Boolean Functions: Influence, threshold and noise

Gil Kalai*

Abstract.
This lecture studies the analysis of Boolean functions and present a few ideas, results, proofs, and problems. We start with the wider picture of expansion in graphs and then concentrate on the graph of the n-dimensional discrete cube Ω_n. Boolean functions are functions from Ω_n to $\{0, 1\}$. We consider the notion of the influence of variables on Boolean functions. The influence of a variable on a Boolean function is the probability that changing the value of the variable changes the value of the function. We then consider Fourier analysis of real functions on Ω_n and some applications of Fourier methods. We go on to discuss connections with sharp threshold phenomena, percolation, random graphs, extremal combinatorics, correlation inequalities, and more.

2010 Mathematics Subject Classification. Primary 05-02; Secondary 05D05, 05C80.
Keywords. Combinatorics, Boolean functions, Fourier methods, random graphs, percolation, influence, isoperimetry.

1. Prologue: German-Jewish mathematicians in the early days of the Hebrew University of Jerusalem

This paper follows closely the author's lecture at the 7ECM in Berlin in July 2016. Being invited to give a plenary lecture at the ECM was a great honor and, as Keren Vogtmann said in her beautiful opening lecture on outer spaces, it was also a daunting task. I am thankful to Günter Ziegler for his introduction. When I ask myself in what way I am connected to the person I was thirty years ago, one answer is that it is my long-term friendship with Günter and other people that makes me the same person. My lecture deals with the analysis of Boolean functions in relation to expansion (isoperimetric) properties of subsets of the discrete n-dimensional cube. The lecture has made a subjective selection of some results, proofs, and problems from this area.

Yesterday, Leonid Polterovich and I were guests of the exhibition "Transcending Tradition: Jewish Mathematicians in German-Speaking Academic Culture." I will start by briefly mentioning the great impact of German-Jewish mathematicians on the early history of the Einstein Institute of Mathematics and Physics at the Hebrew University of Jerusalem, my main academic home since the early seventies.

In Figure 1 you can see some early faces of our Institute. Edmund Landau, the founder and first head of the Institute, moved to Jerusalem from Göttingen

*Partially supported by ERC advanced grant 320924, NSF grant DMS1300120, and BSF grant 2014290.

Figure 1. Early faces at the Institute of mathematics @HUJI

in 1927 and moved back to Göttingen a year and a half later. Abraham (Adolf) Halevi Fraenkel moved to Jerusalem from Kiel in 1928 and he can be seen as the father of logic, set theory, and computer science in Israel. My own academic great-grandfather is Michael Fekete, who immigrated to Jerusalem from Budapest in 1928. In the lecture I went on to describe two remarkable documents written by Landau in 1925, both related to the inauguration ceremony of the Hebrew University of Jerusalem. This part of the lecture is given in the Appendix. In transforming my 7ECM lecture into an article I tried to follow the style of Landau's 1925 article based on his lecture.

2. Introduction: Graphs and expansion

2.1. Expansion in general graphs. A graph is one of the simplest structures in combinatorics. It consists of a set of vertices, together with a set of edges that join some of the pairs of vertices. Figure 2 demonstrates the very basic notion of expansion: given a set of vertices we are interested in the number of edges between the vertices in the set and those outside the set. Formally, given a graph $G = \langle V(G), E(G)\rangle$ and a subset S of $V(G)$, we consider the set of edges $E(S, \bar{S})$ between S and its complement $\bar{S}$, and let $e(S, \bar{S}) = |E(S, \bar{S})|$. If S is a subset of $V(G)$ with $E(S, \bar{S})$ large, then we can view the set S as an "expanding" set (with respect to the graph G). The study of expansions of graphs is very important and has several aspects. Sets of edges of the form $E(S, \bar{S})$ are called *cut sets*.

- Combinatorics: expansion is a refined notion of connectivity, and it is also a

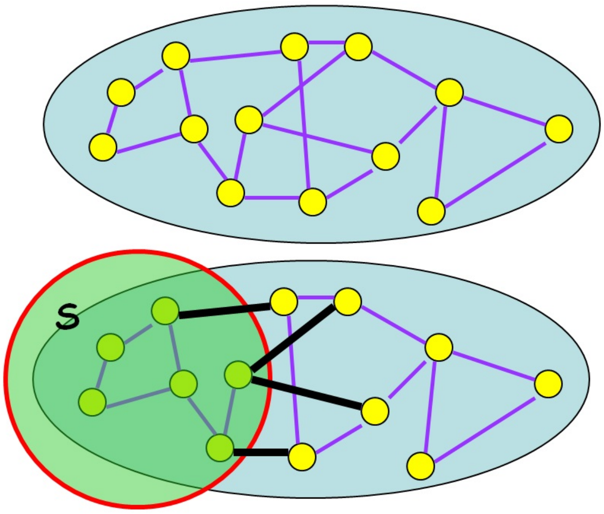

Figure 2. Edge expansion in graphs

part of the study of cut sets in graphs.

- Geometry: expansion is related to the study of graphs as geometric objects and the study of isoperimetric properties of graphs.[1]

- Probability: expansion is related to quick convergence of random walks to their stationary distribution.

- Spectral (Algebra/Analysis): expansion is closely connected to the spectral gap for the Laplacian of a graph.

- Computation: expansion properties of graphs have many applications to and connections with computer science.

2.2. Zooming in: Expansion for the discrete n-dimensional cube. We will zoom in for the rest of this lecture on a very small corner of the study of expansion properties of graphs. Science and mathematics often have a fractal-like nature such that zooming in on a small part of a picture often reveals a beautiful picture on its own with important similarities and connections to the larger picture. We will consider a very special graph, that of the discrete n-dimensional cube.

The *discrete n-dimensional cube* Ω_n is the set of 0-1 vectors of length n. We consider the graph on the set of vertices Ω_n, where two vertices are adjacent if they differ in one coordinate.

[1] The edge-isoperimetric problem for a graph G is to determine, for each k, the minimum value of $E(S, \bar{S})$ over all sets S of size k. There is also the important "vertex-isoperimetric problem," which we do not discuss further.

Theorem 2.1. *The number of edges* $e(A, \bar{A})$ *between a set* A *of vertices of* Ω_n *and its complement* $\bar{A}$ *is at least* $\min(|A|, |\bar{A}|)$.

Proof. The proof actually gives the lower bound

$$e(A, \bar{A}) \geq 2^{-(n-1)}(|A| \cdot |\bar{A}|). \tag{1}$$

Let $E(A, \bar{A})$ be the set of edges between A and $\bar{A}$. Given two vertices x, y of the discrete cube we will consider the *canonical path* between x and y where we flip coordinates of disagreement from left to right.

Observations: 1) Every edge $\{z, u\}$ of Ω_n is contained in precisely 2^{n-1} canonical paths (where u follows z).

Indeed, suppose that z and u differ in the kth coordinate, and that they are contained (in this order) in the canonical path from x to y. Then x must agree with z on the last $n-k$ coordinates, giving 2^{n-k} possibilities for x, and y must agree with u on the last $n-k$ coordinates, giving 2^{k-1} possibilities for y, and altogether 2^{n-1} possibilities for pairs (x, y).

2) Every canonical path from $x \in A$ to $y \in \bar{A}$ contains an edge from $E(A, \bar{A})$.

To complete the proof note that there are $|A| \cdot |\bar{A}|$ canonical paths from vertices in A to vertices in $\bar{A}$, and each edge in $E(A, \bar{A})$ belongs to at most 2^{n-1} of them. $\square$

This proof technique via canonical paths has several important applications.

3. Influence

A *Boolean function* f is a map from Ω_n to $\{0, 1\}$. Let μ denote the uniform probability distribution on the discrete n-cube Ω_n. (We will later discuss other probability measures.) For a Boolean function f, $\mathbb{E}(f) = \mu(\{x : f(x) = 1\})$. Boolean functions are simply the characteristic functions of subsets of the discrete cube.

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

The next definitions will be quite important for the rest of the lecture. Let

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

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

$$I_k(f) = \mu(x \in \Omega_n : f(x) \neq f(\sigma_k(x))).$$

In words, this is the probability that when you flip the value of the kth coordinate the value of f is flipped as well.

