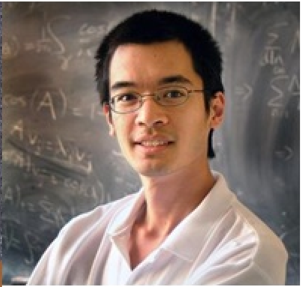
Analysis of Boolean Functions

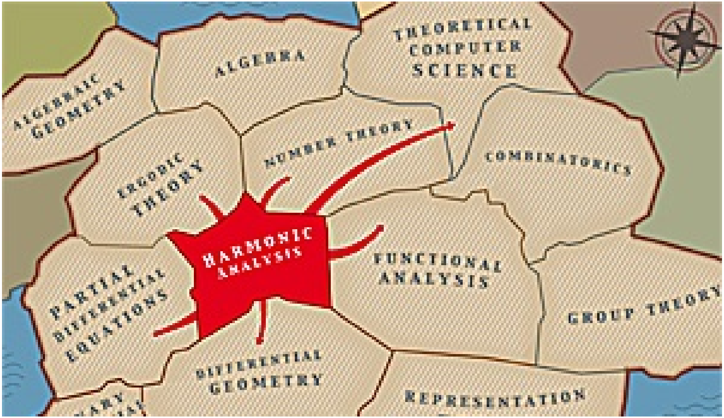
Gil Kalai
Hebrew University of Jerusalem

From Chaos to Harmony, Crafoord days , Lund May 2012.

Congratulations Jean and Terry!!!



Harmonic Analysis for Combinatorics and Theoretical Computer Science



Harmonic Analysis for Combinatorics and Theoretical Computer Science - three directions

- I. **Pseudorandomness and structure**

Example: Roth's theorem on 3-terms arithmetic progressions

Tools: Parseval, Hardy-Littlewood circle method, quadratic and higher Fourier analysis,...

Applications: extremal combinatorics, additive number theory and its many own applications, probability theory, theoretical computer science, cryptography

Harmonic Analysis for Combinatorics and Theoretical Computer Science - three directions(cont.)

- II. **Isoperimetry**

Example: Kahn-K.-Linial (1988) (KKL) theorem

Tools: Parseval, Hypercontractivity/log Sobolev

Applications: extremal combinatorics, probability theory and random graphs, computational complexity, and theoretical computer science, game theory

Harmonic Analysis for Combinatorics and Theoretical Computer Science - three directions (cont.)

- III. **Bounds on Error-Correcting Codes**

Example: McEliece, Rodemich, Rumsey, Welch theorem

Tools: Parseval, Delsarte linear programming method, hypergeometric functions.

Applications: Error-correcting codes, sphere packings, packing and covering in combinatorics and geometry

Harmonic Analysis for Combinatorics and Theoretical Computer Science - three directions (cont.)

Problems:

Harmonic Analysis for Combinatorics and Theoretical Computer Science - three directions (cont.)

Problems:

Are these three areas related?

Harmonic Analysis for Combinatorics and Theoretical Computer Science - three directions (cont.)

Problems:

Are these three areas related?

Is Gowers's "quadratic (and higher) Fourier analysis" of some use on the isoperimetric side? the error-correcting code side?

Harmonic Analysis for Combinatorics and Theoretical Computer Science - three directions (cont.)

Problems:

Are these three areas related?

Is Gowers's "quadratic (and higher) Fourier analysis" of some use on the isoperimetric side? the error-correcting code side?

What can replace Parseval identity for quadratic Fourier analysis?

Harmonic Analysis for Combinatorics and Theoretical Computer Science - three directions (cont.)

Problems:

Are these three areas related?

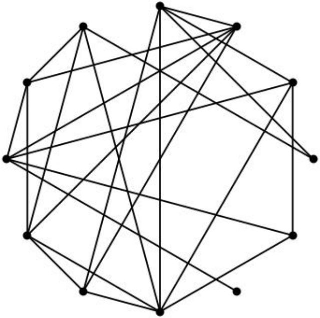
Is Gowers's "quadratic (and higher) Fourier analysis" of some use on the isoperimetric side? the error-correcting code side?

What can replace Parseval identity for quadratic Fourier analysis?

Find other applications of different nature of harmonic analysis to Combinatorics and TCS?

This lecture

This lecture is about the second direction: harmonic analysis applied to discrete isoperimetry. We have several application and potential applications in mind mainly to problems in probability. I will start by mentioning one potential application. It deals with the theory of random graphs initiated by Erdős and Rényi, and the



model $G(n.p)$. In the picture we see a random graph with $n = 12$ and $p = 1/3$.

Threshold and Expectation threshold

Consider a random graph G in $G(n,p)$ and the graph property: G contains a copy of a specific graph H. (Note: H depends on n; a motivating example: H is a Hamiltonian cycle.) Let q be the minimal value for which the expected number of copies of H' in G is at least $1/2$ for every subgraph H' of H. Let p be the value for which the probability that G contains a copy of H is $1/2$.

