ON THE ASYMPTOTIC MINIMUM NUMBER OF MONOCHROMATIC 3-TERM ARITHMETIC PROGRESSIONS

Pablo A. Parrilo
Department of Electrical Engineering and Computer Science,
Massachusetts Institute of Technology, Cambridge, MA 02139
parrilo@mit.edu

Aaron Robertson
Department of Mathematics, Colgate University, Hamilton, NY 13346
aaron@math.colgate.edu

Dan Saracino
Department of Mathematics, Colgate University, Hamilton, NY 13346
dsaracino@mail.colgate.edu

Abstract

Let $V(n)$ be the minimum number of monochromatic 3-term arithmetic progressions in any 2-coloring of $\{1, 2, \ldots, n\}$. We show that

$$\frac{1675}{32768}n^2(1 + o(1)) \leq V(n) \leq \frac{117}{2192}n^2(1 + o(1)).$$

As a consequence, we find that $V(n)$ is strictly greater than the corresponding number for Schur triples (which is $\frac{1}{22}n^2(1 + o(1))$). Additionally, we disprove the conjecture that $V(n) = \frac{1}{16}n^2(1 + o(1))$ as well as a more general conjecture.

1. Introduction

At the Erdős Conference in Budapest in the summer of 1999, Ron Graham proposed the following \$100 problem:

> Let $V(n)$ be the minimum number of monochromatic 3-term arithmetic progressions in any 2-coloring of $[1, n] = \{1, 2, \ldots, n\}$. Given $V(n) = \beta n^2(1 + o(1))$, determine β.

This problem seems to be much more abstruse than the corresponding problem concerning Schur triples (see [D, RZ, Sch]). It was conjectured, and commonly believed, that $\beta = \frac{1}{16}$, in part because of the following "folklore" conjecture.

Conjecture. The minimum number of monochromatic solutions, in any r-coloring of $[1, n]$, of $\sum_{i=1}^{m} c_i x_i = 0$ with $\sum_{i=1}^{m} c_i = 0$ is equal to the value achieved by randomly coloring the integers in $[1, n]$.

In the case of 3-term arithmetic progressions, the equation is $x + y = 2z$ and the value achieved by randomly 2-coloring the integers in $[1, n]$ is $\frac{n^2}{16}(1+o(1))$ since there are $\frac{n^2}{4}(1+o(1))$ 3-term arithmetic progressions in $[1, n]$, of which $\frac{1}{4}$ is the expected fraction of them that are monochromatic under a random 2-coloring.

The conjecture states that $V(n) = \frac{n^2}{16}(1 + o(1))$. We show that this conjecture is false by proving that $V(n) < \frac{n^2}{16}(1 + o(1))$; however, we are uncertain of whether this conjecture is false "in general"– the equation $x + y = 2z$ may be an anomalous one. While we do not determine β conclusively, we are able to offer fairly good upper and lower bounds (the relative difference is approximately 4.5%). We believe that our upper bound is extremely close, if not equal, to $V(n)$.

In [GRR] it is shown that, for n sufficiently large, every 2-coloring of $[1, n]$ admits $\Theta(n^2)$ monochromatic 3-term arithmetic progressions and this fact is assumed throughout this paper. The reader is strongly urged to read [GRR] as it served as the impetus for this article, as well as for [D], [RZ], and [Sch], and will surely be a valuable tool for studying the behavior of other regular linear homogeneous equations.

2. Preliminaries for the Lower Bound

Let $\chi : [1, n] \to \{0, 1\}$ be an arbitrary 2-coloring. Define, for $j = 0, 1$,

$$S_j = \{x : \chi(x) = j, \;\; 1 \le x \le n\}.$$

Let $V(S_0, S_1) = V(n; S_0, S_1)$ be the number of monochromatic 3-term arithmetic progressions in $[1, n]$ under χ.

