}{n}\sum_{k=0}^{n-1}|\mu(T^k A \cap B) - \mu(A)\mu(B)| \longrightarrow 0$$

as $n \to \infty$ because T is weakly mixing. Hence (5) is established.

(5) $\Longrightarrow$ (6): From (5) we have that T must be ergodic, since $T \times \{1\}$ is ergodic, where $\{1\}$ represents the identity transformation on a single point. Therefore, under the assumption of (5), $T \times T$ is also ergodic.

(6) $\Longrightarrow$ (3) : If $A, B \in \mathcal{B}$, then

$$\begin{aligned}
&\frac{1}{n}\sum_{k=0}^{n-1}[\mu(T^k A \cap B) - \mu(A)\mu(B)]^2 \\
= \; &\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^k A \cap B)^2 - 2\mu(A)\mu(B)\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^k A \cap B) + [\mu(A)\mu(B)]^2 \\
= \; &\frac{1}{n}\sum_{k=0}^{n-1}\mu \times \mu[(T \times T)^k (A \times A) \cap (B \times B)] - 2\mu(A)\mu(B)\cdot \\
&\frac{1}{n}\sum_{k=0}^{n-1}\mu \times \mu[(T \times T)^k (A \times X) \cap (B \times X)] + [\mu(A)\mu(B)]^2.
\end{aligned}$$

Since $T \times T$ is ergodic, as $n \to \infty$ this tends to

$$\begin{aligned}
&\mu \times \mu(A \times A)\mu \times \mu(B \times B) \\
- \; &2\mu(A)\mu(B)\mu \times \mu(A \times X)\mu \times \mu(B \times X) + [\mu(A)\mu(B)]^2 \\
= \; &\mu(A)^2\mu(B)^2 - 2\mu(A)^2\mu(B)^2 + \mu(A)^2\mu(B)^2 = 0.
\end{aligned}$$

By the Koopman-von Neumann Lemma 2.3.3, there is a set J of density zero such that

$$\lim_{n\to\infty, n\notin J} |\mu(T^n A \cap B) - \mu(A)\mu(B)|^2 = 0, \text{ which yields (3).}$$

■

Chapter 3

Multiple Recurrence

3.1 Background

From a statistical viewpoint the two most important types of dynamical systems are the weakly mixing systems and the almost periodic systems when one studies recurrence and multiple recurrence properties. In fact, it was shown in [9] that if one would like to prove multiple recurrence for general dynamical systems, one could reduce the problem by only considering the two mentioned cases.

A system $(X, \mathcal{B}, \mu, T)$ is called almost periodic or relatively compact if for any $f \in L^2(X)$ the orbit $\{U^n f\}_{n \geqslant 0}$ is relatively compact in $L^2(X)$, where U is the unitary operator on $L^2(X)$ induced by T.

It can be shown that $(X, \mathcal{B}, \mu, T)$ is almost periodic if and only if one has for any $\varepsilon > 0$ that there exists a relatively dense $E \subset \mathbb{N}$ such that $\|U^n f - f\| < \varepsilon$ for all $n \in E$. This can be seen from the assumed compactness of the orbit closure $\overline{\{U^n f\}} \subset L^2(X, \mathcal{B}, \mu)$. Since $\{U^n f\}$ is totally bounded, there exists a finite subset $\{U^{n_1} f, U^{n_2} f, ..., U^{n_r} f\}$ for which

$$\|U^{n_i} f - U^{n_j} f\| \geq \varepsilon.$$

Also, there exists a greatest r for which this is true. To see this, suppose the contrary. Hence for any $k \in \mathbb{N}$ there exists $U^{n_1^k} f, U^{n_2^k} f, \ldots, U^{n_k^k} f$ such that

$$\left\| U^{n_i^k} f - U^{n_j^k} f \right\| \geq \varepsilon.$$

For the sequence $\{U^{n_1^k} f\}_{k=1}^{\infty}$ we can choose (by compactness) a subsequence $\{U^{n_1^{k_p}} f\}$ which converges to, say, $f_1 \in \overline{\{U^n f\}}$.

Similarly, the sequence $\{U^{n_2^{k_p}} f\}_{k=1}^{\infty}$ has a subsequence which converges to, say, f_2.

In this way we find infinite sets $\mathbb{N} \supset \mathcal{N}_1 \supset \mathcal{N}_2 \supset \ldots$ and $f_1, f_2, \ldots \in \overline{\{U^n f\}}$ such that

$$\lim_{\mathcal{N}_j \ni \ n \to \infty} U^{n_j^n} f = f_j.$$

Then

$$\begin{aligned}\|f_j - f_k\| &= \lim_{\mathcal{N}_k \ni\ n\to\infty} \left\|U^{n_j^n} f - U^{n_k^n} f\right\| \\ &\geq \varepsilon \text{ for all } j < k.\end{aligned}$$

Hence the sequence $\{f_i\} \subset \overline{\{U^n f\}}$ has no convergent subsequence which contradicts the compactness of $\overline{\{U^n f\}}$. Hence the maximality of r is established.

Since there exist a greatest $r \geq 1$ such that there are $n_1 < \cdots < n_r$ in $\mathbb{N}$ with

$$\|U^{n_i} f - U^{n_j} f\| \geq \varepsilon.$$

for $i \neq j$. Since U is a unitary operator, it is also an isometry and we then get that

$$\|U^{n+n_i} f - U^{n+n_j} f\| = \|U^{n_i} f - U^{n_j} f\| \geq \varepsilon$$

for any n and $i \neq j$.

Since r is maximal, we have that

$$\|U^{n+n_i} f - f\| < \varepsilon$$

for at least one $1 \leq i \leq r$.

Hence,

$$\|U^m f - f\| < \varepsilon \text{ for some } n \leq m \leq n + n_r,$$

which shows that this property holds on a relatively dense set.

The converse of our assertion is straightforward.

This characterization of almost periodicity clearly indicates a strong recurrence nature and some multiple recurrence properties for almost periodic systems are obtained as follows:

3.1.1 Theorem ([9], p 536)

If $(X, \mathcal{B}, \mu, T)$ is almost periodic for every $f \in L^\infty(X, \mathcal{B}, \mu)$, $f \geq 0$ but f not 0 a.e, then there exists a relative dense set $E \subseteq \mathbb{N}$ such that

$$\int f\ U^n f\ U^{2n} f\ \cdots\ U^{kn} f\ d\mu > 0$$

for any $n \in E$.

Proof:

Let $a = \int f^{k+1} d\mu$. Then $a > 0$. Since f is essentially bounded, we can assume without loss of generality that $0 \leq f \leq 1$. Choose $\varepsilon < a/(k+1)$. If we choose measurable functions $g_0, g_1, ..., g_k$ with $0 \leq g_i \leq 1$ such that $\|f - g_i\| < \varepsilon,\ i = 0, 1, ..., k$, then, following from the identity

$$\prod_{l=0}^{k} a_l - \prod_{l=0}^{k} b_l = \sum_{j=0}^{k} \left(\prod_{l=0}^{j-1} a_l\right)(a_j - b_j)\left(\prod_{l=j+1}^{k} b_l\right);$$

we have that

$$\begin{aligned}\left|\int \prod_{l=0}^{k} g_l \, d\mu - \int f^{k+1} \, d\mu\right| &= \left|\int \sum_{j=0}^{k} \left(\prod_{l=0}^{j-1} g_l\right)(g_j - f)\left(\prod_{l=j+1}^{k} f\right) \, d\mu\right| \\ &\leq \sum_{j=0}^{k} \int \left(\prod_{l=0}^{j-1} g_l\right) |g_j - f| f^{k-j} \, d\mu \\ &\leq (k+1)\varepsilon < a.\end{aligned}$$

It then follows that

$$\int \prod_{l=0}^{k} g_l \, d\mu \geq a - (k+1)\varepsilon > 0 \tag{1}$$

If we now choose $g_l := U^{ln} f$ for $l = 0, 1, \ldots, k$ and $n \in E$, then from the discussion prior to this theorem, we have that for any $\varepsilon > 0$,

$$\|U^n f - f\| < \frac{\varepsilon}{k}$$

for all n in a relatively dense set E. What remains to be shown is that $\|U^{ln} f - f\| < \varepsilon$ for $l = 0, 1, ..., k$ and for all n in a relatively dense set.

Since T is measure-preserving, we have that

$$\|U^{jn}(U^n f - f)\| = \|U^{(j+1)n} f - U^{jn} f\| < \frac{\varepsilon}{k}$$

for these n, and by the triangle inequality,

$$\begin{aligned}\|U^{kn} f - f\| &\leq \sum_{j=0}^{k-1} \|U^{(j+1)n} f - U^{jn} f\| < k\frac{\varepsilon}{k} = \varepsilon \\ \|U^{(k-1)n} f - f\| &\leq \sum_{j=0}^{k-2} \|U^{(j+1)n} f - U^{jn} f\| < (k-1)\frac{\varepsilon}{k} < \varepsilon \\ &\vdots \\ \|U^n f - f\| &< \tfrac{\varepsilon}{k} < \varepsilon \\ \|f - f\| &= 0 < \varepsilon.\end{aligned}$$

Therefore,

$$\|T^{ln} f - f\| < \varepsilon$$

for $l = 0, 1, ..., k$, which completes the proof by (1). ■

In this instance, if we let $f = \chi_A$, we obtain that for n on a relative dense set and $k \in \mathbb{N}$

$$\mu\left(A \cap T^{-n} A \cap T^{-2n} A \cap \cdots \cap T^{-kn} A\right) > 0$$

if $A \in \mathcal{B}$ and $\mu(A) > 0$.

Example of an almost periodic system

Let $\mathbb{K} = \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle, which can be represented as the reals modulo 1. Let $\mathcal{B}$ be the σ-algebra of Borel subsets of X, and let μ be the Lebesque measure and $T : X \to X$ defined by $T_\alpha x = x + \alpha$ for any fixed α. T_α is measure preserving, as can easily be verified.

Regarded as a map of $\mathbb{K}$,

$$T_\alpha e^{2\pi i\theta} = e^{2\pi i(\theta+\alpha)}.$$

If α is rational, then T_α is periodic, and all orbits will be finite and of the same cardinality.

Hence T_α is most interesting if α is irrational. Clearly for any $\varepsilon > 0$, from the compactness of $\mathbb{K}$, the set $\{n \in \mathbb{N} : \{e^{2\pi i(\theta+n\alpha)}\}$ is ε-separated$\}$ is finite, and it follows by a similar argument than the one earlier in this introduction, that $(X, \mathcal{B}, \mu, T_\alpha)$ is relatively compact, i.e. almost periodic. ■

Multiple recurrence results for weakly mixing systems are much more involved and will be proved in the remainder of this chapter. Let us recall from section 2.3 that we mean by multiple weakly mixing that

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

For $k = 1$ we obtain weakly mixing as given in the definition (Corollary 2.3.7). If $k = 2$, for example, and we let $A_0 = A_1 = A_2 := A$ we get that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\mu(A \cap T^{-n}A \cap T^{-2n}A) - \mu(A)^3\right)^2 = 0,$$

which says that the points of A that return to itself after n and $2n$ time iterations, for large n (i.e. after a sufficiently long period) will have positive measure for most n. This measure can be approximated by the limit in the expression above. Hence there will certainly be a "significant" number of points that return to A.

In the more general case, where the system $(X, \mathcal{B}, \mu, T)$ is weakly mixing of order $(m_1, m_2, \cdots, m_k)$, we let $A := A_1 = A_2 = \cdots = A_k$ and thus we get that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\mu(A \cap T^{-m_1 n}A \cap T^{-m_2 n}A \cap \ldots \cap T^{-m_k n}A) - \mu(A)^{k+1}\right)^2 = 0.$$

This can be seen to say that the set of points of A that returns to A after $m_1 n, m_2 n, \cdots, m_k n$ time steps will also have positive measure for most n (or, what amounts to the same, on average) after enough time has elapsed.

3.2 Van der Corput lemma in Hilbert space

The result that follows is a variant of a well-known result of Van der Corput for uniformly distributed sequences (mod 1) as discussed in [17].

This result will be the backbone of all the results given in this chapter.

3.2.1 LEMMA ([10])

If $\{u_n\}$ is a bounded sequence of vectors in a Hilbert space H, and for each $m \in \mathbb{Z}$ we set

$$\gamma_m = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \langle u_n, u_{n+m} \rangle,$$

then if

$$\lim_{M\to\infty} \frac{1}{M} \sum_{m=0}^{M-1} \gamma_m = 0,$$

we will have that

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

Proof:

If N is much larger than M, then

$$\frac{1}{N} \sum_{n=0}^{N-1} u_n \sim \frac{1}{N} \frac{1}{M} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} u_{n+m},$$

where the $\sim$ sign means the two expressions are close in the norm.

To see this, note that

$$\begin{aligned}
\frac{1}{N}\frac{1}{M} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} u_{n+m} &= \frac{1}{NM} \left[\sum_{n=1}^{M-1} n u_{n-1} + \sum_{n=M-1}^{N-1} M u_n + \sum_{n=N}^{N+M-2} (N+M-n-1) u_n \right] \\
&= \frac{1}{NM} \sum_{n=1}^{M-1} n u_{n-1} + \frac{1}{N} \sum_{n=M-1}^{N-1} u_n + \frac{1}{NM} \sum_{n=N}^{N+M-2} (N+M-n-1) u_n \\
&\sim \frac{1}{N} \sum_{n=M-1}^{N-1} u_n \\
&\sim \frac{1}{N} \sum_{n=0}^{N-1} u_n .
\end{aligned}$$

Note that both the sums $\sum_{n=1}^{M-1} nu_{n-1}$ and $\sum_{n=N}^{N+M-2} (N+M-n-1)u_n$ (above) have less than M terms, and since the u_n are bounded, the sums are close to 0 in the norm for N much larger than M.

Also note that $\frac{1}{N}\sum_{n=M-1}^{N-1} u_n \sim \frac{1}{N}\sum_{n=0}^{N-1} u_n$ since the first M terms become insignificant for N much larger than M.

Hence we consider the convergence of the expression $\left\| \frac{1}{N}\frac{1}{M}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} u_{n+m} \right\|$.

Now,

$$\begin{aligned}
\left\| \frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{M}\sum_{m=0}^{M-1} u_{n+m} \right\|^2 &\leq \frac{1}{N}\sum_{n=0}^{N-1}\left\| \frac{1}{M}\sum_{m=0}^{M-1} u_{n+m} \right\|^2 \\
&= \frac{1}{N}\sum_{n=0}^{N-1}\left\langle \frac{1}{M}\sum_{m=0}^{M-1} u_{n+m}, \frac{1}{M}\sum_{m=0}^{M-1} u_{n+m} \right\rangle \\
&= \frac{1}{NM^2}\sum_{n=0}^{N-1}\sum_{m_1,m_2=0}^{M-1} \langle u_{n+m_1}, u_{n+m_2} \rangle \\
&= \frac{1}{M^2}\sum_{m_1,m_2=0}^{M-1}\frac{1}{N}\sum_{n=0}^{N-1} \langle u_{n+m_1}, u_{n+m_2} \rangle .
\end{aligned}$$

As $N \to \infty$, this converges to

$$\frac{1}{M^2}\sum_{m_1,m_2=0}^{M-1} \gamma_{m_2-m_1}.$$

To see this, note that

$$\begin{aligned}
\gamma_{m_2-m_1} &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} \langle u_n, u_{n+m_2-m_1} \rangle \\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{n=-m_1}^{N-m_1-1} \langle u_{n+m_1}, u_{n+m_2} \rangle \\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} \langle u_{n+m_1}, u_{n+m_2} \rangle .
\end{aligned}$$

The last statement follows from the fact that for any bounded sequence $\{a_n\}, n \in \mathbb{Z}$ of complex numbers,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} a_n = \lim_{N\to\infty}\frac{1}{N}\sum_{n=M}^{N+M-1} a_n$$

for any fixed $M \in \mathbb{Z}$.

Let us prove this statement for positive values of M. The proof for negative values is similar.

We have that

$$\lim_{N\to\infty}\frac{1}{N}\left[\sum_{n=0}^{N-1}a_n-\sum_{n=M}^{N+M-1}a_n\right]=\lim_{N\to\infty}\frac{1}{N}\left[\sum_{n=0}^{M-1}a_n-\sum_{n=N}^{N+M-1}a_n\right].$$

Now both

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{M-1}a_n=0 \text{ and } \lim_{N\to\infty}\frac{1}{N}\sum_{n=N}^{N+M-1}a_n=0$$

since the sums are finite and the sequence $\{a_n\}$ is bounded. Hence the result follows.

