Square-Difference-Free Sets of Size $\Omega(n^{0.7334\cdots})$

Richard Beigel *
Temple University

William Gasarch †
Univ. of MD at College Park

Abstract

A set $A \subseteq \mathsf{N}$ is *square-difference free* (henceforth SDF) if there do not exist $x, y \in A$, $x \neq y$, such that $|x - y|$ is a square. Let $\mathrm{sdf}(n)$ be the size of the largest SDF subset of $\{1, \ldots, n\}$. Ruzsa [10] has shown that proved $\mathrm{sdf}(n) \geq \Omega(n^{\log_{65} 7}) \geq \Omega(n^{0.733077\cdots})$.

$$\mathrm{sdf}(n) = \Omega(n^{0.5(1+\log_{65} 7)}) = \Omega(n^{0.733077\cdots})$$

We improve on the lower bound by showing

$$\mathrm{sdf}(n) = \Omega(n^{0.5(1+\log_{205} 12)}) = \Omega(n^{0.7334\cdots})$$

As a corollary we obtain a new lower bound on the quadratic van der Waerden numbers. We also give the context and history of results of this type.

1 Introduction

Notation 1.1 N is the set $\{1, 2, 3, \ldots\}$. If $n \in \mathsf{N}$, and $n \geq 1$ then $[n] = \{1, \ldots, n\}$.

In 1927 van der Waerden [18] published the following theorem which is now known as van der Waerden's theorem (henceforth VDW's theorem). For a modern treatment see the printed books by Graham et al. [6] or Landman et al. [8]; or the free online book by Gasarch et al. [4].

*Temple University, Dept. of Computer and Information Sciences, 1805 N Broad St Fl 3, Philadelphia, PA 19122. professorB@gmaıl.com

†University of Maryland, College Park, MD 20742. gasarch@cs.umd.edu, Partially supported by NSF grant CCR-01-05413

Theorem 1.2 *For all $k, c \in \mathsf{N}$, there exists $W = W(k,c) \in \mathsf{N}$ such that, for any c-coloring $COL : [W] \to [c]$, there are $a, d \in \mathsf{N}$, $d \neq 0$, such that*

$$a, a+d, a+2d, \ldots, a+(k-1)d \in [W]$$

$$COL(a) = COL(a+d) = COL(a+2d) = \cdots = COL(a+(k-1)d).$$

Van der Waerden's original proof yielded enormous upper bounds on $W(k,c)$. In particular, they were not primitive recursive. Erdos and Turan wanted an alternative proof of this theorem with smaller bounds. To this end they made the following conjectures:

- (ER1) For all k, for all ϵ, for large enough n, for all $A \subseteq \{1, ..., n\}$ such that $|A| \geq \epsilon n$, A has a k-AP.
- (ER2) Let A be a set of natural numbers. If $\sum_{x \in A} 1/x$ diverges then A has arbitrarily long arithmetic sequences. (We do not discuss ER2 but include it for completeness.)

Szemeredi [15] (see also the expositions by Tao[16, 17]) proved ER1; however, the proof used VDW's theorem and hence did not provide better bounds for $W(k,c)$. Furstenberg [3] obtained a different proof of ER1 using ergodic theory. This proof was nonconstructive and hence yielded no bounds for $W(k,c)$. Shelah [14] obtained primitive recursive bounds using purely combinatorial methods. Gowers [5] obtained, using rather difficult mathematics, the best known bound:

$$W(k,c) \leq 2^{2^{c2^{2^{k+9}}}}$$

For now this is where the story of upper bounds on $W(k,c)$ ends. However, the techniques of Furstenberg were then used by Bergelson and Leibman [1] to show the following which we refer to as *the density version of the Polynomial van der Waerden Theorem (henceforth Density PVDW Theorem.)*