Using an approach found in [Sch] and [D], we let

$$f_j = \sum_{s \in S_j} e^{2\pi i s x}, \quad j = 0, 1,$$

which gives us

$$2V(S_0, S_1) = \int_0^1 \left(f_0^2(x)\overline{f_0(2x)} + f_1^2(x)\overline{f_1(2x)} \right) \, dx.$$

We rewrite the integrand as

$$(f_0(x)+f_1(x))^2\left(\overline{f_0(2x)}+\overline{f_1(2x)}\right) \quad -\left(f_0(x)\overline{f_1(2x)}+f_1(x)\overline{f_0(2x)}\right)(f_0(x)+f_1(x)) \\ -\ f_0(x)f_1(x)\left(\overline{f_0(2x)}+\overline{f_1(2x)}\right)$$

and interpret the integral as

$$\begin{aligned} 2V(S_0,S_1) &= \left|\{(a,b,c)\in[1,n]^3 : a+b=2c\}\right| \\ &-\left|\{(a,b)\in(S_0\times S_1)\cup(S_1\times S_0) : 2b-a\in[1,n]\}\right| \\ &-\left|\{(a,b)\in S_0\times S_1 : a+b \text{ is even}\}\right|. \end{aligned}$$

We will now bound the size of these sets, where our equations are valid up to $o(n^2)$.

It is trivial to show that $|\{(a,b,c)\in[1,n]^3 : a+b=2c\}| = \frac{n^2}{2}(1+o(1))$. It is also easy to show that $|\{(a,b)\in S_0\times S_1 : a+b \text{ is even}\}| \le \frac{n^2}{8}(1+o(1))$ as follows. Denote this set by T and let r_o and b_o be the number of odd numbers in $[1,n]$ of color red (in S_0, say) and blue (in S_1), respectively, and let r_e and b_e the number of even numbers in $[1,n]$ of color red and blue, respectively. Then

$$\begin{aligned} |T| &= (r_ob_o+r_eb_e) \\ &= \tfrac{1}{2}\left((r_o+b_o)^2+(r_e+b_e)^2-(r_o^2+r_e^2+b_o^2+b_e^2)\right) \\ &= \tfrac{1}{2}\left(\left(\tfrac{n}{2}\right)^2+\left(\tfrac{n}{2}\right)^2-(r_o^2+b_o^2+r_e^2+b_e^2)\right) \\ &= \tfrac{1}{2}\left(\tfrac{n^2}{2}-(r_o^2+b_o^2+r_e^2+b_e^2)\right) \\ &= \tfrac{n^2}{4}-\tfrac{1}{2}\left(r_o^2+\left(\tfrac{n}{2}-r_o\right)^2+r_e^2+\left(\tfrac{n}{2}-r_e\right)^2\right) \\ &= \tfrac{n}{2}(r_o+r_e)-(r_o^2+r_e^2). \end{aligned}$$

This function attains its maximum of $\frac{n^2}{8}(1+o(1))$ when $r_o=r_e=\frac{n}{4}$.

Next, we define

$$N^+ = \{(a,b)\in(S_0\times S_1)\cup(S_1\times S_0) : 2b-a\in[1,n]\}.$$

Our goal is to find an upper bound for $|N^+|$ and use the following lemma, which follows immediately from the paragraphs above.