Conjecture: [Kahn, K. 2006]

$$p/q = O(\log n).$$

The conjecture can be vastly extended to general Boolean functions, and we will hint on possible connection with harmonic analysis and discrete isoperimetry. (Sneak preview: it will require a far-reaching extension of results by Friedgut, Bourgain and Hatami.)

The discrete n-dimensional cube and Boolean functions

The **discrete n-dimensional cube** Ω_n is the set of 0-1 vectors of length n.

A **Boolean function** f is a map from Ω_n to $\{0, 1\}$.

A boolean function f is **monotone** if f cannot decrease when you switch a coordinate from 0 to 1.

The Bernoulli measure

Let $p, 0 < p < 1$, be a real number. The probability measure μ_p is the product probability distribution whose marginals are given by $\mu_p(x_k = 1) = p$. Let $f : \Omega_n \to \{0,1\}$ be a Boolean function.

$$\mu_p(f) = \sum_{x \in \Omega_n} \mu_p(x) f(x) = \mu_p\{x : \ f(x) = 1\}.$$

The total influence

Two vectors in Ω_n are **neighbors** if they differ in one coordinate.

For $x \in \Omega_n$ let $h(x)$ be the number of neighbors y of x such that $f(y) \neq f(x)$.
The **total influence** of f is defined by

$$I^p(f) = \sum_{x \in \Omega_n} \mu_p(x) h(x).$$

If $p = 1/2$ we will omit p as a subscript or superscript.

Russo's lemma

Russo's lemma: For a monotone Boolean function f,

$$d\mu_p(f)/dp = I^p(f).$$

Very useful in percolation theory and other areas.

The **threshold interval** for a monotone Boolean function f is those values of p so that $\mu_p(f)$ is bounded away from 0 and 1. (Say $0.01 \leq \mu_p(f) \leq 0.99$.)

A typical application of Russo's lemma: If for every value p in the threshold interval $I^p(f)$ is large, then the threshold interval itself is short. This is called a **sharp threshold** phenomenon.

(A version of) Harper's theorem

Harper's theorem: If $\mu_p(f) = t$ then

$$I^p(f) \geq 2t \cdot \log_p t.$$

There is a 3 line proof by induction.
Harmonic analysis proof: without the $\log(1/t)$ factor it follows from Parseval.

Influence of variables on Boolean functions

Let

$$\sigma_k(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) = (x_1, \ldots, x_{k-1}, 1 - x_k, x_{k+1}, \ldots, x_n).$$

The influence of the kth variable on a Boolean function f is defined by:

$$I_k^p(f) = \mu_p(x \in \Omega_n, f(x) \neq f(\sigma_k(x))).$$

KKL's theorem

Theorem (Kahn, K, Linial, 1988; Bourgain Katznelson KKL 1992; Talagrand 1994 Friedgut K. 1996) There exist a variable k such that

$$I_k(f) \geq C\mu(f)(1-\mu(f)) \log n/n.$$

A sharp version (due to Talagrand)

$$\sum I_k^p(f)/\log(e + I_k^p(f)) \geq C(p)\mu_p(f)(1-\mu_p(f)).$$

Hypercontructivity and Harper's theorem:

We assume now $p = 1/2$. $f = \sum \hat{f}(S) W_S$ is the **Fourier-Walsh** expansion of f. Key ideas:

0 Parseval gives $I(f) = 4 \sum \hat{f}(S)|S|$.

1 Bonami-Gross-Beckner hypercontractive inequality.

$$\| \sum \hat{f}(S)(1/2)^{|S|} \|_2 \le \| f \|_{5/4}.$$

2 For Boolean functions the qth power of the q norm is the measure of the support and does not depend on q. If the support is small this means that the q-norm is very different from the r-norm if $r \neq q$.

(See also : Ledoux' book on concentration of measure phenomena)

Part II: Influence and symmetry

Invariance under transitive group

Theorem: If a monotone Boolean function f with n variables is invariant under a transitive group of permutation of the variables, then

$$I^p(f) \geq C\mu_p(f)(1 - \mu_p(f)) \log n.$$

Proof: Follows from KKL's theorem since all individual influences are the same.

Total influence under symmetry of primitive groups

For a transitive group of permutations $\Gamma \subset S_n$, let $I(\Gamma)$ be the minimum influence for a Γ-invariant function Boolean function with n variables.

Theorem: [Bourgain and K. 1998] If Γ is primitive then one of the following possibilities hold.

- $$I(\Gamma) = \theta(\sqrt{n}),$$
- $$(logn)^{(k+1)/k-o(1)} \leq I(\Gamma) \leq C(\log n)^{(k+1)/k},$$
- $I(\Gamma)$ behaves like $(\log n)\mu(n)$, where $\mu(n) \leq \log \log n$ is growing in an arbitrary way.

Jumps in the behavior of $I(\Gamma)$ for primitive groups Γ

If Γ is not A_n and S_n then $I(\Gamma) \leq (\log n)^2$.