To proceed with the main proof, we show that

$$\frac{1}{M^2}\sum_{m_1,m_2=0}^{M-1}\gamma_{m_2-m_1}\to 0 \quad as\ M\to\infty.$$

Now

$$\begin{aligned}&\left|\frac{1}{M^2}\sum_{m_1,m_2=0}^{M-1}\gamma_{m_2-m_1}\right|\\ &=\left|\frac{1}{M^2}(M\gamma_0+(M-1)\gamma_1+\ldots+\gamma_{M-1}+(M-1)\gamma_{-1}+(M-2)\gamma_{-2}+\ldots+\gamma_{1-M})\right|\\ &\le\left|\frac{1}{M}\left(\sum_{m=0}^{M-1}\gamma_m+\sum_{m=1}^{M-1}\gamma_{-m}\right)\right|\end{aligned}$$

But $\gamma_m=\overline{\gamma_{-m}}$ by the following direct calculation:

$$\begin{aligned}\gamma_{-m}=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\langle u_n,u_{n-m}\rangle &=\lim_{N\to\infty}\frac{1}{N}\sum_{n=-m}^{N-m-1}\langle u_{n+m},u_n\rangle\\ &=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+m},u_n\rangle\\ &=\overline{\gamma_m}.\end{aligned}$$

Hence if

$$\lim_{M\to\infty}\frac{1}{M}\sum_{m=0}^{M-1}\gamma_m=0,\ \text{then}\ \lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M-1}\gamma_{-m}=0,$$

therefore

$$\frac{1}{M^2}\sum_{m_1,m_2=0}^{M-1}\gamma_{m_2-m_1}\to 0 \quad as\ M\to\infty.$$

So, finally, we can conclude that

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

■

3.3 Weakly mixing of all orders

H. Furstenberg (1983) proved that for any weakly mixing system $(X,\mathcal{B},\mu,T)$, and sets $A_0, A_1, \ldots, A_k$ in $\mathcal{B}$, we have

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

A slightly different proof, which is simpler and clearer than the one Furstenberg provides, will be given here.

This result will also be generalized to be truly "of all orders", in the sense that for any integers $m_0, m_1, \cdots, m_k \in \mathbb{N}$, we have that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\left(\mu(A_0\cap T^{-m_1n}A_1\cap T^{-m_2n}A_2\cap\ldots\cap T^{-m_kn}A_k)-\mu(A_0)\mu(A_1)\ldots\mu(A_k)\right)^2=0.$$

Let us start by clarifying a technical point regarding U, the unitary operator induced by T. Suppose that $m, n \in \mathbb{N}$ with $n \geq m$, and $f, g \in L^2$. We will frequently use the following fact:

$$U^m fU^n g = U^m\left(fU^{n-m}g\right).$$

This follows from the definition of U as follows:

$$\begin{aligned}(U^m fU^n g)(x) &= f(T^m x)g(T^n x)\\ &= f(T^m x)g(T^{n-m}(T^m x))\\ &= f(T^m x)U^{n-m}g(T^m x)\\ &= (fU^{n-m}g)(T^m x)\\ &= U^m(fU^{n-m}g)(x).\end{aligned}$$

Also, in order to prove the next theorem, we should briefly discuss weakly mixing of product systems. If $(X, \mathcal{B}, \mu, T)$ is weakly mixing, then from Theorem 2.3.9(4) we have that the product system $(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu, T \times T)$ is also weakly mixing. Let f_1 and f_2 be functions on X. Denote the tensor product $f_1 \otimes f_2 : X \times X \to \mathbb{C}$ by

$$f_1 \otimes f_2(x_1, x_2) = f_1(x_1) f_2(x_2).$$

If $f_1, f_2 \in L^2(X, \mathcal{B}, \mu)$, then $f_1 \otimes f_2 \in L^2(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu)$, and we have that

$$\int f_1 \otimes f_2 (U \times U)^n g_1 \otimes g_2 \; d(\mu \times \mu) = \int f_1 U^n g_1 \; d\mu \int f_2 U^n g_2 \; d\mu,$$
$$\int f_1 \otimes f_2 \; d(\mu \times \mu) = \int f_1 \; d\mu \int f_2 \; d\mu,$$
$$\int g_1 \otimes g_2 \; d(\mu \times \mu) = \int g_1 \; d\mu \int g_2 \; d\mu.$$

Now we can state and prove the first result regarding weakly mixing of a higher order. Thereafter we will attempt to extend the result even further. We break the theorem up into two parts.

3.3.1 Theorem ([9], p 533)

If $(X, \mathcal{B}, \mu, T)$ is a weakly mixing system and $A_0, A_1, \ldots, A_k$ are sets in $\mathcal{B}$, then

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

Proof:

If we can show that, for any weakly mixing system $(X, \mathcal{B}, \mu, T)$ and any essentially bounded functions, say $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$, that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k} U^{ln} f_l \; d\mu - \prod_{l=0}^{k} \int f_l \; d\mu \right]^2 = 0 \tag{2[k]}$$

then we are done. To see this, replace each f_l in $2[k]$ by the characteristic function χ_{A_l}.

We then get that

$$\begin{aligned} 0 &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k} U^{ln} f_l \; d\mu - \prod_{l=0}^{k} \int f_l \; d\mu \right]^2 \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k} U^{ln} \chi_{A_l} \; d\mu - \prod_{l=0}^{k} \int \chi_{A_l} \; d\mu \right]^2 \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \chi_{A_0 \cap T^{-n} A_1 \cap T^{-2n} A_2 \cap \ldots \cap T^{-kn} A_k} \; d\mu - \mu(A_0)\mu(A_1)\ldots\mu(A_k) \right]^2 \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\mu(A_0 \cap T^{-n} A_1 \cap T^{-2n} A_2 \cap \ldots \cap T^{-kn} A_k) - \mu(A_0)\mu(A_1)\ldots\mu(A_k) \right)^2 . \end{aligned}$$

We will regard $L^\infty(X, \mathcal{B}, \mu)$ as consisting of real-valued functions. This is sufficient since our final result is a statement regarding real numbers.

We will use an induction argument, with the help of the following condition (which will follow from 2[k-1], as will be proven further on):

If $(X, \mathcal{B}, \mu, T)$ is any weakly mixing system and $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$ are arbitrary, then

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{ln} f_l - \prod_{l=1}^{k} \int f_l \; d\mu \right\|_{L^2(X)} = 0. \tag{3[k]}$$

1. **Show that 3[k] implies 2[k]**

   First we show that 3[k]$\Rightarrow$ 2[k] for $(X, \mathcal{B}, \mu, T)$ and any $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$.

   For this purpose, we assume 3[k] for the moment. Since $(X, \mathcal{B}, \mu, T)$ is weakly mixing we have that the strong convergence in 3[k] implies weakly convergence, hence,

$$\left\langle \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{ln} f_l, f \right\rangle \longrightarrow \left\langle \prod_{l=1}^{k} \int f_l \; d\mu, f \right\rangle$$

   for all $f \in L^\infty(X, \mathcal{B}, \mu)$. Hence

$$\left\langle \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{ln} f_l, f_0 \right\rangle \longrightarrow \left\langle \prod_{l=1}^{k} \int f_l \; d\mu, f_0 \right\rangle$$

   and we obtain that

$$\frac{1}{N} \sum_{n=1}^{N} \int f_0 \prod_{l=1}^{k} U^{ln} f_l \; d\mu \longrightarrow \prod_{l=0}^{k} \int f_l \; d\mu \tag{4}$$

As mentioned above, if $(X, \mathcal{B}, \mu, T)$ is weakly mixing, then the product system $(X\times X, \mathcal{B}\times\mathcal{B}, \mu\times\mu, T\times T)$ is also weakly mixing. Hence, since 3[k] is assumed to hold for all weakly mixing systems and some arbitrary elements of L^∞ as stated above, 3[k] also holds for the weakly mixing dynamical system $(X \times X, \mathcal{B}\times\mathcal{B}, \mu\times\mu, T\times T)$ and $f_0 \otimes f_0, f_1 \otimes f_1, \ldots, f_k \otimes f_k \in L^\infty(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu)$, i.e. 3[k] gives

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{ln} f_l \otimes f_l - \prod_{l=1}^{k} \int_{X\times X} f_l \otimes f_l \; d(\mu\times\mu) \right\|_{L^2(X\times X)} = 0.$$

We then obtain an analogue of (4) if we replace X by $X \times X$, μ by $\mu \times \mu$, f_l by $f_l \otimes f_l$ and T by $T \times T$, i.e.: Since $(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu, T \times T)$ is weakly mixing we have that the strong convergence in 3[k] implies weakly convergence, hence,

$$\left\langle \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{ln} f_l \otimes f_l, f \otimes f \right\rangle \longrightarrow \left\langle \prod_{l=1}^{k} \int_{X\times X} f_l \otimes f_l \; d(\mu\times\mu), f \otimes f \right\rangle$$

for all $f \otimes f \in L^\infty(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu)$. Hence

$$\left\langle \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{ln} f_l \otimes f_l, f_0 \otimes f_0 \right\rangle \longrightarrow \left\langle \prod_{l=1}^{k} \int_{X\times X} f_l \otimes f_l \; d(\mu\times\mu), f_0 \otimes f_0 \right\rangle$$

and we obtain that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int_{X\times X} f_0 \otimes f_0 \prod_{l=1}^{k} (T\times T)^{ln} f_l \otimes f_l \; d(\mu\times\mu) \longrightarrow \prod_{l=0}^{k} \int_{X\times X} f_l \otimes f_l \; d(\mu\times\mu)$$

$$\therefore \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int f_0 \prod_{l=1}^{k} U^{ln} f_l \; d\mu \int f_0 \prod_{l=1}^{k} U^{ln} f_l \; d\mu \longrightarrow \prod_{l=0}^{k} \int f_l \; d\mu \int f_l \; d\mu.$$

Hence

$$\therefore \frac{1}{N} \sum_{n=1}^{N} \left[\int f_0 \prod_{l=1}^{k} U^{ln} f_l \; d\mu \right]^2 \longrightarrow \prod_{l=0}^{k} \left[\int f_l \; d\mu \right]^2 \qquad (5)$$

By Proposition 2.3.5 it follows directly from (4) and (5) that 2[k] holds.

We have that the validity of 3[k] for all weakly mixing systems implies the same for 2[k], in short, 3[k]$\Rightarrow$ 2[k].

2. **Show that 2[1] holds:**

   We need to establish that 2[k] holds for all $k \in \mathbb{N}$. 2[1] consists of the following statement:

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\int f_0 U^n f_1 \; d\mu - \int f_0 \; d\mu \int f_1 \; d\mu \right)^2 = 0,$$

   which is true by the assumption of weakly mixing.

3. **Show that 2[k-1] implies 3[k] (which implies 2[k])**

If we can now show that 2[k-1]⇒ 2[k], we are done. Hence we will show that 2[k-1]⇒ 3[k], which will (as shown above) imply 2[k].

We hence assume 2[k-1], i.e.

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\left[\int\prod_{l=0}^{k-1}U^{ln}f_l\ d\mu-\prod_{l=0}^{k-1}\int f_l\ d\mu\right]^2=0$$

for any weakly mixing system $(X,\mathcal{B},\mu,T)$ and any $f_0,f_1,\ldots,f_k\in L^\infty(X,\mathcal{B},\mu)$.

We will now employ our Van der Corput Lemma 3.2.1 to show that 3[k] holds.

For the purposes of this Lemma, let

$$u_n=\prod_{l=1}^{k}U^{ln}f_l-\prod_{l=1}^{k}\int f_l\ d\mu.$$

For convenience, let us use

$$\kappa=\prod_{l=1}^{k}\int f_l\ d\mu\qquad\left(=\prod_{l=0}^{k-1}\int f_{l+1}\ d\mu\right).$$

It follows that we can write, γ_m mentioned in the Van der Corput lemma as

$$\begin{aligned}\gamma_m&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\langle u_n,u_{n+m}\rangle\\&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\left(\prod_{l=1}^{k}U^{ln}f_l-\kappa\right)\left(\prod_{l=1}^{k}U^{l(n+m)}f_l-\kappa\right)\ d\mu\\&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\left(\prod_{l=1}^{k}U^{ln}f_l-\kappa\right)\left(\prod_{l=1}^{k}U^{l(n+m)}f_l-\kappa\right)\ d\mu.\end{aligned}$$

Now,

$$\begin{aligned}&\int\left(\prod_{l=1}^{k}U^{ln}f_l-\kappa\right)\left(\prod_{l=1}^{k}U^{l(n+m)}f_l-\kappa\right)\ d\mu\\&=\int\left(\prod_{l=1}^{k}U^{ln}f_l\prod_{l=1}^{k}U^{l(n+m)}f_l-\kappa\prod_{l=1}^{k}U^{ln}f_l-\kappa\prod_{l=1}^{k}U^{l(n+m)}f_l+\kappa^2\right)\\&=\int\left(\prod_{l=1}^{k}U^{ln}f_lU^{l(n+m)}f_l-\kappa\prod_{l=1}^{k}U^{ln}f_l-\kappa\prod_{l=1}^{k}U^{l(n+m)}f_l+\kappa^2\right)\\&=\int\prod_{l=1}^{k}U^{ln}f_lU^{l(n+m)}f_l\ d\mu-\kappa\int\prod_{l=1}^{k}U^{ln}f_ld\mu-\kappa\int\prod_{l=1}^{k}U^{l(n+m)}f_l\ d\mu+\kappa^2.\end{aligned}$$

Let us scrutinize each of the three integrals above:

(a) **The first integral**:

$$\int \prod_{l=1}^{k} U^{ln} f_l U^{l(n+m)} f_l \, d\mu = \int \prod_{l=1}^{k} U^{ln} f_l \; U^{l(n+m)} f_l \, d\mu$$
$$= \int \prod_{l=1}^{k} U^{ln} (f_l \; U^{lm} f_l) \, d\mu.$$

Since T is measure-preserving, T^n is also measure-preserving and therefore we can replace U^{ln} by $U^{(l-1)n}$, and obtain

$$\int \prod_{l=1}^{k} U^{ln} (f_l \; U^{lm} f_l) \, d\mu = \int U^n \prod_{l=1}^{k} U^{(l-1)n} (f_l \; U^{lm} f_l) \, d\mu$$
$$= \int \prod_{l=1}^{k} U^{(l-1)n} (f_l \; U^{lm} f_l) \, d\mu$$
$$= \int \prod_{l=0}^{k-1} U^{ln} (f_{l+1} \; U^{(l+1)m} f_{l+1}) \, d\mu,$$

and we notice that the integrals occurring are those that occur in 2[k-1], with the functions f_l replaced by $g_{l,m} = f_{l+1} \; U^{(l+1)m} f_{l+1}$.

Then 2[k-1] and Corollary 2.3.8 gives

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} U^{ln} (f_l \; U^{lm} f_l) \, d\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=0}^{k-1} U^{ln} g_{l,m} d\mu$$
$$= \prod_{l=0}^{k-1} \int g_{l,m} \, d\mu$$
$$= \prod_{l=0}^{k-1} \int f_{l+1} \; U^{(l+1)m} f_{l+1} \, d\mu.$$

(b) **The second integral**

$$\begin{aligned}\int \prod_{l=1}^{k} U^{ln} f_l \, d\mu &= \int U^n \prod_{l=1}^{k} U^{(l-1)n} f_l \, d\mu \\ &= \int \prod_{l=1}^{k} U^{(l-1)n} f_l \, d\mu \\ &= \int \prod_{l=0}^{k-1} U^{ln} f_{l+1} \, d\mu.\end{aligned}$$

We again obtain from 2[k-1] and Corollary 2.3.8 that

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} U^{ln} f_l \, d\mu &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=0}^{k-1} U^{ln} f_{l+1} \, d\mu \\ &= \prod_{l=0}^{k-1} \int f_{l+1} \, d\mu \\ &= \kappa.\end{aligned}$$

(c) **The third integral**

Also here we have

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} U^{l(n+m)} f_l \, d\mu &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} U^{ln} \left(U^{lm} f_l\right) \, d\mu \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int U^n \prod_{l=1}^{k} U^{(l-1)n} \left(U^{lm} f_l\right) \, d\mu \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} U^{(l-1)n} \left(U^{lm} f_l\right) \, d\mu \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=0}^{k-1} U^{ln} \left(U^{(l+1)m} f_{l+1}\right) \, d\mu \\ &= \prod_{l=0}^{k-1} \int U^{(l+1)m} f_{l+1} \, d\mu \quad \text{by 2[k-1]} \\ &= \prod_{l=0}^{k-1} \int f_{l+1} \, d\mu \\ &= \kappa.\end{aligned}$$

Hence we see that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\left(\int\prod_{l=1}^{k}U^{ln}f_lU^{l(n+m)}f_ld\mu-\kappa\int\prod_{l=1}^{k}U^{ln}f_ld\mu-\kappa\int\prod_{l=1}^{k}U^{l(n+m)}f_ld\mu+\kappa^2\right)$$
$$=\prod_{l=0}^{k-1}\int f_{l+1}\;U^{(l+1)m}f_{l+1}\;d\mu-\kappa^2-\kappa^2+\kappa^2 \qquad ,$$
$$=\prod_{l=0}^{k-1}\int f_{l+1}\;U^{(l+1)m}f_{l+1}\;d\mu-\kappa^2.$$

i.e.

$$\gamma_m=\prod_{l=0}^{k-1}\int f_{l+1}\;U^{(l+1)m}f_{l+1}\;d\mu-\kappa^2.$$

We see that

$$\begin{aligned}\lim_{m\to\infty}\gamma_m &= \lim_{m\to\infty}\left(\prod_{l=0}^{k-1}\int f_{l+1}\;U^{(l+1)m}f_{l+1}\;d\mu-\kappa^2\right)\\ &=\prod_{l=0}^{k-1}\left(\int f_{l+1}\;d\mu\int f_{l+1}\;d\mu\right)-\kappa^2 \quad \text{by weakly mixing}\\ &=\kappa^2-\kappa^2\\ &=0.\end{aligned}$$

It now follows by Proposition 2.3.1 that the average

$$\lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}|\gamma_m|=0,$$

and by Proposition 2.3.2 that

$$\lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}\gamma_m=0.$$

Finally, by the Van der Corput Lemma, we have that

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=1}^{N}u_n\right\|=\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=1}^{N}\prod_{l=1}^{k}U^{ln}f_l-\prod_{l=1}^{k}\int f_l\;d\mu\right\|=0.$$

■

In Theorem 1.3.1.2 we proved a Mean Ergodic Theorem for dynamical systems and in Theorem 2.1.3 it was shown that if $(X,\mathcal{B},\mu,T)$ is ergodic then

$$\bar{f}(x)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}f(T^kx)=\int_X fd\mu \quad \text{a.e.}$$

Hence

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k - \int_X f d\mu \right\|_{L^2} (X) \longrightarrow 0$$

as $n \to \infty$.