Lemma 1 *If* $|N^+|\le cn^2(1+o(1))$, *then*

$$V(S_0,S_1)\ge\frac{1}{2}\left(\frac{3}{8}-c\right)n^2(1+o(1)).$$

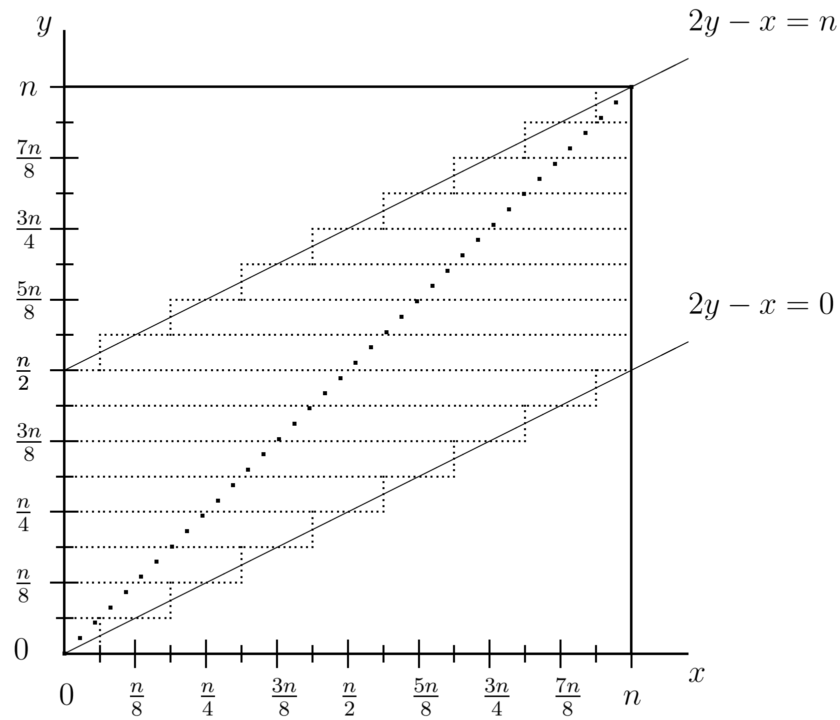

Figure 1: Partition of the square into rectangles, for $L = 16$.

3. Lower Bound Calculations

Our approach will be to consider points in the square $[1, n]^2$. From the definition of N^+, we restrict our attention to those points (x, y) with $0 < 2y - x \leq n$. We also remark that since we are looking for the coefficient of the n^2 term in $V(n)$, we will disregard points that contribute $o(n^2)$ to $V(n)$.

Consider the diagram in Figure 1. We are trying to find the maximum number of achromatic pairs (a, b) that can reside inside the parallelogram bounded by the lines $x = 0$, $x = n$, $2y - x = 0$, and $2y - x = n$. To this end, we cover the parallelogram by L horizontal strips of height $\frac{n}{L}$ and right triangles with dimensions $\frac{n}{2L} \times \frac{n}{L}$ (in Figure 1, we have $L = 16$). As such, we cover more than the parallelogram (we have right triangles outside of the parallelogram). Hence, by maximizing the number of achromatic pairs inside the strips and the right triangles, we have an upper bound on the maximum number of achromatic pairs that can reside inside the parallelogram.

Let $((i-1)\frac{n}{L}, i\frac{n}{L}]$ contain r_i red elements, $i = 1, 2, \ldots, L$. Choosing L to be even, we can easily write down a formula for the number of achromatic pairs that reside in the horizontal strips:

$$\sum_{i=1}^{L/2}\sum_{j=1}^{2i-1}\left(r_i\left(\frac{n}{L}-r_j\right)+\left(\frac{n}{L}-r_i\right)r_j\right)+\sum_{i=L/2+1}^{L}\sum_{j=2i-L}^{L}\left(r_i\left(\frac{n}{L}-r_j\right)+\left(\frac{n}{L}-r_i\right)r_j\right). \quad (1)$$

What remains are the maximum possible number of achromatic points in the L remaining triangles. For these we use the trivial bound of their areas, $L \times \frac{1}{2}\frac{n}{L}\frac{n}{2L} = \frac{n^2}{4L}$. Combining this with (1), we have an upper bound on $|N^+|$:

$$|N^+| \le \frac{n^2}{4L} + \sum_{i=1}^{L/2}\sum_{j=1}^{2i-1}\left(r_i\left(\frac{n}{L}-r_j\right)+\left(\frac{n}{L}-r_i\right)r_j\right) + \sum_{i=L/2+1}^{L}\sum_{j=2i-L}^{L}\left(r_i\left(\frac{n}{L}-r_j\right)+\left(\frac{n}{L}-r_i\right)r_j\right). \quad (2)$$

We present next two different techniques to effectively bound the right-hand side of (2). The first one relies on an explicit enumeration of all the critical points (for $L = 16$), while the second approach uses a procedure based on semidefinite programming.

3.1 Enumeration Bounds for $L = 16$

In this approach, all critical points in $\left(0, \frac{n}{16}\right)^{16}$ are compared against all maximum values at the $3^{16} - 1$ boundary problems. The maximization problem has been programmed into Maple as a small program called `PABLO` and the code is available from the second author's website[1].

