PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 4 (3231–3252)

FROM GRAPH LIMITS TO HIGHER ORDER FOURIER ANALYSIS

BALÁZS SZEGEDY

Abstract

The so-called graph limit theory is an emerging diverse subject at the meeting point of many different areas of mathematics. It enables us to view finite graphs as approximations of often more perfect infinite objects. In this survey paper we tell the story of some of the fundamental ideas in structural limit theories and how these ideas led to a general algebraic approach (the nilspace approach) to higher order Fourier analysis.

1 Introduction

Finite objects are often imperfect approximations of much nicer infinite objects. For example the equations of fluid dynamics or thermodynamics are much simpler if we replace discrete particles by continuous mass. If the particle system is large enough then the continuous model behaves sufficiently similarly to the discrete one in many practical applications. This connection between finite and infinite structures is useful in both directions. Passing to infinite limits can greatly simplify messy calculations with finite objects. Various small quantities (epsilon's), that appear as errors in calculations often disappear in the limit. Beyond getting rid of epsilon's there is a deeper advantage of limit theories. Certain algebraic structures, that are present only in approximate forms in finite structures, appear in a precise form when going to the limit. One of the most surprising discoveries in higher order Fourier analysis is that functions on finite Abelian groups can behave as approximations of functions on inherently non-commutative, topological structures such as nilmanifolds.

The goal of this paper is to take the reader to a journey that starts with a general introduction to structural limits and their applications. We use ergodic theory and graph limit theory to demonstrate a number of fundamental concepts including sampling, quasi-randomness, uniformity norms, convergence, limit space and topologization. We devote

MSC2010: primary 05C99; secondary 37A05, 46B99.

a separate chapter to the non-standard approach which is a powerful tool in limit theories. Finally we turn to higher order Fourier analysis where we explain how nilmanifolds and other even more exotic structures come into play when we look at finite additive structures in the limit.

2 History and basic concepts

The history of structural limits can be traced back all the way to ancient Greeks. Archimedes (287-212 BC) used polygon approximations of the circle to compute its area. Structural limit theories are routinely used in physics. Continuous limits are essential in thermodynamics and fluid dynamics where large but finite particle systems are investigated. On the other hand discrete approximations of continuous objects such as lattice gauge theory play also an important role in physics.

Many of the above limit theories are based on very simple correspondences between finite objects and continuous limit objects. Most of the time the finite approximation is directly related to a continuous space through a prescribed geometric connection. By somewhat abusing the term, we call such limits *scaling limits*. Much more mysterious and surprising limit theories emerged more recently where simple and very general structures are considered such as 0-1 sequences or graphs. In these theories there is no "prescribed" geometry to be approximated. The geometry emerges from the internal "logic" of the structure and thus a great variety of geometric, topological and algebraic structures can appear in the limit. Many of these limit theories are based on taking small random samples from large structures. We call such limit theories *local* limit theories. Some other limit theories are based on observable, large scale properties and we call them *global* limit theories. Furthermore there are hybrid theories such as the local-global convergence of bounded degree graphs Hatami, Lovász, and Szegedy [2014].

Scaling limits of $0-1$ sequences: As an illustration we start with a rather simple (warm up) limit theory for $0-1$ sequences. Later we will see a different and much more complicated theory for the same objects. For $k \in \mathbb{N}$ let $[k] := \{1, 2, \ldots, k\}$. A $0-1$ sequence of length k is a function $f : [k] \to \{0, 1\}$. Assume that we are given a growing sequence $\{f_n\}_{n=1}^{\infty}$ of $0-1$ sequences. *In what sense can we say that these sequences converge?* A simple and natural approach would be to regard the set $[k]$ as a discretization of the $(0, 1]$ interval. This way, for a $0-1$ sequence s of length k we can define the function $\tilde{s} : [0, 1] \to \{0, 1\}$ by $\tilde{s}(x) := s(\lceil kx \rceil)$ (and $\tilde{s}(0) := 0$). Now we can replace the functions f_n by $\tilde{f_n}$ and use one of the readily available convergence notions for functions on $[0, 1]$ such as L^2 or L^1 convergence. Note that they are equivalent for $0-1$ valued functions. The limit object in L^2 is a Lebesgue measurable function $f : [0, 1] \to \{0, 1\}$ with the property that the measure of $f^{-1}(1) \triangle \tilde{f_n}^{-1}(1)$ converges to 0 as n goes to infinity. A much more interesting and flexible limit concept is given by the weak convergence in $L^2([0,1])$. For $0-1$ valued functions this is equivalent with the fact that for every interval $I = [a,b] \subseteq [0,1]$ the measure of $I \cap \tilde{f}_n^{-1}(1)$ converges to some quantity $\mu(I)$ as n goes to infinity. The limit object is a measurable function $f : [0,1] \to [0,1]$ with the property that $\mu(I) = \int_I f d\lambda$ where λ is the Lebesgue measure. If $\tilde{f}_n$ is L^2 convergent then its weak limit is the same as the L^2 limit. However many more sequences satisfy weak convergence.

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of $0-1$ sequences. We say that f_n is **scaling convergent** *if $\{\tilde{f}_n\}_{n=1}^{\infty}$ is a weakly convergent sequence of functions in $L^2([0,1])$. The limit object (scaling limit) is a measurable function of the form $f : [0,1] \to [0,1]$.*

Although scaling convergence is a rather simplistic limit notion we can use it as a toy example to illustrate some of the fundamental concepts that appear in other, more interesting limit theories.

- **Compactness:** *Every sequence of $0-1$ sequences has a scaling convergent subsequence*

- **Uniformity norm:** *Scaling convergence can be metrized through norms. An example for such a norm is the "intervall norm" defined by $\|f\|_{\text{in}} := \sup_I \left|\int_I f d\lambda\right|$. where I runs through all intervals in $(0,1]$. The distance of two $0-1$ sequences f_1 and f_2 (not necessarily of equal lenght) is defined as $\|\tilde{f}_1 - \tilde{f}_2\|_{\text{in}}$.*

- **Quasi randomness:** *A $0-1$ sequence f is ϵ-quasi random with density $p \in [0,1]$ if $\|\tilde{f} - p\|_{\text{in}} \leq \epsilon$. Note that if f_n is a sequence of $0-1$ sequences such that f_n is ϵ_n quasi random with density p and ϵ_n goes to 0 then f_n converges to the constant p function.*

- **Random objects are quasi random:** *Let f_n be a random $0-1$ sequence of length n in which the probabilty of 1 is p. For an arbitrary $\epsilon > 0$ we have that if n is large enough then with probability arbitrarily close to 1 the function f_n is ϵ quasi random.*

- **Low complexity approximation (regularization):** *For every $\epsilon > 0$ there is some natural number N_ϵ such that for every $0-1$ sequence f there is a function $g : [N_\epsilon] \to [0,1]$ such that $\|\tilde{f} - \tilde{g}\|_{\text{in}} \leq \epsilon$. (Note $\tilde{g}$ is defined by the same formula as for $0-1$ sequences and $\tilde{g}$ is a step function on $[0,1]$ with N_ϵ steps.)*

Local limits of $0-1$ sequences: The main problem with scaling convergence is that highly structured sequences such as periodic sequences like $0,1,0,1,0,1,\ldots$ are viewed as quasi random. The above limit concept is based on a prescribed geometric correspondence between integer intervals and the continuous $[0,1]$ interval. A different and much more useful limit concept does not assume any prescribed geometry. It is based on the local statistical properties of $0-1$ sequences. For any given $0-1$ sequence h of length k and f of length $n \geq k$ we define $t(h,f)$ to be the probability that randomly chosen k consecutive bits in f are identical to the sequence h (if $n < k$ then we simply define $t(H,f)$ to be 0).

A sequence $\{f_n\}_{n=1}^{\infty}$ of growing $0-1$ sequences is called **locally** *convergent if for every fix $0-1$ sequence h we have that $\lim_{n\to\infty} t(h,f_n)$ exists.*

This definition was first used by Furstenberg [1977] in his famous correspondence principle stated in the 70's, a major inspiration for all modern limit theories. In Furstenberg's approach finite $0-1$ sequences are regarded as approximations of subsets in certain dynamical systems called measure preserving systems. A measure preserving system is a probability space $(\Omega, \mathfrak{B}, \mu)$ together with a measurable transformation $T : \Omega \to \Omega$ with the property that $\mu(T^{-1}(A)) = \mu(A)$ for every $A \in \mathfrak{B}$.

