Lecture 3 (Nov. 8, 2011)

We will find it more convenient to prove an equivalent formulation of Turán's theorem, where one replaces a graph by its *complement*, i.e.: the graph consisting of the same vertices and those edges missing from the original. We require a definition :

DEFINITION 9 : A collection of vertices in a graph are said to be *independent*, if no two amongst them are joined by an edge. The *independence number* of a graph G, denoted $\alpha(G)$, is the maximum size of an independent set of vertices in G.

The following is then equivalent to Theorem 13 :

Theorem 13' *Let $k \geq 3$ and n be a multiple of $k-1$. Then any graph G with n vertices and strictly fewer than $\binom{n}{2} - \frac{k-2}{k-1} \cdot \frac{n^2}{2} = \frac{n^2}{2(k-1)} - \frac{n}{2}$ edges satisfies $\alpha(G) \geq k$.*

Our proof of this will require three lemmas. The probabilistic component[1] is the first (and most interesting) one, for which we need some more terminology :

DEFINITION 10 : For a vertex v in a graph G, the vertices to which it is joined by an edge are called its *neighbours*. The number of its neighbours is called the *degree* of the vertex v, and is denoted d_v. Two neighbours in a graph are also said to be *adjacent*.

Lemma 14 *For any graph G we have that*

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d_v + 1}. \tag{23}$$

PROOF : Suppose G has n vertices. We consider the probability space (Ω, μ), where Ω is the collection of all possible orderings of the n vertices, hence $|\Omega| = n!$, and μ is uniform measure. For each vertex v, we let X_v be the indicator random variable of the event that v appears before all its neighbours in a randomly chosen ordering. Now since v and its neighbours form a collection of $d_v + 1$ vertices in all, and each of them is equally likely to appear first, it is clear that $\mathbb{E}[X_v] = \frac{1}{d_v+1}$. Let $X = \sum_v X_v$. By Proposition 9, $\mathbb{E}[X] = \sum_v \frac{1}{d_v+1}$. By Proposition 10(i), there is thus at least one ordering of the vertices, call it $\mathcal{O}$, such that $X(\mathcal{O}) \geq \sum_v \frac{1}{d_v+1}$. But now one just needs to observe that, in any ordering whatsoever, those vertices which appear before all their neighbours must form an independent set. This proves

[1] there are other ways to prove this theorem, the standard proof being a kind of double induction on k and n.

(23).

We will need one more simple general fact about graphs.

Lemma 15 *For any graph G we have that*

$$\#\textit{edges in } G = \frac{1}{2}\sum_{v\in V(G)} d_v. \tag{24}$$

PROOF : When we sum up the degress of the vertices, we are summing up the edges emanating from each vertex, and then each edge will be counted twice.

Finally, we need a third fact which is pure algebra/calculus :

Lemma 16 *Let $x_1, x_2, ..., x_n, t$ be positive real numbers. If*

$$x_1 + \cdots + x_n \leq t,$$

then

$$\frac{1}{x_1} + \cdots + \frac{1}{x_n} \geq \frac{n^2}{t},$$

with equality in the latter if and only if $x_1 = \cdots = x_n = t/n$.

PROOF : Equivalently, we need to prove that

$$\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} \frac{1}{x_i}\right) \geq n^2, \tag{25}$$

with equality if and only if $x_1 = \cdots = x_n$. Set

$$y_i := \sqrt{x_i}, \quad z_i := \frac{1}{\sqrt{x_i}},$$

and let $\mathbf{y}, \mathbf{z}$ be the vectors in $\mathbb{R}^n_+$ given by

$$\mathbf{y} := (y_1, ..., y_n), \quad \mathbf{z} := (z_1, ..., z_n).$$

Then the Cauchy-Schwarz inequality in $\mathbb{R}^n$ says that

$$||\mathbf{y}||_2^2 \cdot ||\mathbf{z}||_2^2 \geq | < \mathbf{y}, \mathbf{z} > |^2,$$

with equality if and only if $\mathbf{z} = \lambda\mathbf{y}$, for some $\lambda \in \mathbb{R}$. This is easily checked to be equivalent to the statement that (25) holds, with equality if and only if and only if the x_i are all equal.