Theorem 1.3 *For all $0 < \epsilon < 1$, for all $p_1(x), \ldots, p_k(x) \in Z[x]$ such that $(\forall i)[p_i(0) = 0]$, for almost all n, the following holds:*

$$(\forall A \subseteq [n])[|A| \geq \epsilon n \Rightarrow (\exists a, d \in \mathsf{N})[a, a+p_1(d), a+p_2(d), \ldots, a+p_k(d) \in A]].$$

This theorem has the following corollary which we refer to as *the polynomial van der Waerden Theorem (henceforth PVDW theorem.)*

Theorem 1.4 *For all* $c \in \mathsf{N}$*, for all* $p_1(x), \ldots, p_k(x) \in Z[x]$ *such that* $(\forall i)[p_i(0) = 0]$*, there exists a natural number* $W = W(p_1, \ldots, p_k; c)$ *such that, for all* c*-coloring* $COL : [W] \to [c]$*, there exists* $a, d \in \mathsf{N}$*,* $d \neq 0$*, such that*

$$a, a + p_1(d), a + p_2(d), \ldots, a + p_k(d)) \in [W]$$

$$COL(a) = COL(a + p_1(d)) = COL(a + p_2(d)) = \cdots = COL(a + p_k(d)).$$

Note 1.5

1. The PVDW theorem was proved for $k = 1$ by Furstenberg [3] and (independently) Sárközy [11].

2. Bergelson and Leibman's proof of the Density PVDW Theorem yields the PVDW theorem; however, it does not provide bounds on $W(p_1, \ldots, p_k; c)$. Walters [19] proved the PVDW theorem using purely combinatorial techniques and hence obtained bounds; however, these bounds were not primitive recursive.

3. The original proof of VDW's theorem used an ω^2 induction which is why the bounds are so large. By contrast Walters proof of the PVDW theorem used an ω^ω induction and hence yields much larger bounds.

4. Shelah [13] later obtained primitive recursive (though still large) bounds on $W(p_1, \ldots, p_k; c)$.

5. Note that in the case of VDW's theorem the combinatorial proof came first and the density version (Szemeredi's theorem) came later, while for the PVDW theorem the density version came first and the combinatorial proof came later.

Our interest is in a special case of PVDW and density PVDW.

Def 1.6 A set $A \subseteq \mathsf{N}$ is *square-difference free* (henceforth SDF) if there do not exist $x, y \in A$, $x \neq y$, such that $|x - y|$ is a square.

Def 1.7 Let $\mathrm{sdf}(n)$ be the size of the largest SDF subset of $[n]$.

Theorem 1.3 implies that, for any $0 < \epsilon < 1$, for almost all n,

$$\mathrm{sdf}(n) \leq \epsilon n.$$

The following bounds are known on $\mathrm{sdf}(n)$.

- Sárközy [11] proved

$$\mathrm{sdf}(n) \leq O\Big(\frac{n(\log\log n)^{2/3}}{(\log n)^{1/3}}\Big).$$

- Pintz, Steiger, and Szemerédi [9] proved

$$\mathrm{sdf}(n) \leq \frac{n}{(\log n)^{O(\log\log\log\log n)}}.$$

  (See also an exposition of Wolf [20].)

- Sárközy [12] showed that, for all $\epsilon < 0.5$, $\mathrm{sdf}(n) \geq n^{0.5+\epsilon f(n)}$, where $f(n) = \frac{\log\log\log n}{\log\log n}$.

- Ruzsa [10] proved $\mathrm{sdf}(n) \geq \Omega(n^{\log_{65} 7}) \geq \Omega(n^{0.733077\cdots})$.

We improve on Ruzsa's result by showing

$$\mathrm{sdf}(n) \geq \Omega(n^{\log_{205} 12}) = \Omega(n^{0.7334\cdots}).$$

Our proof is similar to Ruzsa's. We then use the result to get a lower bound on the quadratic VDW number (also know as $W(x^2; c)$ from Theorem 1.4.)

2 Upper Bounds on the Quadratic VDW Number

We present a known upper bound on the Quadratic VDW number so that we can contrast it to the new lower bound we will obtain.

