Report on “Finite Analogs of Szemerédi’s Theorem” by Paul Raff and Doron Zeilberger

October 6, 2009

Some general comments

This paper is concerned with the calculation of the values of certain constants $\alpha_{k,D}$. The authors view these constants as being closely related to Szemerédi’s theorem, and speculate (they say ”dream”) that a careful study of these constants and related quantities might eventually shed additional light on Szemerédi’s theorem, and even give improved lower bounds on the numbers $r_k(n)$, where $r_k(n)$ denotes the size of any largest subset S of $[n] := \{1, 2, \dots, n\}$ such that S contains no k-term arithmetic progression.

The constants $\alpha_{k,D}$ are defined in the following way.

Given an integer $n \geq 1$ and integers $k \geq 3, D \geq 1$, let $R_{k,D}(n)$ denote the size of any largest subset S of $[n] := \{1, 2, \dots, n\}$ such that S contains no subsets of the form

$$\{i, i+d, i+2d, \cdots, i+(k-1)d\},\ (i \geq 1,\ 1 \leq d \leq D).$$

Now, by the definition of $R_{k,D}(n)$, for each fixed k and D, $R_{k,D}(n)$ is obviously a “sub-additive” function of n. This just means that $R_{k,D}(n_1 + n_2) \leq R_{k,D}(n_1) + R_{k,D}(n_2)$ for all n_1, n_2. (The authors don’t mention this.)

It is well-known (see, for example, p. 47 in Graham-Rothschild-Spencer’s Ramsey Theory) that, as for any sub-additive function,

$$\lim_{n\to\infty} \frac{R_{k,D}(n)}{n} = \alpha_{k,D}$$

exists, and furthermore for all $n \geq 1$,

$$\frac{R_{k,D}(n)}{n} \geq \alpha_{k,D}.$$

Early in their paper, the authors state what they call the "Finite version of Szemerédi's theorem":

"Given an integer $n \geq 1$ and integers $k \geq 3, D \geq 1$, let $R_{k,D}(n)$ denote the size of any largest subset S of $[n] := \{1, 2, \ldots, n\}$ for which there are **no** subsets of the form

$$\{i, i+d, i+2d, \cdots, i+(k-1)d\}, \ (i \geq 1, \ 1 \leq d \leq D),$$

Then there exists a rational number $\alpha_{k,D}$ such that

$$\lim_{n \to \infty} \frac{R_{k,D}(n)}{n} = \alpha_{k,D}."$$

Now, I have two comments to make about this statement. First, as remarked above, the sub-additivity of $R_{k,D}(n)$ immediately gives the existence of the numbers $\alpha_{k,D}$. Thus, the only part of their statement of the "Finite version of Szemerédi's theorem" which has any substance, is that the numbers $\alpha_{k,D}$ are rational. But - here is my second comment - surely Szemerédi's theorem is not concerned with the rationality or irrationality of the numbers $\alpha_{k,D}$. Indeed, Szemerédi's Theorem is exactly the statement that $\lim_{D \to \infty} \alpha_{k,D} = 0$, for every $k \geq 1$. (As the authors do remark, later on.)

Thus, it seems to me that this theorem has been improperly named, and it should be made clear that all that is being asserted is the rationality of the numbers $\alpha_{k,D}$. This fact is not without interest, although it is certainly a stretch to assert that it is connected to Szemerédi's theorem. It is proved later on, on pp. 4 and 5. (I have to admit I could not understand this proof, which purportedly proves an additional fact as well. The additional fact is stated on p. 2.)

There is no doubt that the values of the constants $\alpha_{k,D}$ are of interest. The method of calculating these values seems very interesting. To me it is amazing and impressive that the method the authors use actually works, and the method itself is very ingenious.

I would think that calculations of the corresponding numbers for van der Waerden's theorem would be of equal or even greater interest.