Furstenberg's correspondence principle for $\mathbb{Z}$: *Let f_n be a locally convergent sequence of $0-1$ sequences. Then there is a measure preserving system $(\Omega, \mathfrak{B}, \mu, T)$ and a measurable set $S \subseteq \Omega$ such that for every $0-1$ sequence $h : [k] \to \{0,1\}$ the quantity $\lim_{n\to\infty} t(h,f_n)$ is equal to the probability that $(1_S(x), 1_S(x^T), \dots, 1_S(x^{T^{k-1}})) = h$ for a random element $x \in \Omega$.*

Note that originally the correspondence principle was stated in a different and more general form for amenable groups. If the group is $\mathbb{Z}$ then it is basically equivalent with the above statement. A measure preserving system is called **ergodic** if there is no set $A \in \mathfrak{A}$ such that $0 < \mu(A) < 1$ and $\mu(A \triangle T^{-1}(A)) = 0$. Every measure preserving system is the combination of ergodic ones and thus ergodic measure preserving systems are the building blocks of this theory.

We give two examples for convergent $0-1$ sequences and their limits. Let α be a fixed irrational number. Then, as n tends to infinity, the sequences $1_{[0,1/2]}(\{\alpha i\}), i = 1,2,\dots,n$ (where $\{x\}$ denotes the fractional part of x) approximate the semicircle in a dynamical system where the circle is rotated by $2\pi\alpha$ degrees. Both the circle and the semicircle appears in the limit. A much more surprising example (in a slightly different form) is given by Host and Kra [2008a]. Let us take two $\mathbb{Q}$-independent irrational numbers α, β and let $a_i := 1_{[0,1/2]}(\{[i\beta]i\alpha - i(i-1)\alpha\beta/2\})$ where $[x]$ denotes the integer part of x. In this case the limiting dynamical system is defined on a three dimensional compact manifold called *Heisenberg nilmanifold*.

Topologization and algebraization: At this point it is important to mention that Furstenberg's correspondence principle does not immediately give a "natural" topological representation of the limiting measure preserving system. In fact the proof yields a system in which the ground space is the compact set $\{0,1\}^{\mathbb{Z}}$ with the Borel σ-algebra, T is the shift of coordinates by one and μ is some shift invariant measure. The notion of isomorphism between systems allows us to switch $\{0,1\}^{\mathbb{Z}}$ to any other standard Borel space. However in certain classes of systems it is possible to define a "nicest" or "most natural" topology . An old example for such a topologization is given by Kronecker systems Furstenberg [1981]. Assume that the measure preserving map T is ergodic and it has the property that $L^2(\Omega)$ is generated by the eigenvectors of the induced action of T on $L^2(\Omega)$. It turns out that such systems can be represented as rotations in compact abelian groups (called Kronecker systems). The problem of topologization is a recurring topic in limit theories. It often comes together with some form of "algebraization" in the frame of which the unique nicest topology is used to identify an underlying algebraic structure that is intimately tied to the dynamics. Again this can be demonstrated on Kronecker systems where finding the right topology helps in identifying the Abelian group structure. Note that there is a highly successful and beautiful story of topologization and algebraization in ergodic theory in which certain factor-systems of arbitrary measure preserving systems (called characteristic factors) are identified as inverse limits of geometric objects (called nilmanifolds) arising from nilpotent Lie groups Host and Kra [2005],Ziegler [2007]. As this breakthrough was also crucial in the development of higher order Fourier analysis we will give more details in the next paragraph. In many limit theories the following general scheme appears .

discrete objects → measurable objects → topological objects → algebraic objects

The first arrow denotes the limit theory, the second arrow denotes topologization and the third arrow is the algebraization.

Factors: Factor systems play a crucial role in ergodic theory. A factor of a measure preserving system $(\Omega, \mathfrak{B}, \mu, T)$ is a sub σ-algebra $\mathfrak{F}$ in $\mathfrak{B}$ that is T invariant (if $B \in \mathfrak{F}$ then $T^{-1}(B) \in \mathfrak{F}$). Note that if $\mathfrak{F}$ is a factor then $(\Omega, \mathfrak{F}, \mu, T)$ is also a measure preserving system. Often there is a duality between a system of "observable quantities" defined through averages and certain factors, called *characteristic factors*. For example the averages

$$t(f) := \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \int_x f(x) f(T^i(x)) f(T^{2i}(x)) d\mu$$

defined for bounded measurable functions satisfy that $t(f) = t(\mathbb{E}(f|\mathfrak{K}))$ where $\mathfrak{K}$ is the Kronecker factor of the system (the unique largest factor that is a Kronecker system) and $\mathbb{E}(f|\mathfrak{K})$ is the conditional expectation with respect to $\mathfrak{K}$. Since conditional expectation is an elementary operation, this means that properties of $t(f)$ can be completely described in terms of Kronecker systems. The ergodic theoretic proof Furstenberg [1977] of Roth theorem Roth [1953] on 3-term arithmetic progressions is based on this fact and a limiting argument using Furstenberg's correspondence principle. It turns out that in every ergodic measure preserving system there is a sequence of increasing, uniquely defined factors $\mathcal{K}_1 \leq \mathcal{K}_2 \leq \dots$ which starts with the Kronecker factor. Similarly to Roth's theorem, the study of k-term arithmetic progressions can be reduced to $\mathcal{K}_{k-2}$. The results in Host and Kra [2005] and Ziegler [2007] give a complete geometric description for these factors in terms of nilsystems. Let G be a k-step nilpotent Lie group and $\Lambda \leq G$ be a co-compact subgroup. The space $N = \{g\Lambda : g \in G\}$ of left cosets of Λ is a finite dimensional compact manifold on which G acts by left multiplication. It is known that there is a unique G invariant probability measure μ on N. We have that $\{N, \mathcal{B}, \mu, g\}$ is a measure preserving system for every $g \in G$ (where $\mathcal{B}$ is the Borel σ-algebra). If g acts in an ergodic way then it is called a k-step nilsystem. It was proved in Host and Kra [2005] and Ziegler [2007] that for every k the factor $\mathcal{K}_k$ of an ergodic system is the inverse limits of k-step nilsystems.

Local and global limits of graphs: Although Furstenberg's correspondence principle gives the first example for a local limit theory, a systematic study of similar structural limit theories started much later. The general program of studying structures in the limit became popular in the early 2000's when graph limit theory was born Benjamini and Schramm [2001],Lovász and Szegedy [2006],Lovász and Szegedy [2007],Borgs, J. Chayes, Lovász, Sós, Szegedy, and Vesztergombi [2006],Borgs, J. T. Chayes, Lovász, Sós, and Vesztergombi [2008],Borgs, J. T. Chayes, Lovász, Sós, and Vesztergombi [2012]. The motivation to develop an analytic theory for large networks came partially from applied mathematics. The growing access to large networks such as social networks, internet graphs and biological networks like the brain generated a demand for new mathematical tools to understand their approximate structure. Another motivation came from extremal combinatorics where inequalities between subgraph densities are extensively studied. An analytic view of graphs enables the use of powerful methods such as differential calculus to solve extremal problems. Similarly to ergodic theory certain graph sequences approximate infinite structures which can not be perfectly represented by finite objects. It turns out that there are simple extremal problems for graphs which have no precise finite solutions but a nice exact solution appears in the limit. This is somewhat similar to the situation with the inequality $(x^2-2)^2 \geq 0$ which has no precise solution in $\mathbb{Q}$ but it has two solutions in $\mathbb{R}$.

Similarly to $0-1$ sequences graph convergence can be defined through converging sample distributions and thus the convergence notion will depend on the sampling method. Quite surprisingly there are two different natural sampling methods. The first one works well if the graph has a non negligible edge density (such graphs are called *dense*) and the second one is defined only for bounded degree graphs. Note that on n vertices a dense graph has cn^2 edges for some non negligible $c > 0$ whereas a bounded degree graph has

cn edges for some bounded c. This means that dense and bounded degree graphs are at the two opposite ends of the density spectrum. If a graph is neither dense nor bounded degree then we call it *intermediate*.

Let $G = (V, E)$ be a finite graph. In the first sampling method we choose k vertices $v_1, v_2, \dots, v_k$ independently and uniformly from V and take the graph G_k spanned on these vertices. We regard G_k as a random graph on $[k]$. For a graph H on the vertex set $[k]$ let $t^0(H, G)$ denote the probability that $G_k = H$. In dense graph limit theory, a graph sequence $\{G_n\}_{n=1}^{\infty}$ is called convergent if for every fixed graph H the limit $\lim_{i\to\infty} t^0(H, G)$ exists. Another equivalent approach is to define $t(H, G)$ as the probability that a random map from $V(H)$ to $V(G)$ is a *graph homomorphism* i.e. it takes every edge of H to an edge of G. This number is called the homomorphism density of H in G. In a sequence $\{G_n\}_{n=1}^{\infty}$, the convergence of $t(H, G_n)$ for all graphs H is equivalent with the convergence of $t^0(H, G_n)$ for all graphs H. The advantage of using homomorphism densities is that they have nicer algebraic properties such as multiplicativity and reflection positivity Lovász and Szegedy [2006].