The following theorem extends this result for weakly mixing systems to a *multiple Mean Ergodic Theorem.*

3.3.2 Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a weakly mixing system and $f_0, f_1, \ldots, f_k$ be functions in $L^\infty(X, \mathcal{B}, \mu)$. Then

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{ln} f_l - \prod_{l=1}^{k} \int f_l \, d\mu \right\|_{L^2(X)} = 0.$$

Proof:

This result follows by the same induction argument used in the previous theorem. 3[1] is satisfied immediately, following our discussion prior to this theorem. In the previous theorem we have proved that 2[k] holds for all $k \in \mathbb{N}$ and that 2[k-1] implies 3[k] for $k \geq 2$. Hence 3[k] is true for all $k \in \mathbb{N}$. ■

We will now extend Theorem 3.3.1. It will be seen that weakly mixing implies "weakly mixing of all orders". The proof of the next theorem will follow the same general lines as the previous theorem, but we will encounter a few difficulties which did not pose a problem in the case where the "powers" of U were chosen in a regular fashion, i.e. $0, n, 2n, ..., nk$. In this instance we will replace the powers $0, n, 2n, ...$ by any choice of non-negative integers $m_0, m_1, m_2, \cdots, m_k$ such that $0 = m_0 \leq m_1 \leq \cdots \leq m_k$. It will also be seen that the order requirement $m_0 \leq m_1 \leq \cdots \leq m_k$ can be dropped due to the bijectivity of the unitary operator U.

3.3.3 Theorem

If $(X, \mathcal{B}, \mu, T)$ is a weakly mixing system and $A_0, A_1, \ldots, A_k$ are sets in $\mathcal{B}$, and if $m_0, m_1, m_2, \cdots, m_k$ is any choice of non-negative integers such that

$$0 = m_0 < m_1 < m_2 < \cdots < m_k,$$

then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\mu(A_0 \cap T^{-m_1 n} A_1 \cap T^{-m_2 n} A_2 \cap \ldots \cap T^{-m_k n} A_k) - \mu(A_0)\mu(A_1)\ldots\mu(A_k) \right)^2 = 0. \tag{6}$$

Proof:

We will follow the same general plan of the proof of the previous theorem, and give all steps for completeness. Once again we will regard $L^\infty(X, \mathcal{B}, \mu)$ as consisting of real-valued functions.

If we can show that, for any weakly mixing system $(X, \mathcal{B}, \mu, T)$ and any essentially bounded functions, say $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$, that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k} U^{m_l n} f_l \, d\mu - \prod_{l=0}^{k} \int f_l \, d\mu \right]^2 = 0 \qquad (7[\mathrm{k}])$$

for any integers $0 = m_0 < m_1 < m_2 < \cdots < m_k$, then we are done.

To see this, replace each f_l in $7[k]$ by the characteristic function χ_{A_l}.

We then get that

$$\begin{aligned}
0 &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k} U^{m_l n} f_l \, d\mu - \prod_{l=0}^{k} \int f_l \, d\mu \right]^2 \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k} U^{m_l n} \chi_{A_l} \, d\mu - \prod_{l=0}^{k} \int \chi_{A_l} \, d\mu \right]^2 \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \chi_{T^{-m_0 n} A_0 \cap T^{-m_1 n} A_1 \cap T^{-m_2 n} A_2 \cap \ldots \cap T^{-m_k n} A_k} \, d\mu - \mu(A_0)\mu(A_1)\ldots\mu(A_k) \right]^2 \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\mu(T^{-m_0 n} A_0 \cap T^{-m_1 n} A_1 \cap T^{-m_2 n} A_2 \cap \ldots \cap T^{-m_k n} A_k) - \mu(A_0)\mu(A_1)\ldots\mu(A_k) \right)^2 \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\mu(A_0 \cap T^{-m_1 n} A_1 \cap T^{-m_2 n} A_2 \cap \ldots \cap T^{-m_k n} A_k) - \mu(A_0)\mu(A_1)\ldots\mu(A_k) \right)^2
\end{aligned}$$

We will use basically the same induction argument as before, i.e. with the help of the following condition (which will follow from 7[k-1], as will be proven further on):

If $(X, \mathcal{B}, \mu, T)$ is any weakly mixing system and $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$ are arbitrary, then

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{m_l n} f_l - \prod_{l=1}^{k} \int f_l \, d\mu \right\|_{L^2(X)} = 0 \qquad (8[\mathrm{k}])$$

1. **Show that 8[k] implies 7[k]**

First we show that 8[k]⇒ 7[k] for $(X, \mathcal{B}, \mu, T)$ and any $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$.

For this purpose, we assume 8[k] for the moment. Since $(X, \mathcal{B}, \mu, T)$ is weakly mixing we have that the strong convergence in 8[k] implies weak convergence, hence,

$$\left\langle \frac{1}{N}\sum_{n=1}^{N}\prod_{l=1}^{k} U^{m_l} f_l, f \right\rangle \longrightarrow \left\langle \prod_{l=1}^{k}\int f_l \, d\mu, f \right\rangle$$

for all $f \in L^\infty(X, \mathcal{B}, \mu)$. Hence

$$\left\langle \frac{1}{N}\sum_{n=1}^{N}\prod_{l=1}^{k} U^{m_l n} f_l, f_0 \right\rangle \longrightarrow \left\langle \prod_{l=1}^{k}\int f_l \, d\mu, f_0 \right\rangle.$$

Hence

$$\frac{1}{N}\sum_{n=1}^{N}\int f_0 \prod_{l=1}^{k} U^{m_l n} f_l \, d\mu \longrightarrow \int f_0 \, d\mu \prod_{l=1}^{k}\int f_l \, d\mu,$$

and so we obtain that

$$\frac{1}{N}\sum_{n=1}^{N}\int \prod_{l=0}^{k} U^{m_l n} f_l \, d\mu \longrightarrow \prod_{l=0}^{k}\int f_l \, d\mu \tag{9}$$

noting that $m_0 = 0$.

Now, as in the previous theorem, it follows from the fact that the product system is weakly mixing, that we can replace f_l by $f_l \otimes f_l$ and T by $T \times T$. We then obtain that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int_{X\times X} f_0 \otimes f_0 \prod_{l=1}^{k}(T\times T)^{m_l n} f_l \otimes f_l \, d(\mu\times\mu) \longrightarrow \prod_{l=0}^{k}\int_{X\times X} f_l \otimes f_l \, d(\mu\times\mu)$$
$$\therefore \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int f_0 \prod_{l=1}^{k} U^{m_l n} f_l \, d\mu \int f_0 \prod_{l=1}^{k} U^{m_l n} f_l \, d\mu \longrightarrow \prod_{l=0}^{k}\int f_l \, d\mu \int f_l \, d\mu.$$

Hence

$$\therefore \frac{1}{N}\sum_{n=1}^{N}\left[\int f_0 \prod_{l=1}^{k} U^{m_l n} f_l \, d\mu\right]^2 \longrightarrow \prod_{l=0}^{k}\left[\int f_l \, d\mu\right]^2 \tag{10}$$

By Proposition 2.3.5 it follows directly from (9) and (10) that 7[k] holds.

Hence we see that the validity of 8[k] for all weakly mixing systems implies the same for 7[k], i.e., 8[k]⇒ 7[k].

2. **Show that 7[1] holds**

Now back to the induction argument. We need to establish that 7[k] holds for all $k \in \mathbb{N}$. Now, 7[1] consists of the following statement:

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\int U^{m_0 n} f_0 U^{m_1 n} f_1 \; d\mu - \int f_0 \; d\mu \int f_1 \; d\mu \right)^2 = 0,$$

i.e.

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\int f_0 U^{m_1 n} f_1 \; d\mu - \int f_0 \; d\mu \int f_1 \; d\mu \right)^2 = 0,$$

which is immediately true by the assumption of weakly mixing of T, hence of T^{m_1}.

3. **Show that 7[k-1] implies 8[k] (which implies 7[k])**

If we can now show that 7[k-1]⇒ 7[k], we are done. Hence we will show that 7[k-1]⇒ 8[k], which will (as shown above) imply 7[k].

We now assume 7[k-1], i.e.

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k-1} U^{m_l n} f_l \; d\mu - \prod_{l=0}^{k-1} \int f_l \; d\mu \right]^2 = 0,$$

for any k non-negative integers $m_0, m_1, m_2, \cdots, m_k$ such that $0 = m_0 < m_1 < m_2 < \cdots < m_{k-1}$ and any $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$.

We will employ the Van der Corput Lemma 3.2.1 to show that 8[k] holds.

We let

$$u_n = \prod_{l=1}^{k} U^{m_l n} f_l - \prod_{l=1}^{k} \int f_l \; d\mu,$$

and for convenience we use

$$\kappa = \prod_{l=1}^{k} \int f_l \; d\mu \quad \left(= \prod_{l=0}^{k-1} \int f_{l+1} \; d\mu \right).$$

For $r \in \mathbb{N}$, let

$$\begin{aligned} \gamma_r &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle u_n, u_{n+r} \rangle \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \left(\prod_{l=1}^{k} U^{m_l n} f_l - \kappa \right) \left(\prod_{l=1}^{k} U^{m_l (n+r)} f_l - \kappa \right) \; d\mu. \end{aligned}$$

Now

$$
\begin{aligned}
&\int \left(\prod_{l=1}^{k} U^{m_l n} f_l - \kappa\right)\left(\prod_{l=1}^{k} U^{m_l (n+r)} f_l - \kappa\right)\, d\mu \\
&= \int \left(\prod_{l=1}^{k} U^{m_l n} f_l \prod_{l=1}^{k} U^{m_l (n+r)} f_l - \kappa \prod_{l=1}^{k} U^{m_l n} f_l - \kappa \prod_{l=1}^{k} U^{m_l (n+r)} f_l + \kappa^2\right)\, d\mu \\
&= \int \left(\prod_{l=1}^{k} U^{m_l n} f_l U^{m_l (n+r)} f_l - \kappa \prod_{l=1}^{k} U^{m_l n} f_l - \kappa \prod_{l=1}^{k} U^{m_l (n+r)} f_l + \kappa^2\right)\, d\mu \\
&= \int \prod_{l=1}^{k} U^{m_l n} f_l U^{m_l (n+r)} f_l \, d\mu - \kappa \int \prod_{l=1}^{k} U^{m_l n} f_l \, d\mu - \kappa \int \prod_{l=1}^{k} U^{m_l (n+r)} f_l \, d\mu + \kappa^2,
\end{aligned}
$$

the second step in the array above following from commutativity.

Use the notation

$$
a_{n,r} = \int \prod_{l=1}^{k} U^{m_l n} f_l U^{m_l (n+r)} f_l \, d\mu - \kappa \int \prod_{l=1}^{k} U^{m_l n} f_l \, d\mu - \kappa \int \prod_{l=1}^{k} U^{m_l (n+r)} f_l \, d\mu + \kappa^2.
$$

Similar to the proof of the previous theorem, we will scrutinize the three integrals in the last expression each separately.

(a) **The first integral**

$$
\begin{aligned}
\int \prod_{l=1}^{k} U^{m_l n} f_l \; U^{m_l (n+r)} f_l \, d\mu &= \int \prod_{l=1}^{k} U^{m_l n} (f_l \; U^{m_l r} f_l)\, d\mu \\
&= \int U^{m_1 n} \left(\prod_{l=1}^{k} U^{(m_l - m_1) n} (f_l \; U^{m_l r} f_l)\right)\, d\mu \\
&= \int \prod_{l=1}^{k} U^{(m_l - m_1) n} (f_l \; U^{m_l r} f_l)\, d\mu \\
&= \int \prod_{l=0}^{k-1} U^{(m_{l+1} - m_1) n} (f_{l+1} \; U^{m_{l+1} r} f_{l+1})\, d\mu.
\end{aligned}
$$

If we now replace the integers $m_0, m_1, m_2, \cdots, m_{k-1}$ in 7[k-1] by the following integers:

$$
m_0' = m_1 - m_1 = 0,\; m_1' = m_2 - m_1,\; m_2' = m_3 - m_1,\; \ldots,\; m_{k-1}' = m_k - m_1,
$$

we obtain integers $m_0', m_1', m_2', \ldots m_{k-1}'$ that are non-negative with

$$
0 = m_0' < m_1' < m_2' < \cdots < m_{k-1}'.
$$

If we also replace each f_l in 7[k-1] by

$$g_{l,r} = f_{l+1}\ U^{m_{l+1}r} f_{l+1},$$

we obtain that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k-1} U^{m_l' n} \left(f_{l+1}\ U^{m_{l+1}r} f_{l+1} \right)\ d\mu - \prod_{l=0}^{k-1} \int f_{l+1}\ U^{m_{l+1}r} f_{l+1}\ d\mu \right]^2$$
$$= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\int \prod_{l=0}^{k-1} U^{m_l' n} g_{l,r}\ d\mu - \prod_{l=0}^{k-1} \int g_{l,r}\ d\mu \right]^2$$
$$= 0.$$

Thus by 7[k-1] and Corollary 2.3.8 we find that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=0}^{k-1} U^{m_l' n} \left(f_{l+1}\ U^{m_{l+1}r} f_{l+1} \right)\ d\mu \ = \prod_{l=0}^{k-1} \int f_{l+1}\ U^{m_{l+1}r} f_{l+1}\ d\mu.$$

(b) **The second integral**

We proceed in a similar fashion with the second integral, again using 7[k-1] and Corollary 2.3.8.

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} U^{m_l n} f_l\ d\mu \ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int U^{m_1 n} \prod_{l=1}^{k} U^{(m_l - m_1)n} f_l\ d\mu \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} U^{(m_l - m_1)n} f_l\ d\mu \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=0}^{k-1} U^{(m_{l+1} - m_1)n} f_{l+1}\ d\mu \\
&= \prod_{l=0}^{k-1} \int f_{l+1}\ d\mu \\
&= \kappa.
\end{aligned}$$

(c) **The third integral**

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\prod_{l=1}^{k}U^{m_l(n+r)}f_l\ d\mu \\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\prod_{l=1}^{k}U^{m_l n}\left(U^{m_l r}f_l\right)\ d\mu \\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int U^{m_1 n}\prod_{l=1}^{k}U^{(m_l-m_1)n}\left(U^{m_l r}f_l\right)\ d\mu \\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\prod_{l=1}^{k}U^{(m_l-m_1)n}\left(U^{m_l r}f_l\right)\ d\mu \\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\prod_{l=0}^{k-1}U^{(m_{l+1}-m_1)n}\left(U^{m_{l+1} r}f_{l+1}\right)\ d\mu \\
&=\prod_{l=0}^{k-1}\int U^{m_{l+1}r}f_{l+1}\ d\mu \\
&=\prod_{l=0}^{k-1}\int f_{l+1}\ d\mu \\
&=\ \kappa.
\end{aligned}$$

After considering these integrals and averaging them over N we hence obtained that

$$\begin{aligned}
\gamma_r &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}a_{n,r} \\
&=\prod_{l=0}^{k-1}\int f_{l+1}\ U^{m_{l+1}r}f_{l+1}\ d\mu-\kappa^2-\kappa^2+\kappa^2 \\
&=\prod_{l=0}^{k-1}\int f_{l+1}\ U^{m_{l+1}r}f_{l+1}\ d\mu-\kappa^2.
\end{aligned}$$

We now see that by weakly mixing of U, and hence of $U^{m_{l+1}}$, that

$$\begin{aligned}
\lim_{r\to\infty}\gamma_r &= \lim_{r\to\infty}\left(\prod_{l=0}^{k-1}\int f_{l+1}\ U^{m_{l+1}r}f_{l+1}\ d\mu-\kappa^2\right) \\
&=\prod_{l=0}^{k-1}\int f_{l+1}\ d\mu\int f_{l+1}\ d\mu-\kappa^2 \\
&=\kappa^2-\kappa^2 \\
&=0.
\end{aligned}$$

Hence the average

$$\lim_{M\to\infty} \frac{1}{M} \sum_{r=1}^{M} \gamma_r = 0$$

by Propositions 2.3.1 and 2.3.2.

Finally, by the Van der Corput Lemma, we have that

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} u_n \right\| = \lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{m_l n} f_l - \prod_{l=1}^{k} \int f_l \, d\mu \right\| = 0.$$

■

It should be noted that the restriction that $0 = m_0 < m_1 < m_2 < \cdots < m_k$ is introduced for the sake of convenience. If $m_j \neq m_{j'}$ for $j \neq j'$, and if we do not require this ordering, we can rearrange and renumber terms (which is possible due to commutativity) so that we can follow the same argument as we did above to prove weakly mixing of all orders.

Hence the order restriction $0 = m_0 < m_1 < m_2 < \cdots < m_k$ may indeed be dropped.

Of course, Theorem 3.3.2 can be extended, by the same reasoning, to the statement

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} U^{m_l n} f_l - \prod_{l=1}^{k} \int f_l \, d\mu \right\|_{L^2(X)} = 0$$

where the integers m_l are to be understood as in Theorem 3.3.3.