After running for approximately 136 hours on a 2.7GHz G5 Macintosh server, we find that

$$|N^+| \le \frac{579}{2048}n^2(1+o(1)).$$

One coloring that achieves this bound is

$$(r_1, r_2, \ldots, r_{16}) = \left(\frac{7n}{128}, \frac{7n}{128}, 0, \frac{7n}{128}, \frac{n}{16}, 0, 0, 0, \frac{n}{16}, \frac{n}{16}, \frac{n}{16}, \frac{n}{128}, 0, \frac{n}{16}, \frac{n}{128}, \frac{n}{128}\right).$$

Applying Lemma 1, the above result gives us the following theorem.

Theorem 2 $V(n) \ge \frac{189}{4096}n^2(1+o(1))$.

[1] `http://math.colgate.edu/~aaron/programs.html`

3.2 Semidefinite Bounds

A different, more powerful way of bounding $|N^+|$ is based on semidefinite relaxations. For this, consider first the change of variables $r_i := \frac{1+x_i}{2}\frac{n}{L}$, so $r_i \in [0, \frac{n}{L}]$ if and only if $x_i \in [-1, 1]$. Then, equation (2) can be written as

$$\begin{aligned} |N^+| &\leq \frac{n^2}{4L} + \sum_{i=1}^{L/2}\sum_{j=1}^{2i-1} \frac{n^2}{2L^2}(1 - x_i x_j) + \sum_{i=L/2+1}^{L}\sum_{j=2i-L}^{L} \frac{n^2}{2L^2}(1 - x_i x_j) \\ &\leq \frac{n^2}{4L} + \frac{n^2}{4} - \frac{n^2}{4L^2}\, q(\mathbf{x}), \end{aligned}$$

where

$$q(\mathbf{x}) := \sum_{i=1}^{L/2}\sum_{j=1}^{2i-1} 2x_i x_j + \sum_{i=L/2+1}^{L}\sum_{j=2i-L}^{L} 2x_i x_j.$$

Our objective is to bound $|N^+|$ from above. For this, it is clearly enough to obtain a lower bound of the quadratic form $q(\mathbf{x})$ over $[-1, 1]^n$. This quadratic form can be represented as $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, where A is an $L \times L$ symmetric integer matrix, with entries $A_{ij} = B_{ij} + B_{ji}$ and

$$B_{ij} = \begin{cases} 1 & \text{if } j + 1 \leq 2i \leq j + L \\ 0 & \text{otherwise.} \end{cases}$$

A useful bound for quadratic forms on the unit hypercube, used extensively in the combinatorial optimization literature, can be obtained as follows.

Lemma 3 Let A be an $n \times n$ matrix and let $D = \mathrm{diag}(d_1, \ldots, d_n)$ be a diagonal matrix, with $d_i \geq 0$ and such that $A + D$ is positive semidefinite. Then, for all $\mathbf{x} \in [-1, 1]^n$, $\mathbf{x}^T A\mathbf{x}$ is bounded below by $-\sum_{i=1}^n d_i$.

Proof. Consider any vector $\mathbf{x} \in [-1, 1]^n$. Since $A + D$ is positive semidefinite it follows that

$$0 \leq \mathbf{x}^T(A + D)\mathbf{x} = \mathbf{x}^T A\mathbf{x} + \sum_{i=1}^{n} d_i x_i^2.$$

Since $x_i^2 \leq 1$, we have $\mathbf{x}^T A\mathbf{x} \geq -\sum_{i=1}^n d_i x_i^2 \geq -\sum_{i=1}^n d_i$. □

For any finite value of L, a suitable set of d_i can be found by semidefinite programming. For the case $L = 128$ we have found a particular solution (given in the Appendix) using the SDP solver SeDuMi [St], followed by a straightforward rounding procedure (to obtain rational solutions). For such a solution, it can be easily verified on a computer that the 128×128 rational matrix $A + D$ is indeed positive definite. Since we have $\sum_{i=1}^L d_i = 1364$, this gives an upper bound for $|N^+|$ with $c = \frac{1}{4L} + \frac{1}{4} + \frac{1364}{4L^2} = \frac{4469}{16384}$, resulting in the lower bound (via Lemma 1) given in the next theorem.