The *total influence* is defined by

$$I(f) = \sum_{k=1}^{n} I_k(f).$$

If $f = \chi_A$ then the total influence is just a normalized version of the expansion of A, namely, $I(f) = e(A, \bar{A})/2^{n-1}$. Similarly, $I_k(f) = 2^{-n+1} e_k(A, \bar{A})$, where $e_k(A, \bar{A})$ is the number of edges in the "k-direction" between A and $\bar{A}$. What is important here about the graph of the cube is that every edge has a "direction," namely, the coordinate of disagreement between the two endpoints, and you can have a finer statistic of edge-expansion according to the direction of the edge.

We proved that $I(f) \geq 4\mathbb{E}(f)(1 - \mathbb{E}(f))$ and a stronger inequality (going back to Larry Harper [37]) is

$$I(f) \geq 2\mathbb{E}(f) \log_2(1/\mathbb{E}(f)). \tag{2}$$

Thus, when f is supported on a small fraction of vertices of the discrete cube we gain a logarithmic factor.

3.1. KKL's theorem. The following 1988 theorem by Nati Linial, Jeff Kahn, and myself, sometimes referred to as the KKL theorem, will play a central role in this lecture.

Theorem 3.1 (Kahn, Kalai, Linial [40])**.** *For every Boolean function f, there exists a variable k such that*

$$I_k(f) \geq C\mathbb{E}(f)(1 - \mathbb{E}(f)) \log n/n.$$

Here and elsewhere in this paper C (and also c and C') refer to absolute constants. When $\mathbb{E}(f) = 1/2$, namely, when f is supported on half the vertices of the discrete cube, we gain a logarithmic factor in n compared to what we can deduce from the isoperimetric inequality (2). Note that the quantity $\mathrm{Var}(f) = \mathbb{E}(f)(1 - \mathbb{E}(f))$ is the *variance* of f.

This result was conjectured by Michael Ben-Or and Nati Linial in 1985 [7], and we worked hard and finally proved it.

The proof in [40] actually gives that there is always a k with $I_k(f) \geq C^{-I(f)/\mathrm{Var}(f)}$. Michel Talagrand found a sharper version [68]:

$$\sum_{k=1}^{n} I_k(f)/\log(e/I_k(f)) \geq C\mathrm{Var}(f).$$

3.2. The Bernoulli measure. I would like to add at this early stage one more ingredient to the discussion, namely, to talk about a more general probability distribution on the discrete cube. Let $p, 0 < p < 1$, be a real number. The Bernoulli probability measure μ_p is the product probability distribution whose marginals are given by $\mu_p(x_k = 1) = p$. In other words, $\mu_p(x_1, x_2, \ldots, x_n) = p^k(1-p)^{n-k}$, where $k = x_1 + x_2 + \cdots + x_n$.

Let $f : \Omega_n \to \{0, 1\}$ be a Boolean function; we denote

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

The notions of influence, total influence, the edge-expansion theorem, and KKL's theorem all extend to the biased ($p \neq 1/2$) case. The influence of the kth variable on a Boolean function f is defined by

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

The total influence is defined as the sum of individual influences

$$I^p(f) = \sum_{k=1}^{n} I_k^p(f).$$

3.3. Edge-expansion inequalities. We already mentioned the edge-expansion inequality $I(f) \geq 2\mathbb{E}(f)\log(1/\mathbb{E}(f))$. This inequality extends to the case of general $p \leq 1/2$ as follows.

$$I^p(f) \geq \frac{1}{p}\mathbb{E}_p(f)\log_p(\mathbb{E}_p(f)). \tag{3}$$

(Note that both p and $\mathbb{E}_p(f)$ are ≤ 1.)

I would like to give you a sketch of the proof, which is by induction on n.[2] Arguments by induction based on two half cubes are surprisingly powerful for discrete isoperimetric results and in extremal combinatorics. There are examples of very deep theorems obtained by very intricate inductive proofs of this nature.

Proof. (sketch)

Induction on the number of coordinates. Given a monotone Boolean function f on n coordinates, let $\alpha = \mathbb{E}(f|x_n = 0)$ and $\beta = \mathbb{E}(f|x_n = 1)$. Based on the induction hypothesis one needs to prove

$$(1-p)\alpha\log_p\alpha + p\beta\log_p\beta + p(\beta - \alpha)$$

$$-((1-p)\alpha + p\beta)log_p((1-p)\alpha + p\beta) \geq 0.$$

This inequality holds with equality when $\beta = \alpha$. By an easy calculation, the derivative of the left hand side w.r.t. β is nonnegative for all β. □

For $p = 1/2$, even sharper inequalities are known and given $\mathbb{E}(f)$ the precise minimal value of $I(f)$ is known. (It is attained when $f = 1$ for the first $2^n\mathbb{E}(f)$ vertices of Ω_n when you ordered the vertices according to their value as binary integers.) I am not aware of sharper results of this kind for general values of p.

[2]This argument applies to every $0 < p < 1$ when the function is monotone. For general Boolean functions, when $0 < p \leq 1/2$, there is a simple reduction to the monotone case.

3.4. Wanted: Inverse theorems. I would like to mention in this lecture a few basic open problems in this theory and here is the first one. The edge-expansion inequality (general $p \le 1/2$) asserts that $I^p(f) \ge 2\mathbb{E}_p(f)\log_p(1/\mathbb{E}_p(f))$. The question is to understand Boolean functions for which this bound is attained up to a constant multiplicative factor.

Problem 1 [Inverse theorem for edge expansion]: Understand the structure of Boolean functions for which

$$I^p(f) \le K\frac{1}{p}\mathbb{E}_p(f)\log_p(1/\mathbb{E}_p(f)).$$

For the case where both p and $\mathbb{E}_p(f)$ are bounded away from zero and one, Friedgut (1998) [28] proved that such functions are approximately "juntas," namely, they are determined (with high probability) by their values on a fixed bounded set of variables. This is no longer the case when p is small. Major theorems for the case where p is small but $\mathbb{E}_p(f)$ is bounded away from zero and one, were proved by Friedgut (1999) [29], Bourgain (1999) [14], and Hatami (2010) [38]. The case where $\mathbb{E}_p(f)$ is small is wide open and very important; see also [41].[3]

3.5. Examples*[4].

Dictatorship and Juntas. The function $f(x_1, x_, \dots, x_n) = x_k$ is called a dictatorship. Note that $\mathbb{E}(f) = 1/2$, $I_k(f) = 1$, and $I_j(f) = 0$ for $j \ne k$.

AND of variables. Next consider the function $f(x_1, x_2, \dots, x_n) = x_1 \wedge x_2 \wedge \dots \wedge x_r$ (namely, $f = 1$ iff $x_1 = x_2 = \dots = x_r = 1$). Here $\mathbb{E}(f) = 1/2^r$, $I_1(f) = I_2(f) = \dots = I_r(f) = 2^{-r+1}$ and $I_j(f) = 0$ for $j > r$. For this example: $\mathbb{E}_p(f) = p^r$ and $I^p(f) = rp^{r-1}$ ($I_j^p(f) = p^{r-1}$ if $j \le r$ and $I_j^p(f) = 0$ for $j > r$. Thus for this function the isoperimetric inequality (3) is satisfied as equality.

Majority and linear threshold functions. Suppose that the number of variables n is odd. The majority function is defined by $f(x_1, x_2, \dots, x_n) = 1$ iff $x_1 + x_2 + \dots + x_n \ge n/2$. For the majority function, $I_k(f) = \sqrt{\frac{2}{\pi} + o(1)}/\sqrt{n}$. Let $w_1, w_2, \dots, w_n$ be real weights and T be a real number. A *linear threshold function* is is defined by $f(x_1, \dots, x_n) = 1$ iff $\sum w_i x_i \le T$. When all w_is are nonnegative we also call f a weighted majority function.

Recursive majority and tribes. Suppose that $n = 3^m$. Ternary recursive majority $f_m(x_1, x_2, \dots, x_n)$ is defined recursively as the majority of the three values obtained by applying the function f_{m-1} on the first, second, and third groups of 3^{m-1} variables, respectively. Here the influence of each variable behaves like $\Theta(n^{-\log 2/\log 3})$.

[3] We can also ask about cases where the edge-isoperimetric inequalities are sharp up to a multiplicative factor $(1+\epsilon)$ where $\epsilon > 0$ is small. This is also very interesting, see [22].