PROOF OF THEOREM 13' : By Lemma 15, the assumption in the statement of the theorem about the number of edges in G can be written as

$$\frac{1}{2}\sum_{v} d_v < \frac{n^2}{2(k-1)} - \frac{n}{2},$$

which can be rewritten as

$$\sum_v (d_v + 1) < \frac{n^2}{k-1}.$$

Hence, by Lemma 16,

$$\sum_v \frac{1}{d_v + 1} > \frac{n^2}{\frac{n^2}{k-1}} = k - 1.$$

So, by Lemma 14, $\alpha(G) > k - 1$. But $\alpha(G)$ is an integer, thus $\alpha(G) \geq k$, Q.E.D.

Example 3 : Sum-free sets

DEFINITION 10 : Let $(G, +)$ be an abelian group, and A be a subset of G. The sumset $A + A$ is defined as

$$A + A := \{a_1 + a_2 : a_1, a_2 \in A\}. \tag{26}$$

DEFINITION 11 : A subset A of an abelian group $(G, +)$ is said to be *sum-free* if $A \cap (A + A) = \phi$, in other words, if there are no solutions in A to the equation $x = y + z$.

The abelian groups which are of most interest to number theorists are $\mathbb{Z}$ and the groups $\mathbb{Z}_p$, where p is a prime.

EXAMPLE A : Let $n \in \mathbb{N}$ and let A be a sum-free subset of $\{1, ..., n\}$. If a is the largest element of A, and

$$B := \{a - a_1 : a_1 \in A,\ a_1 \neq a\},$$

then A and B are disjoint subsets of $\{1, ..., n\}$. It follows that $|A| \leq \lceil n/2 \rceil$. There are essentially two different examples of a sum-free subset of this size, namely

$$A_1 = \{\text{odd numbers in } [1, n]\}, \quad A_2 = \left(\frac{n}{2}, n\right].$$

EXAMPLE B : Let p be a prime, say $p = 3k+i$, where $k \in \mathbb{N}_0$ and $i \in \{0, 1, 2\}$. If $i \in \{0, 1\}$, then $A := \{k + 1, ..., 2k\}$ is a sum-free set modulo p, whereas if $i = 2$, then $A := \{k + 1, ..., 2k + 1\}$ is sum-free modulo p. Thus, if $p \equiv 2 \pmod 3$, there exists a sum-free set A in $\mathbb{Z}_p$ such that $|A| = \frac{p+1}{3}$. This is best-possible, but a proof is not as simple as in Example A. It is a consequence of the so-called *Cauchy-Davenport theorem*, a special case of which satates that, if p is a prime and A is a subset of $\mathbb{Z}_p$, then

$$|A + A| \geq \min\{p, 2|A| - 1\}.$$

We will now apply a probabilistic argument to prove the following result, which apparently was first proven by Erdős in 1965 and rediscovered by Alon and Kleitman in 1990 :

Theorem 17 *Let S be any finite subset of $\mathbb{Z}$, not containing zero. Then there exists a sum-free subset A of S such that $|A| \geq \frac{|S|+1}{3}$.*

PROOF : Let S be given and choose a prime p satisfying the following two conditions :

(i) $p > \max_{s \in S} |s|$,
(ii) $p \equiv 2 \pmod 3$.