Theorem 4 $V(n) \geq \frac{1675}{32768} n^2(1 + o(1))$.

4. The Upper Bound

Theorem 5 $V(n) \le \frac{117}{2192} n^2 (1 + o(1))$.

Proof. Let $i^m = \underbrace{ii\ldots i}_{m}$, i.e., a string of i's of length m. For any n that is an integer multiple of 548, we construct a coloring, using the colors 0 and 1, given by

$$0^{\frac{28}{548}n}\ 1^{\frac{6}{548}n}\ 0^{\frac{28}{548}n}\ 1^{\frac{37}{548}n}\ 0^{\frac{59}{548}n}\ 1^{\frac{116}{548}n}\ 0^{\frac{116}{548}n}\ 1^{\frac{59}{548}n}\ 0^{\frac{37}{548}n}\ 1^{\frac{28}{548}n}\ 0^{\frac{6}{548}n}\ 1^{\frac{28}{548}n}.$$

If n is not a multiple of 548, we use the coloring above for the first $548\lfloor \frac{n}{548} \rfloor$ elements, and color the few remaining elements arbitrarily. It is tedious – but routine – to show that under this coloring there are $\frac{117}{2192}n^2(1+o(1))$ monochromatic 3-term arithmetic progressions, thereby proving the theorem. □

The above coloring was found using a combination of computational and analytic methods. We briefly describe these next.

As we have seen in the previous sections, the problem can be essentially reduced to the minimization of the quadratic form $q(\mathbf{x})$ over the unit hypercube. To understand the behavior of the solution, we solved instances of this problem for large values of n ($n \approx 2000$). For this, a "good" initial candidate coloring was found using the solution of the semidefinite relaxation described in the previous section, followed by a randomization procedure known as Goemans-Williamson rounding [GW]. The near-optimal solutions found all shared some nice structural features, such as (approximate) anti-symmetry, and being essentially constant over large ranges of n, with a small number of breakpoints (equal to twelve for most solutions).

We then used a continuous approximation to the minimization of $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$, given by

$$\min_{\phi} \int_{-1}^{1} \int_{-1}^{1} k(x,y)\phi(x)\phi(y)dxdy,$$

where the function ϕ must satisfy $|\phi(\cdot)| = 1$ and the kernel $k(x,y)$ is piecewise constant. Based on the numerical solutions for large n, we chose an ansatz where the function ϕ is anti-symmetric ($\phi(-x) = -\phi(x)$) and piecewise constant on twelve different intervals.

Because $k(x,y)$ is piecewise constant, the objective function is a *piecewise quadratic* function of six variables, namely the breakpoints (six variables rather than twelve since we are assuming anti-symmetry). It turns out that, on the partition associated with the solution obtained by numerical computation, this function is strictly convex and its minimum lies inside the partition. Solving for the (local) minimum of this quadratic function, we obtained the breakpoints corresponding to the solution in Theorem 5. The solution presented is thus "locally optimal" in the sense that no small perturbation of the breakpoints will achieve a better value. Of course, in principle the possibility remains that there exist solutions of different structure that achieve even smaller values, so the argument given is not enough to prove global optimality. However, based on our numerical explorations, we conjecture that the bound and coloring pattern presented in Theorem 5 are indeed asymptotically optimal.

The Maple code that computes this quadratic function and performs the minimization is also available at the location cited earlier.

5. Remarks for Further Investigation

Clearly, the parallelogram described at the beginning of Section 3 could be further refined by using larger values of L.

For the enumeration technique in Section 3.1 this would provide sharper bounds, which converge to the optimal constant β. However, since the number of points to be checked grows exponentially with L, there would be an enormous increase in the computational cost (for example, adding two more variables would increase the computing time to approximately 51 days). A possible improvement here could be obtained by finding an upper bound on the triangles for which we have used the trivial bound of their area, although this would not help with the exponential behavior.