[4] Sections with asterisk were added in the written version of the lecture.

Of fundamental importance is the Ben-Or–Linial tribes example. It is obtained by dividing the variables into pairwise disjoint tribes of equal size (approximately) $\log n - \log \ln n + \log \ln 2$. $f = 1$ if there exists a tribe all of whose variables attain the value 1. Here the influence of each variable is $\Theta(\log n/n)$. This example shows that the KKL theorem is sharp up to a constant factor.

Graph properties. A large important family of examples is obtained as follows. Consider a property P of graphs on m vertices. Let $n = m(m-1)/2$, associate Boolean variables with the n edges of the complete graph K_m, and represent every subgraph of K_m by a vector in Ω_n. The property P is now represented by a Boolean function on Ω_n. If the graph property is "G contains a complete subgraph with $C \log n$ vertices." then the influence of every variable is $\Theta(\log^2 n/n)$.

3.5.1. Formulas and circuits. Formulas and circuits allow to build complicated Boolean functions from simple ones and they have crucial importance in computational complexity. Starting with n variables $x_1, x_2, \ldots, x_n$, a *literal* is a variable x_i or its negation $\neg x_i$. Every Boolean function can be written as a formula in conjunctive normal form, namely as AND of ORs of literals. A *circuit* of depth d is defined inductively as follows. A circuit of depth zero is a literal. A circuit of depth one consists of an OR or AND *gate* applied to a set of literals, a circuit of depth k consists of an OR or AND gate applied to the outputs of circuits of depth $k-1$. (We can assume that gates in the odd levels are all OR gates and that the gates of the even levels are all AND gates.) The size of a circuit is the number of gates. The famous **NP** $\neq$ **P**-conjecture (in a slightly stronger form) asserts that the Boolean function described by the graph property of containing a Hamiltonian cycle, cannot be described by a polynomial-size circuit. Formulas are circuits where we allow to use the output of a gate as the input of only one other gates.

3.5.2. symmetry. Finally, a few sentences about symmetry. For a Boolean function $f = f(x_1, x_2, \ldots, x_n)$, define $Aut(f)$, the group of automorphisms of f, as the set of all permutations $\pi \in S_n$ that satisfies

$$f(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}) = f(x_1, x_2, \ldots, x_n).$$

Given a permutation group $\Gamma \subset S_n$ we say that a Boolean function $f(x_1, x_2, \ldots, x_n)$ is Γ-invariant if $\Gamma \subset Aut(f)$.

We recall that a permutation group $\Gamma \subset S_n$ is *transitive* if for every i and j there is $\pi \in \Gamma$ such that $\pi(i) = j$. $\Gamma \subset S_n$ is *imprimitive* if we can divide $[n]$ into blocks $V_1, \ldots, V_k$, $1 \leq k \leq n$ such that every permutation $\pi \in \Gamma$ permutes these blocks, otherwise Γ is *primitive.* The automorphism group of the majority function of n variables is the full symmetric group S_n. The automorphism groups for the tribe example and the ternary recursive majority function are transitive but non-primitive groups of permutations. Boolean functions based on graph properties are Γ-invariant where Γ is the symmetric group on m elements acting on all $\binom{m}{2}$ pairs. Here, Γ is primitive.

4. Fourier

4.1. Fourier–Walsh expansion. We come now to the second principal part of the lecture dealing with Fourier analysis. Fourier analysis on the discrete cube has the pleasant feature that the basic setting is elementary and simple and various subtle difficulties in the continuous case disappear in the discrete case.

Every real function $f : \Omega_n \to \mathbb{R}$ can be expressed in terms of the Fourier–Walsh basis. We write here and for the rest of the paper $[n] = \{1, 2, \dots, n\}$.

$$f = \sum \{\hat{f}(S) W_S : \; S \subset [n]\}, \tag{4}$$

where the *Fourier-Walsh function* W_S is defined by

$$W_S(x_1, x_2, \dots, x_n) = (-1)^{\sum\{x_i : \; i \in S\}}.$$

Note that we have here 2^n functions, one for each subset S of $[n]$. The function W_S is simply the *parity* function for the variables in S. These functions form an orthonormal basis of $\mathbb{R}^{\Omega_n}$ with respect to the inner product

$$\langle f, g \rangle = \sum_{x \in \Omega_n} \mu(x) f(x) g(x).$$

The coefficients $\hat{f}(S) = \langle f, W_s \rangle$, $S \subset [n]$, in (4) are real numbers, called the *Fourier coefficients* of f.

Now, a very basic fact from linear algebra tells you that the inner product of two functions f and g can be expressed in terms of their Fourier coefficients:

Parseval: For real functions f, g on Ω_n,

$$\langle f, g \rangle = \sum_{S \subset [n]} \hat{f}(S) \hat{g}(S).$$

The next step, which is also simple, is to express the influence of variables for Boolean functions in terms of Fourier coefficients. (To a large extent, this step is also part of the large picture of the spectral understanding of expansion in general graphs.) Let $f : \Omega_n \to \{0, 1\}$ be a Boolean function and let $f = \sum\{\hat{f}(S) W_S : \; S \subset [n]\}$ be its Fourier expansion. We easily obtain from Parseval's formula, by looking at the inner product of f with itself, that

$$\mathbb{E}(f) = \|f\|_2^2 = \sum \hat{f}^2(S).$$

When we look at individual influences, we get, again by a similar application of Parseval's formula, a formula for the influence of the k-th variable

$$I_k(f) = 4 \sum_{S \subset [n], k \in S} \hat{f}^2(S).$$

And summing up we obtain that the total influence is expressed in terms of the Fourier–Walsh expansion as follows:

$$I(f) = 4 \sum \hat{f}^2(S)|S|.$$

So every Fourier coefficient (squared) comes with a weight that is the size of the set indexing the Fourier–Walsh function. We can think of $|S|$ as the "frequency" associated with W_S and the Fourier coefficient $\hat{f}(S)$, and of $\hat{f}^2(S)$ as the "energy" of f associated with W_S.

4.2. Expansion and strong spectral expansion. Now, let $t = \mathbb{E}(f)$. We saw that $t = \|f\|_2^2 = \sum_S \hat{f}^2(S)$ and that $I(f) = 4\sum_S \hat{f}^2(S)|S|$.

This gives us again that $I(f) \geq 4\sum_{S\neq\emptyset} \hat{f}^2(S) = 4t(1-t)$, (equation (1)) which we earlier proved using canonical paths.[5] This argument is a special case of the more general lower bound for edge expansion in terms of the spectral gap of the Laplacian of a graph.

Let's consider a probability distribution on subsets S of $\{1, 2, \ldots, n\}$, where the probability of a set S is $t^{-1}\hat{f}^2(S)$. We will refer to this distribution as the *spectral distribution* of f. Recall that the total influence is $4t$ times the expectation of $|S|$ with respect to the spectral distribution. We know that $I(f) \geq 2t \log(1/t)$, and so the expected value of $|S|$ with respect to the spectral distribution is at least $2\log(1/t)$. Thus, when t is small this expected value becomes large.

Now we can ask an interesting question. Can we say that *most* of the "energy" of f comes from levels where $|S| = \Omega(\log(1/t))$? We will refer to this property as *strong spectral expansion.* Strong spectral expansion would follow from an affirmative answer to the following question:

Question: Is it always the case for a Boolean function f that most of the ℓ_2 contribution to $\|f\|_2$ comes from levels where $|S| = \Omega(I(f))$?

However, the answer to this question is negative. For Boolean functions, it is not always true that if the total influence is large, then there is a large contribution to $\|f\|_2$ from Fourier coefficients of high frequencies.

The anomaly of majority. Recall that the majority function is defined by $f(x_1, x_2, \ldots, x_n) = 1$ if $x_1 + x_2 \ldots x_n \geq n/2$. For the majority function f, most of the contribution to $\|f\|^2$ of Fourier coefficients squared $\hat{f}^2(S)$ comes from coefficients where $|S|$ is bounded. This property is referred to as *noise stability* (see Section 4.6). On the other hand, $I(f) = \Theta(\sqrt{n})$. I refer to the discrepancy between the large influence of the majority function and its noise stability as the *anomaly of majority.*

The anomaly of majority is related to the fact that robust classical information and computation are at all possible, in our noisy world.[6]

[5] Some people claim that this is even the same *proof* as the canonical path argument but I do not understand this claim.