Suppose one had at one's disposal a really big table of values of the constants $\alpha_{k,D}$. Besides looking at how fast each column converges to zero, it might also be of interest to examine how fast each column becomes less than, for example, $\frac{1}{2}$. After all, in the paper of Behrend (F. A. Behrend, On sequences of integers containing no arithmetic progression, Časopis Mat. Fys. Praha (Čast Mat.) 67 (1938), 235-239) it is shown that Szemerédi's theorem is *equivalent* to the statement: "For every k there exists $N(k)$ such that if A is any subset of $\{1, 2, \ldots, N(k)\}$ with $|A| > \frac{1}{2}N(k)$, then A contains a k-term arithmetic progression." (In fact, as Behrend shows, the number $\frac{1}{2}$ here can be replaced by any constant less than 1.)

Some specific comments

1. The abstract is missing. (The proof-reader didn't notice there was no abstract? But Zeilberger has at least 5 previous articles in INTEGERS.)

2. The paper seems to have been written in great haste, and proof-read in great haste. For example, in the table on p. 2, which after all gives the values of some of the constants $\alpha_{k,D}$, which are the whole point of the paper, there is an error in the first column. According to the definitions, $\alpha_{k,D}$ is a non-increasing function of D, for each fixed k. But in the first column, $D_{3,12} = \frac{56}{177} < \frac{8}{19} = D_{3,13}$. Perhaps the denominator is correct, and $\frac{8}{19}$ should be $\frac{6}{19}$. But then, shouldn't this typo have been noticed? I'm assuming it is a typo, and not something more serious.

3. p. 1, line 14. "...we can only talk about *finite* analogs,..." [of Szemerédi's theorem] I don't see anywhere in the paper an analog of Szemerédi's theorem.

4. p. 2, first paragraph. The sentence "$R_{k,D}(n)$ is a quasi-linear function of n, ... " is very unclear. What is "quasi-linear?" (Is it implicitly defined here? If so, please say so. If not, please say what the term means.) Is the displayed equation valid for all n?

5. p. 2, line -5. What does "in the sense of the Bible Codes" mean? Some readers will "get it," some readers won't have a clue. Both classes will probably be annoyed. I suppose this phrase is meant to be picturesque, but why not just omit it?

6a. p. 3, first or second paragraph. Please do the reader the courtesy of reminding him or her what the weight-enumerator is.

6b. p. 3, last paragraph. The whole paragraph is unclear. Is $S[P, \emptyset]$

meant to be the set of words not starting with the SPACE symbol? This is a crucial paragraph, and should be written clearly.

7. p. 4, displayed equation. I suppose that x_a stands for what was previously called $x[a]$. Why change the notation arbitrarily?

8. p. 4, last paragraph. Why are there as many equations as there are unknowns? This discussion is too abbreviated. What are the unknowns, anyway? Another crucial point, which is left unclear.

9. p. 5, second paragraph (starting with "Expanding $R(z, t)$ ... "). I could not understand this paragraph.

10. the section "Pipe dreams." This section is vague and mysterious. The statement "Regardless, hopefully we can get some general theorems, ..." seems pretty meaningless. This section should be re-written, and made more concrete. (Or else just omitted entirely.)

11. p. 7, first reference. At the time of first reading, a couple of weeks ago, the given URL did not work. Now, it does work. That is, it works provided the missing symbol tilde (˜) is inserted. Here's another case of careless proof reading.

Some more general comments

It is interesting to approach Szemerédi's theorem via the constants $\alpha_{k,D}$. In general, there should be more clarity in the paper, and it should be more carefully written. In its present form it is certainly not acceptable for INTEGERS.

I agree with the authors that the description of the algorithm should be mostly hand-waving. The idea behind the algorithm is ingenious, and is described fairly well, although the description should be clarified. In other parts of the paper, particularly in the proof that the numbers $\alpha_{k,D}$ are rational, there is too much hand-waving.

I think readers of INTEGERS would like to see the similar results on the corresponding van der Waerden numbers.