Random and Quasirandom Graphs | Instructor: Padraic Bartlett

Lecture 4: Examples of Quasirandom Graphs

Week 4 | *Mathcamp 2012*

In our last lecture, we did three things:

- We defined the concept of a quasirandom graph!
- We tried to start proving our theorem about quasirandom graphs.
- We quickly found out that this was far harder than expected.

This is not so much because of the difficulty of the material: if the ideas on day 2 made sense, then you're actually completely capable of following the remaining proofs in this equivalence! (Up to a bunch of linear algebra.) However, the problem with these proofs (as you saw yesterday) is the sheer volume of **notation** that crops up in the proofs: they're insanely dense, and following any more than one in a lecture is pretty much impossible, just because of all of the symbols.

So: we're not doing that! Instead, we're going to shift gears markedly: for the moment, believe that all of these quasirandom properties are equivalent. What can we **do** with them? In this lecture, we will study a family of quasirandom graphs. Specifically, we will prove that (1) this family is actually quasirandom, and (2) use quasirandomness to prove some statements that are not at all obvious about this family

First, we restate the quasirandomness definition, so that you have it on hand:

Definition. Let $\mathcal{G} = \{G_{k_n}\}_{n=1}^{\infty}$ be any sequence of graphs, each on k_n vertices, where the k_n's are an nondecreasing sequence that tends to infinity. We say that this sequence is **quasirandom** if, roughly speaking, it "looks like a random graph" in a number of quantifiable ways.

To make this rigorous, here's an additional bit of notation. Suppose that G, H are two graphs. Let $N_G^*(H)$ denote the number of labelled occurences of H as an induced subgraph of G. Similarly, let $N_G(H)$ denote the number of labeled occurences of H as a subgraph of G (not necessarily induced.)

We say that our sequence $\mathcal{G}$ is quasirandom if and only if its elements satisfy the following list of asymptotic properties, as the number of vertices in any such element G goes to infinity:

$P_1(s)$: For any graph H_s on s vertices,

$$N_G^*(H_s) = (1 + o(1)) \cdot n^s \cdot 2^{-\binom{s}{2}}.$$

$P_2(t)$: Let C_t denote the cycle of length t. Then

$$e(G) \geq (1 + o(1)) \cdot \frac{n^2}{4}, \text{ and}$$
$$N_G(C_t) \leq (1 + o(1)) \cdot \frac{n^t}{2^t}.$$

P_3: Let $A(G)$ denote the adjacency matrix of G, and $|\lambda_1| \geq \ldots \geq |\lambda_n|$ be the eigenvalues of $A(G)$. Then

$$e(G) \geq (1 + o(1)) \cdot \frac{n^2}{4}, \text{ and}$$
$$\lambda_1 = (1 + o(1)) \cdot \frac{n}{2}, \quad \lambda_2 = o(n).$$

P_4: Given any subset $S \subseteq V$,

$$e(S) = \frac{|S|^2}{4} + o(n^2).$$

P_5: Given any pair of vertices $v, v' \in G$, let $s(v, v')$ denote the number of vertices y such that both (v, y) and (v', y) are either both edges or both nonedges in G. Then

$$\sum_{v,v'} \left| s(v, v') - \frac{n}{2} \right| = o(n^3).$$

1 Paley Graphs

The Paley graphs are defined as follows:

Definition. Let p be an odd prime that is 1 mod 4. Define the Paley graph as follows:

- Our vertex set is $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.
- Connect two elements x, y with an edge $\{x, y\}$ if and only if $x - y$ is a **quadratic residue**: i.e. there is some element $a \in \mathbf{F}_p$ such that $x - y = a^2$.

The first thing that we claim about these Paley graphs is that they're well-defined: i.e. that we've actually made a graph! The only problematic part of the definition above was that we said that $\{x, y\}$ was an edge if and only if $x - y$ is a quadratic residue. The problem with this is that our definitions for edge are different depending on how we order these vertices: if we think of our edge as $\{x, y\}$, we're saying that this holds if and only if $x - y$ is a quadratic residue, while if we think of our edge as $\{y, x\}$, we're saying that this holds if and only if $y - x$ is a quadratic residue! These conditions, on the face of them, can be quite different! In particular, both of these objects are quadratic residues if and only if -1 is a quadratic residue: i.e. that there is some value $a \in \mathbb{F}_p$ such that $a^2 = -1$.

In some finite fields, this is not true: in $\mathbb{F}_7$, for example, there is no element that squares to -1. (Check this!) However, as it turns out, this **is** always true in $\mathbb{F}_p$, whenever p is a prime that's 1 mod 4.

We prove this as follows:

- Take $\mathbb{F}_p$, and remove 0; this is now a group with 0 mod 4 elements with respect to multiplication.