[6] In [43] I argue that the absence of quantum error-correcting codes that are noise-stable accounts for the impossibility of computationally superior quantum computers.

4.3. The proof of KKL's theorem. I would now like to mention a technical tool that allows you to prove strong spectral expansion for sets of small measure, and the KKL theorem. We need two technical ingredients that are interesting and important. The first is the noise operator, and the second is hypercontractive inequalities.

Noise and hypercontractivity. Given a real function f on the discrete cube with Fourier expansion $f = \sum\{\hat{f}(S)W_S :\ S \subset [n]\}$, the noisy version of f, denoted by $T_\rho(f)$, is obtained by suppressing the Fourier coefficients exponentially with $|S|$. It is defined formally as follows:

$$T_\rho(f) = \sum\{\hat{f}(S)(\rho)^{|S|}W_S :\ S \subset [n]\}.$$

We come to a very important inequality, which was discovered by Aline Bonami[7]
The Bonami hypercontractive inequality:

$$\|T_\rho(f)\|_2 \ \leq\ \|f\|_{1+\rho^2}.$$

Here

$$\|f\|_p^p = \sum_{x\in\Omega_n} \mu(x)|f(x)|^p.$$

In what follows we will only use the special case, $\rho = 1/2$ which (when squared) reads $\|T_{1/2}(f)\|_2^2 = \sum \hat{f}^2(S)2^{-2|S|}\ \leq\ \|f\|_{5/4}^2$.

Hypercontractivity and strong spectral expansion.

Theorem 4.1. *Let f be a Boolean function with $\|f\|_2^2 = t$. Then*

$$\sum\{\hat{f}^2(S) : |S| < \frac{1}{10}\log(1/t)\} \leq t^{1.1}. \tag{5}$$

This is the strong spectral expansion: the overall contribution of Fourier coefficients (squared) with small frequencies is small. (Hence, the overall contribution of those with high frequencies must be large.)

Proof. (sketch, without the computation)

0. Parseval gives $I(f) = 4\sum \hat{f}^2(S)|S|$.
1. Bonami hypercontractive inequality:

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

And now comes the punchline.

[7]This inequality was rediscovered by Leonard Gross in relation with Gross's log-Sobolev inequalities. The applications to combinatorics were greatly influenced by a paper of William Beckner. The inequality can be regarded as an extension of classic inequality by Aleksandr Khinchin.

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 $r \neq q$. In our case, $\sum \hat{f}^2(S)(1/2)^{2|S|} \leq \|f\|^2_{5/4} = t^{8/5}$. This implies (5).

$\square$

Hypercontractivity, and the proof of KKL's theorem. The strong spectral expansion theorem asserts that if g is a Boolean function, and $\mathbb{E}(g) = t$, then most of the contribution to $\|g\|_2^2$ comes from Fourier–Walsh coefficients $\hat{g}^2(S)$ where $|S| \geq C \log(1/t)$. As a matter of fact, our proof applies word-for-word even when $g : \Omega_n \to \{-1, 0, 1\}$. (We shall need this slight extension.)

Let $f_k(x) : \Omega_n \to \{-1, 0, 1\}$ defined by

$$f_k(x) = f(x) - f(\sigma_k(x)).$$

We can think of f_k as a partial derivative of f in the kth direction. Thus $\|f_k\|_2^2 = I_k(f)$. Recall that KKL's theorem asserts that there exists a variable k such that

$$I_k(f) \geq C\mathrm{Var}(f) \log n/n.$$

Now, if for some k, $I_k(f) \geq \mathrm{Var}/\sqrt{n}$, then we are done. The spectral edge-expansion theorem tells us that if $I_k(f) < \mathrm{Var}/\sqrt{n}$, then the contribution of Fourier coefficients (squared) of f_k below the $\frac{1}{10}\log n$ level is negligible, and if this holds for every k it easily implies that the contribution to $\|f\|_2^2$ of Fourier coefficients $\hat{f}^2(S)$ where $0 < |S| < \frac{1}{10}\log n$ is also negligible. Therefore, the contribution to $\|f\|_2^2$ of Fourier coefficients $\hat{f}^2(S)$ where $|S| \geq \frac{1}{10}\log n$ is $(1 - o(1))\mathrm{Var}(f)$, and hence the total influence of f is at least $C \log n \mathrm{Var}(f)$. This gives that either we have a variable with very large influence or else the total influence is at least $C \log n$.

4.4. The entropy influence conjecture.

We come to our second problem.

Let $f : \Omega_n \to \{0, 1\}$ be a Boolean function and let $f = \sum\{\hat{f}(S)W_S : \ S \subset [n]\}$, be its Fourier expansion. Define

$$H(f) = \sum_{S \subset [n]} \hat{f}^2(S) \log(1/\hat{f}^2(S)).$$

Problem 2 (The entropy-influence conjecture of Friedgut and Kalai, 1996) [30]: Prove that for some absolute constant $c > 0$, for every Boolean function f,

$$I(f) \geq c \cdot H(f).$$

This conjecture includes KKL's theorem and has various interesting consequences. The motivating application of the conjecture was to the give lower bounds for Γ-symmetric Boolean functions. Such lower bounds were eventually proved by Bourgain and Kalai [12] (see Section 5.2), and I still hope that the techniques of this work may shed light and the conjecture in its full generality. I will briefly mention one additional application.

A conjecture of Mansour. Consider a Boolean function f described by a formula in conjunctive normal form (namely, AND of ORs of literals) of polynomial size in n. (See Section 3.5.1.) Mansour [57] conjectured that most of the Fourier coefficients are concentrated only on polynomial number of coefficients! This conjecture is still open. A theorem of Håstad and Boppana implies that $I(f) = O(\log n)$; thus Mansour's conjecture will follow from the entropy/influence conjecture.

4.5. First-passage percolation. I want to mention another topic where the proof technique of KKL's theorem applies fairly directly. Consider an infinite planar grid where every edge is assigned a length: 1 with probability 1/2 and 2 with probability 1/2 (independently). This model of a random metric on the planar grid is called first-passage percolation.

Question: What is the variance $V(n)$ of the distance from $(0, 0)$ to $(n, 0)$?

Kesten (1993) [48] proved that $V(n) = O(n)$ and Benjamini, Kalai, and Schramm (2003) [6] proved that $V(n) = O(n/\log n)$. The proof is very similar to the proof of KKL's theorem using the same "partial derivatives" of the distance function from $(0, 0)$ to $(n, 0)$, and applying the hypercontractive inequity. (The proof extends to every dimension $D \geq 2$.) Unlike the KKL theorem this bound is not sharp. Our third problem is:

Problem 3: Prove that $V(n) = o(n/\log n)$. More ambitiously, prove that $V(n) = O(n^{1-c})$ for some $c > 0$.

In the plane it is conjectured that $V(n)$ is proportional to $n^{2/3}$. For much more on first-passage percolation see the survey article [2]. For some related important developments, see [15, 16, 63, 1].

4.6. Noise-stability and noise sensitivity. We will consider here (for simplicity) monotone balanced Boolean functions. We say here that f is balanced if $\frac{1}{10} \leq \mathbb{E}(f) \leq \frac{9}{10}$. A class U of such Boolean functions is uniformly *noise stable* if for every $\epsilon > 0$ there is k such that for every $f \in U$,

$$\sum_{S:\ |S| \leq k} \hat{f}^2(S) \geq (1 - \epsilon)\|f\|_2^2.$$

An equivalent definitions goes as follows. Given $x = (x_1, x_2, \ldots, x_n)$ define $y = (y_1, y_2, \ldots, y_n)$ by $y_i = x_i$ with probability $1 - s$ and $y_i = 1 - x_i$ with probability s, independently for $i = 1, 2, \ldots, n$. Define $N_s(f)$ as the probability that $f(y) = 0$ conditioned on $f(x) = 1$ where x is drawn at random from Ω_n and y is described above. Uniform noise stability is equivalent to the statement that for every $\epsilon > 0$ there is $s > 0$ such that for every $f \in U$, $N_s(f) < \epsilon$. A sequence of balanced monotone Boolean functions f_m is *noise sensitive* if for every $k > 0$, $\lim_{n\to\infty} \sum_{S:1<|S|\leq k} \hat{f}^2(S) = 0$. An equivalent formulation is that for every $s > 0$ the correlation between $f(x)$ and $f(y)$ tends to 0. We mention now briefly some basic facts about noise stability and sensitivity.