Notation 2.1 Let $Q(c)$ be the least n such that for all c-colorings of $[n]$ there exists a, d, such that a and $a + d^2$ are the same color. Note that $Q(c)$ exists by Theorem 1.4.

The following is an easy corollary of the result of Pintz, Steiger, and Szemerédi mentioned above.

Corollary 2.2 *For all* $0 < \epsilon < 1$, $Q(c) \leq 2^{c^{f(c,\epsilon)}}$ *where* $f(c,\epsilon) = O(\frac{1}{(\log\log\log c)^\epsilon})$.

Proof:

Let $n = 2^{c^{f(c,\epsilon)}}$. Let COL be a c-coloring of $[n]$. Some color, say RED, must appear at least n/c times. We show that

$$\frac{n}{c} \geq \frac{n}{(\log n)^{O(\log\log\log\log n)}}.$$

By the result of Pintz, Steiger, and Szemerédi mentioned above this will imply that there are two RED points that are a square apart.

$$\begin{aligned}
\tfrac{n}{c} &\geq \tfrac{n}{(\log n)^{O(\log\log\log\log n)}} \\
(\log n)^{\Omega((\log\log\log\log n)} &\geq c \\
c^{\Omega(f(c,\epsilon)\log\log\log c^{f(c,\epsilon)})} &\geq c^1 \\
\Omega(f(c,\epsilon)\log\log\log c^{f(c,\epsilon)}) &\geq 1 \\
\Omega(f(c,\epsilon)\log\log(f(c,\epsilon)\log c)) &\geq 1 \\
\Omega(f(c,\epsilon)\log(\log(f(c,\epsilon)) + \log(\log c))) &\geq 1 \\
\Omega(f(c,\epsilon)\log(\log(\log c))) &\geq 1 \\
\Omega((\log\log\log(c)))^{1-\epsilon}) &\geq 1
\end{aligned}$$

For large enough c this inequality will hold. ∎

3 An SDF set of size $\geq \Omega(n^{0.5})$

We present the result $\mathrm{sdf}(n) \geq n^{0.5}$, since it is easy and, while known [12], is not online and seems hard to find. We do not need this result; however, it is very nice.

Recall Bertrand's Postulate[1] which we state as a lemma.

Lemma 3.1 *For all* n *there is a prime* p *such that* $n \leq p \leq 2n$.

[1]Bertrand's Postulate was actually proven by Chebyshev's. Bertrand conjectured that, for all $n > 3$, there is a prime between n and $2n-2$. Bertrand proved it for all $n < 3 \times 10^6$. Chebyshev proved it completely in 1850. It is usually stated as we do below. A proof due to Erdös can be found either in the Classic Number Theory text of Hardy and Wright [7] or on the Wikipedia entry on Bertrand's Postulate.

Theorem 3.2

$$\text{sdf}(n) = \Omega(n^{0.5}).$$

Proof: By Bertrand's Postulate there exists a prime p such that

$$\frac{n^{0.5}}{2} \leq p \leq n^{0.5}.$$

Let

$$A = \{p, 2p, 3p, \ldots, p^2\}.$$

Clearly, $|A| = p \geq \Omega(\sqrt{n})$. We show that A is SDF.
Let ip and jp be two elements of A. Note that

$$jp - ip = (j - i)p.$$

We can assume that $i < j$, so

$$1 \leq i < j \leq p.$$

Thus we have $j - i < p$. Hence $(j - i)p$ has only one factor of p, so $jp - ip$ cannot be a square. ∎

4 An SDF set of size $\geq \Omega(n^{0.7334\cdots})$

To obtain large SDF sets, we will first work with SDF sets with respect to various moduli.

Convention 4.1 Throughout this section when we deal with mod m we will use the set $[m] = \{1, \ldots, m\}$ rather than the more traditional $\{0, \ldots, m-1\}$. In calculations we may use 0 instead of m for clarity. For example, if we have that $b_1 \equiv b_2 \pmod{m}$ then we will feel free to write $b_1 - b_2 \equiv 0 \pmod{m}$.