For the semidefinite bounds in Section 3.2, it is relatively straightforward (and computationally feasible) to provide slightly better lower bounds by increasing the value of L. However, even if we let $L \to +\infty$, the obtained bounds will likely *not* converge to the optimal value of β, as there seems to be an "irreducible" gap between the original problem and its corresponding semidefinite relaxation. While this issue is relatively well-understood for finite problems, it would be of interest to fully understand the situation in this infinite limit.

Given our belief that the bound presented in Theorem 5 is sharp, perhaps the most promising approach would be to attempt to directly prove the (asymptotic) global optimality of the corresponding solution. Given numerical evidence, it seems that the optimal coloring is anti-symmetric, although we have been unable to prove that this must be the case.

Finally, the techniques in this paper can be applied to any 3-term equation. The semidefinite programming technique easily generalizes, while the methods in Section 2 (for the lower bound) may not be as straightforward. For example, if we consider the equation $ax+by = cz$, then for $V(S_0, S_1)$, as defined in Section 2, we would have

$$2V(S_0, S_1) = \int_0^1 \left(f_0(ax) f_0(bx) \overline{f_0(cx)} + f_1(ax) f_1(bx) \overline{f_1(cx)} \right) dx,$$

which may not lend itself to the type of analysis that is done in Section 2.

Acknowledgment

We thank Ron Graham for some background information, but more importantly for providing the combinatorics community with interesting and challenging problems. The first author thanks Bernd Sturmfels for forwarding the email (from Doron Zeilberger) that spurred this collaboration.

References

[D] B. Datskovsky, On the Number of Monochromatic Schur Triples, *Advances in Applied Math.* **31** (2003), 193-198.

[GW] M. X. Goemans and D. P. Williamson, Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming, *Journal of the ACM* **42(6)** (1995), 115-1145.

[GRR] R. Graham, V. Rödl, and A. Ruciński, On Schur Properties of Random Subsets of Integers, *J. Number Theory* **61** (1996), 388-408.

[RZ] A. Robertson and D. Zeilberger, A 2-Coloring of $[1, N]$ Can Have $N^2/22 + O(N)$ Monochromatic Schur Triples, But Not Less!, *Elect. J. of Comb.* **5(1)** (1998), R19.

[Sch] T. Schoen, On the Number of Monochromatic Schur Triples, *Europ. J. of Comb.* **20(8)** (1999), 855-866.

[St] J.F. Sturm, Using SeDuMi 1.02, A MATLAB Toolbox for Optimization Over Symmetric Cones, *Optimization Methods and Software* **11-12** (1999) 625-653.

Appendix

A particular solution for the d_i in Lemma 3 is given by the numbers below.

$$d = \frac{1}{4}\left[\begin{array}{cccccccccccccccc}
27 & 22 & 14 & 14 & 13 & 11 & 5 & 2 & 9 & 14 & 20 & 24 & 26 & 29 & 28 & 26 \\
26 & 26 & 26 & 25 & 24 & 23 & 23 & 21 & 22 & 21 & 27 & 30 & 37 & 41 & 48 & 50 \\
54 & 53 & 53 & 53 & 53 & 55 & 59 & 65 & 70 & 76 & 79 & 83 & 84 & 86 & 84 & 81 \\
74 & 69 & 61 & 53 & 49 & 50 & 56 & 61 & 66 & 65 & 61 & 51 & 46 & 46 & 41 & 37 \\
37 & 41 & 46 & 46 & 51 & 61 & 65 & 66 & 61 & 56 & 50 & 49 & 53 & 61 & 69 & 74 \\
81 & 84 & 86 & 84 & 83 & 79 & 76 & 70 & 65 & 59 & 55 & 53 & 53 & 53 & 53 & 54 \\
50 & 48 & 41 & 37 & 30 & 27 & 21 & 22 & 21 & 23 & 23 & 24 & 25 & 26 & 26 & 26 \\
26 & 28 & 29 & 26 & 24 & 20 & 14 & 9 & 2 & 5 & 11 & 13 & 14 & 14 & 22 & 27
\end{array}\right].$$