Benjamini, Kalai, and Schramm (1999) [5] proved that noise stable monotone balanced Boolean functions f must have positive (bounded away from zero) correlation with a *weighted majority function*. They also showed that noise stability for monotone Boolean functions implies that $\sum_{k=1}^n I_k^2(f)$ is bounded away from zero. One of the most important results regarding the analysis of Boolean functions is the "majority is stablest" theorem by Mossel, O'Donnell, and Oleszkiewicz [59, 51], which, as its name suggests, asserts that among all Boolean functions with diminishing maximum influence, the majority function asymptotically maximizes the stability to noise.

5. Threshold

We now come to the third and last principal part of the lecture. Let f be a monotone Boolean function. Write $\mathbb{E}_p(f) = \mu_p\{x : \Omega_n : f(x) = 1\}$. $\mathbb{E}_p(f)$ equals the probability that $f(x) = 1$ for a vector in Ω_n randomly drawn according to the Bernoulli probability distribution μ_p. It is easy to prove that $\mathbb{E}_p(f)$ is monotone in p. A more detailed information is given by Russo's lemma.

Theorem 5.1 (Russo's lemma). *For a monotone Boolean function f,*

$$d\mathbb{E}_p(f)/dp = I^p(f).$$

This result is very useful in percolation theory and other areas. The *threshold interval* for a monotone Boolean function f is those values of p so that $\mathbb{E}_p(f)$ is bounded away from 0 and 1. (Say $0.01 \le \mathbb{E}_p(f) \le 0.99$.)

A typical application of Russo's lemma goes like this: If for every value p in the threshold interval the total influence $I^p(f)$ is large, then the threshold interval itself is short. This is called a *sharp threshold* phenomenon. Russo's 1982 zero-one law [65] asserts that if all variables have small influence then the threshold interval is small. KKL's theorem and its refinements allow proving strong results of this kind [68, 30, 42].

5.1. Invariance under transitive group. We will study now how the total influence and hence the length of the threshold interval depends on the symmetry of the Boolean function. Roughly speaking, stronger symmetry implies larger total influence and hence sharper threshold. We refer the reader to Section 3.5.2 for the definition of transitive and primitive permutation groups.

Theorem 5.2 (Friedgut and Kalai 1996 [30]). *If a monotone Boolean function f with n variables is invariant under a transitive group of permutations of the variables, then its threshold interval is of length $O(1/\log n)$.*

Proof. The relation

$$I^p(f) \ge C\mathbb{E}_p(f)(1 - \mathbb{E}_p(f)) \log n,$$

follows from KKL's theorem (extended to the Bernoulli case) as follows: The theorem asserts that for some k, $I_k^p(f) \ge C\mathbb{E}_p(f)(1-\mathbb{E}_p(f)) \log n/n$. The symmetry implies that for every p all individual influences are the same and therefore $I^p(f) \ge C\mathbb{E}_p(f)(1-\mathbb{E}_p(f)) \log n$. We conclude that as long as $\mathbb{E}_p(f)$ is bounded away from zero and one the total influence is at least $C' \log n$. □

5.2. Total influence under symmetry of primitive groups. A subsequent (much harder) result was proved by Jean Bourgain and myself. Our aim was to find methods allowing one (in certain cases) to cross the $\log n$ barrier. For a group of permutations $\Gamma \subset S_n$, let $I(\Gamma)$ be the minimum total influence for a balanced Γ-invariant function Boolean function with n variables.

The main theorem of [12] gives, for an arbitrary permutation group Γ, a lower bound for $I(\Gamma)$ in terms of the sizes of orbits of Γ acting on subsets of $[n]$. (This is complemented by an easy upper bound for $I(G)$ also in terms of orbit sizes.) This theorem combined with further study of orbits in permutation groups[8] leads to the following good understanding of $I(\Gamma)$ for primitive groups.

Theorem 5.3 (Bourgain and Kalai 1997 [12])**.** *If Γ is primitive then one of the following possibilities hold.*

(i) $I(\Gamma) = \Theta(\sqrt{n})$,

(ii) $(logn)^{(k+1)/k-o(1)} \leq I(\Gamma) \leq C(\log n)^{(k+1)/k}$, *(here $k \geq 1$ is an integer).*

(iii) $I(\Gamma)$ *behaves like* $(\log n)\mu(n)$, *where* $\mu(n) \leq \log\log n$ *is growing in an arbitrary way.*

For graph properties we obtained:

Theorem 5.4 (Bourgain and Kalai 1997 [12])**.** *For every $\delta > 0$, the total influence of a Boolean function f with n variables based on a graph property is at least*

$$C_\delta(\log n)^{2-\delta}\mathrm{Var}(f).$$

5.3. Two unexpected applications*. Let me now mention two recent unexpected applications of the theorems about thresholds.

Suppose that you want to transmit a message of n bits; however, for every bit there is a probability p that the receiver will obtain a question-mark sign rather than the value of the bit. This situation is referred to as the *erasure channel.* Shannon proved an upper bound on the amount of information (capacity) one can transmit in such a noisy channel.

Recently Kumar and Pfister [53] used Theorem 5.3 to show that the Reed–Muller codes achieve capacity on erasure channels. (A somewhat weaker result based on Theorem 5.2 was achieved independently by Kudekar, Mondelli, Sasoglu, and Urbanke, [52].)

The second application is due to Ellis and Narayanan [23]. It settles a conjecture by Peter Frankl from 1981.

Theorem 5.5. *Let $\mathcal{F}$ be a family of subsets of $[n]$. Suppose that:*

1) $\mathcal{F}$ is invariant under a transitive group Γ of permutations of $[n]$,

2) Every three sets in $\mathcal{F}$ have a point in common.

Then $|\mathcal{F}| = o(2^n)$.

[8]based on the O'Nan-Scott theorem on primitive permutation groups, and thus also on the classification theorem for finite simple groups,

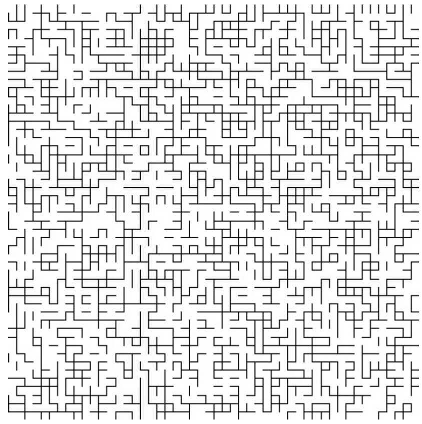

Figure 3. Planar percolation

5.4. The expectation threshold conjecture. In what follows, $G(n, p)$ refers to a random graph on a set of n vertices where the probability for two vertices to be adjacent is p. Consider a random graph G in $G(n, p)$ and the graph property: G contains a copy of a specific graph H. (Note: H may depend 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 $1/2$. It is easy to see that $q \leq p$.

Problem 4 (Conjecture by Kahn and Kalai from 2006 [41]): Show that

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

In random graph theory we are mainly interested in the value of p, and the value of q is a very "cheap" approximation to it. When H is a perfect matching then p behaves like $\log n/n$ and q behaves like $1/n$. This example shows that the $\log n$ factor cannot be reduced. This (very bold) conjecture can be vastly extended to general Boolean functions and there are mysterious connections with Problem 1; see [41] and [72].

6. Connections and applications

In the final part of the lecture I want to mention briefly a few connections and applications.

6.1. Bernoulli percolation in 2D and 3D. The d-dimensional grid graph has $\mathbb{Z}^d$ as its vertices, and two vertices are adjacent if they are unit distance apart. Bernoulli percolation is the study of random subgraphs of the grid graph where every edge is taken with probability p (independently). In this theory, results and insights for the infinite and finite models interplay. The crossing event for percolation refers to a finite axis-parallel box in $\mathbb{Z}^d$ (rectangle in the plane) and to the event that there is a path crossing from the left side to the right side of the box. Note that the crossing event can be described by a Boolean function of n variables where n is the number of edges in the box (rectangle).