- Fun fact from group theory / that you could also probably prove directly: because this is a group with order divisible by 4, there is some element x such that $x^4 \equiv 1 \mod p$, but $x^3, x^2, x \not\equiv 1 \mod p$.

- Rewrite this as $x^4 - 1 \equiv 0 \mod p$. Factor this into the two polynomials $(x^2+1)(x^2-1) \equiv 0 \mod p$. Because $x^2 \neq 1$, this is a nonzero number, and we can cancel it out. This leaves $x^2 + 1 \equiv 0 \mod p$.

- But this is just precisely $x^2 \equiv -1 \mod p$! Which is what we were looking for.

We've proven that the Paley graphs are well-defined! We now claim it's quasirandom. To do this, we again use property P_5: this is because for graphs where we know the individual elements well but not the overall structure, this is like the only property we can get a good handle on.

To do this, take any two elements $x, y \in \mathbb{F}_p$. When is a third element z either adjacent to both of these elements or not adjacent to either element? Well, this happens when

$$z - x, z - y$$

are either both quadratic residues or both quadratic nonresidues.

Fun facts about quadratic residues:

- If you look at $\mathbb{F}_p$ and remove the zero, precisely half of the elements in this collection are quadratic residues. This is because if $x^2 = a$, we also have $(-x)^2 = a$, and furthermore if we ever have $x^2 \equiv y^2 \mod p$, we have $x^2 - y^2 \equiv (x+y)(x-y) \equiv 0 \mod p$. This holds if and only if one of $(x+y), (x-y)$ are 0, because we're in a field and therefore don't have zero divisors: i.e. if and only if $x = \pm y$.

- Also, notice that the product of any two quadratic residues is a quadratic residue: this is pretty immediate, because having $x^2 = a$ and $y^2 = b$ gives you $(xy)^2 = ab$.

- Similarly, the product of a quadratic residue and a quadratic nonresidue is a quadratic nonresidue. To see why, suppose that $a = x^2$ is the quadratic residue and b is the quadratic nonresidue. If $ab = y^2$ for some y, then we would have $b = \left(\frac{y}{x}\right)^2$ (which we can do because $\mathbb{F}_p$ is a field), which makes b a quadratic residue, a contradiction!

- Consider a multiplication table for $\mathbb{F}_p$ without 0. In every row and column, we have all of the elements of this group show up exactly once: therefore, in the entire table, all of the elements show up the same number of times! What does this mean, given the results we just came up with above?

  Well: if the product of a nonresidue and a residue is a nonresidue, while the product of a residue and a nonresidue is also a nonresidue, these products account for (on one hand) half of the elements in our table, but (on the other hand) **all** of the nonresidues, because exactly half of the elements are nonresidues!

  This means that any of the other products we make must be residues! In particular, this means that the product of any two nonresidues magically **must** become a residue!

If we apply this to our problem, this means that

$$z - x, z - y$$

are either both quadratic residues or both quadratic nonresidues if and only if

$$\frac{z-x}{z-y}$$

is a quadratic residue! (Because if exactly one of them was a residue and the other was not, then this ratio would fail to be a residue.)

But, for any of the quadratic residues a^2 other than 1, there is a unique z such that

$$\frac{z-x}{z-y} = 1 + \frac{y-x}{z-y} \equiv a^2 \mod p;$$

this is because this is a field, so we can multiply by $(z-y)$ and shuffle coefficients together to get a linear function in z

$$\begin{aligned} z - y + y - x &\equiv a^2(z-y) \mod p \\ \Leftrightarrow z(1-a^2) - (x - a^2 y) &\equiv 0 \mod p. \end{aligned}$$

which (again, because it's a field) has exactly one root when thought of as a function of z.

So: this means that there are precisely as many vertices z that are either both common neighbors and common non-neighbors to x, y as there are non-1 quadratic residues: i.e. $\frac{q-1}{2} - 1 = \frac{q-3}{2}$.

Consequently, this means that $s(x,y) = \frac{q-3}{2}$, and therefore that

$$\sum_{x,y \in \mathbb{F}_p} \left| s(x,y) - \frac{q}{2} \right| = \sum_{x,y \in \mathbb{F}_p} \frac{3}{2} = \frac{3n^2}{2},$$

which is definitely $o(n^3)$! Therefore, this sequence of graphs is quasirandom, as claimed, and therefore has all of the other quasirandom graph properties. In particular, this means that in any given subset of, say, $q/25$ of the vertices, we'd start seeing about half of those elements differing by a quadratic residue: this is because quasirandom graphs have roughly the right number of edges in any subset of their edges, and so in particular these subsets will eventually have to start having the right number of edges. This is not necessarily obvious: you might think that maybe you could pick a set of values such that their differences would mostly all be quadratic residues, or all be quadratic nonresidues! But our results on quasirandom graphs says that this is completely impossible.

Beautiful, right?