For the second sampling method let $\mathcal{G}_d$ denote the set of finite graphs with maximum degree at most d. Let furthermore $\mathcal{G}_d^r$ denote the set of graphs of maximum degree at most d with a distinguished vertex o called the *root* such that every other vertex is of distance at most r from o. Now if $G = (V, E)$ is in $\mathcal{G}_d$ then let v be a uniform random vertex in V. Let $N_r(v)$ denote the v-rooted isomorphism class of the radius r-neighborhood of v in G. We have that $N_r(v)$ is an element in $\mathcal{G}_d^r$ and thus the random choice of v imposes a probability distribution $\mu(r, G)$ on $\mathcal{G}_d^r$. A graph sequence $\{G_n\}_{n=1}^{\infty}$ is called Benjamini-Schramm convergent if $\mu(r, G_n)$ is convergent in distribution for every fixed r. The convergence notion was introduced in the paper Benjamini and Schramm [2001] to study random walks on planar graphs. Colored and directed versions of this convergence notion can be also introduced in a similar way. Benjamini-Schramm convergence provides a rather general framework for many different problems. Note that it generalizes the local convergence of $0-1$ sequences because one can represent finite $0-1$ sequences by directed paths with 0 and 1 labels on the nodes. Limit objects for Benjamini-Schram convergent sequences are probability distributions on infinite rooted graphs with a certain measure preserving property that generalizes the concept of measure preserving system. Note that Benjamini-Schramm convergence is closely related to group theory. A finitely presented group is called Sofic if its Cayley graph is the limit of finite graphs in which the edges are directed and labeled by the generators of the group. Sofic groups are much better understood than general abstract groups. The study of sofic groups is a fruitful interplay between graph limit theory and group theory.

Global aspects of graph limit theory arise both in the dense and the bounded degree frameworks. In case of dense graph limit theory the local point of view is often not strong enough. Although it turns out that one can represent convergent sequences by so-called graphons Lovász and Szegedy [2006] i.e. symemtric measurable functions of the form $W : [0,1]^2 \to [0,1]$ a stronger theorem that connects the convergence with Szemerédi's regularity lemma is more useful. Szemerédi's famous regularity lemma is a structure theorem describing the large scale structure of graphs in terms of quasi rendom parts. A basic compactness result in dense graph limit theory Lovász and Szegedy [2007] (see also Theorem 1) connects the local and global point of views. This is used in many applications including property testing Lovász and Szegedy [2010b] and large deviation principles Chatterjee and Varadhan [2011].

The Benjamini-Schramm convergence is inherently a local convergence notion and thus it is not strong enough for many applications. For example random d-regular graphs are locally tree-like but they have a highly non-trivial global structure that has not been completely described. The formalize this problem one needs a refinement of Benjamini-Schramm convergence called local-global convergence Hatami, Lovász, and Szegedy [2014]. The concept of local-global convergence was successfully used in the study of eigenvectors of random regular graphs. It was proved by complicated analytic, information theoretic and graph limit methods in Backhausz and Szegedy [2016] that almost eigenvectors of random regular graphs have a near Gaussian entry distribution. This serves as an illustrative example for the fact that deep results in graph theory can be obtained through the limit approach.

We have to mention that the branch of graph limit theory that deals with intermediate graphs (between dense and bounded degree) is rather underdeveloped. There are numerous competing candidates for an intermediate limit theory Borgs, J. T. Chayes, Cohn, and Zhao [n.d.],Borgs, J. T. Chayes, Cohn, and Zhao [2018],Szegedy [n.d.(c)],Kunszenti-Kovács, Lovász, and Szegedy [2016],Nesetril and Mendez [2013], Frenkel [2018] but they have very few applications so far. The hope is that at least one of these approaches will become a useful tool to study real life networks such as connections in the brain or social networks. These networks are typically of intermediate type.

Limits in additive combinatorics and higher order Fourier analysis: Let A be a finite Abelian group and S be a subset in A. Many questions in additive combinatorics deal with the approximate structure of S. For example Szemerédi's theorem can be interpreted as a result about the density of arithemtic progressions of subsets in cyclic groups. It turns out that limit approaches are natural in this subject. Let $M \in \mathbb{Z}^{m\times n}$ be an integer matrix such that each element in M is coprime to the order of A. Then we can define the density of M in the pair (A, S) as the probability that $\sum M_{i,j} x_j \in S$ holds for every i with random uniform independent choice of elements $x_1, x_2, \ldots, x_n \in A$. For example the density of 3 term arithmetic progressions in S is the density of the matrix $((1,0),(1,1),(1,2))$ in S. We say that a sequence $\{(A_i, S_i)\}_{i=1}^{\infty}$ is convergence if the density of all coprime matrix M in the elements of the sequence converges. This type of convergence was first investigated in Szegedy [n.d.(b)] and limit objects were also constructed. The subject is deeply connected to Gowers norms and the subject of higher order Fourier analysis. We give more details in chapter 5.

3 Dense graph limit theory

A *graphon* is a measurable function of the form $W : [0,1]^2 \to [0,1]$ with the property $W(x,y) = W(y,x)$ for every $x, y \in [0,1]$. Let $\mathcal{W}$ denote the set of all graphons. If G is a finite graph on the vertex set $[n]$ then its graphon representation W_G is defined by the formula

$$W(x,y) = 1_{E(G)}(\lceil nx \rceil, \lceil ny \rceil).$$

For a graph H on the vertex set $[k]$ let

$$t(H,W) := \int_{x_1,x_2,\dots,x_k \in [0,1]} \prod_{(i,j)\in E(H)} W(x_i,x_j)\, dx_1 dx_2 \dots dx_k.$$

The quantity $t(H,W)$ is an analytic generalization of the so called homomorphism density defined for finite graphs. This is justified by the easy observation that $t(H,G) = t(H,W_G)$. We will need the so-called cut norm $\|.\|_\square$ on $L^\infty([0,1]^2)$. Let $F : [0,1]^2 \to \mathbb{R}$ be a bounded measurable function. Then

$$\|F\|_\square := \sup_{A,B \subseteq [0,1]} \left| \int_{A\times B} F(x,y)\, dx dy \right|$$

where A and B run through all measurable sets in $[0,1]$. Using this norm we can introduce a measure for "similarity" of two graphons U and W by $\|U - W\|_\square$. However this is not the similarity notion that we use for convergence. We need to factor out by graphon ismorphisms. If $\psi : [0,1] \to [0,1]$ is a measure preserving transformation the we define $W^\psi(x,y) := W(\psi(x), \psi(y))$. It is easy to check that this transformation on graphons preserves the homomorphism densities: $t(H,W) = t(H,W^\psi)$ holds for every finite graph H. The next distance was introduced in Lovász and Szegedy [2007] :

$$\delta_\square(U,W) := \inf_{\phi,\psi:[0,1]\to[0,1]} \|U^\phi - W^\psi\|_\square$$

where ϕ and ψ are measure preserving transformations. It is easy to check that $\delta_\square$ is a pseudometrics i.e. it satisfies all axioms except that $d(x,y) = 0$ does not necessarily imply that $x = y$. In order to get an actual metrics we have to factor out by the equivalence relation $\sim_{\delta_\square}$ defined by $x \sim_{\delta_\square} y \Leftrightarrow d(x,y) = 0$. Let $\mathcal{X} := \mathcal{W}/\sim_{\delta_\square}$. Since

$\delta_\square(U,W) = 0$ implies that $t(H,U) = t(H,W)$ holds for every graph H we have that $t(H,-)$ is well defined on $\mathfrak{X}$. The following result Lovász and Szegedy [2006],Lovász and Szegedy [2007] in graph limit theory is fundamental in many applications.

Theorem 1. *We have the following statements for the metric space* $(\mathfrak{X}, \delta_\square)$.

1. *The metric* $\delta_\square$ *defines a compact, Hausdorff, second countable topology on* $\mathfrak{X}$.
2. *The function* $X \to t(H,X)$ *is a continuous function on* $\mathfrak{X}$ *for every finite graph* H.

Two important corollaries are the following.

Corollary 3.1. *Assume that* $\{G_i\}_{i=1}^{\infty}$ *is a sequence of graphs such that* $f(H) := \lim_{i\to\infty} t(H,G_i)$ *exists for every finite graph* H. *Then there is a graphon* $W \in \mathfrak{W}$ *such that* $f(H) = t(H,W)$ *holds for every* H.

