Applications of Dynamical Systems in Combinatorial Number Theory

Yihan Zhang

January 3, 2017

Contents

Fürstenberg(1977) gave a new proof of Szemerédi's theorem, bringing topological dynamics into combinatorial number theory, promoting a new specialization. Dynamical systems have various branches, e.g., topological dynamical systems, ergodic dynamical systems, smooth dynamical systems, random dynamical systems, Hamiltonian dynamical systems, etc. Nowadays, as for techniques in dynamical systems, it is topological dynamics and ergodic theory that are most used in combinatorial number theory. In this note, we consider applications of dynamical systems in combinatorial number theory, including four sections.

1. A brief introduction to dynamical systems including basic concepts and results, in which we will see that dynamical systems bear a strong relationship with combinatorial number theory;

2. applications of dynamical systems in combinatorial number theory, Fürstenberg's new proof of van der Waerden's theorem via topological dynamics and his new proof of Szemerédi's theorem via ergodic theory;

3. a brief sketch of Fürstenberg's proof of Szemerédi's theorem;

4. basic ideas of Gowers' proof of Szemerédi's theorem via higher order Fourier analysis.

1 Introduction to dynamical systems

Dynamical systems is a subject concerning with qualitative properties of group actions on spaces, where group actions are maps subject to

1. $e(x) = x$;

2. $(g_1 g_2)(x) = g_2(g_1(x))$.

Consider actions of semigroup $\mathbb{Z}^+$ and group $\mathbb{Z}$.

1. $\mathbb{Z}^+$-action. $T : X \to X$, $T^0 = Id, T^1 = T, T^2 = T \circ T, \cdots$.

2. $\mathbb{Z}$-action. $T : X \to X$, $\cdots, T^{-1}, T^0 = Id, T^1 = T, T^2 = T \circ T, \cdots$.

Definition 1 (t.d.s., Birkhoff)**.** *(X, G) is called a t.d.s. if X is a compact metric space and G is a topological group or semigroup acting on X.*

For example, $G = \mathbb{Z}^+, (X, T) \coloneqq (X, \mathbb{Z}^+)$, $T : X \to X$ is a continuous map. $G = \mathbb{Z}, (X, T) \coloneqq (X, \mathbb{Z})$, $T : X \to X$ is a homeomorphism. We will also need $\mathbb{Z}^d$-action subject to $T_i \circ T_j = T_j \circ T_i, 1 \leq i \leq j \leq d$.

Definition 2 (m.d.s., Poincaré, Birkhoff, von Neumann)**.** *$(X, \mathcal{B}, \mu, G)$ is called an m.d.s. if X is a set, $\mathcal{B}$ is a σ-algebra on X, G is a group or semigroup, and $g : X \to X$ is a measure-preserving transformation subject to $\mu(B) = \mu(g^{-1}(B)), \forall B \in \mathcal{B}, g \in G$.*

T.d.s. and m.d.s. are twins. Many concepts and results are parallel to each other. For one thing, for any t.d.s., there exists an invariant measure on Borel σ-algebra. That is to say, given any t.d.s. (X, T), there is a measure-preserving system $(X, \mathcal{B}_X, \mu, T)$ corresponding to it. For another thing, any ergodic system on Lebesgue space has a topological representation, which means for any $(X, \mathcal{B}_X, \mu, T)$, we can treat it as $(Y, \mathcal{B}_Y, \nu, S)$.

Now we're going to introduce some basic concepts in t.d.s. and m.d.s.

Definition 3. *Assume X is a compact metric space, $T : X \to X$ is a continuous map. $x \in X$ is called a recurrent point if $\exists \{n_i\} \subset \mathbb{Z}^+$ s.t. $n_i \to +\infty$ and $T^{n_i} x \to x$.*

It's trivial to verify that fixed points and periodic points are special recurrent points. Recurrent points can be easily constructed in symbolic dynamics. Let $A = \{0, 1, \cdots, k-1\} (k \geq 2)$ and $X_k = A^{\mathbb{N}}$. In fact, $X_k = \{(x_1, x_2, \cdots) : x_i \in A, i \in \mathbb{N}\}$. Define a metric $\rho((x_1, \cdots), (y_1, \cdots)) = (\min\{i : x_i \neq y_i\})^{-1}$, which means the distance between any two symbols is the reciprocal of the first position where their words are different. Define a map $T : X_k \to X_k$ subject to $T(x_1, x_2, x_3, \cdots) = (x_2, x_3, \cdots)$. This shift eliminates the first word of a symbol. It's easy to verify that this map is continuous in this metric space. Thus (X_k, T) becomes a symbolic dynamical system. We have the following theorem.