If $I(\Gamma) \leq (\log n)^{1.99}$ then $I(\Gamma) \leq (\log n)^{3/2}$

If $I(\Gamma) \leq (\log n)^{3/2-\epsilon}$ then $I(\Gamma) \leq (\log n)^{4/3}$

If $I(\Gamma) \leq (\log n)^{4/3-\epsilon}$ then $I(\Gamma) \leq (\log n)^{5/4}$

...

If $I(\Gamma) \leq (\log n)^{1+o(1)}$ then $I(\Gamma) \leq \log n \cdot \log \log n$

Threshold behavior for random graphs

The case that Γ is S_n acting on unordered pairs from $[n] = \{1, 2, \ldots, n\}$ describes graph properties. The conclusion is that the threshold interval for graph properties is at most

$$1/\log^{2-o(1)} n.$$

Hypercontractivity and the lower bounds

Both the upper bounds and the lower bounds depend on finding invariants of the group which causes the threshold to go above $\log n$. Giving constructions for the upper bounds requires a detailed understanding of primitive permutation groups based on the classification theorem and O'Nan-Scott theorem.

The lower bounds are based on delicate and complicated harmonic analysis.

Step I: hypercontractivity + random restriction argument + clever inequalities takes you in the graph case from $\log n$ to $\log n^{3/2}$.

Step II: Extremely subtle "bootstrap" to amplify the outcome.

The Entropy Influence conjecture (Friedgut + K. 1996)

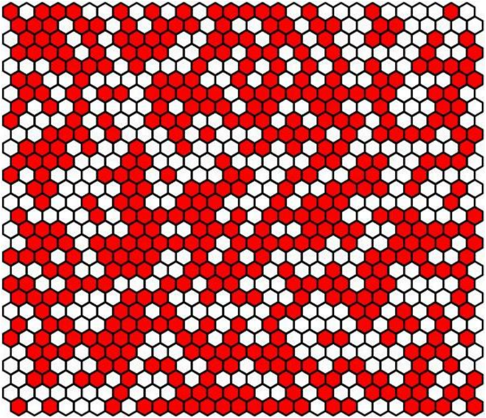
If the Fourier-Walsh expansion of f is $f = \sum \hat{f}(S) W_S$ define

$$E(f) = \sum \hat{f}(S) f^2(S) \log(1/\hat{f}^2(S)).$$

Conjecture: For some absolute constant C,

$$I(f) \geq C \cdot E(f).$$

Scaling-limit symmetry, critical exponents, spectral distribution,...



Prelude: A necessary and sufficient condition for sharp threshold window.

The **Shapley value** of the kth variable is defined by

$$\psi_k(f) = \int_0^1 I_k^p(f)dp.$$

Theorem: (K. 2005)
A necessary and sufficient condition for diminishing threshold window is that the maximum of the Shapley values tends to 0.

Problem: Close the exponential gap in this theorem.

Part III: Stability of Harper's theorem: from Harper to Hatami and beyond

Low Influence and Juntas

A **dictatorship** is a Boolean function depending on one variable. A K-**junta** is a Boolean function depending on K variables.

Theorem: (Friedgut, follows easily from KKL) If p is bounded away from 0 and 1 and $I^p(f) < C$ then f is close to a $K(C)$-Junta.

This works if $\log p / \log n = o(1)$ the most interesting applications would be when p is a power of n. There the theorem is not true.

The works of Friedgut and Bourgain (1999)

Suppose that f is a Boolean function and

$$I^p(f) < pC,$$

then

Friedgut's theorem (1999): If f represent a monotone graph property then f is close to a a "locally defined" function g.

Bourgain's theorem (1999): Unconditionally, f has a substantial "locally defined" ingredient.

Hatami's theorem: Pseudo-juntas

Suppose that for every subset of variables S, we have a function $J_S : \{0,1\}^S -> \{0,1\}$ which can be viewed as a constraint over the variables with indices in S. Now there are two conditions:
A Boolean function is a K-**psudo-junta** if
(1) the expected number of variables in satisfied constraints is bounded by a constant K.
(2) $f(x) = f(y)$ if the variables in satisfied constraints and also their values are the same for x and y.

Hatami's theorem: For every C there is $K(C)$, such that if

$$I^p(f) < pC,$$

then f is close to a $K(C)$-pseudo-junta.

A conjectural extension of Hatami's theorem

Conjecture: Suppose that $\mu_p(f) = t$ and

$$I(f) \leq C \log(1/t)t$$

then f is close to a $O(\log(1/t))$-pseudo-junta.

Stability versions of Harper's theorems

	t bounded away from 0	t small
p bounded away from 0	Friedgut's easy theorem (based on KKL) Many applications to PCP and other areas Some applications to percolation	
p s small	Friedgut's hard THM Bourgain's THM Hatami's THM Many applications for proving sharp threshold behavior: k-SAT, connectivity, Ramsey	Super Hatami Conjecture. Potential applications to finding the location of the threshold, and to other things.

Stability versions for Harper's theorems