6.1.1. Much is known in two dimensions. Much is known for planar Bernoulli percolation: Kesten proved that the critical probability of planar percolation is 1/2, namely, for $p < 1/2$ there is no infinite connected component and when $p > 1/2$ there is one. This relies on the fundamental Russo–Seymour–Welsh theorem about the probability of the crossing event in rectangular grids. At the critical probability $p = 1/2$ the probability of an infinite component is zero. The influence of the crossing event described as a Boolean function on n variables at the critical probability is larger than n^α and smaller than n^β for $0 < \alpha < \beta < 1/2$. See [46, 47, 49]. Smirnov's celebrated result asserts that the crossing probabilities are conformally invariant.[9] Schramm identified SLE_6 as the scaling limit of the exploration path for critical planar percolation. Closer to the theme of this paper, the crossing event for percolation was proved to be noise-sensitive [5] and even very noise-sensitive [66]. Garban–Pete–Schramm [32] gave a remarkable spectral description of the crossing event in planar percolation. The critical probability for percolation was computed even for the more general Potts model by Beffara and Duminil-Copin [9].

6.1.2. Little is known in three dimensions. Matters change greatly when we move to higher dimensions.

(Famous) Problem 5: Understand 3-D percolation.

In particular,

- Show that there is no percolation at the critical probability p_c.
- Prove an analog for the Russo–Seymour–Welsh theorem.
- Give bounds for influences of the crossing event.
- We can also ask for a proof of noise-sensitivity, for existence and description of a scaling limit, and for description of Fourier probability distribution for the crossing event.

Let me (informally) mention a recent result.

Theorem 6.1 (Kalai and Kozma 2016)**.** *For the crossing event in 3D percolation, regarded as a Boolean function with n variables, either $I \geq n^\beta$, $\beta > 0$, or the energy on "well-separated sets S," with $|S| \leq \log^{3/2} n$ is negligible.*

[9] For a certain variant of percolation based on the triangular planar grid. This is still open for the rectangular grid.

Suppose that we consider the crossing event for a $L \times L \times L$ box. (So $n \sim L^3$.) A set S is"well-separated" if for every ball B of radius $L^{1/10}$, there is at most one element in S representing an edge in B. We would like to dispense with this "well-separated" condition altogether and this would yield a much desired lower bound above $\log^{1+\epsilon} n$, $\epsilon > 0$, for the total influence of the crossing event.

6.2. Projection. I now want to briefly describe some questions recently addressed in a paper [13] with Jean Bourgain and Jeff Kahn.

Problem 6: For a set $A \subset \Omega_n$ and $T \subset [n]$, let A_T denote the projection of A on the coordinates in T. Given b, $0 < b < 1$, and $t = \mu(A)$ what is the maximum value guaranteed for $\mu(A_T)$ over all subsets T of $[n]$ of size bn?

The famous Sauer–Shelah lemma asserts that if $|A| > \sum_{i=0}^{k} \binom{n}{i}$ then for some T, $|T| = i + 1$, $\mu(A_T) = 1$. This implies that when $b < 1/2$, if $t > 2^{H(\alpha)}n$ where $H(\alpha)$ is the entropy function, we can guarantee that for some T, $|T| = bn$, $A_T = \{0,1\}^T$. While interesting special cases of Problem 6 remains also for $b < 1/2$ we will restrict our attention to $b \geq 1/2$ and will consider three special cases of Problem 6.

Problem 6.1: Let $b, 1/2 < b < 1$ be a real number. For a subset A of measure $1/2$ of the discrete cube Ω_n, what can be said about the maximum value of $\mu(A_T)$, for $T \subset [n]$, and $|T| = bn$?

We have now a good understanding of Problem 6.1. Given $A \subset \Omega_n$, $\mu(A) = 1/2$, it follows from KKL's theorem that for some constant $c > 0$, and for every $b < 1$, there exists T with $T = bn$ such that $\mu(A_T) \geq 1 - n^{c(1-b)}$. On the other hand, we proved in [13] that for every $\delta > 0$ there is $C = C_\delta$ and $A \subset \Omega_n$ with $\mu(A) = 1/2$ such that for every $T = bn$, $b = (1/2 + \delta)n$ we have that $\mu(A_T) \geq 1 - n^{-C}$.

Problem 6.2: Given $b, 1/2 \leq b < 1$, what is the minimum value of t so that every set $A \subset \Omega_n$ of measure t will have a projection A_T for some subset T, $|T| = bn$ with $\mu(A_T) \geq 1/2$?

It follows from KKL's theorem that for some $c > 0$ if $t \geq 1/n^{c/(1-b)}$, we can guarantee a subset T of bn coordinates such that $\mu(A_T) \geq 1/2$.. There have been some recent progress on both lower and upper bounds : One the one hand we proved that for every $C > 0$ there is some $b > 1/2$ such that if $t \geq n^{-C}$ then there is a set T of size bn, with $\mu(A_T) \geq 1/2$. On the other hand, there is $b > 1/2$, and a set A, $\mu(A) = \exp(-n^{1-\alpha})$ such that $\mu(A_T) \leq exp(-n^{\beta})$. Here $\alpha, \beta > 0$ are small real numbers. The gap between the lower and upper bounds is still very large.

Problem 6.3: Given $\mu(A)$ what can be said about the maximum of $\mu(A_T)$ for $|T| = n/2$?

6.3. Correlation. Let f and g be monotone Boolean functions. The covariance of f and g is defined by $cov(f, g) = \mathbb{E}(fg) - \mathbb{E}(f)\mathbb{E}(g)$. A famous theorem of Harris and Kleitman asserts that $cov(f, g) \geq 0$.

Problem 7: Given two monotone functions f and g, what can be said about $cov(f, g)$?

Talagrand's correlation formula: If f and g are monotone Boolean functions with n variables, then

$$cov(f, g) \geq K \sum_{k=1}^{n} I_k(f) I_k(g) / \log(1/ \sum_{k=1}^{n} I_k(f) I_k(g)).$$

Talagrand's formula is very beautiful, and referring to Landau's 1925 comments one could say that the formula and its study can be a solace. For some subsequent developments, see [45, 44].

Kahn's correlation conjecture: If f and g are monotone Boolean functions and g is an odd function (i.e., $g(1-x) = 1 - g(x)$), then

$$cov(f, g) \geq \sum_{k=1}^{n} I_k(f) \sum_{S:k\in S} \frac{1}{|S|} \hat{g}^2(S).$$

This beautiful conjecture (reproduced and studied in [31]) dates back to the late 80s (but the Fourier-theoretic formulation came a few years later). It implies the following well-known conjecture of Chvátal in extremal combinatorics.

A family $\mathcal{F}$ of sets is called an *ideal* (a.k.a. a down-family, a.k.a. a simplicial complex) if it satisfies the property that $S \in \mathcal{F}$ and $R \subset S$ implies $R \in \mathcal{F}$. A family of sets $\mathcal{G}$ is *intersecting* if for every S and T in $\mathcal{G}$, $S \cap T \neq \emptyset$.

Chvátal's conjecture: Let $\mathcal{F}$ be an ideal of subsets of a set V and let $\mathcal{G}$ be an intersecting subfamily of $\mathcal{F}$. Then there is an element $v \in V$ such that

$$|\mathcal{G}| \leq |\{S \in \mathcal{F} : v \in S\}|.$$

6.4. Intersection*. Let $\mathcal{F}$ be an intersecting family of subsets of $[n]$. Since $\mathcal{F}$ cannot contain a set and its complement, we obtain that

$$|\mathcal{F}| \leq 2^{n-1}.$$

This is sharp as seen by the family of sets containing a single element (dictatorship). It is surprising how slight changes of this problem are already deep and difficult. For example, if we ask that every two sets in the family have a nonempty intersection and also the complements of every two sets have a nonempty intersection, then

$$|\mathcal{F}| \leq 2^{n-2}.$$

This result of Kleitman answered a question by Erdős that stood open for several years and its solution is based on the Harris–Kleitman inequality. Erdős, Ko, and Rado (EKR) proved that if $\mathcal{F}$ is an *intersecting* family of k-subsets of $[n]$ and $2k \leq n$ then $|F| \leq \binom{n-1}{k-1}$. EKR-type results and problems consist of a wide area of extremal combinatorics. Chvátal's conjecture, which we mentioned in the previous section, can be seen as a vast generalization of Erdős–Ko–Rado's theorem.