Chapter 4

Non-commutative ergodic theory

In Chapter 4, a basic familiarity with operator algebras and in particular C^*-algebra theory is assumed. An introduction to these theories may be found in [18] as well as in *Fundamentals of the theory of operator algebras*, by Kadison and Ringrose.

4.1 The Gelfand-Naimark-Segal (GNS) construction

A powerful tool for the study of ergodic theory in non-commutative dynamical systems as discussed in [18], p 93, will be introduced next.

We will see that the GNS construction provides us with a way to regard every C^*-algebra as a C^*-subalgebra of $B(H)$ for some Hilbert space H (with $B(H)$ denoting the set of all bounded linear maps $H \to H$).

A *representation* of a C^*-algebra A is a pair (H, π) where H is a Hilbert space and $\pi : A \to B(H)$ is a $*$-homomorphism. (H, π) is said to be *faithful* if π is injective.

We will now set out to construct a representation for any given C^*-algebra.

Now, if H is an inner product space (pre-Hilbert space), then there is a unique inner product on the Banach space completion $\hat{H}$ of H extending the inner product of H and having as its associated norm the norm of $\hat{H}$ ([16], Theorem 3.2.3). We call $\hat{H}$ endowed with this inner product the *Hilbert space completion* of H.

With each positive linear functional we can associate a representation. Suppose that φ is a positive linear functional on a C^*-algebra A.

If

$$N_\varphi = \{a \in A \,|\varphi(a^*a) = 0\},$$

then N_φ is a closed left ideal of A and the map

$$(A/N_\varphi) \times (A/N_\varphi) \to \mathbb{C}, \quad \langle a + N_\varphi, b + N_\varphi\rangle \mapsto \varphi(b^*a),$$

is a well-defined inner product on A/N_φ. We now denote the Hilbert space completion of A/N_φ by H_φ.

If $a \in A$, then an operator $\pi(a) \in B(A/N_\varphi)$ can be defined by

$$\pi(a)(b + N_\varphi) = ab + N_\varphi.$$

The inequality $\|\pi(a)\| \leq \|a\|$ holds since we have

$$\|\pi(a)(b + N_\varphi)\|^2 = \varphi(b^*a^*ab) \leq \|a\|^2\varphi(b^*b) = \|a\|^2 \|b + N_\varphi\|^2 .$$

The operator $\pi(a)$ has a unique extension to a bounded operator $\pi_\varphi(a)$ on H_φ. The map

$$\pi_\varphi : A \to B(H_\varphi), \quad a \mapsto \pi_\varphi(a),$$

is a $*$-isomorphism.

The representation (H_φ, π_φ) of A is the *Gelfand-Naimark-Segal representation* (or *GNS representation*) associated to φ.

Some remarks on notation

As in [20], for any u in a Hilbert space H, ω_u stands for the linear functional $B(H) \ni x \mapsto \langle xu, u\rangle$. We let u_φ denote the canonical cyclic vector of π_φ, satisfying $\varphi = \omega_{u_\varphi} \circ \pi_\varphi$. We then see that $\|u_\varphi\|^2 = \|\varphi\|$ ($= 1$ if φ is a state). If the C^*-algebra $\mathcal{U}$ is unital, then $\pi_\varphi(1_\mathcal{U}) = 1_{H_\varphi}$, and so $\pi_\varphi(1_\mathcal{U})u_\varphi = u_\varphi$ and $\varphi(1_\mathcal{U}) = \|\varphi\|$ ($= 1$ if φ is a state).

4.2 C^*-dynamical systems

Let us proceed to introduce the basic notions of ergodic theory in a C^*-algebra theoretical framework.

In the framework of C^*-algebra theory, the measure-preserving transformation T acting on the probability space $(X, \mathcal{B}, \mu)$ will be replaced by a $*$-endomorphism Φ with

$$\Phi : L^\infty(X, \mathcal{B}, \mu) \to L^\infty(X, \mathcal{B}, \mu), \quad \Phi(f) = f \circ T.$$

If we define the map

$$\varphi : L^\infty(X, \mathcal{B}, \mu) \to \mathbb{C}, \quad \varphi(f) = \int_X f \; d\mu,$$

it follows that φ is a faithful normal state.

The fact that φ is faithful can easily be seen from the basic properties of integrals, i.e. if $f > 0$ then $\varphi(f) = \int_X f \; d\mu > 0$, noting that $f\overline{f} = \overline{f}f$ shows that φ is normal.

The map φ also leaves Φ invariant since

$$\varphi\left(\Phi(f)\right) = \int \Phi(f) \; d\mu = \int f \circ T \; d\mu = \int f \; d\mu = \varphi(f).$$

Let $\mathcal{U}$ be a C^*-algebra, φ a state on $\mathcal{U}$ and $\Phi : \mathcal{U} \to \mathcal{U}$ a positive linear map such that

$$\varphi \circ \Phi = \varphi \text{ and } \varphi\left(\Phi(x)^*\Phi(x)\right) \leq \varphi(x^*x) \text{ for every } x \in \mathcal{U}.$$

We can then define the notions ergodicity and weakly mixing as follows: We say that

- Φ is *ergodic* with respect to φ if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left(\varphi(y\Phi^n(x)) - \varphi(y)\varphi(x)\right) = 0 \quad \text{for all } x, y \in \mathcal{U};$$

- Φ is *weakly mixing* with respect to φ if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left|\varphi(y\Phi^n(x)) - \varphi(y)\varphi(x)\right| = 0 \quad \text{for all } x, y \in \mathcal{U}.$$

It is clear that these definitions of ergodicity and weakly mixing are extensions of the commutative case where $\mathcal{U} = L^\infty(X, \mathcal{B}, \mu)$ and x and y replaced by $\chi_A, \chi_B \in L^\infty(X, \mathcal{B}, \mu)$ respectively, i.e.

$$\begin{aligned}
\varphi(y\Phi^n(x)) - \varphi(y)\varphi(x) &= \int_X \chi_B(x)\chi_A(T^n x)\, d\mu - \int_X \chi_B(x)\, d\mu \int_X \chi_A(x)\, d\mu \\
&= \int_X \chi_{B\cap T^{-n}A}(x)\, d\mu - \mu(B)\mu(A) \\
&= \mu(T^{-n}A \cap B) - \mu(A)\mu(B).
\end{aligned}$$

A pair $(\mathcal{U}, \Phi)$ where $\mathcal{U}$ is a C^*-algebra and Φ a $*$-homomorphism $\Phi : \mathcal{U} \to \mathcal{U}$ is called a *C^*-dynamical system.*

If, in addition, we have a state φ such that $\varphi \circ \Phi = \varphi$, i.e. we have a C^*-dynamical system that leaves a state invariant, then we call the triplet $(\mathcal{U}, \varphi, \Phi)$ a *state preserving C^*-dynamical system.*

We say that the state preserving C^*-dynamical system $(\mathcal{U}, \varphi, \Phi)$ is *ergodic* if Φ is ergodic with respect to φ and *weakly mixing* if Φ is weakly mixing with respect to φ, as defined above.

We see that any measure-preserving dynamical system $(X, \mathcal{B}, \mu, T)$ can be thought of as a state preserving C^*-dynamical system $(L^\infty(\mu), \varphi, \Phi)$ by letting, as before,

$$\Phi(f) = f \circ T \text{ and } \varphi(f) = \int_X f\, d\mu.$$

4.3 Non-commutative Khintchine and Poincairé Recurrence Theorem

Using Hilbert space techniques, we can prove the following non-commutative generalization of the Khintchine recurrence theorem. From this result we will then derive a similar generalization for the Poincairé Recurrence Theorem.

4.3.1 THEOREM ([19], p 2)

Let $\mathcal{U}$ be an algebra with unit $\mathbf{1}$, and let $\Phi : \mathcal{U} \to \mathcal{U}$ be a C^*-morphism of unital algebras.

If there exists a state $\varphi \in \mathcal{U}^*$ such that $\varphi \circ \Phi = \varphi$, then for every $a \in \mathcal{U}$ and every $\varepsilon > 0$ there exists a relatively dense subset E of $\mathbb{N}$ such that

$$\operatorname{Re} \varphi(\Phi^n(a^*)a) \geq |\varphi(a)|^2 - \varepsilon$$

for every $n \in E$.

To prove this theorem, we will first prove the following, more general result, which is similar to Theorem 1.3.1.3.

4.3.2 LEMMA ([19], p 3)

Let H be a vector space endowed with a hermitian form $\langle \cdot, \cdot \rangle$ and let $\| \cdot \|$ be the corresponding norm. Suppose that there exists a linear operator $U : H \to H$ such that

$$\begin{aligned} \|Ux\| &= \|x\| \text{ for every } x \in H \\ Uv &= v \text{ for some } v \in H, \ \|v\| = 1. \end{aligned}$$

Then, for every $x \in H$ and every $\varepsilon > 0$ there exists a relatively dense subset E of $\mathbb{N}$ such that

$$\operatorname{Re} \langle U^n x, x \rangle \geq |\langle x, x \rangle|^2 - \varepsilon$$

for every $n \in E$.

Proof:

By the Mean Ergodic Theorem there exists a projection P on $\{z|Uz = z\}$ such that for every $\varepsilon > 0$ and every $x \in H$ we can find an $N \in \mathbb{N}$ such that

$$\left\| \frac{1}{N} \sum_{k=0}^{N-1} U^k x - Px \right\|^2 \leq \frac{\varepsilon}{2}.$$

Let us write

$$x_N = \frac{1}{N}\sum_{k=0}^{N-1} U^k x.$$

Since $\|Ux\| = \|x\|$ for every $x \in H$, i.e. U is a contraction, and since $UP = P$, we also have that

$$\left\|U^l x_N - Px\right\|^2 \leq \frac{\varepsilon}{2}$$

for every $l \in \mathbb{N}$. We then obtain that

$$\begin{aligned}
\left\|U^l x_N - x_N\right\|^2 &= \left\|U^l x_N - Px + Px - x_N\right\|^2 \\
&\leq \left(\left\|U^l x_N - Px\right\| + \|x_N - Px\|\right)^2 \\
&= \left\|U^l x_N - Px\right\|^2 + 2\left\|U^l x_N - Px\right\| \|x_N - Px\| + \|x_N - Px\|^2 \\
&\leq \tfrac{\varepsilon}{2} + 2\tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} \\
&= 2\varepsilon
\end{aligned}$$

for every $l \in \mathbb{N}$.

Since

$$\begin{aligned}
\left\|U^l x_N - x_N\right\|^2 &= \langle U^l x_N - x_N, U^l x_N - x_N\rangle \\
&= \|U^l x_N\| - \langle U^l x_N, x_N\rangle - \langle x_N, U^l x_N\rangle + \|x_N\|^2 \\
&= 2\|x_N\|^2 - 2Re\langle U^l x_N, x_N\rangle,
\end{aligned}$$

we have that

$$\text{Re } \left\langle U^l x_N, x_N\right\rangle \geq \|x_N\|^2 - \varepsilon$$

for every $l \geq N-1$.

On the other hand, we have that

$$\begin{aligned}
0 \leq \|x_N - \langle x_N, v\rangle v\|^2 &= \langle x_N - \langle x_N, v\rangle v, x_N - \langle x_N, v\rangle v\rangle \\
&= \|x_N\|^2 - \langle\langle x_N, v\rangle v, x_N\rangle - \langle x_N, \langle x_N, v\rangle v\rangle + \langle\langle x_N, v\rangle v, \langle x_N, v\rangle v\rangle \\
&= \|x_N\|^2 - |\langle x_N, v\rangle|^2,
\end{aligned}$$

and

$$\langle x_N, v\rangle = \left\langle \frac{1}{N}\sum_{k=0}^{N-1} U^k x, v\right\rangle = \frac{1}{N}\sum_{k=0}^{N-1}\left\langle U^k x, v\right\rangle = \frac{1}{N}\sum_{k=0}^{N-1}\left\langle x, (U^*)^k v\right\rangle = \frac{1}{N}\sum_{k=0}^{N-1}\langle x, v\rangle = \langle x, v\rangle$$

since $U^* v = v$ ([22], p 408). This can be seen from the following:

We know that

$$1 = \|U\| = \|U^*\|$$

and hence that

$$1 = \|v\|^2 = \langle v, v\rangle = \langle Uv, v\rangle = \langle v, U^*v\rangle \leq \|v\|\|U^*v\| \leq \|v\|\|U^*\|\|v\| = \|v\|^2 = 1.$$

Therefore

$$\langle v, U^*v\rangle = \|v\|\|U^*v\|$$

and

$$\|v\|\|U^*v\| = \|v\|^2, \quad \text{i.e } \|U^*v\| = \|v\|,$$

which implies that

$$\|v - U^*v\|^2 = \|v\|^2 - \langle v, U^*v\rangle - \langle U^*v, v\rangle + \|U^*v\|^2 = 0.$$

Thus we conclude that $U^*v = v$.

Now back to the main line of proof.

We have that

$$|\langle x, v\rangle|^2 \leq \|x_N\|^2 \leq \varepsilon + \text{Re } \left\langle U^l x_N, x_N\right\rangle.$$

Since we have that

$$\begin{aligned}
\text{Re } \left\langle U^l x_N, x_N\right\rangle &= \text{Re } \left\langle \frac{1}{N}\sum_{k=0}^{N-1} U^{k+l}x, \frac{1}{N}\sum_{j=0}^{N-1} U^j x\right\rangle \\
&= \frac{1}{N^2}\sum_{j,k=0}^{N-1} \text{Re } \left\langle U^{l+k}x, U^j x\right\rangle \\
&= \frac{1}{N^2}\sum_{j,k=0}^{N-1} \text{Re } \left\langle U^{l+k-j}x, x\right\rangle
\end{aligned}$$

and hence for each integer $n \in \mathbb{N}^*$ there exist integers $j(n), k(n) \in [0, N-1]$ such that

$$\text{Re } \left\langle U^{nN+k(n)-j(n)}x, x\right\rangle \geq \|x_N\|^2 - \varepsilon.$$

Furthermore, the set $E = \{nN + k(n) - j(n) | n \in \mathbb{N}^*\}$ is relatively dense in $\mathbb{N}$ because

$$(n-1)N \leq nN + k(n) - j(n) \leq (n+1)N$$

and hence E contains an element in every interval of length $2N$. ■

Proof of Theorem 4.3.1:

If we endow $\mathcal{U}$ with the GNS-hermitian product $\langle x, y\rangle = \varphi(y^*x)$, and choose as U the mapping Φ and as v the unit $\mathbf{1}$ of $\mathcal{U}$, then Theorem 4.3.1 follows from Lemma 4.3.2. ■

This theorem can be further extended for arbitrary C^*-algebras $\mathcal{U}$ and linear mappings $\Phi : \mathcal{U} \to \mathcal{U}$ such that

$$\varphi(\Phi(a)^*\Phi(a)) \leq \varphi(a^*a) \text{ for every } a \in \mathcal{U}.$$

In that instance, the proof of this assertion is much more involved [20].

For the purposes of the next theorem, let us introduce the terms upper and lower density of sets ([8], p 73 and [20], p 46). If S is a finite set, we denote its cardinality by $|S|$ or by $Card(S)$. We define the *upper density* $D^*(S)$ and the *lower density* $D_*(S)$ of some set $S \subset \mathbb{N}$ by

$$D^*(S) = \limsup_{n\to\infty} \frac{1}{n+1}\sum_{k=0}^{n} \chi_S(k) = \limsup_{n\to\infty} \frac{Card(S \cap [0,n])}{n+1},$$

and

$$D_*(S) = \liminf_{n\to\infty} \frac{1}{n+1}\sum_{k=0}^{n} \chi_S(k) = \liminf_{n\to\infty} \frac{Card(S \cap [0,n])}{n+1}.$$

The set S is said to have *density* $D(S)$ if $D^*(S) = D_*(S)$.

We recall that we say that a set $S \subset \mathbb{N}$ is relatively dense if the set has bounded gaps, i.e there is an $L > 0$ such that every interval of natural numbers with length L or higher will intersect S. Hence if S is relatively dense, then if $n > L$, $Card(S \cap [0,n]) \geq \frac{n}{L} - 1$, i.e. $D_*(S) \geq \liminf_{n\to\infty} \dfrac{n-L}{L(n+1)} = \dfrac{1}{L} > 0$. Thus we have that relatively dense sets have positive lower density.

We may note at this stage that it is not true that every set of positive lower density is relatively dense. A counter-example was supplied by R. Duvenhage. Let $E = \{k_j\}_{j=1}^{\infty}$ be such that $k_1 = 1$ and the sequence of differences $b_j = k_{j+1} - k_j$ is given by

$$\{b_1, b_2, \ldots\} = \{2, 1, 3, 1, 1, 4, 1, 1, 1, 5, 1, 1, 1, 1, 6, 1, 1, 1, 1, 1, 7, \ldots\},$$

i.e.

$$E = \{1, \quad 3, 4, \quad 7, 8, 9, \quad 13, 14, 15, 16, \quad 21, 22, 23, 24, 25, \quad 31, 32, 33, 34, 35, 36, \quad 43, \ldots\}.$$

This example was constructed to ensure that the number of elements of E grows at the same rate as the magnitude of the total number of *gaps* or *jumps* in the sequence. We find that all terms

$$a_n := \frac{1}{n}\sum_{k=1}^{n} \chi_E(k) \geq \frac{1}{2}, \quad n = 1, 2, \ldots.$$

In particular, each term $a_{n(n+1)} = \frac{1}{2}$, $n = 1, 2, \ldots$. This means that

$$\liminf_{n\to\infty} a_n = \frac{1}{2},$$

which means that $D_*(E) = \frac{1}{2} > 0$. Since it is clear that E does not have bounded gaps, the counter-example is complete.