Def 4.2 Let $n \in \mathsf{N}$. A set $A \subseteq [n]$ is *square-difference free mod n* (henceforth SDFMOD(n)) if there do not exist $x, y \in A$, $x \neq y$, such that $x - y$ is a square mod n.

Def 4.3 Let sdfmod(n) be the size of the largest SDFMOD(n) set.

Note that sdfmod(n) $\leq$ sdf(n). We will obtain lower bounds for sdf(n) by obtaining lower bounds for sdfmod(n). The next lemma shows how to construct such sets.

Lemma 4.4 *Assume m is squarefree, $k \geq 1$, and B, X, S, Y are sets such that the following hold:*

1. *S is an* SDFMOD(m) *subset of* $[m]$.
2. *X is an* SDFMOD(m^{2k-2}) *subset of* $[m^{2k-2}]$.
3. *$B = \{mz + b \mid z \in \{0, \ldots, m-1\} \wedge b \in S\}$. Note that $B \subseteq [m^2]$ and $|B| = m|S|$.*
4. *$Y = \{m^2x + s \pmod{m^{2k}} \mid x \in X \wedge s \in B\}$. Since Y is defined* (mod m^{2k})*, when we use Convention 4.1, we have $Y \subseteq [m^{2k}]$. Note that $|Y| = |X||B| = m|S||X|$.*

Then Y is an SDFMOD(m^{2k}) *subset of* $[m^{2k}]$.

Proof: Suppose, by way of contradiction, that there exist two elements of Y, y_1 and y_2, whose difference is a square mod m^{2k}. By the definition of Y, we can write those elements as

- $y_1 = m^2x_1 + s_1$, where $x_1 \in X$ and $s_1 \in B$.
- $y_2 = m^2x_2 + s_2$, where $x_2 \in X$ and $s_2 \in B$.

Since $s_1, s_2 \in B$,

- $s_1 = mz_1 + b_1$, where $z_1 \in \{0, \ldots, m-1\}$ and $b_1 \in S$.
- $s_2 = mz_2 + b_2$, where $z_2 \in \{0, \ldots, m-1\}$ and $b_2 \in S$.

Hence

- $y_1 = m^2x_1 + mz_1 + b_1$, where $x_1 \in X$, $z_1 \in \{0, \ldots, m-1\}$, and $b_1 \in S$.
- $y_2 = m^2x_2 + mz_2 + b_2$, where $x_2 \in X$, $z_2 \in \{0, \ldots, m-1\}$, and $b_2 \in S$.

Since $y_1 - y_2$ is a square mod m^{2k} there exists a, L such that

$$y_1 - y_2 = a^2 + L_1 m^{2k}$$

$$m^2(x_1 - x_2) + m(z_1 - z_2) + (b_1 - b_2) = a^2 + Lm^{2k}.$$

Reducing this equation mod m, we obtain

$$b_1 - b_2 \equiv a^2 \pmod{m}.$$

By the definition of S, $b_1 = b_2$, so we have

$$a^2 \equiv 0 \pmod{m}.$$

Since m divides a^2, and m is squarefree, m divides a. Hence $a = cm$, so $a^2 = c^2 m^2$. Thus we have

$$m^2(x_1 - x_2) + m(z_1 - z_2) = c^2 m^2 + Lm^{2k}.$$

Reducing this equation mod m^2, and using the fact that $k \geq 1$, we obtain

$$m(z_1 - z_2) \equiv 0 \pmod{m^2}.$$

Since $0 \leq z_1, z_2 \leq m - 1$, we have $m\,|z_1 - z_2| < m^2$, hence $z_1 = z_2$.
Since $b_1 = b_2$ and $z_1 = z_2$, we now have

$$m^2(x_1 - x_2) = c^2 m^2 + Lm^{2k}$$

Dividing by m^2, we obtain

$$(x_1 - x_2) = c^2 + Lm^{2k-2}.$$

Recall that $x_1, x_2 \in X$. By the condition on X, there do not exist two elements of X whose difference is a square mod m^{2k-2}. Since the last equation states that the difference of two elements of X is a square mod m^{2k-2}, this is a contradiction. ∎

Lemma 4.5 *For all $k \geq 1$,* $\text{sdfmod}(m^{2k}) \geq m \cdot \text{sdfmod}(m) \cdot \text{sdfmod}(m^{2k-2})$.