Dirichlet's theorem on the existence of primes in arithmetic progressions guarantees the existence of such a prime[2]. Say $p = 3k + 2$ and let $C := \{k+1, ..., 2k+1\}$. As noted in Example B above, the set C is sum-free modulo p. We shall work in the probability space (Ω, μ), where $\Omega = \{1, 2, ..., p-1\}$ and μ is uniform measure. For each $s \in S$ let $f_s : \Omega \to \Omega$ be the map given by

$$f_s : \omega \mapsto \omega s \pmod p.$$

The choice of p (property (i)) guarantees that each of the maps f_s is one-to-one. Let $X_s := \mathcal{X}_{f_s, C}$. Then, by (20), for every s we have

$$\mathbb{E}[X_s] = \frac{|C|}{p-1} > \frac{1}{3}.$$

Let $X = \sum_{s \in S} X_s$. By linearity of expectation,

$$\mathbb{E}[X] > \frac{|S|}{3}.$$

Thus, by Proposition 10(i), there exists some $\omega \in \Omega$ such that $X(\omega) > |S|/3$. But, unwinding the definitions, we see that

$$X(\omega) = \#\{s \in S : \omega s \pmod p \in C\}. \tag{27}$$

Let A be the subset of S on the right of (27). This is a sum-free subset of S, since a dilation of it lies, modulo p, entirely within C, and hence is sum-free. Since $|A| > |S|/3$ and $|A|$ is an integer, we must have $|A| \geq (|S| + 1)/3$, Q.E.D.

Remark 1 One can reformulate the above argument in non-probabilistic language, in which case it basically employs the well-known method in combinatorics of *counting pairs*. In the proof, we are basically counting in two different ways the ordered pairs (ω, s) which satisfy (i) $\omega \in \Omega$ (ii) $s \in S$ (iii) $\omega s \in C \pmod p$. I leave it as an exercise to the reader to fill out the details.

Remark 2 As shown in Example B, the set C employed in the above proof is a sum-free subset of $\mathbb{Z}_p$ of maximum size. Hence, it is natural to conjecture that Theorem 17 cannot be improved upon. It turns out that this is

[2]One can prove by much more simple means that there exist infinitely many primes congruent to 2 (mod 3). I'll leave it as an exercise.

not the case, but it seems to be non-trivial to show it. In a long and difficult paper, Bourgain [1] showed that, for any finite $S \subseteq \mathbb{Z}$, not containing zero, one can always find a sum-free subset A of S such that $|A| \geq \frac{|S|+2}{3}$. Nothing better than this is known, I think.

For upper bounds, it suffices to find examples of sets $S \leq \mathbb{N}$ without large sum-free subsets. I believe the current record is due to Lewko [2], who found, via computer search, a set of 28 positive integers with no sum-free subset of size 12. From such a single example, one can construct (I leave it as another exercise to determine how) arbitrarily large, finite sets $S \subseteq \mathbb{N}$ for which there are no sum-free subsets of size exceeding $\frac{11}{28}|S|$. The gap between 1/3 and 11/28 is a significant open problem.

Example 4 : Shannon's theorem on error-correcting codes

See the handout given in class from Alon-Spencer. Note that there are at least two error in their text. In the 10th line on the first page, the phrase "$3p^2 + p^3 = 0.031$" should read "$3p^2(1-p) + p^3 = 0.028$". In the fifth line of the proof of Theorem 14.1.1, the phrase

"x is the unique vector in $\{0,1\}^m$ within $n(p+\delta)$ of $f(x)$"

should read

"$f(x)$ is the unique vector in $f(\{0,1\}^m) \subseteq \{0,1\}^n$ within $n(p+\delta)$ of y".

References

[1] J. Bourgain, *Estimates related to sumfree subsets of sets of integers*, Israel J. Math. **97** (1997), no.1, 71–92.

[2] M. Lewko, *An improved upper bound for the sum-free subset constant*, J. Integer Seq. **13** (2010), no.8, Article 10.8.3, 15pp (electronic).

Lecture 4 (Nov. 10, 2011)

In the proof of Shannon's Theorem handed out in class, some things are glossed over pretty quickly. So here are some notes to help you read that proof.

DEFINITION 12 : The *entropy function*[3] $H : [0,1] \to [0,1]$ is defined as follows :

$$H(p) := \begin{cases} -p\log_2 p - (1-p)\log_2(1-p), & \text{if } p \notin \{0,1\}, \\ 0, & \text{if } p \in \{0,1\}. \end{cases} \tag{28}$$

It is easy to see that $H(p) = H(1-p)$, that $H(p)$ is continuous and strictly increasing for $p \in (0, 1/2)$ and that $H(1/2) = 1$.