From the non-commutative Khintchine theorem follows the non-commutative analogue of Poincaré's Recurrence Theorem.

4.3.3 Theorem ([19], p 3)

Let $(\mathcal{U}, \varphi, \Phi)$ be a state preserving C^*-dynamical system. Then

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \left|\varphi(\Phi^k(a^*)a)\right| > 0$$

for every $a \in \mathcal{U}$ with $\varphi(a) \neq 0$.

Proof:

We will show that for any bounded sequence $\{a_k\}_{k\geqslant 1}$ of nonnegative integers, the following two conditions are equivalent:

1. $\liminf_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} a_k > 0$;
2. $\liminf_{N\to\infty} \frac{Card\{k \in [0,N] | a_k > \varepsilon\}}{N} > 0$ for some $\varepsilon > 0$.

Hence we prove that

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} a_k = 0$$

if and only if

$$\liminf_{N\to\infty} \frac{Card\{k \in [0,N] | a_k > \varepsilon\}}{N} = 0 \text{ for all } \varepsilon > 0.$$

Note: For every nonnegative sequence $\{a_k\}_{k\geqslant 0}$ we have the following: If $\liminf_{N\to\infty} a_k = 0$, then we can construct a subsequence (using the definition of the limit inferior repeatedly) $\{a_{n_k}\}_{k\geqslant 0}$ with $1 \leq n_0 < n_1 < n_2 < \ldots$ such that $\lim_{k\to\infty} a_{n_k} = 0$. Conversely, since $\liminf_{k\to\infty} a_k$ is the smallest number that can be obtained as a limit of a subsequence of $\{a_k\}_{k\geqslant 0}$, we have that $\lim_{k\to\infty} a_{n_k} = 0$ implies that $\liminf_{k\to\infty} a_k = 0$, since the sequence is nonnegative.

From this, our required equivalence follows: We have that

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} a_k = 0$$

if and only if there is a subsequence $\{a_{n_k}\}_{k\geqslant 0}$ with $1 \leq n_0 < n_1 < n_2 < \ldots$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} a_{n_k} = 0$$

if and only if (by Lemma 2.3.3)

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \chi_{E'_\varepsilon}(k) = 0 \quad \text{for all } \varepsilon > 0, \quad \text{with } E'_\varepsilon = \{n_k : a_{n_k} > \varepsilon\}$$

if and only if

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} \chi_{E_\varepsilon}(k) = 0 \quad \text{for all } \varepsilon > 0, \quad \text{with } E_\varepsilon = \{k : a_k > \varepsilon\}$$

(since $\{\chi_{E'_\varepsilon}(k)\}$ is a subsequence of $\{\chi_{E_\varepsilon}(k)\}$). Hence the equivalence above, being the contrapositive of the equivalence of 1 and 2 earlier above, is proved.

Now, from the generalized Khintchine theorem, since there is a relatively dense subset F of $\mathbb{N}$ such that

$$\text{Re } \varphi\left(\Phi^k(a^*)a\right) \geq |\varphi(a)|^2 - \varepsilon$$

for every $k \in F$, and for every $\varepsilon > 0$, we also have (following the remark preceding this theorem) that the set

$$\{k \in \mathbb{N} : \text{Re } \varphi\left(\Phi^k(a^*)a\right) > \varepsilon\}$$

has positive lower density, for some $\varepsilon > 0$, i.e.

$$\liminf_{N\to\infty} \frac{Card\left\{k \in [0, N] \ : \ Re\ \varphi\left(\Phi^k(a^*)a\right) > \varepsilon\right\}}{N} > 0 \text{ for some } \varepsilon > 0,$$

which implies that

$$\liminf_{N\to\infty} \frac{Card\left\{k \in [0, N] \ : \ \left|\varphi\left(\Phi^k(a^*)a\right)\right| > \varepsilon\right\}}{N} > 0 \text{ for some } \varepsilon > 0.$$

The theorem now follows by the equivalence stated above. ■.

4.4 Multiple Weakly Mixing

We will conclude this dissertation by illustrating some remarkable results by Niculescu, Ströh and Zsidó [20], where it is illustrated that many of the multiple recurrence and multiple weakly mixing results discussed in Chapter 3 can be extended to non-commutative dynamical systems under rather weak restrictions.

In this dissertation it should be noted that all discussions are restricted to C^*-algebras with unity. However, Niculescu, Ströh and Zsidó [20] have shown that the requirement that the C^*-algebra contains a unit element, can be dropped.

4.4.1 Preliminary concepts

We state a few definitions and results that will serve as background to the definitions and results for non-commutative dynamical systems.

1. **Van der Corput type results**

   As mentioned earlier, we now construct a zero density set similar to the construction in the Koopman-Von Neumann Lemma.

   4.4.1.1 Theorem ([20], p 46)

   Let $E_1, E_2, ... \subset \mathbb{N}$ be a sequence of zero density subsets. Then there exists a zero density subset $E \subset \mathbb{N}$ such that $E_j \backslash E$ is finite for every $j \geq 1$.

   Proof:

   Choose by induction a sequence $1 \leq n_1 < n_2 < ...$ of integers such that

   $$\frac{1}{n+1}\sum_{k=0}^{n}\chi_{E_j}(k) \leq \frac{1}{p^2} \text{ for } n \geq n_p \text{ and } j = 1, 2, ..., p$$

   and let $E = \bigcup_{j=1}^{\infty}(E_j \cap (n_j, +\infty))$. We then have that $E_j \backslash E \subset [0, n_j], \quad j = 1, 2, \ldots$ and hence $E_j \backslash E$ is finite.

   For $n_p < n \leq n_{p+1}$ we have that $\chi_E(k) \leq \sum_{j=1}^{p}\chi_{E_j}(k)$. This is so since if $k \in E$, then also $k \in \bigcup_{j=1}^{p} E_j$. We then have that

   $$\frac{1}{n+1}\sum_{k=0}^{n}\chi_E(k) \leq \frac{1}{n+1}\sum_{k=0}^{n}\sum_{j=1}^{p}\chi_{E_j}(k) = \sum_{j=1}^{p}\frac{1}{n+1}\sum_{k=0}^{n}\chi_{E_j}(k) \leq p\frac{1}{p^2} = \frac{1}{p},$$

   so E has density zero. ■

4.4.1.2 Theorem ([20], p 47)

If $\{u_n\}_{n\geq 0}$ is a bounded sequence in a Hilbert space H then the following statements are equivalent:

(a) $D-\lim\limits_{n\to\infty}\langle u_n,u\rangle=0$ *for all* $u\in H$;

(b) $D-\lim\limits_{n\to\infty} u_n=0$ *with respect to the weak topology of* H.

Proof:

$(b)\Rightarrow(a)$ follows immediately by [16], 3.8-1 p.188.

We prove $(a)\Rightarrow(b)$.

Let H_0 denote the closed linear span of $\{u_n : n\geq 0\}$. By (a) there are subsets $E_j\subset\mathbb{N}$ of density zero such that $\lim\limits_{n\to\infty,\ n\notin E_j}\langle u_n,u_j\rangle=0$ for all $j\geq 0$ and by Theorem 4.4.1.1 there exists a subset $E\subset\mathbb{N}$ of density zero with $E_j\backslash E$ finite for all $j\geq 0$. Then $\lim\limits_{n\to\infty,\ n\notin E}\langle u_n,u_j\rangle=0$ for all $j\geq 0$, so the closed linear subspace $\{u\in H:\lim\limits_{n\to\infty,\ n\notin E}\langle u_n,u\rangle=0\}\subset H$ contains the sequence $\{u_j\}_{j\geq 0}$, hence all of H_0. Since it trivially contains $H\backslash H_0$, it is equal to H. Hence, (b) holds. ■

4.4.1.3 Proposition ([20], p 43)

If $1\leq n\geq h\geq d\geq 0$ are natural numbers and $a_{1-h},...,a_{n+h-d}$ are elements of an additive semigroup then

$$\sum_{k=1}^{n+h}\sum_{j=k-h}^{k-d}a_j=(h-d+1)\sum_{j=1-d}^{n}a_j+\sum_{j=1}^{h-d}(h-d+1-j)(a_{1-d-j}+a_{n+j}).$$

Proof:

By rewriting and re-assigning indexes, we get that

$$\sum_{k=1}^{n+h}\sum_{j=k-h}^{k-d}a_j=\sum_{j=1-h}^{-d}\sum_{k=1}^{j+h}a_j+\sum_{j=1-d}^{n}\sum_{k=j+d}^{j+h}a_j+\sum_{j=n+1}^{n+h-d}\sum_{k=j+d}^{n+h}a_j$$

$$=\sum_{j=1-h}^{-d}(j+h)a_j+\sum_{j=1-d}^{n}(h-d+1)a_j+\sum_{j=n+1}^{n+h-d}(n+h-j-d+1)a_j.$$

We also have that

$$\begin{aligned}\sum_{j=1-h}^{-d}(j+h)a_j &= \sum_{j=-(-d)}^{-(1-h)}(-j+h)a_{-j}\\ &= \sum_{j=d}^{h-1}(h-j)a_{-j}\\ &= \sum_{j=1}^{h-d}(h-d+1-j)a_{1-d-j}\end{aligned}$$

and

$$\sum_{j=n+1}^{n+h-d}(n+h-j-d+1)a_j = \sum_{j=1}^{h-d}(h-d+1-j)a_{n+j}.$$

Therefore,

$$\sum_{k=1}^{n+h}\sum_{j=k-h}^{k-d}a_j = (h-d+1)\sum_{j=1-d}^{n}a_j + \sum_{j=1}^{h-d}(h-d+1-j)(a_{1-d-j}+a_{n+j}).$$

■

Proposition 4.4.1.3 yields the following:

4.4.1.4 Proposition ([20], p 43)

If $1 \le n \ge h \ge d \ge 0$ *are natural numbers and* $a_1, ..., a_n$ *are elements of an additive semigroup with neutral element, then putting* $a_j = 0$ *for* $j \le 0$ *and for* $j \ge n+1$*, we have*

$$\sum_{k=1}^{n+h}\sum_{j=k-h}^{k-d}a_j = (h-d+1)\sum_{j=1}^{n}a_j.$$

We also have a counterpart of Proposition 4.4.1.3 for double sums:

4.4.1.5 PROPOSITION ([20], p 43)

If $1 \leq n \geq h \geq 0$ *are natural numbers and* $a_{j,j'}, 1-h \leq j, j' \leq n+h$ *are elements of an additive semigroup then*

$$\sum_{k=1}^{n+h} \sum_{j,j'=k-h}^{k} a_{j,j'} = (h+1)\sum_{j=1}^{n} a_{j,j'} + \sum_{d=1}^{h}(h-d+1) \sum_{j=1-d}^{n} (a_{j,j'+d} + a_{j+d,j})$$
$$+\sum_{d=1}^{h}(h-d+1)(a_{1-d,1-d} + a_{n+d,n+d})$$
$$+\sum_{d=1}^{h}\sum_{j=1}^{h-d}(h-d+1-j)(a_{1-d-j,1-j} + a_{1-j,1-d-j} + a_{n+j,n+d+j} + a_{n+d+j,n+j}).$$

Proof:

We show this by direct computation. First we separate the "diagonal" terms in the sum, and rewrite the rest. We have that

$$\begin{aligned}
\sum_{k=1}^{n+h} \sum_{j,j'=k-h}^{k} a_{j,j'} &= \sum_{k=1}^{n+h} \sum_{j=k-h}^{k} a_{j,j} + \sum_{k=1}^{n+h} \sum_{j,j'=k-h,\ j<j'}^{k} (a_{j,j'} + a_{j',j}) \\
&= \sum_{k=1}^{n+h} \sum_{j=k-h}^{k} a_{j,j} + \sum_{k=1}^{n+h}\sum_{d=1}^{h}\sum_{j=k-h}^{k-d} (a_{j,j+d} + a_{j+d,j}) \\
&= \sum_{k=1}^{n+h} \sum_{j=k-h}^{k} a_{j,j} + \sum_{d=1}^{h}\sum_{k=1}^{n+h}\sum_{j=k-h}^{k-d} (a_{j,j+d} + a_{j+d,j}).
\end{aligned}$$

By Proposition 4.4.1.3 we have that

$$\sum_{k=1}^{n+h} \sum_{j=k-h}^{k} a_{j,j} = (h+1)\sum_{j=1}^{n} a_{j,j} + \sum_{j=1}^{h}(h+1-j)(a_{1-j,1-j} + a_{n+j,n+j})$$

and

$$\sum_{k=1}^{n+h}\sum_{j=1}^{k-d}(a_{j,j+d} + a_{j+d,j}) = (h-d+1) \sum_{j=1-d}^{n} (a_{j,j+d} + a_{j+d,j})$$
$$+\sum_{j=1}^{h-d}(h-d-j+1)(a_{1-d-j,1-j} + a_{1-j,1-d-j} + a_{n+j,n+d+j} + a_{n+d+j,n+j}).$$

This yields the required equality. ■

By Proposition 4.4.1.5 we have the following special case:

4.4.1.6 Proposition ([20], p 44)

If $1 \leq n \geq h \geq 0$ *are natural numbers and* $a_{j,j'}, 1 \leq j, j' \leq n$ *are elements of an additive semigroup with neutral element then, putting* $a_{j,j'} = 0$ *for* j *or* $j' \leq 0$ *and for* j *or* $j' \geq n+1$*, we have*

$$\sum_{k=1}^{n+h} \sum_{j,j'=k-h}^{k} a_{j,j'} = (h+1) \sum_{j=1}^{n} a_{j,j} + \sum_{d=1}^{h} (h-d+1) \sum_{j=1}^{n} (a_{j,j+d} + a_{j+d,j}).$$

■

The following result gives a counterpart for the Cauchy inequality for normed vector spaces.

4.4.1.7 Proposition ([20], p 44)

If $n \geq 1$ *is a natural number and* $a_1, ..., a_n$ *are elements of a* $*$*-algebra then*

$$\left(\sum_{k=1}^{n} a_k\right)^* \left(\sum_{k=1}^{n} a_k\right) \leq n \sum_{k=1}^{n} a_k^* a_k. \tag{4.1}$$

Proof:

We prove by induction.

The case where $n = 1$ follows immediately. Assume, then, that it holds for some $n \geq 1$ and let $a_1, ..., a_n, a_{n+1}$ be elements of a $*$-algebra.

Note:

$$\begin{aligned} a_k^* a_{n+1} + a_{n+1}^* a_k &= a_k^* a_k + a_{n+1}^* a_{n+1} - (a_k - a_{n+1})^*(a_k - a_{n+1}) \\ &\leq a_k^* a_k + a_{n+1}^* a_{n+1}, \end{aligned}$$

since $(a_k - a_{n+1})^*(a_k - a_{n+1})$ is a positive element.

Then

$$\begin{aligned}\left(\sum_{k=1}^{n+1} a_k\right)^* \left(\sum_{k=1}^{n+1} a_k\right) &= \left(\sum_{k=1}^{n} a_k + a_{n+1}\right)^* \left(\sum_{k=1}^{n} a_k + a_{n+1}\right) \\ &= \left(\sum_{k=1}^{n} a_k\right)^* \left(\sum_{k=1}^{n} a_k\right) + \sum_{k=1}^{n}(a_k^* a_{n+1} + a_{n+1}^* a_k) + a_{n+1}^* a_{n+1} \\ &\leq n\sum_{k=1}^{n} a_k^* a_k + \sum_{k=1}^{n}(a_k^* a_k + a_{n+1}^* a_{n+1}) + a_{n+1}^* a_{n+1} \\ &= n\sum_{k=1}^{n} a_k^* a_k + \sum_{k=1}^{n} a_k^* a_k + n a_{n+1}^* a_{n+1} + a_{n+1}^* a_{n+1} \\ &= (n+1)\sum_{k=1}^{n+1} a_k^* a_k,\end{aligned}$$

where the inequality follows by using the induction assumption and the note above.

■

4.4.1.8 Proposition ([20], p 45)

If $1 \leq n \geq h \geq 0$ *are natural numbers and* $a_1, ..., a_n$ *are elements of a* $*$*-algebra then*

$$\begin{aligned}&(h+1)^2 \left(\sum_{j=1}^{n} a_j\right)^* \left(\sum_{j=1}^{n} a_j\right) \\ &\leq (n+h)(h+1)\sum_{j=1}^{n} a_j^* a_j + 2(n+h)\sum_{d=1}^{h}(h-d+1) Re \sum_{j=1}^{n} a_j^* a_{j+d},\end{aligned}$$

where $Re(a) = \frac{1}{2}(a + a^*)$.

