18.317 Combinatorics, Probability, and Computations on Groups	19 September 2001

Lecture 5

Lecturer: Igor Pak — *Scribe: Dennis Clark*

Proof of a Lemma

Our plan here is to prove the following lemma, needed in the fixed proof of Erdős-Turan's theorem:

Lemma 1 *Suppose* $A = \{1 \le a_1 < a_2 < \cdots < a_r \le n\}$. *Then:*

$$Pr(\sigma \in S_n \textit{has no cycles in } A) < \frac{1}{\sum_{i=1}^{r} a_i} \tag{1}$$

But we'll need a few sublemmas first.

Sublemma 2

$$E(\# \textit{ of } l\textit{-cycles in } \sigma \in S_n) = \frac{1}{l} \tag{2}$$

Proof:

$$E = \binom{n}{l} \frac{(l-1)!(n-l)!}{n!} = \frac{1}{l} \tag{3}$$

■

Sublemma 3 *Assume* $l \neq m$, *and* $l + m \le n$. *Then:*

$$E(\# \textit{ of } l\textit{-cycles} \cdot \# \textit{ of } m\textit{-cycles in } \sigma \in S_n) = \frac{1}{l\,m} \tag{4}$$

Proof: First, we take the sum in the first equation over the number of ways to split n into an l-subset, and m-subset, and an $n - m - l$ subset.

$$E = \sum Pr(\text{cycle at } l\text{-subset and cycle at } m\text{-subset}) \tag{5}$$

Then, recognize that all of the probabilities will be symmetric, so we can just multiply appropriately:

$$\begin{aligned} E &= \binom{n}{l, m, n-l-m} \cdot \frac{(l-1)!(m-1)!(n-m-l)!}{n!} && (6)\\ &= \frac{n!(l-1)!(m-1)!(n-m-l)!}{l!m!(n-m-l)!n!} && (7)\\ &= \frac{1}{lm} && (8) \end{aligned}$$

■

As a corollary, we get:

Corollary 4

$$E(\#\ of\ cycles\ in\ \sigma) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \log n + \gamma + O(\frac{1}{n}) \tag{9}$$

where γ is a constant.

Proof:

$$E = \sum_l E(\#\text{ of } l\text{-cycles}) = \sum \frac{1}{l} \tag{10}$$

■

Another sublemma, here departing from Erdős-Turan:

Sublemma 5

$$E\big((\#\ of\ l\text{-}cycles)^2\big) = \begin{cases} \frac{1}{l} & if\ 2l > n \\ \frac{1}{l} + \frac{1}{l^2} & if\ 2l \leq n \end{cases} \tag{11}$$

Proof:

To start, observe that

$$E(\#\text{ of } l\text{-cycles}) = \frac{1}{l}\sum_{i=1}^{n} Pr(i \in l\text{-cycle}) = \frac{1}{l} \tag{12}$$

from the proof of Sublemma 2. This gives:

Corollary 6 *$Pr(i \in l$-cycle$) = \frac{1}{n}$.*

Therefore the total number of cycles is approimately $\log n$. Similarly, Erdős-Turan obtain

$$\log(\prod \text{cycle length}) \approx \frac{1}{2}\log^2 n \tag{13}$$

So then we can observe that the square of the number of l-cycles is precisely the number of ordered pairs of elements belonging to l-cycles divided by l^2, which yields:

$$\begin{aligned} E &= \frac{1}{l^2}\sum_{i=1}^{n}\sum_{ij1}^{n} Pr(i \in l\text{-cycle and} j \in l\text{-cycle}) && (14) \\ &= \frac{1}{l^2}(\sum_{i=1}^{n} Pr(i \in l\text{-cycle}) + \sum_{i \neq j}^{n} Pr(i \in l\text{-cycle}) Pr(j \in l\text{-cycle}|i \in l\text{-cycle})) && (15) \end{aligned}$$

We consider the sums one at a time. First, the first sum is equal to:

$$\frac{1}{l^2} \cdot n \cdot \frac{1}{n} = \frac{1}{l^2} \tag{16}$$

from Corollary 6.

Let us consider now the second sum, which is equal to

$$\frac{1}{l^2} \cdot n \cdot \frac{1}{n} \cdot (n-1)\left(\frac{l-1}{n-1} + \frac{n-l}{n-1} \cdot \frac{1}{n-l}\right) \tag{17}$$

The justification for the fractions inside the parentheses is as follows: consider that the element i is already in an l-cycle. We want the probability that j is in an l-cycle. First, the probability that j is in the same l-cycle as i is simply $\frac{l-1}{n-1}$, since j can be in any of $n-1$ places, $l-1$ of which are what we're looking for. The second summand is the probability that j is among the remaining points but is in an l-cycle anyway, which can only happen if $l+l \leq n$. So we get the following:

$$E((\#\text{ of } l\text{-cycles})^2) = \begin{cases} \frac{1}{l^2} + \frac{1}{l^2} \cdot (l-1) & \text{if } 2l > n \\ \frac{1}{l^2} + \frac{1}{l^2} \cdot l & \text{if } 2l \leq n \end{cases} \tag{18}$$

Which then gives us:

$$E = \begin{cases} \frac{1}{l} & \text{if } 2l > n \\ \frac{1}{l} + \frac{1}{l^2} & \text{if } 2l \leq n \end{cases} \tag{19}$$

as needed. ■

Then, we return at last to the proof of Lemma 1:

Proof:

We first estimate the expected value of the number of A-cycles:

$$E(\#\text{ of } A\text{-cycles}) = \sum_{i=1}^{r} E(\#\text{ of } a_i\text{-cycles}) = \sum_{i=1}^{r} \frac{1}{a_i} \tag{20}$$

This gives:

$$E((\#\text{ of } A\text{-cycles})^2) = \sum_{i \neq j}^{r} E(\# a_i\text{-cycles} \cdot \# a_j \text{cycles}) + \sum_{i=1}^{r} E((\# a_i\text{-cycles})^2) \tag{21}$$

Applying the two main sublemmas to the two sums yields:

$$\leq 2 \sum_{1 \leq i < j \leq r} \frac{1}{a_i a_j} + \sum_i \left(\frac{1}{a_i^2} + \frac{1}{a_i} \right) = \sum_i \frac{1}{a_i} + \left(\sum_i \frac{1}{a_i} \right)^2 \tag{22}$$

Now, let X be a random variable describing the number of A-cycles. We have

$$Var(\#\text{ of } A\text{-cycles}) = Var(X) = E(X^2) - (E(X))^2 \tag{23}$$

$$= \left(\sum_i \frac{1}{a_i} + \left(\sum_i \frac{1}{a_i} \right)^2 \right) - \left(\sum_i \frac{1}{a_i} \right)^2 = \sum_i \frac{1}{a_i} \tag{24}$$

From here and Chebyshev inequality we get the following:

$$\begin{aligned} Pr(X=0) &= Pr(X + E(X) \leq E(X)) && (25) \\ &\leq \frac{Var(X)}{(E(X)^2)} && (26) \\ &\leq \frac{\sum \frac{1}{a_i}}{(\sum \frac{1}{a_i})^2} && (27) \\ &\leq \left(\sum_i \frac{1}{a_i} \right)^{-1} && (28) \end{aligned}$$

which proves the lemma. ■