Corollary 3.2. *Szemerédi's regularity lemma Szemerédi [1978] (even in stronger forms) follows from Theorem 1.*

Note that, although Corollary 3.1 may be deduced from earlier results on exchangeability Aldous [1985], Theorem 1 combines both the local and global aspects of convergence and so it is a stronger statement. In some sense it can be regarded as a common generalization of both Szemerédi's regularity lemma Szemerédi [1978] and a result on exchangeability Aldous [1985].

Topologization of graph limit theory: In the definition of a graphon $W : [0,1]^2 \to [0,1]$ the $[0,1]$ interval on the left hand side is replaceable by any standard probability space (Ω,μ). In general we need that (Ω,μ) is atomless but for certain special graphons even atoms maybe allowed. Note that the values of W represent probabilities and so the $[0,1]$ interval is crucial on the right hand side. Thus the general form of a graphon is a symmetric measurable function $W : \Omega\times\Omega \to [0,1]$. Homomorphism densities $t(H,W)$ are defined for all such general graphons and two of them are equivalent if all homomorphism densities are the same. The folowing question arises: *Given a graphon* W. *Is there a most natural topological space* X *and Borel measure* μ *on* X *such that* W *is equivalent with a graphon of the form* $W' : X^2 \to [0,1]$? An answer to this question was given in Lovász and Szegedy [2010a]. For a general graphon $W : \Omega\times\Omega \to [0,1]$ there is a unique purified version of W on some Polish space X with various useful properties. The language of topologization induced a line of exciting research in extremal combinatorics. Here we give a brief overview on applications of graphons in extremal graph theory.

Extremal graphs and graphons: The study of inequalities between subgraph densities and the structure of extremal graphs is an old topic in extremal combinatorics. A classical example is Mantel's theorem which implies that a triangle free graph H on $2n$ vertices maximizes the number of edges if H is the complete bipartite graph with equal color classes. Another example is given by the Chung-Graham-Wilson theorem Chung, Graham, and Wilson [1989]. If we wish to minimize the density of the four cycle in a graph H with edge density $1/2$ then H has to be sufficiently quasi random. However the perfect minimum of the problem (that is $1/16$) can not be attained by any finite graph but one can get arbitrarily close to it. Both statements can be conveniently formulated in the framework of dense graph limit theory. In the first one we maximize $t(e, G)$ in a graph G with the restriction that $t(C_3, G) = 0$ (where e is the edge and C_3 is the triangle). In the second one we fix $t(e, G)$ to be $1/2$ and we minimize $t(C_4, G)$. Since the graphon space is the completion of the space of graphs it is very natural to investigate these problems in a way that we replace G by a graphon W. If we fix finite graphs $H_1, H_2, \ldots, H_k$ then all possible inequalities between $t(H_1, W), t(H_2, W), \ldots, t(H_k, W)$ are encoded in the k-dimesional point set

$$\mathfrak{L}(H_1, H_2, \ldots, H_k) := \{(t(H_1, W), t(H_2, W), \ldots, t(H_k, W)) : W \in \mathfrak{W}\}.$$

Note that this is a closed subset in $[0, 1]^k$. As an example let e be a single edge and let P_2 denote the path with two edges. It is easy to prove that $t(P_2, W) \geq t(e, W)^2$. This inequality is encoded encoded in $\mathfrak{L}(e, P_2)$ is the form that $\mathfrak{L}(e, P_2) \subseteq \{(x, y) : y \geq x^2\}$. We have however that $\mathfrak{L}(e, P_2)$ carries much more information. The shape of $\mathfrak{L}(H_1, H_2, \ldots, H_k)$ is know in very few instances. It took decades of research to completely describe the two dimensional shape $\mathfrak{L}(e, C_3)$ which gives all possible inequalities between $t(e, W)$ and $t(C_3, W)$. The characterization of $\mathfrak{L}(e, C_3)$ was completed by Razborov [2008] partially using limit methods (a certain differentiation on the graph limit space). Another direction of research investigates the structure of a graphon W with given subgraph densities. *A graphon W is called finitely forcible Lovász and Szegedy [2011] if there are finitely many graphs $H_1, H_2, \ldots, H_k$ such that if $t(H_i, W') = t(H_i, W)$ holds for $i = 1, 2, \ldots, k$ for some $W' \in \mathfrak{W}$ then W' is equivalent with W.* The motivation to study finitely forcible graphons is that they represent a large family of extremal problems with unique solution. It is very natural to ask how complicated can extremal graph theory get at the structural level. Originally it was conjectured that finitely forcible graphons admit a step function structure which is equivalent with the fact that the topologization of the graphon is a finite space. This was disproved in Lovász and Szegedy [ibid.] and various examples were given with more interesting underlying topolgy. However the topology in all of these examples is compact and finite dimensional. It was asked in Lovász and Szegedy [ibid.] whether this is always the case. Quite surprisingly both conjectures turned out to be false. Extremal problems with strikingly complicated topologies were constructed in Glebov, Klimosova, and Kral [2014], Cooper, Kaiser, Noel, et al. [2015]. This gives a very strong justification of graph limit theory in extremal cobinatorics by showing that complicated infinite structures are somehow encoded into finite looking problems. The marriage between extremal graph theory and graph limit theory has turned into a growing subject with surprising results. It brought topology and analysis into graph theory and gave a deep insight into the nature and complexity of extremal structures.

4 The ultra limit method

The use of ultra products in structural limit theory Elek and Szegedy [2012],Szegedy [n.d.(b)] was partially motivated by the great complexity of the proofs in hypergraph regularity theory Nagle, Rödl, and Schacht [2006],Rödl and Schacht [2007],Rödl and Skokan [2004] and later in higher order Fourier analysis. For example the hypergraph removal lemma is a simple to state and beautiful theorem, but its combinatorial proofs are extremely complicated. This may not be surprising in the light of the fact that it implies Szemerédi's famous theorem Szemerédi [1975] on arithmetic progressions even in a multi-dimensional form Solymosi [2004]. However it was observed in Elek and Szegedy [2012] that great simplification can be made to these proofs if one works in a limiting setting. The limit theory which is particularly useful here is based on ultra products of measure spaces. Without going to technical details we give an overview of a scheme that was successfully used in hypergraph theory Elek and Szegedy [ibid.] and additive combinatorics Szegedy [n.d.(b)]. *This scheme is based on the philosophy that if there is any "reasonable" limit of a sequence of structures* $S_1, S_2, \ldots$ *then it has to appear somehow on the ultra product space* $\mathbf{S} := \prod_\omega S_i$ *where* ω *is a non-principal ultra filter. Usually the limit object appears as a factor space of* $\mathbf{S}$ *endowed with some structure obtained from* $\mathbf{S}$. In the followings we give a strategy that unifies some of the applications of the ultra limit method without aiming for full generality.

Introducing a limit theory:

1. **Structures:** *Let* $\mathfrak{F}$ *be a family of structures (for example finite graphs, hypergraphs or subsets in finite or more generally compact Abelian groups).*

2. **Function representation:** *Represent each element* $F \in \mathfrak{F}$ *as a function on some simpler structure* $Q \in \mathfrak{Q}$. *We assume that each structure* $Q \in \mathfrak{Q}$ *is equipped with a probability measure* μ_Q. *Let* $\mathfrak{R}$ *denote the representation function* $\mathfrak{R} : \mathfrak{F} \mapsto \cup_{Q\in\mathfrak{Q}} L^\infty(Q, \mu_Q)$.

   For example, in case of graphs, $\mathfrak{Q}$ is the family of finite product sets of the form $V \times V$ and $\mu_{V\times V}$ is the uniform measure. The representation function for a graph $G = (V, E)$ is given by $\mathfrak{R}(G) := 1_E$ on the product set $V \times V$. In other words $\mathfrak{R}(G)$ is the adjacency matrix of G. For k uniform hypergraphs $\mathfrak{Q}$ is the set of power sets of the form V^k. If $\mathfrak{F}$ is the set of measurable subsets in compact Abelian groups then $\mathfrak{Q}$ is the set of compact Abelian groups with the Haar measure.

3. **Moments:** *Define a set of moments* $\mathfrak{M}$ *such that for each* $m \in \mathfrak{M}$ *and* $Q \in \mathfrak{Q}$ *there is a functional* $m : L^\infty(Q, \mu_Q) \to \mathbb{C}$.