Proof: Let S be an SDFMOD(m) set of size sdfmod(m) and let X be an SDFMOD(m^{2k-2}) set of size sdfmod(m^{2k-2}). By Lemma 4.4 there exists Y, an SDFMOD(m^{2k}) set of size $m \cdot \text{sdfmod}(m) \cdot \text{sdfmod}(m^{2k-2})$. Hence

$$\text{sdfmod}(m^{2k}) \geq m \cdot \text{sdfmod}(m) \cdot \text{sdfmod}(m^{2k-2}).$$

∎

Lemma 4.6 *Assume that there exists a squarefree m and a set $S \subseteq [m]$ such that S is* SDFMOD(m). *Then* $\text{sdf}(n) \geq \Omega(n^{0.5(1+\log_m |S|)})$. *(The constant implicit in the Ω depends on m.)*

Proof: By the premise, $\text{sdfmod}(m) \geq |S|$. By Lemma 4.5

$$(\forall k \geq 1)[\text{sdfmod}(m^{2k}) \geq m|S|\text{sdfmod}(m^{2k-2})].$$

Hence

$$\text{sdfmod}(m^{2k}) \geq (m|S|)^k \text{sdfmod}(1) = (m|S|)^k$$

Let $n = m^{2k}$, so $k = \log_m \sqrt{n}$. Then

$$\begin{aligned} (m|S|)^k &= m^k(|S|)^k \\ &= m^k(|S|)^{\log_m \sqrt{n}} \\ &= n^{0.5}|S|^{\log_m \sqrt{n}} \end{aligned}$$

Note that

$$|S|^{\log_m \sqrt{n}} = (\sqrt{n})^{\log_m |S|} = n^{0.5 \log_m |S|}.$$

Hence, for $n = m^{2k}$,

$$\text{sdfmod}(m^{2k}) \geq (m|S|)^k = n^{0.5} n^{0.5 \log_m |S|} = n^{0.5(1+\log_m |S|)}.$$

Thus we have

$$\text{sdfmod}(n) \geq \Omega(n^{0.5(1+\log_m |S|)}).$$

∎

Theorem 4.7 $\text{sdf}(n) \geq \Omega(n^{\log_{205} 12}) \geq \Omega(n^{0.7334\cdots})$.

Proof: Let $m = 205$ and $S = \{0, 2, 8, 14, 77, 79, 85, 96, 103, 109, 111, 181\}$. Clearly, m is square free. An easy calculation shows that there are no two elements of S whose difference is a square mod m. Note that $\log_m |S| = \log_{205} 12 > 0.4668$. Hence, by Lemma 4.6,

$$\text{sdf}(n) \geq \Omega(n^{0.5(1+\log_m |S|)}) \geq \Omega(n^{0.5(1+0.4668)}) \geq \Omega(n^{0.7334\cdots}).$$

∎

Note 4.8 Ruzsa used $m = 65$ and a set S of size 7 to obtain his results. He did not specify his set S; however, $S = \{0, 2, 5, 22, 24, 43, 46\}$ will suffice.

5 Square-Difference-Free Colorings

Once we have a large Square-free difference set can we use it to obtain a lower bound on $Q(c)$. YES!

Def 5.1 Let $A \subseteq [n]$. B is a *translate of A relative to n* if there exists t such that

$$B = \{x + t : x \in A\} \cap [n].$$

We will omit the "relative to n" when n is clear from context.

The following lemma is by Lipton, Chandra, and Furst [2]. We provide their proof for completeness.