Proof:

Put $a_j = 0$ for $j \leq 0$ and for $j \geq n+1$. Using successively Proposition 4.4.1.4, Proposition 4.4.1.7 and Proposition 4.4.1.6, we get

$$(h+1)^2\left(\sum_{j=1}^{n}a_j\right)^*\left(\sum_{j=1}^{n}a_j\right)=\left(\sum_{k=1}^{n+h}\sum_{j=k-h}^{k}a_j\right)^*\left(\sum_{k=1}^{n+h}\sum_{j=k-h}^{k}a_j\right)$$

$$\leq (n+h)\sum_{k=1}^{n+h}\left(\sum_{j=k-h}^{k}a_j\right)^*\left(\sum_{j=k-h}^{k}a_j\right)=(n+h)\sum_{k=1}^{n+h}\sum_{j,j'=k-h}^{k}a_j^*a_{j'}$$

$$=(n+h)\left((h+1)\sum_{j=1}^{n}a_j^*a_j+\sum_{d=1}^{h}(h-d+1)\sum_{j=1}^{n}(a_j^*a_{j+d}+a_{j+d}^*a_j)\right)$$

$$=(n+h)\left((h+1)\sum_{j=1}^{n}a_j^*a_j+\sum_{d=1}^{h}(h-d+1)\sum_{j=1}^{n}(a_j^*a_{j+d}+(a_j^*a_{j+d})^*)\right)$$

$$=(n+h)(h+1)\sum_{j=1}^{n}a_j^*a_j+(n+h)\sum_{d=1}^{h}(h-d+1)\sum_{j=1}^{n}2Re(a_j^*a_{j+d}).$$

■

A Van der Corput type inequality in Hilbert space then follows:

4.4.1.9 PROPOSITION ([20], p 46)

If $1\leq n\geq h\geq 0$ *are natural numbers and* $u_1,...,u_n$ *are vectors in a Hilbert space* H, *then*

$$(h+1)^2\left\|\sum_{j=1}^{n}u_j\right\|^2$$
$$\leq (n+h)(h+1)\sum_{j=1}^{n}\|u_j\|^2+2(n+h)\sum_{d=1}^{h}(h-d+1)Re\sum_{j=1}^{n}\langle u_j,u_{j+d}\rangle.$$

Proof:

Choose bounded linear operators a_j with $u_j=a_ju_1$ in the following way:

Let H_{u_1} be the close linear span of u_1. Let $H=H_{u_1}\oplus H_{u_1}^{\perp}$. For any $u\in H$, let $u=\alpha u_1+y$, where $y\in H_{u_1}^{\perp}$. Now define $a_ju_1=u_j$ on H_{u_1} and $a_jy=0$ on $H_{u_1}^{\perp}$. We then have that

$$a_ju=\alpha a_ju_1=\alpha u_j.$$

It is clear that each operator a_j is bounded (since each u_j is bounded) and linear, since for each $u,v\in H$, with $u=\alpha u_1+y$ and $v=\beta u_1+z$, where $y,z\in H_{u_1}^{\perp}$,

$$a_j(\gamma u+v)=(\gamma\alpha+\beta)u_j=\gamma a_ju+a_jv.$$

Hence $a_j \in B(H),\ j = 1, 2, ..., n$.

Now apply Proposition 4.4.1.8 to these $a_1, ..., a_n$, and take the value of the positive linear operator ω_{u_1} on both sides of the inequality. We then get that

$$\omega_{u_1}\left((h+1)^2\left(\sum_{j=1}^{n} a_j\right)^*\left(\sum_{j=1}^{n} a_j\right)\right)$$
$$\leq \omega_{u_1}\left((n+h)(h+1)\sum_{j=1}^{n} a_j^* a_j + 2(n+h)\sum_{d=1}^{h}(h-d+1)Re\sum_{j=1}^{n} a_j^* a_{j+d}\right)$$

i.e.

$$(h+1)^2\left\|\sum_{j=1}^{n} a_j\right\|^2$$
$$\leq (n+h)(h+1)\sum_{j=1}^{n}\|a_j\|^2 + 2(n+h)\sum_{d=1}^{h}(h-d+1)Re\sum_{j=1}^{n}\langle a_{j+d}, a_j\rangle.$$

Since we have that

$$\|a_j\|^2 = \omega_{u_1}(a_j^* a_j) = \langle a_j u_1, a_j u_1\rangle = \langle u_j, u_j\rangle = \|u_j\|^2$$

and similarly that

$$\left\|\sum_{j=1}^{n} a_j\right\|^2 = \left\|\sum_{j=1}^{n} u_j\right\|^2,$$

the result follows.

■

4.4.1.10 Theorem ([20], p 37)

Let $u_1, u_2, ...$ be a bounded sequence in a Hilbert space H. *If*

$$\lim_{h\to\infty}\limsup_{n\to\infty}\frac{1}{h}\sum_{d=1}^{h}\frac{1}{n}\sum_{k=1}^{n}|Re\langle u_{k+d}, u_k\rangle| = 0, \tag{4.2}$$

then

$$\lim_{n\to\infty}\left\|\frac{1}{n}\sum_{j=1}^{n} u_{k_j}\right\| = 0$$

for every relatively dense set $1 \leq k_1 < k_2 < ... \in \mathbb{N}$.

Proof:

First we have to prove that a strictly increasing sequence $\{k_j\}_{j\geqslant 1}$ is relatively dense if and only if $L = \sup_{j\geqslant 1}(k_{j+1} - k_j) < +\infty$.

If $\{k_j\}_{j\geqslant 1}$ is relatively dense, then there is a positive integer K such that every interval of natural numbers of length $\geqslant K$ contains some element of $\{k_j\}_{j\geqslant 1}$. Hence $L \leq K < +\infty$. Conversely, if $M < +\infty$, then every interval of length M will contain some element of $\{k_j\}_{j\geqslant 1}$, and hence it is relatively dense.

Now we continue with the main proof. Let $1 \leq k_1 < k_2 < ...$ in $\mathbb{N}$ be relatively dense. Let $k_0 = 0$ and let $L = \sup_{j\geqslant 1}(k_{j+1} - k_j)$. Then we have that $d \leq k_{j+d} - k_j \leq L \cdot d$ for all $j, d \geq 0$, in particular $j \leq k_j \leq L \cdot j$.

Denote $c = \sup \|u_k\| < +\infty$. By Proposition 4.4.1.9 we get for any natural numbers $n \geq h \geq 1$:

$$\begin{aligned}
\left\| \frac{1}{n} \sum_{j=1}^{n} u_{k_j} \right\|^2 &\leq \frac{n+h}{n^2(h+1)} \sum_{j=1}^{n} \|u_{k_j}\|^2 + \frac{2(n+h)}{n^2(h+1)} \sum_{d=1}^{h} \frac{h-d+1}{h+1} Re \sum_{j=1}^{n} \langle u_{k_{j+d}}, u_{k_j} \rangle \\
&\leq \frac{n+h}{n(h+1)} c^2 + 2\left(1 + \frac{h}{n}\right) \frac{1}{n(h+1)} \sum_{l=1}^{L\cdot h} \sum_{j=1}^{n} |Re\langle u_{k_j+l}, u_{k_j}\rangle| \\
&\leq \left(\frac{1}{n} + \frac{1}{h+1}\right) c^2 + \frac{4}{n(h+1)} \sum_{l=1}^{L\cdot h} \sum_{k=1}^{L\cdot n} |Re\langle u_{k+l}, u_k\rangle| \,.
\end{aligned}$$

Given any $\varepsilon > 0$, choose an integer h_ε such that $\frac{c^2}{h_\varepsilon} \leq \varepsilon$. Then for any $h \geq h_\varepsilon$ we have that

$$\limsup_{n\to\infty} \left\| \frac{1}{n} \sum_{j=1}^{n} u_{k_j} \right\|^2 \leq \varepsilon + \limsup_{n\to\infty} \frac{4}{h} \sum_{d=1}^{L\cdot h} \frac{1}{n} \sum_{k=1}^{L\cdot n} |Re\langle u_{k+d}, u_k\rangle| \,.$$

Hence

$$\limsup_{n\to\infty} \left\| \frac{1}{n} \sum_{j=1}^{n} u_{k_j} \right\|^2 \leq \varepsilon + 4 \lim_{h\to\infty} \limsup_{n\to\infty} \frac{1}{h} \sum_{d=1}^{L\cdot h} \frac{1}{n} \sum_{k=1}^{L\cdot n} |Re\langle u_{k+d}, u_k\rangle| = \varepsilon$$

by (4.2).

Therefore

$$\limsup_{n\to\infty} \left\| \frac{1}{n} \sum_{j=1}^{n} u_{k_j} \right\| = 0.$$

Since

$$0 \leq \liminf_{n\to\infty} \left\| \frac{1}{n} \sum_{j=1}^{n} u_{k_j} \right\| \leq \limsup_{n\to\infty} \left\| \frac{1}{n} \sum_{j=1}^{n} u_{k_j} \right\| = 0,$$

we see from the equality of the limit inferior and the limit superior that

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{j=1}^{n} u_{k_j} \right\| = 0.$$

■

Weakly mixing and uniformly weakly mixing sequences

Let H be a Hilbert space and $\{u_k\}_{k\geqslant 1}$ a bounded sequence in H. Then we say that $\{u_k\}_{k\geqslant 1}$ is *weakly mixing to zero* if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} |\langle u_k, u\rangle| = 0 \Leftrightarrow D - \lim_{k\to\infty} \langle u_k, u\rangle = 0 \text{ for all } u \in H.$$

We say that $\{u_k\}_{k\geqslant 1}$ is *uniformly weakly mixing to zero* if

$$\lim_{N\to\infty} \left(\sup_{\|u\|\leqslant 1} \left(\frac{1}{N} \sum_{k=1}^{N} |\langle u_k, u\rangle| \right) \right) = 0.$$

4.4.1.11 PROPOSITION

Let H be a Hilbert space and $\{u_k\}_{k\geqslant 1}$ a bounded sequence in H. If $\{u_k\}_{k\geqslant 1}$ is uniformly weakly mixing to zero, then it is also weakly mixing to zero.

Proof:

Suppose that $\{u_k\}_{k\geqslant 1}$ is uniformly weakly mixing to zero. Let $u \in H$. If $u = 0$ then trivially $\frac{1}{N}\sum_{k=1}^{N} |\langle u_k, u\rangle| = 0$. If $u \neq 0$, let $u' = \frac{u}{\|u\|}$, so that $\|u'\| = 1$.

Since $\{u_k\}_{k\geqslant 1}$ is uniformly weakly mixing to zero, and

$$0 \leq \frac{1}{N} \sum_{k=1}^{N} |\langle u_k, u'\rangle| \leq \sup_{\|u\|\leqslant 1} \left(\frac{1}{N} \sum_{k=1}^{N} |\langle u_k, u\rangle| \right),$$

it follows by the squeeze theorem that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} |\langle u_k, u'\rangle| = 0.$$

Therefore,

$$0 = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \left| \left\langle u_k, \frac{u}{\|u\|} \right\rangle \right| = \frac{1}{\|u\|} \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} |\langle u_k, u\rangle|,$$

and so

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} |\langle u_k, u\rangle| = 0.$$

Since u is arbitrary, we have that $\{u_k\}_{k\geqslant 1}$ is weakly mixing to zero. ■

4.4.1.12 THEOREM ([20], p 49)

For a bounded sequence $\{u_k\}_{k\geqslant 1}$ *in a Hilbert space* H*, the following are equivalent:*

(i) $\{u_k\}_{k\geqslant 1}$ *is uniformly weakly mixing to zero.*

(ii) *For every sequence* $1 \leq k_1 < k_2 < \ldots$ *in* $\mathbb{N}$ *of positive lower density,*

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{j=1}^{N} u_{k_j} \right\| = 0.$$

(iii) *For every relatively dense sequence* $1 \leq k_1 < k_2 < \ldots$ *in* $\mathbb{N}$*,*

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{j=1}^{N} u_{k_j} \right\| = 0.$$

Hence we see that for every bounded sequence $u_1, u_2, \ldots$ in a Hilbert space H, (4.2) implies that $\{u_k\}_{k\geqslant 1}$ is uniformly weakly mixing to zero (hence weakly mixing to zero).

The proof of this result is given in [23]. This paper is being submitted for publication and hence the proof will not be given here. The result above is a slight extension of the following interesting characterization for uniform weak convergence (not in density):

4.4.1.13 PROPOSITION ([2], p 240)

For a bounded sequence $\{u_k\}_{k\geqslant 1}$ in a Hilbert space H, the following are equivalent:

(a) *For any $\varepsilon > 0$, there exists a $K = K(\varepsilon)$ such that for any $v \in H$ with $\|v\| \leq 1$ there exists a set $P \subset \mathbb{N}$ with $|P| \leqslant K$ such that*

$$|\langle u_k, v\rangle| \geq \varepsilon \Rightarrow k \in P.$$

(b) *For any increasing sequence $\{k_j\} \subset \mathbb{N}$*

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{j=1}^{N} u_{k_j} \right\| = 0.$$

Proof:

First we note that from (a) we have that for any $\varepsilon > 0$, there exists a $K = K(\varepsilon)$ such that for any $m \in \mathbb{N}$ there exists a set $P \subset \mathbb{N}$ with $|P| \leqslant K$ such that

$$|\langle u_k, u_m\rangle| \geq \varepsilon \Rightarrow k \in P.$$

To see this, let $\varepsilon > 0$ be arbitrary and choose any $m \in \mathbb{N}$. Define $u'_m = \frac{u_m}{\|u_m\|}$, so that $\|u'_m\| = 1$. Then, by (a) there is a K as well as a set $P \subset \mathbb{N}$ with $|P| \leqslant K$ such that

$$|\langle u_k, u'_m\rangle| \geq \frac{\varepsilon}{\|u_m\|} \Rightarrow k \in P,$$

i.e.

$$\left|\left\langle u_k, \frac{u_m}{\|u_m\|}\right\rangle\right| \geq \frac{\varepsilon}{\|u_m\|} \Rightarrow k \in P,$$

so

$$|\langle u_k, u_m\rangle| \geq \varepsilon \Rightarrow k \in P.$$

Since ε and m are arbitrary, the implication is proved.

Now we proceed to show that (a)$\Rightarrow$(b):

Let $\varepsilon > 0$. Take $K = K(\frac{\varepsilon^2}{2})$ as discussed above, and $L > \frac{2K}{\varepsilon^2}$. If $N > L$ and $k_1 < k_2 < \cdots < k_N$, then

$$\begin{aligned} \left\| \frac{1}{N} \sum_{j=1}^{N} u_{k_j} \right\|^2 &\leq \frac{1}{N^2} \sum_{i,j=1}^{N} |\langle u_{k_i}, u_{k_j}\rangle| \\ &\leq \frac{1}{N^2} \left(KN + \frac{N^2\varepsilon^2}{2} \right) \\ &< \frac{K}{L} + \frac{\varepsilon^2}{2} \;<\; \varepsilon^2 \end{aligned}$$

and thus

$$\left\| \frac{1}{N} \sum_{j=1}^{N} u_{k_j} \right\| < \varepsilon.$$

(b)⇒(a): Suppose that (1) does not hold. From the first part of this proof we then see that we can find an $\varepsilon > 0$ such that for each $m \in \mathbb{N}$ there exists a $v_m \in H$ with $\|v_m\| \leq 1$ and a set $P_m \subset \mathbb{N}$ such that $\lim_{m\to\infty} |P_m| \to \infty$ and

$$|\langle u_k, v_m \rangle| \geq \varepsilon, \quad k \in P_m.$$

For each m the set P_m is finite and hence we can replace P_m by a subset thereof and assume that all the numbers $\langle u_k, v_m \rangle$ lie in the same quadrant in the complex plane for $k \in P_m$, and so that

$$\begin{aligned} \left\| \frac{1}{|P_m|} \sum_{k\in P_m} u_k \right\| &= \sup_{\|v\|\leqslant 1} \left| \left\langle \frac{1}{|P_m|} \sum_{k\in P_m} u_k, v \right\rangle \right| \\ &\geq \frac{1}{|P_m|} \left| \sum_{k\in P_m} \langle u_k, v_m \rangle \right| \\ &= \frac{1}{|P_m|} \sum_{k\in P_m} |\langle u_k, v_m \rangle| \\ &\geq \tfrac{\varepsilon}{4}. \end{aligned}$$

The sets P_m may be constructed in such a way that the sequence $\{|P_m|\}_{m=1}^{\infty}$ grows arbitrarily fast and that $\max(P_m) < \min(P_{m-1})$ for each m. Now let

$$P = \bigcup_{m=1}^{\infty} P_m.$$

We then have that for each $N \in \mathbb{N}$ and each sequence $k_1, k_2, \ldots, k_N$ there exists an m such that the said sequence is completely contained in the set P_m. Hence we have that

$$\left\| \frac{1}{N} \sum_{j=1}^{N} u_{k_j} \right\| \geq \frac{\varepsilon}{4}$$

for some $\varepsilon > 0$ as discussed above. This contradicts (2). ■

4.4.1.14 Corollary ([20], p 39)

Let $u_1, u_2, ...$ be a bounded sequence in a Hilbert space H such that

$$D - \lim_{d\to\infty}\left(D - \limsup_{n\to\infty}|Re\langle u_{n+d}, u_n\rangle|\right) = 0.$$

Then $\{u_k\}_{k\geqslant 1}$ is uniformly weakly mixing to zero, hence weakly mixing to zero.

Proof:

Since

$$D - \lim_{d\to\infty}\left(D - \limsup_{n\to\infty}|Re\langle u_{n+d}, u_n\rangle|\right) = 0$$

implies (4.2), which implies (by Theorem 4.4.1.10) that

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{j=1}^{N}u_{k_j}\right\| = 0$$

for every relatively dense sequence $1 \leq k_1 < k_2 < ...$ in $\mathbb{N}$, it follows from Theorem 4.4.1.12 that $\{u_k\}_{k\geqslant 1}$ is uniformly weakly mixing to zero, and by Proposition 4.4.1.11 that it is also weakly mixing to zero.