   Note that the name ”moment” refers to the fact that elements in $L^\infty(Q, \mu)$ are random variables. Here we use generalizations of classical moments which make use of the underlying structure of Q. For example if $\mathfrak{F}$ is the family of finite graphs, then $\mathfrak{M}$ is the set of finite directed graphs. Each finite directed graph $H = ([k], F)$ defines a moment by $t(H, W) := \mathbb{E}_{x_1,x_2,\dots,x_k \in V}(\prod_{(i,j)\in F} W(x_i, x_j))$ for an arbitrary function $W : V \times V \to \mathbb{C}$. (If W is symmetric then $t(H, W)$ does not depend on the direction on H.) Note that if we allow multiple edges in H then the n-fold single edge corresponds to the n-th classical moment $\mathbb{E}(W^n)$. For functions on Abelian groups, moments are densities of additive patterns. For example if $f : A \to \mathbb{C}$ then the 3 term arithmetic progression density in f is defined by $\mathbb{E}_{x,t} f(x)f(x+t)f(x+2t)$. Similarly the parallelogram density is defined by $\mathbb{E}_{x,t_1,t_2} f(x)f(x+t_1)f(x+t_2)f(x+t_1+t_2)$.

4. **Convergence:** *Define a limit notion in* $\mathfrak{F}$ *in the following way. A sequence* $\{F_i\}_{i=1}^\infty$ *in* $\mathfrak{F}$ *is called convergent if for every* $m \in \mathfrak{M}$ *we have that* $\lim_{i\to\infty} m(\mathfrak{R}(F_i))$ *exists.*

   Note that this convergence notion naturally extends to functions on structures in $\mathfrak{Q}$. This allows us to define convergent sequences of matrices, multidimensional arrays (functions on product sets) or functions on Abelian groups.

5. **Quasi randomness and similarity** *Define a norm* $\|.\|_U$ *on each function space* $L^\infty(Q, \mu_Q)$ *that measures quasi randomness such that if* $\|f\|_U$ *is close to* 0 *then* f *is considered to be quasi random. We need the property that for every* $m \in \mathfrak{M}$ *and* $\epsilon > 0$ *there is* $\delta > 0$ *such that if* $f, g \in L^\infty(Q, \mu_Q)$ *satisfy* $|f|, |g| \le 1$ *and* $\|f - g\|_U \le \delta$ *then* $|m(f) - m(g)| \le \epsilon$.

   In case of graphs we can use the four cycle norm $\|f\|_U := t(C_4, f)^{1/4}$ or an appropriately normalized version of the cut norm. For hypergraphs we can use the so-called octahedral norms. On Abelian groups we typically work with one of the Gowers norms Gowers [2001],Gowers [1998] depending on the set of moments we need to control.

The ultra limit method:

1. **The ultra limit space** *Let* $\mathbf{Q}$ *denote the ultra product of some sequence* $\{Q_i\}_{i=1}^\infty$ *in* $\mathfrak{Q}$. *There is an ultra product* σ*-algebra* $\mathfrak{A}$ *and an ultra product measure* μ *on* $\mathbf{Q}$ *that comes from the measure space structures on* Q_i *by a known construction. Each uniformly bounded function system* $\{f_i \in L^\infty(Q_i)\}_{i=1}^\infty$ *has an ultra limit function* $f \in L^\infty(\mathbf{Q}, \mathfrak{A}, \mu)$ *and each function* $g \in L^\infty(\mathbf{Q}, \mathfrak{A}, \mu)$ *arises this way*

(up to 0 measure change). We can also lift the moments in $\mathfrak{M}$ and the norm $\|.\|_U$ to $L^\infty(\mathbf{Q}, \mathfrak{A}, \mu)$. Note that $\|.\|_U$ usually becomes a seminorm on $L^\infty(\mathbf{Q}, \mathfrak{A}, \mu)$ and thus it can take 0 value.

2. **Characteristic factors of the ultra limit space:** *Similarly to ergodic theory our goal here is to identify a sub σ-algebra $\mathfrak{A}_U$ in $\mathfrak{A}$ such that $\|.\|_U$ is a norm on $L^\infty(\mathbf{Q}, \mathfrak{A}_U, \mu)$ and $\|f - \mathbb{E}(f|\mathfrak{A}_U)\|_U = 0$ holds for every $f \in L^\infty(\mathbf{Q}, \mathfrak{A}, \mu)$. With this property we also obtain that $m(f) = m(\mathbb{E}(f|\mathfrak{A}_U))$ and $m(f - \mathbb{E}(f|\mathfrak{A}_U)) = 0$ holds.*

   These equations imply that the information on the values of the moments is completely encoded in the projection to $\mathfrak{A}_U$. Once we identify this σ-algebra, the goal is to understand its structure. Note that $\mathfrak{A}_U$ is a huge non-separable σ-algebra. The next step is to reduce it to a separable factor.

3. **Separable realization:** *Let us fix a function in $f \in L^\infty(\mathbf{Q}, \mathfrak{A}_U, \mu)$. Our goal here is to find a separable (countable based) sub σ-algebra $\mathfrak{A}_f$ in $\mathfrak{A}_U$ which respects certain operations that come from the algebraic structure of* $\mathbf{Q}$ *but at the same time $f \in L^\infty(\mathbf{Q}, \mathfrak{A}_f, \mu)$.*

   Note that f itself generates a separable sub σ-algebra in $\mathfrak{A}$. However this σ-algebra does not automatically respects the algebraic structure on $\mathbf{Q}$. For example in case of graphs $\mathbf{Q} = \mathbf{V} \times \mathbf{V}$ where $Q_i = V_i \times V_i$ and $\mathbf{V}$ is the ultraproduct of $\{V_i\}_{i=1}^\infty$. Here we look for a separable σ-algebra that respects this product structure i.e. it is the "square" of some σ-algebra on $\mathbf{V}$.

4. **Topologization and algebraization (the separable model):** *The set* $\mathbf{Q}$ *is naturally endowed with a σ-topology Szegedy [n.d.(b)]. Our goal here is to find a compact, Hausdorff, separable factor topology with factor map $\pi : \mathbf{Q} \to X$ such that the σ algebra generated by π is $\mathfrak{A}_f$. We also wish to construct an algebraic structure on X such that π is a morphism in an appropriate category. This way we can find a Borel measurable function $f' : X \to \mathbb{C}$ such that $f' \circ \pi = f$ holds almost everywhere. Now we regard (X, f') as a separable model for the non standard object* $\mathbf{Q}$ *together with f.*

   This is the part of the method where we came back from the non standard universe to the world of reasonable, constructible structures. Note however that many times the algebraic structure on X is (has to be) more general than the structures in $\mathbb{Q}$. It is in a class $\overline{\mathbb{Q}}$ containing $\mathbb{Q}$. In other words the non standard framework "teaches" us how to extend the class $\mathbb{Q}$ to get a limit closed theory. This was very beneficial in case of higher order Fourier analysis where the non standard framework "suggested" the class of nilspaces Camarena and Szegedy [2010].

5. **Using the separable model for limit theory and regularization:** *There are two main application of the separable model of a function on* $\mathbf{Q}$. *The first one is that if* $\{f_i : Q_i \to \mathbb{R}\}_{i=1}^{\infty}$ *is a convergent sequence of uniformly bounded functions then the separable model for their ultra limit is an appropriate limit object for the sequence. The second application is to prove regularity lemmas in* $\mathfrak{F}$ *or more generally for functions on elements in* $\mathfrak{Q}$.

   Let $Q \in \mathfrak{Q}$ and $f \in L^\infty(Q, \mu_Q)$ with $|f| \leq 1$. A regulaity lemma is a decomposition theorem of f into a structured part and a quasi random part.

5 Higher order Fourier analysis: limit theory and nilspaces

Fourier analysis is a very powerful tool to study the structure of functions on finite or more generally compact Abelian groups Rudin [1990]. If $f : A \to \mathbb{C}$ is a measurable function in $L^2(A, \mu_A)$ (where μ_A is the Haar measure on A) then there is a unique decomposition of the form $f = \sum_{\chi \in \hat{A}} c_\chi \chi$ converging in L^2 where $\hat{A}$ is the set of linear characters of A and the numbers $c_\chi \in \mathbb{C}$ are the Fourier coefficients. Note that for finite A the Haar measure is the uniform probability measure on A. The uniqueness of the decomposition follows from the fact that $\hat{A}$ is a basis in the Hilbert space $L^2(A, \mu_A)$. It is also an important fact that the characters them self form a commutative group, called *dual group*, with respect to point-wise multiplication.