Theorem 1. *$x = (x_1, x_2, x_3, \cdots) \in X_k$ is a recurrent point iff every word in x appears infinitely.*

For example, let $B_1 = 1, B_2 = B_1 0^1 B_1, \cdots, B_{n+1} = B_n 0^n B_n$, then $x = \lim B_n$ is a recurrent point.

Definition 4. *$F \subset \mathbb{Z}^+$ is called an IP-set if $\exists \{p_i\} \subset \mathbb{N}$ s.t. $\{\varepsilon_1 p_1 + \cdots + \varepsilon_n p_n : \varepsilon_i = 0, 1, n \in \mathbb{N}\} \subset F$.*

The following theorem establishes a connection between a concept in dynamical systems(recurrent point) and an interesting set(IP-set) we might concern in number theory.

Theorem 2 (Fürstenberg)**.** *Assume (X, T) is a t.d.s. Then x is a recurrent point iff for any open neighborhood U of x, $N(x, U) \coloneqq \{n \in \mathbb{Z}^+ : T^n x \in U\}$ is an IP-set.*

We can imagine $N(x, U)$ to indicate how long it takes for x to go back into its neighborhood after iterative actions of T.

Definition 5. *$A \subset X$ is said to be invariant if $T(A) \subset A$.*

Definition 6. *X or (X, T) is said to be minimal if there exists no nonempty proper invariant subsets.*

We can determine whether a system is minimal using the following theorem concerning with orbits.

Theorem 3. *(X, T) is minimal iff $\forall x \in X, orb(x, T) \coloneqq \{x, Tx, T^2x, \cdots\}$ is dense in X.*

Definition 7. *x is called a minimal point if $orb(x, T)$ is minimal.*

Definition 8. *$\mathbb{Z}^+ \supset F = \{a_1 < a_2 < \cdots\}$ is said to be syndetic or relatively dense if $\exists M$ s.t. $a_{i+1} - a_i \leq M, \forall i \in \mathbb{N}$.*

That is to say, a set in ascending order is syndetic if there is a uniform gap between every two adjacent numbers. Again, the following theorem also shows a strong relationship between concepts in dynamical systems and number theory.

Theorem 4 (1940s)**.** *Assume (X, T) is a t.d.s., $x \in X$. Then x is a minimal point iff for any neighborhood U of x, $N(x, U)$ is syndetic.*

Minimal points can also be obtained in symbolic dynamics. We have the famous Morse sequence.

Theorem 5 (Morse sequence)**.** *$x \in (x_1, x_2, \cdots) \in X_k$ is minimal iff every word in x appears syndetically.*

For instance, let $B_1 = 1, B_2 = 10, \cdots, B_{n+1} = B_n B_n'$, where B_n' is the duality of B_n. Then $x = \lim B_n$ is minimal. In fact, $x = 1001011001101001 \cdots$. More definitions we might use are listed below.

Definition 9. *Upper and lower density of $S \subset \mathbb{Z}$ are defined as*

$$D^*(S) = \limsup_{n\to\infty} \frac{|S \cap [-n,n]|}{2n+1}$$

and

$$D_*(S) = \liminf_{n\to\infty} \frac{|S \cap [-n,n]|}{2n+1}.$$

The above definitions indicate roughly how dense a set is in $[-n, n]$.

Definition 10. *Upper and lower Banach density of $S \subset \mathbb{Z}$ are defined as*

$$BD^*(S) = \limsup_{|I|\to\infty} \frac{|S \cap I|}{|I|}$$

and

$$BD_*(S) = \liminf_{|I|\to\infty} \frac{|S \cap I|}{|I|}.$$

The above definitions look at every subset of $\mathbb{Z}$ to see how dense the set is in that segment. Upper Banach density of $\mathbb{Z}$ and piecewise syndetic subset can also delineate dynamical properties concerning with recurrence. We give two recurrence theorems, respectively t.d.s. version and m.d.s. version., which connect dynamical properties with number properties.