Let me mention the recent resolution of two beautiful conjectures, where Fourier methods played a crucial rule.

Theorem 6.2 (Ellis, Friedgut, and Pilpel 2011 [19], conjectured by Deza and Frankl 1977)**.** *For every k if n is sufficiently large, any family of permutations on an n-element set such that every two permutations in the family agree in at least k places, contains at most $(n-k)!$ permutations.*

Theorem 6.3 (Ellis, Filmus, and Friedgut 2012 [20], conjectured by Simonovits and Sós 1976)**.** *Let $\mathcal{F}$ be a family of graphs on the vertex set $[n] = \{1, 2, \ldots, n\}$. Suppose that for every two graphs in the family there is a triangle included in both. Then*

$$|\mathcal{F}| \leq \frac{1}{8} 2^{\binom{n}{2}}.$$

The first result uses "non–Abelian" Fourier analysis on the symmetric group, i.e., the representation theory of the symmetric group. The spectral analysis of graphs and the Fourier analysis of Boolean functions have played an interesting role (along with various other methods) for other Erdős–Ko–Rado type theorems. One line of applications (see [21] and papers cited there) is stability theorems for Erdős–Ko–Rado results in the spirit of the classic results by Hilton–Milner [39] and Frankl [27].

6.5. Computing. The combinatorics, analysis, and geometry of Boolean functions, Fourier expansion, and noise, have strong connections with and many applications to various areas of the theory of computing: algorithms, computational complexity, derandomization, distributed computing, cryptography, error-correcting codes, computational learning theory, and quantum information and computation.

7. Conclusion

Fourier and combinatorics. In this lecture I mentioned one area in which Fourier analysis is connected to combinatorics, namely, the study of discrete isoperimetric relations. The study of discrete isoperimetric inequalities is a rich area and exploring further connections to Fourier analysis within this area is also of great interest. Another direction is that of pseudorandomness both in additive combinatorics and computer science. Yet another important direction is that of finding upper bounds for the size of error-correcting codes and sphere packing, and last but not least is discrepancy theory. It would be very nice to find further applications to these and other areas of discrete mathematics, and also some connections between the use of Fourier analysis in these areas.

Generalization. Having zoomed in on the graph of the discrete cube we can zoom out again and ask for generalizations. We can consider other probability product spaces and even more general probability distributions. We can consider more general real functions on Ω_n. We can consider other graphs.[10] We can

[10] Hypercontractivity and the closely related log-Sobolev inequalities were studied for general graphs by Diaconis and Saloff-Coste [18]. One way to think about it is to consider for a class of

study other groups and representations: here we considered essentially the group $Z/2Z$. Ω_n is the graph of the n-dimensional cube and we can consider expansion and spectral properties of graphs of other (simple) polytopes. Finally, we can study high-dimensional expansion [26, 54], which is an exciting recent notion extending the expansion property of graphs to higher-dimensional cellular objects. Understanding high-dimensional expansion for the two-dimensional skeleton of the n-dimensional cube could be of interest.

Recommended books. I will end my lecture by mentioning three books. One book is by Michel Ledoux, "The Concentration of Measure Phenomenon" [55]. It describes the larger area of discrete isoperimetric inequalities and relations with probability theory. The next two books deal directly with the topic of this lecture and they nicely complement each other: "Analysis of Boolean Functions" by Ryan O'Donnell [61] and "Noise Sensitivity of Boolean Functions and Percolation" by Cristoph Garban and Jeffrey Steif [34].

Thank you very much!

8. Appendix: Two remarkable documents by Edmund Landau

I would like to say a few words about two remarkable documents written by Landau in 1925, both related to the inauguration ceremony of the Hebrew University of Jerusalem. You can read more about them in the paper "Zionist internationalism through number theory: Edmund Landau at the Opening of the Hebrew University in 1925" by Leo Corry and Norbert Schappacher [17]. The first document is Landau's toast for the opening ceremonies. Let me quote two sentences:

> "May great benefit emerge from this house dedicated to pure science, which does not know borders between people and people. And may this awareness emerge from Zion and penetrate the hearts of all those who are still far from this view."

The second document, also from 1925, is probably the first mathematical paper written in Hebrew in modern times. It is devoted to twenty-three problems in number theory and here are its concluding sentences.

> "At this number of twenty-three problems I want to stop, because twenty-three is a prime number, i.e., a very handsome number for us. I am certain that I should not fear to be asked by you, for what purpose does one deal with the theory of numbers and what applications may it have. For we deal with science for the sake of it, and dealing with it

graphs an inequality of the form

$$\|e^{-\Delta} f\|_2 \le \|f\|_p,$$

for some fixed constant $p < 2$. Here, Δ represents the Laplacian of a graph in the family.

> was a solace in the days of internal and external war that as Jews and as Germans we fought and still fight today."

I wish to make two remarks: First, note that Landau moved from the very ambitious hopes and program of science as a bridge that eliminates borders between nations to a more modest and realistic hope that science and mathematics give comfort in difficult times. Juggling between very ambitious programs and sober reality is in the nature of our profession and we are getting paid both for the high hopes and aims, as well as for the modest results. Second, Landau is famous for his very rigorous and formal mathematical style but his 1925 lecture is entertaining and playful. I don't know if his move to Jerusalem was the reason for this apparent change of style. Parts of Landau's lecture almost read like stand-up comedy. Here is, word for word, what Landau wrote about the twin prime conjecture: "Satan knows [the answer]. What I mean is that besides God Almighty no one knows the answer, not even my friend Hardy in Oxford." These days, ninety years after Landau's lecture, we can say that besides God Almighty no one knows the answer and not even our friend James Maynard from Oxford. We can only hope that the situation will change before long.

Landau's hopeful comments were made only nine years after the end of the terrible First World War. He himself died in 1938 in Berlin, after having been stripped of his teaching privileges a few years earlier. I don't know to what extent the beauty of mathematics was a source of comfort in his last years, but we can assume that this was indeed the case. My life, like the lives of many others of my generation, was overshadowed by the Second World War and the Holocaust and influenced by the quest to come to terms with those horrible events.

References

[1] D. Ahlberg, C. Hoffman, Random coalescing geodesics in first-passage percolation, arXiv:1609.02447.

[2] A. Auffinger, J. Hanson, and M. Damron, 50 years of first-passage percolation. arXiv:1511.03262.

[3] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, New York (1992).

[4] N. Alon, I. Dinur, E. Friedgut, and B. Sudakov, Graph products, Fourier analysis and spectral techniques, *Geom. Funct. Anal.* 14 (2004), 913–940.

[5] I. Benjamini, G. Kalai, and O. Schramm, Noise sensitivity of Boolean functions and applications to percolation, *Publ. I.H.E.S.* 90 (1999), 5–43.

[6] I. Benjamini, G. Kalai, O. Schramm, First passage percolation has sublinear distance variance, *Ann. Probab.* 31 (2003), 1970–1978.

[7] M. Ben-Or and N. Linial, Collective coin flipping, in *Randomness and Computation* (S. Micali, ed.), New York, Academic Press (1990), pp. 91–115. Earlier version: Collective coin flipping, robust voting games, and minima of Banzhaf value, *Proc. 26th IEEE Symp. on the Foundation of Computer Science* (1985), pp. 408–416.

[8] W. Beckner, Inequalities in Fourier analysis, *Annals of Math.* 102 (1975), 159–182.

[9] V. Beffara, and H. Duminil-Copin, The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$, *Probability Theory and Related Fields* 153 (2012), 511–542.

[10] B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* 7 (1987), 35–38.

[11] A. Bonami, Etude des coefficients Fourier des fonctiones de $L^p(G)$, *Ann. Inst. Fourier* 20 (1970), 335–402.

[12] J. Bourgain and G. Kalai, Influences of variables and threshold intervals under group symmetries, *Geom. Funct. Anal.* 7 (1997), 438–461.

[13] J. Bourgain, J. Kahn, and G. Kalai, Influential coalitions for Boolean functions, to appear in *Th. of Computation.* arXiv:1409.3033.

[14] J. Bourgain, On sharp thresholds of monotone properties, *J. Amer. Math. Soc.* 12 (1999), 1051–1054.

[15] S. Chatterjee, Chaos, concentration, and multiple valleys, arXiv:0810.4221.