In 1953 Roth used Fourier analysis to prove a lower bound for the number of 3 term arithmetic progressions in subsets of cyclic groups Roth [1953]. In particular it implies that positive upper density sets in $\mathbb{Z}$ contain non trivial 3-term arithmetic progressions. The same problem for k-term arithmetic progressions was conjectured by Erdős and Turán in 1936 and solved by Szemerédi [1975] in 1974. Szemerédi's solution is completely combinatorial. It is quite remarkable that despite of the strength of Fourier analysis it is less useful for higher than 3 term arithmetic progressions (although it was extended for 4 term progressions in 1972 Roth [1972]). A deep reason for this phenomenon was discovered by Gowers [2001], Gowers [1998] in 1998. His results gave a new insight into how densities of additive patterns behave in subsets of Abelian groups by revealing a hierarchy of structural complexity classes governed by the so-called Gowers norms. Roughly speaking, at the bottom of the hierarchy there is the universe of structures, or observable quantities that can be detected by the dominant terms in Fourier decompositions. In particular the density of 3-term arithmetic progressions belongs to this part of the hierarchy. However it turns out that Fourier analysis does not go deep enough into the structure of a function (or characteristic function of a set) to clearly detect 4 or higher term arithmetic progressions: this information may be "dissolved" into many small Fourier terms. The Gowers norms $U_2, U_3, \ldots$ provide an increasingly fine way of comparing functions from a structural point of view. The U_2 norm is closely connected to Fourier analysis. Gowers formulated the following very far reaching hypothesis: *for every natural number k there is a k-th order version of Fourier analysis that is connected to the U_{k+1} norm.* In particular k-th order Fourier analysis should be the appropriate theory to study $k+2$ term arithmetic progressions.

Gowers coined the term *"higher order Fourier analysis"* and he developed a version of it that was enough to improve the bounds in Szemerédi's theorem. However the following question was left open: *Is there some structural decomposition theorem in k-th order Fourier analysis that relates the U_k norm to some algebraically defined functions similar to characters?* The intuitive meaning behind these norms is that $\|f\|_{U_k}$ is small if and only if f is quasi random in $k-1$-th order Fourier analysis. A way of posing the previous question is the following: For each k let us find a set of nice enough functions (called structured functions) such that for $|f| \leq 1$ we have that $\|f\|_{U_k}$ is non negligible if and only if f has a non negligible correlation with one of these functions. (In case of the U_2 norm the set of linear characters satisfy this property.) Such a statement is called an *inverse theorem for the U_k norm*. Despite of that fact that an inverse theorem is seemingly weaker than a complete decomposition theorem, known techniques can be used to turn them into Szermerédi type regularity lemmas.

There are several reasons why higher order Fourier analysis can't be as exact and rigid as ordinary Fourier analysis. One obvious reason is that linear characters span the full Hilbert space $L^2(A, \mu_A)$ and thus there is no room left for other basis elements. Quite surprisingly this obstacle disappears in the limit. If we have an increasing sequence of finite Abelian groups A_i, then there are very many function sequences $f_i : A_i \to \mathbb{C}$ such that $\|f_i\|_2 = 1$ and f_i is more and more orthogonal to every character χ i.e. $\|\hat{f_i}\|_\infty$ goes to 0. On the ultra limit group $\mathbf{A}$ we find that ultra limits of linear characters generate only a small part of the Hilbert space $L^2(\mathbf{A}, \mu)$. This leaves more than enough room for higher order terms. In the rest of this chapter we give a short introduction to Gowers norms and explain how they lead to exact higher order Fourier decompositions in the limit. Then we explain an even deeper theory describing the algebraic meaning of these decompositions in terms of nilspace theory Host and Kra [2008b],Camarena and Szegedy [2010]. This leads to general inverse theorems and regularity lemmas for the Gowers norms on arbitrary compact abelian groups Szegedy [n.d.(b)]. Note that another but not equivalent approach to inverse theorems was developed by Green, Tao, and Ziegler [2012],Tao and Ziegler [2012] for various classes on abelian groups. They were particularly interested in inverse theorems from Gowers norms for integer sequences Green, Tao, and Ziegler [2012] since it leads to spectacular number theoretic applications developed by and Green and Tao [2010]. It is important to mention that in Ergodic theory, the Host-Kra seminorms Host and Kra [2005] play a similar role in measure preserving systems as Gowers norms do on compact Abelian groups. Thus a tremendous amount of great ideas were transported from ergodic theory (especially from the works of Host and Kra [ibid.],Host and Kra [2008b],Host and Kra [2008a] and Ziegler [2007]) to higher order Fourier analysis.

Let $\Delta_t : L^\infty(A, \mu_A) \to L^\infty(A, \mu_A)$ denote the multiplicative "differential operator" defined by $(\Delta_t f)(x) := f(x)\overline{f(x+t)}$. Since $\Delta_{t_1}(\Delta_{t_2}(f)) = \Delta_{t_2}(\Delta_{t_1}(f))$ we can simply use multi indices $\Delta_{t_1,t_2,\ldots,t_k}$. The U_k norm of a function f is defined by

$$\|f\|_{U_k} := \left(\mathbb{E}_{x,t_1,t_2,\ldots,t_k}(\Delta_{t_1,t_2,\ldots,t_k} f)(x)\right)^{1/2^k}.$$

Note that U_k is only a norm if $k \geq 2$. If $k = 1$ then $\|f\|_{U_1} = |\mathbb{E}(f)|$ and thus it is a seminorm. It was observed by Gowers that $\|f\|_{U_2}$ is the l_4 norm of the Fourier transform of f showing the connection between the U_2 norm and Fourier analysis.

We say that $\{A_i\}_{i=1}^\infty$ is a growing sequence of compact Abelian groups if their sizes tend to infinity. (If the size of a group A is already infinite then the constant sequence $A_i = A$ satisfies this.) Let $\mathbf{A}$ be the ultra product of a growing sequence of Abelian groups. The Gowers norms are also defined for functions in $L^\infty(\mathbf{A}, \mathfrak{A}, \mu)$. Quite surprisingly, all the Gowers norms become seminorms in this non-standard framework. For each U_k the set $W_k = \{f : \|f\|_{U_k} = 0\}$ is a linear subspace in $L^\infty(\mathbf{A}, \mathfrak{A}, \mu)$. It turns out that the orthogonal space of W_k in L^2 is equal to $L^2(\mathbf{A}, \mathfrak{F}_{k-1}, \mu)$ for some sub σ-algebra $\mathfrak{F}_{k-1}$ in $\mathfrak{A}$. Intuitively, $\mathfrak{F}_k$ is the σ-algebra of the k-th order structured sets. The next theorem from Szegedy [n.d.(a)] (and proved with different methods in Szegedy [n.d.(b)]) uses these σ-algebras to define higher order Fourier decompositions.

Theorem 2. *Let* $\mathbf{A}$ *be as above. Then*

1. *For each natural number k there is a unique* σ*-algebra* $\mathfrak{F}_k \subset \mathfrak{A}$ *such that* U_{k+1} *is a norm on* $(\mathbf{A}, \mathfrak{F}_k, \mu)$ *and* $\|f - \mathbb{E}(f|\mathfrak{F}_k)\|_{U_{k+1}} = 0$ *for every* $f \in L^\infty(\mathbf{A}, \mathfrak{A}, \mu)$.

2. *We have that* $L^2(\mathbf{A}, \mathfrak{F}_k, \mu) = \bigoplus_{W \in \hat{\mathbf{A}}_k} W$ *where* $\hat{\mathbf{A}}_k$ *is the set of shift invariant rank one modules in* $L^2(\mathbf{A}, \mathfrak{F}_k, \mu)$ *over the algebra* $L^\infty(\mathbf{A}, \mathfrak{F}_{k-1}, \mu)$ *and the sum is an orthogonal sum.*

3. *Every function* $f \in L^\infty(\mathbf{A}, \mathfrak{A}, \mu)$ *has a unique decomposition in the form*

$$f = f - \mathbb{E}(f|\mathfrak{F}_k) + \sum_{W \in \hat{\mathbf{A}}_k} P_W(f)$$

converging in L^2 *where* P_W *is the projection of* f *to the rank one module* $\hat{\mathbf{A}}_k$. *Note that only countably many terms in the sum are non-zero.*

4. *The set* $\hat{\mathbf{A}}_k$ *is an Abelian group (called k-th order dual group) with respect to point wise multiplication (using bounded representatives chosen from the modules).*

Linear characters are in shift invariant one dimensional subspaces on $L^2(A, \mu_A)$ and each one dimensional subspace is a module over L^∞ of the trivial σ-algebra. This way the above theorem is a direct generalization of ordinary Fourier decomposition. The fact that higher order generalizations of the dual group appear in the limit already shows the algebraic benefits of the limit framework however, as we will see soon, this is just the surface of an even deep algebraic theory. Theorem 2 can also be turned back to statements on usual compact Abelian groups using standard methods, however the exact nature disappears and various errors appear.