Theorem 6 (Birkhoff recurrence)**.** *Every compact t.d.s. has a minimal point, thus has a recurrent point.*

Definition 11. *$F \subset \mathbb{Z}$ is called a Birkhoff recurrent set if for any t.d.s. (X,T), $\exists \{n_i\} \subset F, n_i \to \infty, x \in X,$ s.t. $T^{n_i}x \to x$.*

Theorem 7. *F is a Birkhoff recurrent set iff for any syndetic subset A, $F \cap (A - A) \neq \phi$, where $A - A \coloneqq \{a - b : a, b \in A\}$.*

Theorem 8 (Poincaré recurrence)**.** *Assume $(X, \mathcal{B}, T, \mu)$ is an m.d.s. If $A \in \mathcal{B}$ and $\mu(A) > 0$, then $\exists n \in \mathbb{N}$ s.t. $\mu(A \cap T^{-n}A) > 0$.*

This theorem can be directly shown by pigeonhole principle. As $A, T^{-1}A, T^{-2}A, \cdots$ are all in X, if none of them intersects, then total measure will approach infinity. Similarly, we can define Poincaré recurrent set and have the following theorem.

Theorem 9. *F is a Poincaré recurrent set iff for any set A with positive upper Banach density, $F \cap (A - A) \neq \phi$.*

2 Dynamical systems vs. number theory

In this section we will introduce two significant theorems in combinatorial number theory and see roughly how they are proved via dynamical systems. First let's consider van der Waerden's theorem.

Theorem 10 (van der Waerden)**.** *Assume $d \in \mathbb{N}$ and $\mathbb{N} = N_i \cup \cdots \cup N_d$. Then $\exists 1 \leq i \leq d$ s.t. N_i contains arbitrarily long arithmetic progressions, i.e. $\forall k \in \mathbb{N}, \exists a = a(k), b = b(k)$, s.t. $a, a+b, a+2b, \cdots, a+kb \in N_i$.*

We will use the following theorem in topological dynamics, or to be precise, its reduced version, to prove the above theorem.

Theorem 11 (Fürstenberg-Weiss topological multiple recurrence)**.** *Assume $d \in \mathbb{N}$, $T_1, \cdots, T_d : X \to X$ are continuous maps from compact metric space X to X subject to $T_i \circ T_j = T_j \circ T_i, \forall 1 \leq i \leq j \leq d$. Then $\exists x \in X$ and $n_i \to \infty$ s.t. $\forall 1 \leq j \leq d, T_j^{n_i} x \to x, i \to \infty$.*

Compared with singular recurrence theorem, the above theorem gives a common x and n_i for every T_j.

Theorem 12 (reduced topological multiple recurrence)**.** *Assume $d \in \mathbb{N}$, $T : X \to X$ is a continuous map from compact metric space X to X. Then $\exists x \in X$ and $n_i \to \infty$ s.t. $\forall 1 \leq j \leq d, T^{jn_i} x \to x, i \to \infty$.*

The reduced version can be obtained by taking T_j in the original theorem to be T^j. Now let's roughly see how van der Waerden's theorem is proved using the above theorem.

Proof. Assume the decomposition of $\mathbb{N}$ is pairwise disjoint. We convert a number sequence to a symbol, $z = (z_1, z_2, \cdots) \in \{1, \cdots, d\}^{\mathbb{N}}$ s.t. $z_i = j \iff i \in N_j$. We will tackle this problem in symbolic dynamics. Let $X_d = \{1, \cdots d\}^{\mathbb{N}}$ and $T : X_d \to X_d$ to be a shift. Define metric $\rho((x_1, \cdots), (y_1, \cdots)) = (\min\{i : x_i \neq y_i\})^{-1}$. Notice that $\rho((x_1, \cdots), (y_1, \cdots)) < 1 \iff x_1 = y_1$. Let $X = \overline{orb(z, T)}$. Notice $T^{jn_i} x \to x$ and apply reduced topological multiple recurrence theorem, we have $x \in X$ and $n \in \mathbb{N}$ s.t. $\rho(T^n x, x) < 1, \rho(T^{2n} x, x), \cdots, \rho(T^{kn} x, x) < 1$. This indicates that $x_1 = x_{n+1} = \cdots = x_{kn+1} = i$. As $\exists m_j$ s.t. $T^{m_j} z \to x$, we can find an m_j s.t. $z_{m_j+1} = z_{m_j+n+1} = \cdots = z_{m_j+kn+1} = i$. Refer to the definition of z, immediately we have $m_j + 1, m_j + n + 1, \cdots, m_j + kn + 1 \in N_i$. According to the above process, for a fixed k, we find an N_i. As the decomposition is finite, there must be infinitely many i in some N_i which is exactly what we need. $\square$

A pretty concise proof, right? Now let's see Szemerédi's theorem.