Lemma 5.2 *Let $A \subseteq [n]$. There exist $c \leq O(\frac{n \log n}{|A|})$ and sets $A_1, \ldots, A_c$ that are translates of A such that $[n] = A_1 \cup \cdots \cup A_c$. (Note that the lemma holds for any set A; however, we will apply it when A is* SDF.*) $S = \{0, 2, 5, 22, 24, 43, 46\}$ works.*

Proof:

Pick a translation of A by picking $t \in \{-n, -n+1, \ldots, n\}$. The probability that $x \in A + t$ is $\frac{|A|}{2n+1} \geq \frac{|A|}{3n}$. Hence probability that $x \notin A + t$ is at most

$$1 - \frac{|A|}{3n}.$$

If we pick s translations $t_1, \ldots, t_s$ at random (s to be determined later) then the expected number of x that are not covered by any $A + t_i$ is

$$T\left(1 - \frac{|A|}{3n}\right)^s \leq Te^{-s\frac{|A|}{3n}}.$$

We need to pick s such that this quantity is < 1 We take $s = 4\frac{n \ln n}{|A|}$ which yields

$$ne^{-s\frac{|A|}{3n}} = ne^{-(4 \ln n/3)} = n^{-1/3} < 1.$$

∎

Lemma 5.3 *Let c, n be such that $c \leq O(\frac{n \log n}{\text{sdf}(n)})$, and Then $Q(c) \geq n$.*

Proof: Let $A \subseteq [n]$ be an SDF set of size $\text{sdf}(n)$. By Lemma 5.2, there exist $c \leq O(\frac{n \log n}{\text{sdf}(n)})$ translates of A such that the union of the translates covers all of $[n]$. Call the translates $A_1, \ldots, A_c$. Let χ be the c-coloring of $[n]$ that maps a number x to the least i such that $x \in A_i$. For $1 \leq i \leq c$ let C_i be the set of numbers that are colored i. Since each C_i is an SDF, this coloring has no $x, y \in [n]$, $x \neq y$, such that $\chi(x) = \chi(y)$ and $|x - y|$ is a square. Hence $Q(c) \geq n$. ∎

Theorem 5.4 $Q(c) \geq \Omega(c^{3.75})$.

Proof: Fix c. We want to find an n as small as possible such that

$$c \leq O\left(\frac{n \log n}{\text{sdf}(n)}\right).$$

Since $\text{sdf}(n) \geq \Omega(n^{0.7334\cdots})$

$$\left(\frac{n \log n}{n^{0.7334\cdots}}\right) \leq O(n^{0.2666}).$$

Hence it will suffice to find an n as small as possible such that

$$c \leq O(n^{0.2666}).$$

We can take

$$n \geq \Omega(c^{1/0.2666}) \geq \Omega(c^{3.75}).$$

Hence $Q(c) \geq \Omega(c^{3.75})$.

∎

6 Open Problem

Combining the upper bound of Pintz, Steiger, Szemerédi with our lower bound we have:

$$\Omega(n^{0.7334\cdots}) \leq \mathrm{sdf}(n) \leq \frac{n}{\log^{O(\log\log\log\log n)} n}$$

Combining the lower bound of Theorem 5.4 with the upper bound of Corollary 2.2 we obtain

$$\Omega(c^{3.75}) \leq Q(c) \leq 2^{c^{O(1/(\log\log\log c)^{\epsilon})}}.$$

The open problem is to close these gaps. One way to raise the lower bounds on $\mathrm{sdf}(n)$ is to find values of m and $|S|$ that satisfy the premise of Lemma 4.6 with a larger value of $\log_m |S|$ then we obtained. The best upper bounds on Q are from density results that use sophisticated methods. It would be interesting to obtain upper bounds on Q using purely combinatorial means.

7 Acknowledgments

We would like to thank Georgia Martin and Edward Gan for proofreading and commentary. We would like to thank Boris Bukh, Ben Green, and Nikos Frantzikinakis for pointing us to the papers of Ruzsa [10] and Pinter-Steiger-Szemerédi [9].