Theorem 2 does not give a full explanation of the algebraic nature of higher order Fourier analysis. It does not provide a structural description of how the rank one modules look like. To obtain a complete algebraic characterization the new theory of nilspaces was needed which generalizes to notion of Abelian groups. This theory was developed in Camarena and Szegedy [2010] but it was initiated in a different form in Host and Kra [2008b]. More detailed lecture notes on Camarena and Szegedy [2010] is Candela [2017b],Candela [2017a]. Here we give a brief description of nilspace theory and explain how it appears in higher order Fourier analysis.

A combinatorial cube of dimension n is the product set $\{0,1\}^n$. A morphism between two combinatorial cubes is a map $f : \{0,1\}^n \to \{0,1\}^m$ such that it extends to an affine homomorphism (a homomorphism and a shift) from $\mathbb{Z}^n$ to $\{0,1\}^m$. A combinatorial description of morphisms is the following: each coordinate of $f(x_1, x_2, \ldots, x_n)$ is one of 1, 0, x_i and $1-x_i$ for some $i \in [n]$. For example $f(x_1, x_2, x_3, x_4) := (1, x_1, x_1, 1-x_1, x_2, 0)$ is morphism from $\{0,1\}^4$ to $\{0,1\}^6$. An abstract nilspace is a set N together with maps (also called morphisms) from cubes to N satisfying three simple axioms. For each k we denote by $C^k(N) \subseteq N^{\{0,1\}^n}$ the set of morphisms from $\{0,1\}^n$ to N.

1. **Composition axiom:** If $f : \{0,1\}^n \to \{0,1\}^m$ is a morphism and $g \in C^m(N)$ then $g \circ f \in C^n(N)$.

2. **Ergodicity axiom:** $C^1(N) = N^{\{0,1\}}$ (Every map $f : \{0,1\} \to N$ is a morphism.)

3. **Completion axiom:** Let $f : \{0,1\}^n \setminus \{1^n\} \to N$ be a function such that its restriction to every $n-1$ dimensional face of $\{0,1\}^n$ is a morphism. Then f can be extended to a function $\tilde{f} : \{0,1\}^n \to N$ such that $\tilde{f}$ is a morphism.

If the completion in the last axiom is unique for some n then N is called an $n-1$ *step nilspace*. One step nilspaces are Abelian groups such that $f : \{0,1\}^n \to A$ is a morphism if and only if it can be extended to a map $\mathbb{Z} \to A$ which is an affine morphism. We give a general family of examples for nilspaces.

The group construction: Let G be an at most k-nilpotent group. Let $\{G_i\}_{i=1}^{k+1}$ be a central series with $G_{k+1} = \{1\}$, $G_1 = G$ and $[G_i, G_j] \subseteq G_{i+j}$. We define a cubic structure on

G which depends on the given central series. The set of n morphisms $f : \{0,1\}^n \to G$ is the smallest set satisfying the following properties.

1. *The constant* 1 *map is a cube,*
2. *If $f : \{0,1\}^n \to G$ is a cube and $g \in G_i$ then the function f' obtained from f by multiplying the values on some $(n-i)$-dimensional face from the left by g is a cube.*

Let $\Lambda \leq G$ be a subgroup in G and let $N = \{g\Lambda : g \in G\}$ denote the set of left cosets of Λ in G. We define $C^n(N)$ to be the set of morphisms $f : \{0,1\}^n \to G$ composed with the map $g \mapsto g\Lambda$. It can be verified that N with this cubic structure is a nilspace.
Example 1.) higher degree abelian groups: In the previous construction if G is abelian, $G_1, G_2, \ldots, G_k = G$, $G_{k+1} = \{1\}$ and $\Lambda = \{1\}$ then G with the above cubic structure is called a k-degree Abelian group. It is true that 1-degree Abelian groups are exactly the one step nilspaces. Every k-degree Abelian group is a k-step nilspace. However there are many more k-step nilspaces for $k \geq 2$.
Example 2.) nilmanifolds: Let G be a connected nilpotent Lie group with some filtration and assume that Λ is a discrete co-compact subgroup. Then the left coset space of Λ is a compact manifold. The above construction produces a nilspace structure on N.

We can talk about topological or compact nilspaces. Assume that N is a topological space an $C^n(N) \subseteq N^{\{0,1\}^n}$ is closed in the product topology. Then we say that N is a **topological nilspace**. If N is a compact (Hausdorff and second countable) topological nilspace then we say that N is a **compact nilspace**. Nilspaces form a cetegory. A morphism between two nilspaces is a function $f : N \to M$ such that for every n and $g \in C^n(N)$ we have that $f \circ g$ is in $C^n(M)$. In the category of compact nilspaces we assume that morphisms are continuous. The next theorem Szegedy [n.d.(b)] gives an algebraic description of the σ-algebras on $\mathbf{A}$. Note that the Abelian group $\mathbf{A}$ is a one step nilspace.

Theorem 3. *Let $f \in L^\infty(\mathbf{A}, \mathfrak{A}, \mu)$. Then f is measurable in $\mathfrak{F}_k$ if and only if there is a morphism $\gamma : \mathbf{A} \to N$ into a k-step compact nilspace such that*

1. *γ is continuous in the σ-topology on $\mathbf{A}$.*
2. *There is a Borel function $g : N \to \mathbb{C}$ such that $f = g \circ \gamma$ almost surely.*

The above theorem shows that in the limit, the k degree structured functions are exactly those that factor through k-step compact nilspaces. This statement also implies inverse theorems for the Gowers norms on compact Abelian groups, however they are more technical. The reason for the difficulty is that the clean qualitative separation between complexity classes that we detected in the limit on $\mathbf{A}$ becomes a more quantitative issue for concrete compact Abelian groups. For this reason we need to involve notions such as the complexity of a nilspace and a function on it. We mention that the structured functions that appear in the inverse theorem for the U_{k+1} norm (see Szegedy [n.d.(b)]) have the form $g \circ \gamma$ where γ is a morphism from A to a bounded complexity, finite dimensional nilspace N and g is a continuous function with bounded Lipschitz constant.

References

David J. Aldous (1985). "Exchangeability and related topics". In: *École d'été de probabilités de Saint-Flour, XIII—1983*. Vol. 1117. Lecture Notes in Math. Springer, Berlin, pp. 1–198. MR: 883646 (cit. on p. 3240).

Ágnes Backhausz and Balázs Szegedy (2016). "On the almost eigenvectors of random regular graphs". arXiv: 1607.04785 (cit. on p. 3238).

Itai Benjamini and Oded Schramm (2001). "Recurrence of distributional limits of finite planar graphs". *Electron. J. Probab.* 6, no. 23, 13. MR: 1873300 (cit. on pp. 3236, 3237).

C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi (2008). "Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing". *Adv. Math.* 219.6, pp. 1801–1851. MR: 2455626 (cit. on p. 3236).

– (2012). "Convergent sequences of dense graphs II. Multiway cuts and statistical physics". *Ann. of Math. (2)* 176.1, pp. 151–219. MR: 2925382 (cit. on p. 3236).

Christian Borgs, Jennifer T. Chayes, Henry Cohn, and Yufei Zhao (n.d.). "An L^p theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions". arXiv: 1401.2906 (cit. on p. 3238).

– (2018). "An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence". *Ann. Probab.* 46.1, pp. 337–396. arXiv: 1408.0744. MR: 3758733 (cit. on p. 3238).

Christian Borgs, Jennifer Chayes, László Lovász, Vera T. Sós, Balázs Szegedy, and Katalin Vesztergombi (2006). "Graph limits and parameter testing". In: *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*. ACM, New York, pp. 261–270. MR: 2277152 (cit. on p. 3236).

Omar Antolin Camarena and Balázs Szegedy (2010). "Nilspaces, nilmanifolds and their morphisms". arXiv: 1009.3825 (cit. on pp. 3244, 3246, 3248).

Pablo Candela (2017a). "Notes on compact nilspaces". *Discrete Anal.* Paper No. 16, 57. MR: 3695479 (cit. on p. 3248).

– (2017b). "Notes on nilspaces: algebraic aspects". *Discrete Anal.* Paper No. 15, 59. MR: 3695478 (cit. on p. 3248).

Sourav Chatterjee and S. R. S. Varadhan (2011). "The large deviation principle for the Erdős-Rényi random graph". *European J. Combin.* 32.7, pp. 1000–1017. MR: 2825532 (cit. on p. 3238).

F. R. K. Chung, R. L. Graham, and R. M. Wilson (1989). "Quasi-random graphs". *Combinatorica* 9.4, pp. 345–362. MR: 1054011 (cit. on p. 3241).