We now give a formal definition of the binomial distribution.

DEFINITION 13 : Let $p \in [0,1]$. Let (Ω, μ) be the probability space given by $\Omega = \{0,1\}$, $\mu(\{1\}) = p$, $\mu(\{0\}) = 1-p$. Let χ_1 be the indicator of the event $\{1\}$, i.e.:

$$\chi_1(\omega) = \begin{cases} 1, & \text{if } \omega = 1, \\ 0, & \text{if } \omega = 0. \end{cases}$$

In words, χ_1 is the indicator of the event that a head is obtained when a biased coin, with a probability p of heads, is tossed once. Henceforth, such a coin will simply be called *p-biased* to save space.

Now let $n \in \mathbb{N}$. Let (Ω_n, μ_n) be the probability space where $\Omega_n = \{0,1\}^n$ and μ_n is defined by

$$\mu_n(\omega_1, ..., \omega_n) = \prod_{i=1}^{n} \mu(\omega_i).$$

In other words, if the string $\omega = (\omega_1, ..., \omega_n)$ contains k ones and $n-k$ zeroes, then $\mu_n(\omega) = p^k(1-p)^{n-k}$. The pair (Ω_n, μ_n) is called the *n-fold binomial distribution with parameter p* and is usually denoted $B(n,p)$.

When studying the binomial distribution, the basic random variable of interest is the variable $X = X^n$ defined by

$$X^n((\omega_1, ..., \omega_n)) = \#\{i : \omega_i = 1\}. \tag{29}$$

In words, X^n is the number of heads one obtains after tossing a p-biased coin a total of n times, assuming that the results of distinct tosses are independent of one another. It is convenient to write X as a sum of indicator variables $X^n = \sum_{i=1}^n X_i^n$, where

$$X_i^n((\omega_1, ..., \omega_n)) = \begin{cases} 1, & \text{if } \omega_i = 1, \\ 0, & \text{if } \omega_i = 0. \end{cases} \tag{30}$$

[3]It would be more correct to say that this is an example of an entropy function, but we do not wish to discuss the concept of entropy in any more general terms here.

It is natural to think of each X_i^n as being a random variable on $\Omega = \Omega_1$, since it records the result of just one coin toss, namely the i:th one, in which case it is identical to the variable χ_1 above[4]. One then speaks of X^n as being a sum of *independent, identically distributed (i.i.d.)* indicator variables. There is a comprehensive theory of such random variables, which we will get a glimpse of later in the course when we study Chernoff-type inequalities. The main point is that such variables are very *strongly concentrated* about their mean. Since, for each $i = 1, ..., n$ one clearly has $\mathbb{E}[X_i^n] = p$, by linearity of expectation we know that $\mathbb{E}[X^n] = np$. In other words, the expected number of heads after n tosses is just np. That X^n is "strongly concentrated" about its mean means, intuitively, that it is highly unlikely to attain values far from np - in other words, if you make n tosses, and n is large, then the number of heads recorded is very unlikely to stray too far from np. The following proposition makes this idea precise :

Proposition 18 *Let $p \in [0, 1]$. Then, for any fixed $\delta > 0$,*

$$\mathbb{P}(|X^n - np| \geq \delta n) \to 0, \quad \text{as } n \to \infty. \tag{31}$$

Indeed, the probability goes to zero exponentially as a function of n, for each fixed $\delta > 0$.

Proof Idea : This result falls out from a Chernoff-type inequality (see later), but it can also be proven "with one's bare hands", so to speak. One begins by noting that, for each $k \in \{0, ..., n\}$,

$$\mathbb{P}(X^n = k) = \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k}. \tag{32}$$

Let $f(k)$ denote the function on the right-hand side of (32). One easily checks that, for any $k < n$,

$$\frac{f(k+1)}{f(k)} = \left(\frac{n-k}{k+1} \right) \left(\frac{p}{1-p} \right). \tag{33}$$

From there it is easy to deduce that $f(k)$ attains a maximum at $k = np$, and that for any fixed $\delta > 0$, there exists $\delta' > 0$, depending on δ, such that if $n(p + \delta) \leq k \leq n$, then $\frac{f(k+1)}{f(k)} \leq 1 - \delta'$, whereas if $0 \leq k \leq n(p - \delta)$, then $\frac{f(k+1)}{f(k)} \geq 1 + \delta'$. From there, one can deduce the exponential decay with n of the probability in (31).