References

[1] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *Journal of the American Mathematical Society*, pages 725–753, 1996. `http://www.math.ohio-state.edu/~vitaly/` or `http://www.cs.umd.edu/~gasarch/vdw/vdw.html`.

[2] A. Chandra, M. Furst, and R. Lipton. Multiparty protocols. In *Proceedings of the Fifteenth Annual ACM Symposium on the Theory of Computing,* Boston MA, pages 94–99, 1983. http://portal.acm.org/citation.cfm?id=808737.

[3] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi's on arithmetic progressions. *Journal d'Analyse Mathematique*, 31:204–256, 1977. http://www.cs.umd.edu/~gasarch/vdw/furstenbergsz.pdf.

[4] W. Gasarch, C. Kruskal, and A. Parrish. Van der Waerden's theorem: Variants and applications. www.gasarch.edu/~gasarch/~vdw/vdw.html.

[5] W. Gowers. A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, 11:465–588, 2001. http://www.dpmms.cam.ac.uk/~wtg10/papers/html or http://www.springerlink.com.

[6] R. Graham, B. Rothchild, and J. Spencer. *Ramsey Theory.* Wiley, 1990.

[7] G. Hardy and E. Wright. *An introduction to the theory of numbers.* Clarendon Press, Oxford, 1979. Fifth Edition. The first edition was in 1938.

[8] B. Landmann and A. Robertson. *Ramsey Theory over the integers.* AMS, 2003.

[9] J. Pintz, W. Steiger, and E. Szemeredi. On sets of natural numbers whose difference set contains no squares. *Journal of the London Mathematical Society*, 37:219–231, 1988. http://jlms.oxfordjournals.org/.

[10] I. Ruzsa. Difference sets without squares. *Periodica Mathematica Hungarica*, pages 205–209, 1984. http://www.cs.umd.edu/~gasarch/vdw/sqdiff-ruzsa.pdf.

[11] A. Sárközy. On difference sets of sequences of integers I. *Acta Math. Sci. Hung.*, 31:125–149, 1977. http://www.cs.umd.edu/~gasarch/vdw/sarkozyONE.pdf.

[12] A. Sárközy. On difference sets of sequences of integers II. *Annales Universitatis Scientiarum Budapestinensis De Reolando Eotvos Nominatete*, 21, 1978.

[13] Shelah. A partition theorem. *Scientiae Math Japonicae*, pages 413–438, 2002. Paper 679 at the Shelah Archive: `http://shelah.logic.at/short600.html`.

[14] S. Shelah. Primitive recursive bounds for van der Waerden numbers. *Journal of the American Mathematical Society*, pages 683–697, 1988. `http://www.jstor.org/view/08940347/di963031/96p0024f/0`.

[15] E. Szeméredi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:299–345, 1975. `http://www.cs.umd.edu/~gasarch/vdw/szdensity.pdf`.

[16] T. Tao. Szemerédi's proof of Szemerédi's theorem. `http://www.math.ucla.edu/~tao/preprints/acnt.html`.

[17] T. Tao. The ergodic and combinatorial approaches to Szemerédi's theorem, 2006. `http://arxiv.org/abs/math.CO/0604456`.

[18] B. van der Waerden. Beweis einer Baudetschen Vermutung. *Nieuw Arch. Wisk.*, 15:212–216, 1927.

[19] M. Walters. Combinatorial proofs of the polynomial van der Waerden theorem and the polynomial Hales-Jewett theorem. *Journal of the London Mathematical Society*, 61:1–12, 2000. `http://jlms.oxfordjournals.org/cgi/reprint/61/1/1` or `http://jlms.oxfordjournals.org/` or or `http://www.cs.umd.edu/~gasarch/vdw/vdw.html`.

[20] J. Wolf. Sets whose differences set is square-free, 2008. Unpublished manuscript.