■

2. **Weakly mixing and symmetrically weakly mixing in non-commutative dynamical systems**

We continue this section by introducing a few new concepts. Some of the concepts are introduced in [20] to simplify the proofs, which may become quite tedious without these inventions. In the proofs of Theorems 3.3.1 and 3.3.3 regarding weakly mixing of higher orders we worked in commutative spaces, where we were able to write

$$\prod_{l=1}^{k}U^{m_l n}f_l\prod_{l=1}^{k}U^{m_l(n+r)}f_l = \prod_{l=1}^{k}U^{m_l n}f_l U^{m_l(n+r)}f_l.$$

Commutativity was a key factor in the strategy used to solve those problems. In the present situation, where we study non-commutative spaces, we need to devise ways to work around non-commutativity. Hence we will introduce a new concept *symmetrically weakly mixing.* We will later also introduce a weak form of commutativity, i.e. *asymptotic abelianness in density*, under which we will see that we can extend Theorems 3.3.1 and 3.3.3.

We first introduce the ideas of Multiple Weakly Mixing (WM) and Symmetrically Weakly Mixing (SWM) as introduced in [20].

Let $(\mathcal{U}, \varphi, \Phi)$ be a state preserving C^*-dynamical system. For any integers $k \geq 1$ and nonnegative integers $m_1, ..., m_k \geq 1,\ m_j \neq m_{j'}$ for $j \neq j'$, we define:

- $(\mathrm{WM}_{m_1,...,m_k})$ $(\mathcal{U}, \varphi, \Phi)$ is *weakly mixing of order* $(m_1, ..., m_k)$ if

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left| \varphi\left(\prod_{l=0}^{k} \Phi^{m_l n}(x_l) \right) - \prod_{l=0}^{k} \varphi(x_l) \right| = 0$$

for all $x_0, x_1, ..., x_k \in \mathcal{U}$ and with $m_0 = 0$;

For any integer $k \geq 1$ and $m_0, m_1, ..., m_k \in \mathbb{N},\ m_j \neq m_{j'}$ for $j \neq j'$, we define

- $(\mathrm{SWM}_{m_0,...,m_k})$ $(\mathcal{U}, \varphi, \Phi)$ is *symmetrically weakly mixing of order* $(m_1, ..., m_k)$ if

$$D - \lim_{n\to\infty} \varphi\left(\prod_{l=-k}^{k} \Phi^{m_{|l|} n}(x_l) \right) = \varphi(x_0) \prod_{l=1}^{k} \varphi(x_{-l} x_l)$$

for all $x_0, x_{\pm 1}, ..., x_{\pm k} \in \mathcal{U}$.

As discussed in Section 4.1 we use the GNS representation associated to φ denoted by

$$\pi_\varphi : \mathcal{U} \to B(H_\varphi)$$

and denote by u_φ its canonical cyclic vector. Hence if $a \in \mathcal{U}$, then $\pi_\varphi(a) u_\varphi \in H_\varphi$.

Hence we see by Proposition 4.4.1.11 that $(\mathrm{WM}_{m_1,...,m_k})$ is implied if the bounded sequence

$$u_n = \pi_\varphi\left(\prod_{l=1}^{k} \Phi^{m_l n}(x_l) \right) u_\varphi - \prod_{l=1}^{k} \varphi(x_l) u_\varphi \in H_\varphi$$

is weakly mixing to zero. We will give more detail about this in the proof of Theorem 4.4.2.1.

We now state a few results which will be used in the proof of our subsequent results.

4.4.1.15 Remark

Let $\mathcal{U}$ be a unital C^*-algebra and $(\mathcal{U}, \varphi, \Phi)$ a state preserving C^*-dynamical system. For any integers $k \geq 1$ and $m_0, m_1, ..., m_k \in \mathbb{N},\ m_j \neq m_{j'}$ for $j \neq j'$, we have that:

(a) $(\mathrm{SWM}_{m_0,...,m_k})$ implies that

$$D - \lim_{n\to\infty} \varphi\left(\Phi^{m_k n}(x_{-k}) \left(\prod_{l=-k+1}^{k-1} \Phi^{m_{|l|} n}(x_l) \right) \Phi^{m_k n}(x_k) \right)$$
$$= \varphi(x_0) \prod_{l=1}^{k} \varphi(x_{-l} x_l)$$

for all $x_0, x_{\pm 1}, ..., x_{\pm k} \in \mathcal{U}$.

(b) $(\mathrm{SWM}_{m_0,\ldots,m_k}) \Rightarrow (\mathrm{SWM}_{m_0,\ldots,m_q})$ for every $1 \leq q \leq k$.

Proof:

The proof of (a) is trivial, since it is an immediate repetition of the definition of $(\mathrm{SWM}_{m_0,\ldots,m_k})$.

Since the result in (a) holds for all $x_0, x_{\pm 1}, \ldots, x_{\pm k} \in \mathcal{U}$, we let $x_{-k} = x_k = 1_{\mathcal{U}}$ in (a).

Hence $(\mathrm{SWM}_{m_0,\ldots,m_k})$ implies that

$$D\!-\lim_{n\to\infty} \varphi\left(\Phi^{m_k n}(1_{\mathcal{U}})\left(\prod_{l=-k+1}^{k-1} \Phi^{m_{|l|}n}(x_l)\right)\Phi^{m_k n}(1_{\mathcal{U}})\right) = \varphi(x_0)\left(\prod_{l=1}^{k-1}\varphi(x_{-l}x_l)\right)\varphi(1_{\mathcal{U}}),$$

i.e.

$$D-\lim_{n\to\infty} \varphi\left(\prod_{l=-k+1}^{k-1} \Phi^{m_{|l|}n}(x_l)\right) = \varphi(x_0)\left(\prod_{l=1}^{k-1}\varphi(x_{-l}x_l)\right),$$

which is $(\mathrm{SWM}_{m_0,\ldots,m_{k-1}})$. Hence the result is established. ■

Note: By $(\mathrm{SWM}_{m_1,\ldots,m_k})$ we mean that all the terms involving m_0 and x_0 are simply omitted from the expressions in the definition of SWM, and it follows directly from the Remark above that $(\mathrm{SWM}_{m_1,\ldots,m_k})$ implies that

$$D-\lim_{n\to\infty} \varphi\left(\Phi^{m_k n}(x_{-k})\left(\prod_{l=-k+1,\ l\neq 0}^{k-1} \Phi^{m_{|l|}n}(x_l)\right)\Phi^{m_k n}(x_k)\right) = \prod_{l=1}^{k}\varphi(x_{-l}x_l).$$

Also $(\mathrm{SWM}_{m_1,\ldots,m_k}) \Rightarrow (\mathrm{SWM}_{m_1,\ldots,m_q})$ for any $1 \leq q \leq k$.

This note will become relevant in a later result.

Next we will state and prove a result involving a weak form of commutativity, i.e. the property called *norm-asymptotically abelian in density.* $(\mathcal{U}, \Phi)$ is said to be norm-asymptotically abelian in density if

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}\left\|\left[\Phi^k(x), y\right]\right\| = 0$$

for all $x, y \in \mathcal{U}$, with the notation $[x, y] = xy - yx$ denoting the commutator. Note that we also have that

$$D-\lim_{n\to\infty}\left\|\left[\Phi^k(x), y\right]\right\| = 0$$

by Corollary 2.3.4.

4.4.1.16 LEMMA ([20], p 40)

Let $(\mathcal{U}, \Phi)$ *be a* C^**-algebra which is norm-asymptotically abelian in density. Let* $k \geq 1$ *and* $m_1, ..., m_k \in \mathbb{N}$, $m_j \neq m_{j'}$ for $j \neq j'$. Then

$$D - \lim_{n\to\infty} \left\| \prod_{l=-k}^{k} \Phi^{m_{|l|}n}(x_l) - \Phi^{m_0 n}(x_0) \prod_{l=1}^{k} \Phi^{m_l n}(x_{-l}x_l) \right\| = 0$$

for all $x_0, x_{\pm 1}, ..., x_{\pm k} \in \mathcal{U}$.

Proof:

(a) Use the notation $T_l = \Phi^{m_{|l|}n}(x_l)$ for simplicity.

We know that

$$\Phi^{m_l n}(x_{-l}x_l) = \Phi^{m_l n}(x_{-l})\Phi^{m_l n}(x_l) = T_{-l}T_l,$$

for $l = 1, 2, ..., k$, hence

$$\Phi^{m_0 n}(x_0) \prod_{l=1}^{k} \Phi^{m_l n}(x_{-l}x_l) = T_0 \prod_{l=1}^{k} (T_{-l}T_l)$$

and hence we must check whether

$$D - \lim_{n\to\infty} \left\| \prod_{l=-k}^{k} T_l - T_0 \prod_{l=1}^{k} (T_{-l}T_l) \right\| = 0.$$

The following calculations are used to show this:

We have that, for any $1 \leq q \leq k$,

$$\prod_{l=-q}^{q} T_l - \left(\prod_{l=-q+1}^{q-1} T_l \right) T_{-q}T_q = \sum_{s=-q+1}^{q-1} \left(\prod_{l=-q+1}^{s-1} T_l \right) [T_{-q}, T_s] \left(\prod_{l=s+1}^{q} T_l \right).$$

This can be verified easily for the cases $q = 1, 2, 3, ..$ and in general we find that the first term of the sum on the right hand side always yields

$$\prod_{l=-q}^{q} T_l$$

and the last term yields

$$-\left(\prod_{l=-q+1}^{q-1} T_l \right) T_{-q}T_q,$$

whereas the middle terms all cancel out by addition.

We then obtain that

$$\begin{aligned}
&\prod_{l=-k}^{k} T_l - \left(\prod_{l=-q+1}^{q-1} T_l\right)\left(\prod_{l=q}^{k}(T_{-l}T_l)\right)\\
&= \prod_{l=-k}^{-q-1} T_l \prod_{l=-q}^{q} T_l \prod_{l=q+1}^{k} T_l - \left(\prod_{l=-q+1}^{q-1} T_l\right)(T_{-q}T_q)\left(\prod_{l=q+1}^{k}(T_{-l}T_l)\right)\\
&= \sum_{r=q}^{k}\sum_{s=-r+1}^{r-1}\left(\prod_{l=-r+1}^{s-1} T_l\right)[T_{-r},T_s]\left(\prod_{l=s+1}^{r} T_l\right)\left(\prod_{l=r+1}^{k}(T_{-l}T_l)\right),
\end{aligned}$$

which can be verified in the same way as, and with the help of, the previous equation.

If we let $q = 1$ in the equation above, we get

$$\begin{aligned}
&\prod_{l=-k}^{k} T_l - T_0\prod_{l=1}^{k}(T_{-l}T_l)\\
&= \sum_{r=1}^{k}\sum_{s=-r+1}^{r-1}\left(\prod_{l=-r+1}^{s-1} T_l\right)[T_{-r},T_s]\left(\prod_{l=s+1}^{r} T_l\right)\left(\prod_{l=r+1}^{k}(T_{-l}T_l)\right).
\end{aligned}$$

Now let $c = \max_{-k\leq l\leq k} \|x_l\|$, and note that

$$\|T_l\| = \|\Phi^{m_{|l|}n}(x_l)\| \leq \|\Phi^{m_{|l|}n}\|\|x_l\| = \|x_l\| \leq c$$

since $\|\Phi\| = \Phi(1) = 1$ and since Φ is bounded. We now see that

$$\begin{aligned}
&\left\|\prod_{l=-k}^{k} T_l - T_0\prod_{l=1}^{k}(T_{-l}T_l)\right\|\\
&= \left\|\sum_{r=1}^{k}\sum_{s=-r+1}^{r-1}\left(\prod_{l=-r+1}^{s-1} T_l\right)[T_{-r},T_s]\left(\prod_{l=s+1}^{r} T_l\right)\left(\prod_{l=r+1}^{k}(T_{-l}T_l)\right)\right\|\\
&\leq \sum_{r=1}^{k}\sum_{s=-r+1}^{r-1}\left(\prod_{l=-r+1}^{s-1}\|T_l\|\right)\|[T_{-r},T_s]\|\left(\prod_{l=s+1}^{r}\|T_l\|\right)\left(\prod_{l=r+1}^{k}\|T_{-l}\|\|T_l\|\right)\\
&\leq \sum_{r=1}^{k}\sum_{s=-r+1}^{r-1} c^{(s-1+r-1)+(r-s-1)+(2k-2r-2)}\|[T_{-r},T_s]\|\\
&= c^{2k-3}\sum_{r=1}^{k}\sum_{s=-r+1}^{r-1}\|[T_{-r},T_s]\|.
\end{aligned}$$

It is also seen, since Φ is a positive linear homomorphism, that

$$\begin{aligned}
\|[T_{-r}, T_s]\| &= \|[\Phi^{m_r n}(x_{-r}), \Phi^{m_s n}(x_s)]\| \\
&= \|\Phi^{m_r n}(x_{-r})\Phi^{m_s n}(x_s) - \Phi^{m_s n}(x_s)\Phi^{m_r n}(x_{-r})\| \\
&= \left\|\Phi^{m_s n}\left(\Phi^{(m_r - m_s)n}(x_{-r}) \cdot x_s - x_s \Phi^{(m_r - m_s)n}(x_{-r})\right)\right\| \\
&\leq \left\|\left[\Phi^{(m_r - m_s)n}(x_{-r}), x_s\right]\right\|.
\end{aligned}$$

Therefore, from the norm-asymptotic abelianness in density assumption for $(\mathcal{U}, \Phi)$ and the squeeze theorem we get that

$$D - \lim_{n\to\infty} \|[T_{-r}, T_s]\| = 0.$$

As shown above, we have that

$$\begin{aligned}
0 &\leq \left\|\prod_{l=-k}^{k} T_l - T_0 \prod_{l=1}^{k} (T_{-l}T_l)\right\| \\
&\leq c^{2k-3} \sum_{r=1}^{k} \sum_{s=-r+1}^{r-1} \|[T_{-r}, T_s]\|.
\end{aligned}$$

Hence

$$D - \lim_{n\to\infty} \left\|\prod_{l=-k}^{k} T_l - T_0 \prod_{l=1}^{k} (T_{-l}T_l)\right\| = 0.$$

■

4.4.2 Multiple weakly mixing results

We are now in a position to state and prove two important multiple weakly mixing results. We will first prove a strong supporting theorem from which our final conclusion will follow. Once again it may be mentioned that [20] gave a proof for a more general case (where the existence of a unit element is not assumed).

4.4.2.1 Theorem ([20], p 40)

Let $\mathcal{U}$ be a unital C^*-algebra and $(\mathcal{U}, \varphi, \Phi)$ a weakly mixing state preserving C^*-dynamical system. For any integers $k \geq 1$ and $m_1, ..., m_k \in \mathbb{N}$, $m_j \neq m_{j'}$ for $j \neq j'$, we have that:

$$(\mathrm{SWM}_{m_1,...,m_k}) \Rightarrow (\mathrm{WM}_{m_1,...,m_k}).$$

Furthermore, if $(\mathcal{U}, \Phi)$ is norm-asymptotically abelian in density then

$$(\mathrm{WM}_{m_1,...,m_k}) \Leftrightarrow (\mathrm{SWM}_{0,m_1,...,m_k}).$$

Proof:

Let $x_0, x_1, \ldots, x_k \in \mathcal{U}$ be arbitrary.

We will use the GNS representation associated with φ, and prove that the following sequence

$$\begin{aligned} u_n &= \pi_\varphi\left(\prod_{l=1}^{k} \Phi^{m_l n}(x_l)\right) u_\varphi - \left(\prod_{l=1}^{k} \varphi(x_l)\right) u_\varphi \in H_\varphi \\ &= \pi_\varphi\left(\Phi^{m_1 n}(x_1)\Phi^{m_2 n}(x_2)\cdots\Phi^{m_k n}(x_k)\right) u_\varphi - \varphi(x_1)\varphi(x_2)\cdots\varphi(x_k) u_\varphi \end{aligned}$$

is uniformly weakly mixing (hence weakly mixing) to zero under the assumption that the system satisfies $(\mathrm{SWM}_{m_1,...,m_k})$.

We now introduce a new symbol which will enable us to use the Van der Corput result to show the uniformly weakly mixing required above.

For every $1 \leq q \leq k$, let

$$\begin{aligned} u_n^{(q)} &= \pi_\varphi\left(\Phi^{m_1 n}(x_1)\cdots\Phi^{m_q n}(x_q)\right) u_\varphi - \varphi(x_q)\pi_\varphi\left(\Phi^{m_1 n}(x_1)\cdots\Phi^{m_{q-1} n}(x_{q-1})\right) u_\varphi \\ &= \pi_\varphi\left(\prod_{l=1}^{q} \Phi^{m_l n}(x_l)\right) u_\varphi - \varphi(x_q)\pi_\varphi\left(\prod_{l=1}^{q-1} \Phi^{m_l n}(x_l)\right) u_\varphi \\ &= \pi_\varphi\left(\left(\prod_{l=1}^{q-1} \Phi^{m_l n}(x_l)\right)\left(\Phi^{m_q n}(x_q) - \varphi(x_q)1_{\mathcal{U}}\right)\right) u_\varphi \\ &= \pi_\varphi\left(\left(\prod_{l=1}^{q-1} \Phi^{m_l n}(x_l)\right)\Phi^{m_q n}\left(x_q - \varphi(x_q)1_{\mathcal{U}}\right)\right) u_\varphi, \end{aligned}$$

the last step following from the fact that Φ is assumed to be state-preserving.