Theorem 13 (Szemerédi)**.** *If a subset A of $\mathbb{N}$ has positive Banach density, then it contains arbitrarily long arithmetic progressions.*

Proof of Szemerédi theorem involves reduced multiple ergodic recurrence theorem.

Theorem 14 (Fürstenberg multiple ergodic recurrence)**.** *Assume* $d \in \mathbb{N}$*,* $(X, \mathcal{B}, \mu, T)$ *is a measure-preserving system subject to* $T_i \circ T_j = T_j \circ T_i$*,* $A \in \mathcal{B}$*,* $\mu(A) > 0$*. Then* $\forall d \geq 1$*,*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T_1^{-n} A \cap T_2^{-n} A \cap \cdots \cap T_d^{-n} A) > 0.$$

Theorem 15 (reduced Fürstenberg multiple ergodic recurrence)**.** *Assume* $d \in \mathbb{N}$*,* $(X, \mathcal{B}, \mu, T)$ *is a measure-preserving system,* $A \in \mathcal{B}$*,* $\mu(A) > 0$*. Then* $\forall d \geq 1, \exists n > 0$ *s.t.*

$$\mu(A \cap T^{-n} A \cap T^{-2n} A \cap \cdots \cap T^{-dn} A) > 0.$$

Similar to topological recurrence theorem, this reduced version can be obtained by taking T_j in the original theorem to be T^j. We provide a topological form of the above theorem.

Theorem 16 (reduced Fürstenberg multiple ergodic recurrence)**.** *Assume* (X, T) *is a minimal t.d.s. Then for any nonempty open subset* U *of* X*,* $\forall d \geq 1, \exists n > 0$ *s.t.*

$$U \cap T^{-n} U \cap \cdots \cap T^{-dn} U \neq \phi.$$

Then we introduce Fürstenberg correspondence principle, including two cases.

Theorem 17 (Fürstenberg correspondence principle, measure-preserving case)**.** *Assume* $E \subset \mathbb{Z}$*,* $BD^*(E) > 0$*. Then there exists an m.d.s.* $(X, \mathcal{X}, \mu, T)$ *and* $A \in \mathcal{X}$*,* $\mu(A) = BD^*(E)$ *s.t.* $\forall \alpha \in \mathscr{F}(\mathbb{Z})$*,*

$$BD^* \left(\bigcap_{n\in\alpha} (E - n) \right) \geq \mu \left(\bigcap_{n\in\alpha} T^{-n} A \right).$$

Assume $(X, \mathcal{X}, \mu, T)$ *is an m.d.s. and* $A \in \mathcal{X}$*,* $\mu(A) > 0$*. Then* $\exists E \subset \mathbb{Z}, D^*(E) \geq \mu(A)$ *s.t.*

$$\left\{ \alpha \in \mathscr{F}(\mathbb{Z}) : \bigcap_{n\in\alpha} (E - n) \neq \phi \right\} \subset \left\{ \alpha \in \mathscr{F}(\mathbb{Z}) : \mu \left(\bigcap_{n\in\alpha} T^{-n} A \right) > 0 \right\}.$$

This case shows that there is an intimate relationship between a set with positive Banach density and a set with positive measure in a measure-preserving system.

Theorem 18 (Fürstenberg correspondence principle, topological case)**.** *Assume* $E \subset \mathbb{Z}$ *is a syndetic set. Then there exists a minimal system* (X, T) *and a nonempty open subset* U *of* X *s.t.*

$$\left\{ \alpha \in \mathscr{F}(\mathbb{Z}) : \bigcap_{n\in\alpha} T^{-n} U \neq \phi \right\} \subset \left\{ \alpha \in \mathscr{F}(\mathbb{Z}) : \bigcap_{n\in\alpha} (E - n) \neq \phi \right\}.$$