Finally, the proof of Shannon's Theorem invloves some estimates relating binomial coefficients to the entropy function :

[4]Formally, this involves what is called *conditioning*. We ignore a formal definition of this here.

Proposition 19 *(i) Let $\xi \in [0,1]$. Then*

$$\left(\begin{array}{c} n \\ \xi n \end{array} \right) = 2^{n(H(\xi)+o_n(1))}. \tag{34}$$

(ii) Moreover, if $\xi \leq 1/2$, then

$$\sum_{i=0}^{\xi n} \left(\begin{array}{c} n \\ i \end{array} \right) = 2^{n(H(\xi)+o_n(1))}. \tag{35}$$

PROOF : The first estimate follows from Stirling's approximation $n! \sim n^n e^{-n}\sqrt{2\pi n}$ and a little computation. The second estimate follows from the first and the observation that, if $\xi \in (0, 1/2]$, then

$$\left(\begin{array}{c} n \\ \xi n \end{array} \right) \leq \sum_{i=0}^{\xi n} \left(\begin{array}{c} n \\ i \end{array} \right) \leq (1+\xi n) \left(\begin{array}{c} n \\ \xi n \end{array} \right), \tag{36}$$

since the binomial coefficient $\left(\begin{array}{c} n \\ i \end{array} \right)$ increases with i. The point is then that the linear factor $1 + \xi n$ on the right of (36) does not affect the exponential estimate coming from (34) - it only changes the $o_n(1)$ function.

The following theorem was then proven in class :

Theorem 20 (Shannon 1948) *Let $p \in (0, 1/2)$ and $\epsilon > 0$. Then for all n sufficiently large, depending on ϵ, and for all $m < 1 - H(p) - \epsilon$, there exists an (m, n)-Coding Scheme such that, for transmission across a noisy channel, with noise parameter p, the probability of error is less than ϵ.*

PROOF : See the handout. In particular, see the handout or consult your own handwritten lecture notes if you are unsure of the meaning of any of the terminology in the statement of the theorem.

Local Coloring

Jeff started lecturing on this topic, and another sheet was handed out in class. You should refer to this and your own handwritten notes. The theorem that we started, but did not finish proving, was

Theorem 21 *For each $k \in \mathbb{N}$ there exists $\epsilon_k > 0$ such that, for all $n \gg_{\epsilon_k} 0$, there exist graphs G on n vertices with $\chi(G) > k$ and yet $\chi(G|_S) \leq 3$ for every set S of vertices of size at most $\epsilon_k n$.*

PROOF OF THEOREM : See handout. Note that the proof uses the following two lemmas, which are not explicitly proven on the handout :

Lemma 22 *Let $k \leq n$ be positive pintegers. Then*

$$\left(\begin{array}{c} n \\ k \end{array} \right) \leq \left(\frac{ne}{k}\right)^k. \tag{37}$$

PROOF : Clearly, $\left(\begin{array}{c} n \\ k \end{array} \right) \leq \frac{n^k}{k!}$, so it suffices to show that $e^k \geq \frac{k^k}{k!}$. But the right-hand side of this inequality is just the k:th term in the Taylor expansion of e^k.

Lemma 23 *Let G be a graph on n vertices and $k \in \mathbb{N}$. If $\chi(G) \leq k$ then $\alpha(G) \geq \lceil n/k \rceil$.*

PROOF : Consider a k-coloring of G. By the pigeonhole principle, at least one of the colors must be used on at least $\lceil n/k \rceil$ of the vertices. But the set of vertices with any prescribed color must form an independent set in G.