Use $\tilde{x}_q = x_q - \varphi(x_q)1_{\mathcal{U}}$ for simplicity. Now we consider the inner product terms relevant to the Van der Corput type results, culminating in Corollary 4.4.1.14:

$$\begin{aligned}
\langle u_{n+m}^{(q)}, u_n^{(q)} \rangle &= \varphi\left((u_n^{(q)})^* \cdot u_{n+m}^{(q)} \right) \\
&= \varphi\left(\left(\left(\prod_{l=1}^{q-1} \Phi^{m_l n}(x_l) \right) \Phi^{m_q n}(\tilde{x}_q) \right)^* \cdot \left(\prod_{l=1}^{q-1} \Phi^{m_l(n+m)}(x_l) \right) \Phi^{m_q(n+m)}(\tilde{x}_q) \right) \\
&= \varphi\left(\Phi^{m_q n}(\tilde{x}_q^*) \left(\prod_{l=1}^{q-1} \Phi^{m_l n}(x_l^*) \right) \cdot \left(\prod_{l=1}^{q-1} \Phi^{m_l(n+m)}(x_l) \right) \Phi^{m_q(n+m)}(\tilde{x}_q) \right) \\
&= \varphi\left(\Phi^{m_q n}(\tilde{x}_q^*) \left(\prod_{l=-q+1}^{-1} \Phi^{m_{|l|} n}(x_{-l}^*) \right) \cdot \left(\prod_{l=1}^{q-1} \Phi^{m_l(n+m)}(x_l) \right) \Phi^{m_q(n+m)}(\tilde{x}_q) \right) \\
&= \varphi\left(\Phi^{m_q n}(\tilde{x}_q^*) \left(\prod_{l=-q+1}^{-1} \Phi^{m_{|l|} n}(x_{-l}^*) \right) \cdot \left(\prod_{l=1}^{q-1} \Phi^{m_l n}\left(\Phi^{m_l m}(x_l)\right) \right) \Phi^{m_q n}\left(\Phi^{m_q m}(\tilde{x}_q)\right) \right).
\end{aligned}$$

Now, since we have ($\mathrm{SWM}_{m_1,\ldots,m_k}$), we also have ($\mathrm{SWM}_{m_1,\ldots,m_q}$).

We can now use Remark 4.4.1.15, and the note after the proof of the same Remark: Let $x_{-q} \rightleftharpoons \tilde{x}_q^*$, $x_q \rightleftharpoons \Phi^{m_q m}(\tilde{x}_q)$, and for each $-q+1 \leq l \leq -1$ let $x_l \rightleftharpoons \tilde{x}_{-l}^*$ and finally for each $1 \leq l \leq q-1$ let $x_l \rightleftharpoons \Phi^{m_l m}(\tilde{x}_l)$.

Then we obtain from the above calculation, from ($\mathrm{SWM}_{m_1,\ldots,m_q}$) and from Remark 4.4.1.15 that

$$\begin{aligned}
D - \lim_{n\to\infty} \langle u_{n+m}^{(q)}, u_n^{(q)} \rangle &= \left(\prod_{l=1}^{q-1} \varphi(\tilde{x}_{-(-l)}^* \Phi^{m_l m}(\tilde{x}_l)) \right) \cdot \varphi(\tilde{x}_q^* \Phi^{m_q m}(\tilde{x}_q)) \\
&= \left(\prod_{l=1}^{q-1} \varphi(\tilde{x}_l^* \Phi^{m_l m}(\tilde{x}_l)) \right) \cdot \varphi(\tilde{x}_q^* \Phi^{m_q m}(\tilde{x}_q)).
\end{aligned}$$

Also, since Φ is assumed to be weakly mixing with respect to φ, we have that Φ^{m_q} is weakly mixing with respect to φ as well. Therefore

$$D - \lim_{m\to\infty} \prod_{l=1}^{q-1} \varphi(\tilde{x}_l^* \Phi^{m_l m}(\tilde{x}_l)) = \prod_{l=1}^{q-1} \varphi(\tilde{x}_l^*)\varphi(\tilde{x}_l),$$

and

$$\begin{aligned}
D - \lim_{m\to\infty} \varphi(\tilde{x}_q^* \Phi^{m_q m}(\tilde{x}_q)) &= D - \lim_{m\to\infty} \varphi\left((x_q - \varphi(x_q)1_{\mathcal{U}})^* \Phi^{m_q m}(x_q - \varphi(x_q)1_{\mathcal{U}}) \right) \\
&= D - \lim_{m\to\infty} \varphi\left(x_q^* - \overline{\varphi(x_q)}1_{\mathcal{U}})(\Phi^{m_q m}(x_q) - \varphi(x_q)1_{\mathcal{U}} \right)
\end{aligned}$$

$$= D-\lim_{m\to\infty}\varphi\left((x_q^*\Phi^{m_qm}(x_q)-x_q^*\varphi(x_q)1_{\mathcal{U}}-\overline{\varphi(x_q)}1_{\mathcal{U}}\Phi^{m_qm}(x_q)+\overline{\varphi(x_q)}\varphi(x_q)1_{\mathcal{U}}\right)$$

$$= D-\lim_{m\to\infty}\left(\varphi(x_q^*\Phi^{m_qm}(x_q))-\varphi(x_q^*\varphi(x_q)1_{\mathcal{U}})-\varphi(\overline{\varphi(x_q)}1_{\mathcal{U}}\Phi^{m_qm}(x_q))+\varphi(\overline{\varphi(x_q)}\varphi(x_q)1_{\mathcal{U}})\right)$$

$$= D-\lim_{m\to\infty}\varphi\left(x_q^*\Phi^{m_qm}(x_q)\right)-\overline{\varphi(x_q)}\varphi(x_q)$$

$$= 0.$$

Thus we have that

$$D-\lim_{m\to\infty}\left(D-\lim_{n\to\infty}\langle u_{n+m}^{(q)},u_n^{(q)}\rangle\right)=0,$$

and by Corollary 4.4.1.14 it follows that $\{u_n^{(q)}\}_{n\geq 1}$ is uniformly weakly weakly mixing to zero.

The choice of the terms $u_n^{(q)}$ becomes apparent if we notice that

$$u_n=\sum_{q=1}^{k}\left(\sum_{l=q+1}^{k}\varphi(x_l)\right)u_n^{(q)}.$$

We then see that

$$\begin{aligned}\sup_{\|u\|\leqslant 1}\left(\frac{1}{N}\sum_{n=0}^{N-1}|\langle u_n,u\rangle|\right) &= \sup_{\|u\|\leqslant 1}\left(\frac{1}{N}\sum_{n=0}^{N-1}\left|\sum_{q=1}^{k}\left(\sum_{l=q+1}^{k}\varphi(x_l)\right)\langle u_n^{(q)},u\rangle\right|\right)\\ &\leq \sup_{\|u\|\leqslant 1}\left(\sum_{q=1}^{k}\left|\sum_{l=q+1}^{k}\varphi(x_l)\right|\frac{1}{N}\sum_{n=0}^{N-1}|\langle u_n^{(q)},u\rangle|\right)\\ &\longrightarrow 0\end{aligned}$$

as $N\to\infty$. Therefore, the sequence $\{u_n\}_{n\geq 1}$ is also uniformly weakly mixing, hence weakly mixing to zero.

Thus, we have that, with $u=\pi_\varphi(x_0^*)u_\varphi$, and $m_0=0$,

$$\begin{aligned}&\frac{1}{N}\sum_{n=0}^{N-1}\left|\varphi\left(\prod_{l=0}^{k}\Phi^{m_ln}(x_l)\right)-\prod_{l=0}^{k}\varphi(x_l)\right|\\ &=\frac{1}{N}\sum_{n=0}^{N-1}\left|\varphi\left(x_0\prod_{l=1}^{k}\Phi^{m_ln}(x_l)-x_0\prod_{l=1}^{k}\varphi(x_l)\right)\right|\\ &=\frac{1}{N}\sum_{n=0}^{N-1}\left|\varphi\left(x_0\left(\prod_{l=1}^{k}\Phi^{m_ln}(x_l)-\prod_{l=1}^{k}\varphi(x_l)\right)\right)\right|\\ &=\frac{1}{N}\sum_{n=0}^{N-1}|\langle u_n,\pi_\varphi(x_0^*)u_\varphi\rangle|\\ &\longrightarrow 0\end{aligned}$$

which yields $(\mathrm{WM}_{m_1,\ldots,m_k})$ immediately.

Hence we have shown that

$$(\mathrm{SWM}_{m_1,\ldots,m_k}) \Rightarrow (\mathrm{WM}_{m_1,\ldots,m_k}).$$

Finally, it must be shown that, if $(\mathcal{U}, \Phi)$ is norm-asymptotically abelian in density then

$$(\mathrm{WM}_{m_1,\ldots,m_k}) \Leftrightarrow (\mathrm{SWM}_{0,m_1,\ldots,m_k}).$$

But this is a consequence of Lemma 4.4.1.16. ■

Next follows a multiple weakly mixing result, which will be seen to be a non-commutative extension of the "weakly mixing of all orders" result, Theorem 3.3.3, which was proved in the commutative case.

To extend this property (weakly mixing of all orders) we still need some measure of commutativity, i.e. we will require that $(\mathcal{U}, \Phi)$ is norm-asymptotically abelian in density.

4.4.2.2 Theorem ([20], p 42)

Let $\mathcal{U}$ be a unital C^*-algebra and $(\mathcal{U}, \varphi, \Phi)$ a weakly mixing state preserving C^*-dynamical system. Assume that $(\mathcal{U}, \Phi)$ is norm-asymptotically abelian in density. Let $k \geq 1$ and let $m_1, \ldots, m_k \in \mathbb{N}$ satisfying $m_1 < \cdots < m_k$. Then $(\mathcal{U}, \varphi, \Phi)$ is weakly mixing of order $(m_1, \ldots, m_k)$.

Proof:

We prove this by induction.

For any integer $m_1 \geq 1$, we have that

$$D - \lim_{n\to\infty} \varphi\left(\Phi^{m_1 n}(x_{-1})\Phi^{m_1 n}(x_1)\right) = D - \lim_{n\to\infty} \varphi\left(\Phi^{m_1 n}(x_{-1}x_1)\right) = \varphi(x_{-1}x_1)$$

for all $x_{-1},\ x_1 \in \mathcal{U}$, and we see that (SWM_{m_1}) is trivially satisfied. It then follows from Theorem 4.4.2.1 that (WM_{m_1}) is satisfied.

The theorem is then true for the case of one arbitrary integer $m_1 \geq 1$.

Now assume that the theorem is true for some $k \geq 1$ and any integers $m_1 < \cdots < m_k$. If we can conclude from this that the theorem also holds for any $k+1$ integers $n_1 < \cdots < n_{k+1}$, the induction proof is complete.

Hence assume that $(\mathrm{WM}_{m_1,\cdots,m_k})$ holds for some $k \geq 1$ and integers $m_1 < \cdots < m_k$. Then, it follows from this assumption, for any integers $n_1 < \cdots < n_{k+1}$, that $(\mathrm{WM}_{n_2-n_1,\cdots,n_{p+1}-n_1})$ is satisfied.

We now use Theorem 4.4.2.1 to get that

$$\begin{aligned}(\mathrm{WM}_{n_2-n_1,\cdots,n_{p+1}-n_1}) &\Rightarrow (\mathrm{SWM}_{0,n_2-n_1,\cdots,n_{p+1}-n_1})\\ &\Leftrightarrow (\mathrm{SWM}_{n_1,n_2,\cdots,n_{p+1}})\\ &\Rightarrow (\mathrm{WM}_{n_1,n_2,\cdots,n_{p+1}}).\end{aligned}$$

■

It is trivial to check that any commutative system is norm-asymptotically abelian in density, since all terms $[x, y] = 0$ whenever x and y commute. This shows that Theorem 3.3.3 is a special case of the last Theorem.

Bibliography

[1] R. Alicki & M. Fannes. *Quantum Dynamical Systems.* Oxford University Press: London, 2001.

[2] D. Berend & V. Bergelson. Mixing Sequences in Hilbert Spaces. *Proc. Amer. Math. Soc.* **98**, 2 (1986), 239-246.

[3] F. Blume. Ergodic Theory. *Handbook of Measure Theory, Volume 2.* E. Pap (ed). Elsevier Science B.V.: Amsterdam, 2002.

[4] D.L. Cohn. *Measure Theory.* Birkhäuser: Boston, 1980.

[5] G. De Barra. *Measure Theory and Integration.* John Wiley & Sons: New York, 1981.

[6] N. Dunford & J.T. Schwartz. *Linear Operators - Part I: General Theory.* John Wiley & Sons: New York, 1988.

[7] R. Duvenhage & A. Ströh. Recurrence and Ergodicity in Unital $*$-algebras. *J. Math. Anal. and Appl.* **287** (2003), 430-443.

[8] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory.* Princeton University Press: Princeton, 1981.

[9] H. Furstenberg, Y. Katznelson & D. Ornstein. The Ergodic Theoretical Proof of Szemerédi's Theorem. *Bull. Amer. Math. Soc.* **7**, 3 (1982), 527-552.

[10] H. Furstenberg. Nonconventional Ergodic Averages. *The Legacy of John von Neumann - Proceedings of Symposia in Pure Mathematics, Volume 50.* J. Glimm, J. Impagliazzo & I. Singer (eds). American Mathematical Society: Providence, 1990.

[11] P.R. Halmos. *Lectures on Ergodic Theory.* The Mathematical Society of Japan: 1956.

[12] B.E. Johnson. Non-Commutative Generalizations of Mathematics. *Bull. London Math. Soc.* **14** (1982), 465-471.

[13] L.K. Jones. A mean ergodic theorem for weakly mixing operators. *Advances in Math.***7**, (1971), 211-216.

[14] L.K. Jones & M. Lin. Ergodic Theorems of Weak Mixing Type. *Proc. Amer. Math. Soc.* **57**, 1 (1976), 50-52.

[15] U. Krengel. *Ergodic Theorems.* Walter de Gruyter: Berlin, 1985.

[16] E. Kreyszig. *Introductory Functional Analysis with Applications.* John Wiley & Sons: New York, 1978.

[17] L. Kuipers & H. Niederreiter. *Uniform Distribution of Sequences.* John Wiley & Sons: New York, 1974.

[18] G.J. Murphy. *C^*-Algebras and Operator Theory.* Academic Press: San Diego, 1990.

[19] C.P. Niculescu & A. Ströh. A Hilbert Space Approach of Poincaré Reccurence Theorem. *Rev. Roumaine Math. Pures Appl.* **44**(1999), 799-805.

[20] C.P. Niculescu, A. Ströh & L. Zsidó. Noncommutative Extensions of Classical and Multiple Recurrence Theorems. *J. Operator Theory.* **50** (2003), 3-52.

[21] K. Petersen. *Ergodic Theory.* Cambridge University Press: Cambridge, 1983.

[22] F. Riesz & B. Sz.-Nagy. *Functional Analysis.* Dover Publications: New York, 1990.

[23] L. Zsidó. Weak mixing properties for vector sequences, to appear.

A Hilbert space approach to multiple recurrence in ergodic theory

by

F.J.C. Beyers

Supervisor : Professor A. Ströh
Department : Mathematics and Applied Mathematics
Degree : Magister Scientiae

ABSTRACT

The use of Hilbert space theory became an important tool for ergodic theoreticians ever since John von Neumann proved the fundamental Mean Ergodic theorem in Hilbert space. Recurrence is one of the corner stones in the study of dynamical systems. In this dissertation some extended ideas besides those of the basic, well-known recurrence results are investigated. Hilbert space theory proves to be a very useful approach towards the solution of multiple recurrence problems in ergodic theory.

Another very important use of Hilbert space theory became evident only relatively recently, when it was realized that non-commutative dynamical systems become accessible to the ergodic theorist through the important Gelfand-Naimark-Segal (GNS) representation of C^*-algebras as Hilbert spaces. Through this construction we are enabled to invoke the rich catalogue of Hilbert space ergodic results to approach the more general, and usually more involved, non-commutative extensions of classical ergodic-theoretical results.

In order to make this text self-contained, the basic, standard, ergodic-theoretical results are included in this text. In many instances Hilbert space counterparts of these basic results are also stated and proved. Chapters 1 and 2 are devoted to the introduction of these basic ergodic-theoretical results such as an introduction to the idea of measure-theoretic dynamical systems, citing some basic examples, Poincaré's recurrence, the ergodic theorems of Von Neumann and Birkhoff, ergodicity, mixing and weakly mixing. In Chapter 2 several rudimentary results, which are the basic tools used in proofs, are also given.

In Chapter 3 we show how a Hilbert space result, i.e. a variant of a result by Van der Corput for uniformly distributed sequences modulo 1, is used to simplify the proofs of some multiple recurrence problems. First we use it to simplify and clarify the proof of a multiple recurrence result by Furstenberg, and also to extend that result to a more general case, using the same Van der Corput lemma. This may be considered the main result of this thesis, since it supplies an original proof of this result. The Van der Corput lemma helps to simplify many of the tedious terms that are found in Furstenberg's proof.

In Chapter 4 we list and discuss a few important results where classical (commutative) ergodic results were extended to the non-commutative case. As stated before, these extensions are mainly due to the accessibility of Hilbert space theory through the GNS construction. The main result in this section is a result proved by Niculescu, Ströh and Zsidó, which is proved here using a similar Van der Corput lemma as in the commutative case. Although we prove a special case of the theorem by Niculescu, Ströh and Zsidó, the same method (Van der Corput) can be used to prove the generalized result.

DECLARATION

I, the undersigned, hereby declare that the dissertation submitted herewith for the degree Magister Scientiae to the University of Pretoria contains my own, independent work and has not been submitted for any degree at any other university.

Signature:

Name: FJC Beyers Date: 2004-12-09