For any minimal system (X, T) *and nonempty open subset* U *of* X*, there exists a syndetic subset* E *of* X *s.t.*

$$\left\{ \alpha \in \mathscr{F}(\mathbb{Z}) : \bigcap_{n\in\alpha} (E - n) \neq \phi \right\} \subset \left\{ \alpha \in \mathscr{F}(\mathbb{Z}) : \bigcap_{n\in\alpha} T^{-n} U \neq \phi \right\}.$$

This case shows that a syndetic set is closely related to a minimal system. As the proof of Szemerédi's theorem via ergodic theory is relatively nontrivial, we won't delve much into it. As shown above, techniques in topological dynamics and ergodic theory play a strong role in combinatorial number theory. Now let's go through other proofs of Szemerédi's theorem.

1. Gowers, proof regarding finite case, giving a nice bound, promoting the branch higher order Fourier analysis.

2. Tao, proof via hypergraph.

Other results or open problems concerning related topics are shown below.

Theorem 19 (Green-Tao)**.** *The set of prime numbers contains arbitrarily long arithmetic progressions.*

Conjecture 1 (Erdös-Turán)**.** *If* $\{p_i\} \subset \mathbb{N}$ *and* $\sum\limits_i \frac{1}{p_i} = +\infty$*, then* $\{p_i\}$ *contains arbitrarily long arithmetic progressions.*

It's a pretty hard problem, even for $k = 3$.

Conjecture 2 (Littlewood)**.** $\forall x, y \in \mathbb{R}, \inf \{n\|nx\|\|ny\| : 0 \neq n \in \mathbb{Z}\} = 0.$

A hitherto known result is shown below.

Theorem 20 (Einsiedler-Katok-Lindenstrauss)**.** *Hausdorff dimension of the set of* $(x, y) \in \mathbb{R}^2$ *which doesn't fit Littlewood conjecture is* 0*.*

The idea of the above theorem comes from Fürstenberg's $\times 2 \times 3$ problem.

3 Fürstenberg's proof of Szemerédi's theorem

Assume (X, T) is a t.d.s., generally people believe it can be a combination of a *simple* system and a *complex* system. Equicontinuity can be a character of *simpleness*.

Definition 12. (X, T) *is said to be equicontinuous if* $\forall \varepsilon > 0, \exists \delta > 0$ *s.t. if* $\rho(x, y) < \delta$*, then* $\rho(T^n x, T^n y) < \varepsilon, \forall n \in \mathbb{Z}$*.*

That is to say, a system is equicontinuous if any two points won't get quite apart from each other after iterations of T.

Definition 13. (X, T) *is said to be distal if* $\forall x \neq y, \inf\limits_{n \in \mathbb{Z}} d(T^n x, T^n y) > 0$*.*

Distal is a concept nearly opposite to equicontinuous which indicates that orbits of any two points will never get too close. Distal systems are originally believed to be very similar to equicontinuous systems.

Definition 14. *(X, T) is said to be transitive if $\exists x \in X$ s.t. $orb(x)$ is dense in X.*

Definition 15. *(X, T) is said to be weakly mixing if $(X \times X, T \times T)$ is transitive.*

The above definition indicates that orbits of a point pair in a weakly mixing systems can traverse nearly everywhere in product space. This type of systems are believed to be *complex*. Fürstenberg(1963) gave a counterexample which shows that minimal distal systems are not necessarily equicontinuous. He also made it clear how a minimal distal system is constructed. A minimal distal system is the inverse limit of equicontinuous extensions starting from a point system.

$$\{pt\} \xleftarrow{\theta_1} X_1 \xleftarrow{\theta_2} X_2 \longleftarrow \cdots \longleftarrow \cdots X$$

For general minimal systems, things are much more sophisticated. Ellis-Glasner-Shapiro(1975) gave a structure theorem which involves equicontinuous extensions, proximal extensions and weakly mixing extensions.

$$\begin{tikzcd} Y \arrow[d, "\pi"'] & Y_1 \arrow[l, "\theta_1"'] & Y_2 \arrow[l, "\theta_2"'] & \cdots \arrow[l] & \cdots Y_\infty \arrow[l] \\ X \end{tikzcd}$$

Inspired by structure theorem for minimal systems, Fürstenberg also proved a structure theorem for ergodic measure-preserving systems which has a more concise form involving only equicontinuous extensions and weakly mixing extensions.