[16] S. Chatterjee, *Superconcentration and Related Topics,* Springer Monographs in Mathematics. Springer, Berlin-Heidelberg, 2014.

[17] L. Corry and N. Schappacher, Zionist internationalism through number theory: Edmund Landau at the opening of the Hebrew University in 1925, *Science in Context* 23 (2010) 427–471.

[18] P. Diaconis and L. Saloff-Coste, Logarithmic Sobolev inequalities for finite Markov chains, *Ann. Appl. Prob.* 6 (1996), 695–750.

[19] D. Ellis and E. Friedgut, and H. Pilpel, Intersecting families of permutations. *J. of the Amererican Math. Soc.* 24 (2011), 649–682.

[20] D. Ellis, Y. Filmus, and E. Friedgut, Triangle-intersecting families of graphs, *J. of the European Math. Soc.* 14 (2012), 841–885.

[21] D. Ellis, N. Keller, and N. Lifschitz, Stability versions of Erdős–Ko–Rado type theorems, via isoperimetry. arxiv.1604.02160.

[22] D. Ellis, N. Keller, and N. Lifschitz, On the structure of subsets of the discrete cube with small edge boundary, arXiv:1612.06680.

[23] D. Ellis and B. Narayanan, On symmetric 3-wise intersecting families. To appear in *Proc. of the American Math. Soc.*, arxive 1608.03314.

[24] P. Erdős and A. Renyi, On Random graphs I, *Publ. Math. Debrecen* 6 (1959), 609–627.

[25] P. Erdős, C. Ko, and R. Rado, Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser. 2*, 12 (1961), 313–320.

[26] J. Fox, M. Gromov, V. Lafforgue, A. Naor, and J. Pach, Overlap properties of geometric expanders, *Journal fur die reine und angewandte Mathematik.*, 671 (2012), 49–83.

[27] P. Frankl, Erdős–Ko–Rado theorem with conditions on the maximal degree, *J. Combin. Th. Series A 46* (1987), 252–263.

[28] E. Friedgut, Boolean functions with low average sensitivity depend on few coordinates, *Combinatorica* 18 (1998), 27–35.

[29] E. Friedgut, Sharp thresholds of graphs properties, and the k-sat problem, *Jour. Amer. Math. Soc.* 12 (1999), 1017–1054.

[30] E. Friedgut and G. Kalai, Every monotone graph property has a sharp threshold, *Proc. Amer. Math. Soc.* **124** (1996), 2993–3002.

[31] E. Friedgut, J. Kahn, G. Kalai, and N. Keller, Chvátal's conjecture and correlation inequalities. arXiv:1608.08954.

[32] C. Garban, G. Pete, and O. Schramm The Fourier spectrum of critical percolation *Acta Mathematica,* 205, (2010), 19–104.

[33] C. Garban, Oded Schramm's contributions to noise sensitivity *Annals of Probability,* 39, (2011), 1702–1767.

[34] C. Garban and J. E. Steiff, *Noise Sensitivity of Boolean Functions and Percolation,* Cambridge University Press, (2014).

[35] G. Grimmett, *Percolation*, Springer-Verlag, Berlin (1989).

[36] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.* 97 (1975) 1061–1083.

[37] L. Harper, A necessary condition on minimal cube numberings, *J. Appl. Prob.*, 4 (1967) 397–401.

[38] H. Hatami, A structure theorem for Boolean functions with small total influences, *Annals of Mathematics*, 176 (2012), 509–533. .

[39] A. J. W. Hilton and E. C. Milner, Some intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser. 2* 18 (1967), 369–384.

[40] J. Kahn, G. Kalai, and N. Linial, The influence of variables on Boolean functions, in *Proc. 29-th Annual Symposium on Foundations of Computer Science* (1988), 68–80.

[41] J. Kahn and G. Kalai, Thresholds and expectation thresholds, *Combin. Probab. Comput.* 16 (2007), 495–502.

[42] G. Kalai, Social indeterminacy, *Econometrica,* 72 (2004), 1565–1581.

[43] G. Kalai, The quantum compute puzzle, *Notices of the AMS* 63 (2016), 508–516. Expanded version in arXiv:1605.00992.

[44] G. Kalai, N. Keller, and E. Mossel, On the correlation of increasing families. To appear *J. Combin. Th. Ser. A.* arXiv:1511.04600.

[45] N. Keller, E. Mossel, and A. Sen, Geometric influences II: Correlation inequalities and noise sensitivity, *Ann. Inst. Henri Poincare 50* (2014), 1121–1139.

[46] H. Kesten, The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$, *Comm. Math. Phys.* 74 (1980), 41–59.

[47] H. Kesten, Scaling relations for $2D$-percolation, *Comm. Math. Phys.* 109 (1987), 109–156.

[48] H. Kesten, On the speed of convergence in first-passage percolation, *Ann. Appl. Probab.*, 3 (1993), 296–338.

[49] H. Kesten and Y. Zhang, Strict inequalites for some critical exponents in 2D-percolation, *J. Statist. Phys.* 46 (1987), 1031–1055.

[50] A. Khintchine, Über dyadische Brüche, *Math. Z.* 18 (1923), 109–116.

[51] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell, Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comp.* 37 (2007), 319–357.

[52] S. Kudekar, M. Mondelli, E. Sasoglu, and R. Urbanke, Reed–Muller codes achieve capacity on the binary erasure channel under MAP decoding, arXiv:1505.05831.

[53] S. Kumar and H. D. Pfister, Reed–Muller codes achieve capacity on erasure channels, arXiv:1505.05123.

[54] N. Linial, and R. Meshulam, Homological connectivity of random 2-dimensional complexes, *Combinatorica,* 26 (2006), 475–487.

[55] M. Ledoux, *The Concentration of Measure Phenomenon,* Mathematical Surveys and Monographs, 89, American Mathematical Society, Providence, RI (2001).

[56] N. Linial, Y. Mansour and N. Nisan, Constant depth circuits, Fourier transform, and learnability, *J. Assoc. Comput. Mach.* 40 (1993), 607–620.

[57] Y. Mansour, An $O(n^{\log\log n})$ learning algorithm for DNF under the uniform distribution. *J. Comput. System Sci.* 50 (1995), 543–550.

[58] G. Margulis. Probabilistic characteristics of graphs with large connectivity (in Russian), *Probl. Pered. Inform.* 10 (1974), 101–108.

[59] E. Mossel, R. O'Donnell and K. Oleszkiewicz, Noise stability of functions with low influence: Invariance and optimality, *Annals of Mathematics* 171 (2010), 295–341.

[60] E. Nelson, The free Markov field, *J. Functional Analysis* 12 (1973), 211–227.

[61] R. O'Donnell, *Analysis of Boolean Functions,* Cambridge University Press (2014).

[62] R. O'Donnell, M. Saks, O. Schramm, and R. Servedio, Every decision tree has an influential variable, FOCS 2005. arXiv:cs/0508071.

[63] R. Rossignol, Noise-stability and central limit theorems for effective resistanc e of random electric networks, *Annals of Probability* 44 (2016), 1053–1106.

[64] L. Russo, A note on percolation, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 43 (1978), 39–48.

[65] L. Russo, An approximate zero-one law, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 61 (1982), 129–139.

[66] O. Schramm and J. Steif, Quantitative noise sensitivity and exceptional times for percolation, *Annals of Mathematics* 171 (2010), 619–672.

[67] M. Talagrand, Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem, *Geom. and Funct. Anal.* 3 (1993), 295–314.

[68] M. Talagrand, On Russo's approximate zero-one law, *Annals of Probability*, 22 (1994), 1576–1587.

[69] M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces, *Publ. I.H.E.S.* 81 (1995), 73–205.

[70] M. Talagrand, How much are increasing sets positively correlated? *Combinatorica* 16 (1996), 243–258.

[71] M. Talagrand, On boundaries and influences, *Combinatorica* 17 (1997), 275–285.

[72] M. Talagrand, Are many small sets explicitely small, *STOC* 2009.

[73] B. Tsirelson and A. Vershik, Examples of nonlinear continuous tensor products of measure spaces and non-Fock factorizations, *Rev. Math. Phys.* 10 (1998), 81–145.

Gil Kalai, Einstein Institute of Mathematics, Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel
and Departments of Mathematics and Computer Science, Yale University, New Haven, CT
E-mail: kalai@math.huji.ac.il