Finite Versions of Szemerédi's Theorem on Arithmetical Progressions

Paul RAFF[1] *and Doron ZEILBERGER*[1]

Introduction

One of the crowning achievements of combinatorics and number theory ([S]) is

Szemerédi's Theorem: For any positive integer k, and any positive integer n, let $S_k(n)$ be the size of the largest subset of $[1, n]$ that *avoids* arithmetical progressions of size k. Then

$$\lim_{n\to\infty} \frac{S_k(n)}{n} = 0 \quad .$$

The exact rate of convergence is not known, but quite a few Fields-medalists (Bourgain, Gowers, Tao) are trying to make it quantitative, but their results are (probably) far from sharp.

In this article, we offer yet-another-approach, that we believe does have the potential, eventually, to tackle the "deep questions", and at any rate, is of independent interest. Let's study "*finite*" versions of Szemerédi's theorem.

Fix k (the size of the arithmetic progression) and d, a positive integer, and let $f_{k,d}(n)$ be the cardinality of the largest subset of $\{1, 2, \dots, n\}$ that avoids arithmetic progressions of size k whose *difference* is $\le d$. In other words $f_{k,d}(n)$ is the size of largest subset of $\{1, 2, \dots, n\}$, that does not contain any subset of the form

$$\{a, a + j, a + 2j, \dots, a + (k - 1)j\}$$

and $j \le d$. We will see below that $f_{k,d}(n)$ is eventually periodic, in the sense that there exist a period $P = P_{k,d}$ and an integer $Q = Q_{k,d}$ such that, starting at a $N_{k,d}$

$$f_{k,d}(Pm + i) = Qm + \alpha_i \quad , \quad i = 1 \dots P$$

with certain α_i with $0 \le \alpha_i \le Q$. We will present fully implemented algorithms to conjecture $P_{k,d}$, $Q_{k,d}$, and the α_i's, and then prove that the conjecture is true (completely automatically). In particular, we would have

$$C_{k,d} := \lim_{n\to\infty} \frac{f_{k,d}(n)}{n} = \frac{Q_{k,d}}{P_{k,d}} \quad .$$

Another way of stating Szemerédi's theorem is

$$\lim_{d\to\infty} C_{k,d} = 0 \quad .$$

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `[praff,zeilberg] at math dot rutgers dot edu` ,
`http://www.math.rutgers.edu/~[praff,zeilberg]` . First version: June 3, 2009. Accompanied by Maple package `ENDRE` and Mathematica package downloadable from the authors' websites. Supported in part by the NSF.

Hopefully more extensive computations would give, for specific $k = 3, 4, \ldots$ (and perhaps eventually for general k) a *trend*, that may give experimental insight about the rate of convergence of $C_{k,d}$. We also believe that our approach may suggest a novel rigorous approach that may eventually yield rigorous lower and upper bounds on $C_{k,d}$, by studying, in general terms, the appropriate system of equations defined below, that determine $P_{k,d}$, $Q_{k,d}$, and hence $C_{k,d}$. This, in turn, may enable to improve the current lower and upper bounds for the behaviour of $S_k(n)$ itself.

References

[S] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27** 199-245.