$$\{pt\} \xleftarrow{\theta_1} X_1 \xleftarrow{\theta_2} X_2 \longleftarrow \cdots \longleftarrow \cdots X$$

After proving multiple ergodic recurrence theorem, Fürstenberg posed a question: assume $(X, \mathcal{B}, \mu, T)$ is a measure-preserving system and $f_i \in L^\infty, 1 \leq i \leq d$, is

$$\frac{1}{N} \sum_{n=0}^{N-1} f_1(T^n x) \cdots f_d(T^{dn} x)$$

convergent in L^2 or pointwise sense? Notice that Birkhoff's ergodic theorem and von Neumann's mean ergodic theorem only concern with one term of the above sum. This is one of the most significant questions in ergodic theory in recent three decades. Characteristic factors play an important role when tackling this question. Faced with multiple ergodic average $\frac{1}{N} \sum_{n=0}^{N-1} f_1(T^n x) \cdots f_d(T^{dn} x)$, we wish to convert f_i to a conditional expectation on a simpler σ-algebra Z_d. So we consider $\frac{1}{N} \sum_{n=0}^{N-1} E(f_1|Z_d)(T^n x) \cdots E(f_d|Z_d)(T^{dn} x)$ and expect convergence of these two sums are equivalent. In fact, this question is solved in L^2 sense.

1. $d = 1$, von Neumann's mean ergodic theorem(1940s);
2. $d = 2$, Fürstenberg(1977);
3. $d = 3$, Conze-Lesigne(1984, 1987, 1988) under special conditions, Host-Kra(2001) under general conditions;
4. $d = 4$, Ziegler(2002) under special conditions;
5. Host-Kra(2005, *Ann. of Math.*), Ziegler(2007, *JAMS*).

Host-Kra's approach has vast importance. Assume G is a group, $g, h \in G$. Let $[g, h] := ghg^{-1}h^{-1}$ and $[A, B]$ to be the subgroup spanned by $\{[a, b] : a \in A, b \in B\}$. Subgroup $G_j, j \geq 1$ is defined by recursion $G_1 = G, G_{j+1} = [G_j, G]$.

Definition 16. *G is called a k-step nilpotent group if $G_{k+1}, k \geq 1$ is trivial.*

Definition 17. *Assume G is a k-step nilpotent Lie group and Γ is a discrete cocompact subgroup of G. Then compact manifold $X = G/\Gamma$ is called a k-step nilpotent manifold.*

G can act on X by left shift, denoted by $(g, x) \mapsto gx$. Haar measure μ on X is the only invariant probability measure under this action.

Definition 18. *Assume $\tau \in G$, T is a shift on X s.t. $x \mapsto \tau x$. Then (X, T, μ) or (X, T) is called a k-step nilpotent system.*

In fact, 1-step nilpotent is equivalent to equicontinuous. Let's see some simple examples of nilpotent systems. $d = 1$, rotation on $X = \mathbb{T}^1$ is a 1-step nilpotent system; $d = 2$, assume $X = \mathbb{T}^2$, define $T : X \to X$ s.t. $T(x, y) = (x + \alpha, x + y), \alpha \in \mathbb{R}\backslash\mathbb{Q}$, then (X, T) is a 2-step nilpotent system. Some other results regarding the above question are shown below.

1. Host-Kra constructed a factor denoted by Z_d for any ergodic system $(X, \mathcal{B}, \mu, T)$ and any $d \in \mathbb{N}$ which is the inverse limit of a d-step nilpotent system. They proved that Z_d is the characteristic factor of $\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x) \cdots f_d(T^{dn} x)$ and its convergence in use of Leibman's results.
2. Tao(2008), convergence of $\frac{1}{N}\sum_{n=0}^{N-1} f_1(T_1^n x) \cdots f_d(T_d^n x)$ in L^2 sense, where $T_i \circ T_j = T_j \circ T_i$.
3. Towsner(2009), proof of the above result via nonstandard analysis.
4. Austin(2010), Host(2010), proof of the above result via traditional ergodic theory.
5. Walsh(2012, *Ann. of Math.*), proof of more general convergence in L^2 sense.

Conjecture 3. $\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x) \cdots f_d(T^{dn} x)$ *is convergent in pointwise sense.*

This is a pretty hard problem. Major progresses are listed below.

1. $d = 1$, Birkoff's theorem(1940s);

2. $d = 2$, Bourgain(1989).

