Vacation Scholarship Experiences: Combinatorics and Graph Theory

David Wakeham – BSc

Introduction

During my vacation scholarship, I worked with Associate Professor David Wood on some problems in combinatorics and graph theory. In particular, we investigated the maximum density of a set of natural numbers subject to certain multiplicative constraints.

This involved elementary number theory, some graph theory, computer programming, and techniques from "experimental" mathematics.

Basic Concepts

Take a set R of rational numbers greater than 1. A set of natural numbers X is called *R-multiplicative* if, for any two elements $x, y \in X$ with $x > y$, the quotient $x/y \notin R$. We think of R as a set of forbidden quotients.

Define $[n] := \{1, \ldots, n\}$. Given R, we want to see how big we can make an R-multiplicative set X in the asymptotic limit. We measure the size of $X \subseteq \mathbb{N}$ using *density*, which is defined by

$$\delta(X) := \lim_{n\to\infty} \frac{|X \cap [n]|}{n}.$$

We are interested in the related limit

$$\delta_R := \lim_{n\to\infty} \frac{\max_{X\subseteq[n]} |X|}{n}$$

where X is R-multiplicative. This is called the *maximum density* of an R-multiplicative set.

For example, if $R = \{2\}$, then X is R-multiplicative if $2x \notin X$ for every $x \in X$. In this case, X is called *double-free*, and $\delta_R = 2/3$. The *even subpowers of 2*

$$E = \{2^{2i} a \mid i \in \mathbb{N} \cup \{0\}, a \in \mathbb{N}, 2 \nmid a\}$$

form a double-free set, and $\delta(E) = \delta_R$. Generally, δ_R is known when $|R| = 1$.

Approach

Our strategy is to model the problem as a graph $G = (\mathbb{N}, E)$, where the vertex set is the set of natural numbers and edges join elements with forbidden quotients; that is, $xy \in E$ if and only if $x/y \in R$.

It follows that R-multiplicative sets are independent sets in G, and

$$\delta_R = \lim_{n\to\infty} \frac{\alpha(G_n)}{n},$$

where G_n is the subgraph of G induced by $[n]$ and $\alpha(G_n)$ is the size of a maximum independent set (MIS) in G_n.

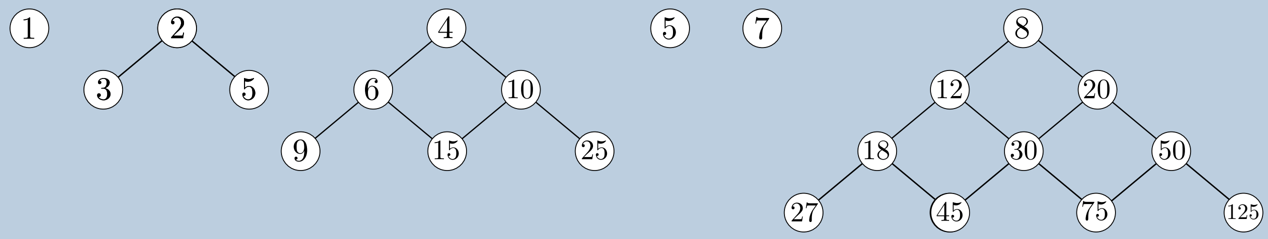
Results

Our main result is as follows. Let a, b, c be pairwise coprime, with $1 < a < b < c$, and

$$R = \left\{\frac{b}{a}, \frac{c}{a}\right\}.$$

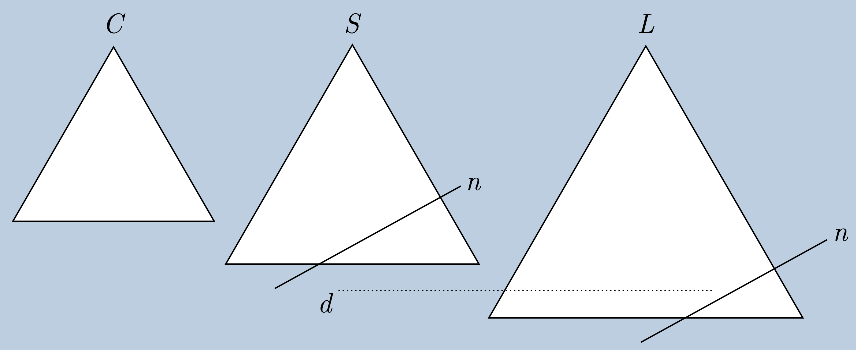
For all fixed $\epsilon > 0$, there is a constant-time algorithm which computes δ_R to within ϵ.

When $a = 2$, $b = 3$, and $c = 5$, the first few components of G look like this:



In general, the components of G for arbitrary a, b, c are finite triangular lattices like these. We call the number of rows in a component its *height*. Components of G_n are lattices "cut off" by straight lines whose gradient depends on a, b, and c, and whose y-intercept depends on n.

Pick a fixed maximum height d. The main result is proved by splitting the components of G_n into three types: complete components (C), "small" incomplete components (S) with height less than d when complete, and "large" incomplete components (L) with height greater than d. These are pictured below:



Using graph-theory, we can calculate the density of an MIS in complete components analytically, and the density of an MIS in small incomplete components, for arbitrary d, using a computer. Finally, we proved that d can be chosen to make the density of an MIS in large incomplete components arbitrarily small. This means we can calculate δ_R to arbitrary accuracy. We suspect that δ_R has no closed-form expression and have some "experimental" evidence to support this.

We wrote a Python script to do the computations. Here are some approximate values for small a, b, c:

a	b	c	δ_R
2	3	5	0.729
2	3	7	0.740
3	4	5	0.790
2	5	7	0.823
3	5	7	0.823

Unfortunately, the program is very slow for higher precisions. It can calculate two decimal places of accuracy in under a minute, but takes hours to calculate four!

We are looking at higher-dimensional generalisations of this result, other situations when $|R| = 2$, and even cases where R is infinite. We hope to publish our findings in the near future.

Experience

I found my vacation work immensely rewarding. I learned analytic techniques from graph theory, asymptotic analysis, and combinatorics, although computational methods also played a central role. Collaborating with David Wood gave me valuable insights into the life of a mathematician and the research process.

For students considering a career in research mathematics, I would unhesitatingly recommend the program.