Jacob W Cooper, Tomáš Kaiser, Jonathan A Noel, et al. (2015). "Weak regularity and finitely forcible graph limits". *Electronic Notes in Discrete Mathematics* 49, pp. 139–143 (cit. on p. 3241).

Manfred Einsiedler and Thomas Ward (2011). *Ergodic theory with a view towards number theory*. Vol. 259. Graduate Texts in Mathematics. Springer-Verlag London, Ltd., London, pp. xviii+481. MR: 2723325.

Gábor Elek and Balázs Szegedy (2012). "A measure-theoretic approach to the theory of dense hypergraphs". *Adv. Math.* 231.3-4, pp. 1731–1772. MR: 2964622 (cit. on p. 3242).

Péter E. Frenkel (2018). "Convergence of graphs with intermediate density". *Trans. Amer. Math. Soc.* 370.5, pp. 3363–3404. arXiv: 1602.05937. MR: 3766852 (cit. on p. 3238).

H. Furstenberg (1981). *Recurrence in ergodic theory and combinatorial number theory*. M. B. Porter Lectures. Princeton University Press, Princeton, N.J., pp. xi+203. MR: 603625 (cit. on p. 3235).

Harry Furstenberg (1977). "Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions". *J. Analyse Math.* 31, pp. 204–256. MR: 0498471 (cit. on pp. 3234, 3235).

Roman Glebov, Tereza Klimosova, and Daniel Kral (2014). "Infinite dimensional finitely forcible graphon". arXiv: 1404.2743 (cit. on p. 3241).

W. T. Gowers (1998). "Fourier analysis and Szemerédi's theorem". In: *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*. Extra Vol. I, pp. 617–629. MR: 1660658 (cit. on pp. 3243, 3245).

– (2001). "A new proof of Szemerédi's theorem". *Geom. Funct. Anal.* 11.3, pp. 465–588. MR: 1844079 (cit. on pp. 3243, 3245).

Ben Green, Terence Tao, and Tamar Ziegler (2012). "An inverse theorem for the Gowers $U^{s+1}[N]$-norm". *Ann. of Math. (2)* 176.2, pp. 1231–1372. arXiv: 1009.3998. MR: 2950773 (cit. on p. 3246).

Benjamin Green and Terence Tao (2010). "Linear equations in primes". *Ann. of Math. (2)* 171.3, pp. 1753–1850. MR: 2680398 (cit. on p. 3246).

Hamed Hatami, László Lovász, and Balázs Szegedy (2014). "Limits of locally-globally convergent graph sequences". *Geom. Funct. Anal.* 24.1, pp. 269–296. MR: 3177383 (cit. on pp. 3232, 3238).

Bernard Host and Bryna Kra (2005). "Nonconventional ergodic averages and nilmanifolds". *Ann. of Math. (2)* 161.1, pp. 397–488. MR: 2150389 (cit. on pp. 3235, 3236, 3246, 3247).
– (2008a). "Analysis of two step nilsequences". *Ann. Inst. Fourier (Grenoble)* 58.5, pp. 1407–1453. MR: 2445824 (cit. on pp. 3234, 3247).
– (2008b). "Parallelepipeds, nilpotent groups and Gowers norms". *Bull. Soc. Math. France* 136.3, pp. 405–437. MR: 2415348 (cit. on pp. 3246–3248).
Dávid Kunszenti-Kovács, László Lovász, and Balázs Szegedy (2016). "Measures on the square as sparse graph limits". arXiv: 1610.05719 (cit. on p. 3238).
L. Lovász and B. Szegedy (2011). "Finitely forcible graphons". *J. Combin. Theory Ser. B* 101.5, pp. 269–301. MR: 2802882 (cit. on p. 3241).
László Lovász (2012). *Large networks and graph limits*. Vol. 60. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, pp. xiv+475. MR: 3012035.
László Lovász and Balázs Szegedy (2006). "Limits of dense graph sequences". *J. Combin. Theory Ser. B* 96.6, pp. 933–957. MR: 2274085 (cit. on pp. 3236–3238, 3240).
– (2007). "Szemerédi's lemma for the analyst". *Geom. Funct. Anal.* 17.1, pp. 252–270. MR: 2306658 (cit. on pp. 3236, 3238–3240).
– (2010a). "Regularity partitions and the topology of graphons". In: *An irregular mind*. Vol. 21. Bolyai Soc. Math. Stud. János Bolyai Math. Soc., Budapest, pp. 415–446. MR: 2815610 (cit. on p. 3240).
– (2010b). "Testing properties of graphs and functions". *Israel J. Math.* 178, pp. 113–156. MR: 2733066 (cit. on p. 3238).
– (2015). "The automorphism group of a graphon". *J. Algebra* 421, pp. 136–166. MR: 3272377.
Brendan Nagle, Vojtěch Rödl, and Mathias Schacht (2006). "The counting lemma for regular k-uniform hypergraphs". *Random Structures Algorithms* 28.2, pp. 113–179. MR: 2198495 (cit. on p. 3242).
Jaroslav Nesetril and Patrice Ossona De Mendez (Mar. 2013). "A unified approach to structural limits, and limits of graphs with bounded tree-depth". arXiv: 1303.6471 (cit. on p. 3238).
Alexander A. Razborov (2008). "On the minimal density of triangles in graphs". *Combin. Probab. Comput.* 17.4, pp. 603–618. MR: 2433944 (cit. on p. 3241).
Vojtěch Rödl and Mathias Schacht (2007). "Regular partitions of hypergraphs: regularity lemmas". *Combin. Probab. Comput.* 16.6, pp. 833–885. MR: 2351688 (cit. on p. 3242).
Vojtěch Rödl and Jozef Skokan (2004). "Regularity lemma for k-uniform hypergraphs". *Random Structures Algorithms* 25.1, pp. 1–42. MR: 2069663 (cit. on p. 3242).
K. F. Roth (1953). "On certain sets of integers". *J. London Math. Soc.* 28, pp. 104–109. MR: 0051853 (cit. on pp. 3235, 3245).

– (1972). "Irregularities of sequences relative to arithmetic progressions. IV". *Period. Math. Hungar.* 2. Collection of articles dedicated to the memory of Alfréd Rényi, I, pp. 301–326. MR: 0369311 (cit. on p. 3245).

Walter Rudin (1990). *Fourier analysis on groups*. Wiley Classics Library. Reprint of the 1962 original, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, pp. x+285. MR: 1038803 (cit. on p. 3245).

J. Solymosi (2004). "A note on a question of Erdős and Graham". *Combin. Probab. Comput.* 13.2, pp. 263–267. MR: 2047239 (cit. on p. 3242).

Balazs Szegedy (n.d.[a]). "Higher order Fourier analysis as an algebraic theory I". arXiv: 0903.0897 (cit. on p. 3247).

– (n.d.[b]). "On higher order Fourier analysis". Preprint. arXiv: 1203.2260 (cit. on pp. 3239, 3242, 3244, 3246, 3247, 3249, 3250).

– (n.d.[c]). "Sparse graph limits, entropy maximization and transitive graphs". arXiv: 1504.00858 (cit. on p. 3238).

Balázs Szegedy (2011). "Limits of kernel operators and the spectral regularity lemma". *European J. Combin.* 32.7, pp. 1156–1167. MR: 2825541.

E. Szemerédi (1975). "On sets of integers containing no k elements in arithmetic progression". *Acta Arith.* 27. Collection of articles in memory of JuriĭVladimirovič Linnik, pp. 199–245. MR: 0369312 (cit. on pp. 3242, 3245).

Endre Szemerédi (1978). "Regular partitions of graphs". In: *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*. Vol. 260. Colloq. Internat. CNRS. CNRS, Paris, pp. 399–401. MR: 540024 (cit. on p. 3240).

Terence Tao (2012). *Higher order Fourier analysis*. Vol. 142. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. x+187. MR: 2931680.

Terence Tao and Tamar Ziegler (2012). "The inverse conjecture for the Gowers norm over finite fields in low characteristic". *Ann. Comb.* 16.1, pp. 121–188. MR: 2948765 (cit. on p. 3246).

Evan Warner (2012). "Ultraproducts and the Foundations of Higher Order Fourier Analysis". PhD thesis. Bachelor thesis, Princeton University.

Tamar Ziegler (2007). "Universal characteristic factors and Furstenberg averages". *J. Amer. Math. Soc.* 20.1, pp. 53–97. MR: 2257397 (cit. on pp. 3235, 3236, 3247).

Received 2017-12-04.

Balázs Szegedy
szegedyb@gmail.com