Recall that for m.d.s., Fürstenberg gave a structure theorem and Host-Kra gave a fine structure theorem(For any m.d.s and $d \in \mathbb{N}$, they constructed a factor which is the inverse limit of a d-step nilpotent system.). As t.d.s. and m.d.s. are twins, we will naturally ask that is there a corresponding factor in t.d.s.? Historically, it's a hard problem(even for $d = 1$, first proved by Veech). Major results are shown below.

1. Host-Maass(2007), proof under minimal distal condition for $d = 2, 3$.

2. Host-Kra-Maass(2010), proof for any d.

3. Host-Kra-Maass defined a relation $RP^{[d]} \subset X \times X$ and proved that it's a closed invariant equivalence relation in minimal distal systems. Furthermore, $X/RP^{[d]}$ is proved to be a maximal d-step nilfactor.

4. Shao-Ye(2012), assume (X, T) is a minimal system and $d \in \mathbb{N}$, then $RP^{[d]}$ is an equivalence relation and $X/RP^{[d]}$ is a maximal d-step nilfactor of (X, T).

Recall previous contents of this note, the thread running through the development of dynamical systems actually lies in the generalization of structure theorem. We list significant generalizations for various systems in chronological order.

1. Minimal distal systems;

2. general minimal systems;

3. ergodic systems;

4. fine structure of ergodic systems;

5. fine structure of topological dynamical systems.

4 Gowers' proof of Szemerédi's theorem

In fact, Szemerédi's theorem has a finite form.

Theorem 21 (Szemerédi)**.** *Assume $k \in \mathbb{N}$ and $\delta > 0$. $\exists N = N(k, \delta)$ s.t. if $A \subset \{1, 2, \cdots, N\}, |A| \geq \delta N$, then A contains arithmetic progressions of arbitrary length k.*

Estimation of $N(k, \delta)$ is an essential project in combinatorial number theory. Although the proof via ergodic theory is concise enough, it gives no bound for $N(k, \delta)$. However, Gowers' theorem gives a nice bound.

Theorem 22 (Gowers, 2001). *$\forall k \in \mathbb{N}$, there exists a constant $c = c(k) > 0$ s.t. any subset of size $N(\log\log N)^{-c}$ contains arithmetic progressions of arbitrary length k. Moreover, we can take $c = 2^{-2^{k+9}}$.*

This bound for $k = 3$ is then improved to $\mathcal{O}\left(\frac{N}{\log^{1-o(1)} N}\right)$ by Sanders(2011, *Ann. of Math.*). Gowers evaluated different proofs in his article *A new proof of Szemerédi's theorem.*

1. Szemerédi's proof gives a rough bound;

2. Fürstenberg's proof gives no bound;

3. Roth's proof(3-progression) promotes the branch Fourier analysis;

4. Gowers' proof gets started the research of higher order Fourier analysis.

Here I cite a paragraph in T. Tao's article: *Traditionally, Fourier analysis has been focused the analysis of functions in terms of linear phase functions such as the sequence $n \to \mathrm{e}(an) = \mathrm{e}^{2\pi \mathrm{i} an}$. In the recent years, though, applications have arose - particularly in connection with problems involving linear patterns such as arithmetic progressions - in which it has been necessary to go beyond the linear phase, replacing them to higher order functions such as $n \to \mathrm{e}(an^2)$. This has given rise to the subject of higher order Fourier analysis. However, the modern theory of higher order Fourier analysis is very recent indeed(and still incomplete to some extent), beginning with the breakthrough work of Gowers and also heavily influenced by the parallel work in ergodic theory, in particular the seminal work of Host and Kra. This area was also quickly seen to have much in common with areas of theoretical computer science, applications of this theory were given to asymptotics for various linear patterns in prime numbers.*

Finally, I give a quick sketch of Gowers' proof.

1. Define pseudorandomness.

2. Any pseudorandom subset of $\{1, 2, \cdots, N\}$ contains the arithmetic progression of length k that we need.

3. If $A \subset \{1, \cdots, N\}, |A| = \delta N$, A is not pseudorandom, then there exists arithmetic progression $P \subset \{1, \cdots, N\}, p \to \infty (N \to \infty)$ s.t. $|A \cap P| \geq (\delta + \varepsilon)|P|, \varepsilon = \varepsilon(\delta, k) > 0$.