Multiparty Communication Complexity: A Serious Approach

0.1 Points

In the last chapter we considered the following problem:

> Alice, Bob, and Carol each have a number between 0 and $2^n - 1$ on their forehead, written in binary; so everyone has exactly n bits on their forehead. Everyone can see the two numbers that are NOT on their own forehead. Call the three numbers a, b, c. They wish to determine if $a + b + c = 2^{n+1} + 1$. At the end of the protocol they should all know. Each player can shout information so the others hear them. Here is one solution: Alice shouts a, which is n bits, and then Bob can compute $a + b + c$ and should YES if $a + b + c = 2^n - 1$, NO otherwise. Can you do this with fewer than n bits of communication?

This problem was FUN for Olivia since there is an elementary solution using only $\frac{n}{2} + O(1)$. I suspect Olivia could find this solution. She can certainly understand the solution.

In this chapter we give a much better upper bound that Olivia could not come up with. We also give matching upper and lower bounds – sort of. We'll see later what that means. We will also look at the k-party version of the problem.

The problem, and the original definition of multiparty communication complexity, and a solution are due to Chandra, Furst, and Lipton ([Chandra *et al.* (1983)]). There were later improvements to the solution by Beigel, Gasarch, and Glenn ([Beigel *et al.* (2006)]). In Section 0.8 we discuss who did what, the motivations of the authors, and why, of the papers I have co-authored, Beigel-Gasarch-Glenn is one of my favorites.

The last chapter was FUN whereas this chapter is SERIOUS. One proof of that: the last chaper had 0 references, this one has 17.

0.2 A Serious Solution

We need a byte of notation.

Definition 0.1. Let $n \in \mathbb{N}$. $f_3 : \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \to \{\text{YES}, \text{NO}\}$ is defined by:

$$f_3(a,b,c) = \begin{cases} \text{YES} & \text{if } a+b+c = 2^{n+1}-1 \\ \text{NO} & \text{if } a+b+c \neq 2^{n+1}-1 \end{cases} \tag{0.1}$$

Definition 0.2. Let Alice, Bob, and Carol be as in Section 0.1. Let $d(f_3)$ be the number of bits that Alice, Bob, and Carol need to communicate so that at the end they all know $f_3(a, b, c)$.

Informally, a proper coloring of $[T] \times [T]$ does not have be three vertices that (a) are the same color, and (b) are the corners of a right isosceles triangle with legs parallel to the axes and hypotenuse parallel to the line $y = -x$. Formally:

Definition 0.3. Let $c, T \in \mathbb{N}$.

(1) A *proper c-coloring of* $[T] \times [T]$ is a function $\mathrm{COL} : [T] \times [T] \rightarrow [c]$ such that there do not exist $a, b \in [T]$ and $\lambda \in [T]$ with

$$\mathrm{COL}(a, b) = \mathrm{COL}(a + \lambda, b) = \mathrm{COL}(a, b + \lambda)$$

(2) Let $\chi(T)$ be the least c such that there is a proper c-coloring of $[T] \times [T]$.

There are tight bounds on $d(f_3)$, up to an additive constant.

Theorem 0.1. *Let* $n \in \mathbb{N}$. $\lg(\chi(2^n)) + \Omega(1) \leq d(f_3) \leq \lg(\chi(2^n)) + O(1)$.

These bounds seem very good except (1) we don't know what the bound is, and (2) we don't even know if we have beaten the $d(f_3) \leq \frac{n}{2} + O(1)$ bound!

Upper Bound: We will show $\lg(\chi(2^n)) + O(1) \leq O(\sqrt{n})$; hence $d(f_3) \leq O(\sqrt{n})$. So we do indeed beat $d(f_3) \leq \frac{n}{2} + O(13$.

Lower Bound: We will show that $\lg(\chi(2^n)) \geq \Omega(\log \log n))$; hence $d(f_3) \geq \Omega(\log \log n)$.

The upshot is

$$\Omega(\log \log n) \leq d(f_3) \leq O(\sqrt{n}).$$

There is a big gap between the upper and lower bounds. Normally if we have this kind of gap one might look for a better protocol. But this option will not work – we already have the optimal protocol! So what we really need is better bounds on $\chi(2^n)$. This seems to be a hard problem in Ramsey Theory which we discuss in Section 0.5.

0.3 The Upper Bound on $d(f_3)$

Theorem 0.2. *Let* $n \in \mathbb{N}$. $d(f_3) \leq \lg(\chi(2^n)) + O(1)$.

Proof. Let COL be a proper coloring of $[2^n - 1] \times [2^n - 1]$ using $\chi(2^n - 1)$ colors. We represent elements of $[\chi(2^n - 1)])]$ by bit strings of length $\lg(\chi(2^n)) + O(1)$. Alice, Bob, and Carol will all know COL ahead of time.

We present a protocol for this problem for which the communication is $\lg(\chi(2^n)) + O(1)$. We will then show that it is correct.

(1) Alice has b, c. Bob has a, c. Carol has a, b. $0 \le a, b, c \le 2^n - 1$.
(2) Alice calculates a' such that $a' + b + c = 2^{n+1} - 1$. If $a' \ge 2^n$ then Alice broadcasts NO and the protocol is over. Else Alice broadcasts $\sigma_1 = \mathrm{COL}(a', b)$.
(3) Bob calculates b' such that $a + b' + c = 2^{n+1} - 1$. If $b' \ge 2^n$ then Bob broadcasts NO and the protocol is over. Else Bob broadcasts 1 if $\sigma_2 = \mathrm{COL}(a, b')$, 0 otherwise.
(4) Carol looks up $\sigma_3 = \mathrm{COL}(a, b)$. Carol broadcasts YESif $\sigma_1 = \sigma_3$ (and hence both equal σ_2 since Bob said that $\sigma_1 = \sigma_2$) and NO otherwise.

Claim 1: If $f_3(a, b, c) = \mathrm{YES}$ then Carol will shout YES.

Proof: If $f_3(a, b, c) = \mathrm{YES}$ then $a' = a$, $b' = b$. Hence $\sigma_1 = \sigma_2 = \sigma_3$ and Carol shouts YES.
End of proof of Claim 1.

Claim 2: If Carol shouts YES then $f_3(a, b, c) = \mathrm{YES}$.

Proof: Assume that Carol shouts YES.

Hence

$$\mathrm{COL}(a, b) = \mathrm{COL}(a', b) = \mathrm{COL}(a, b').$$

We call this **The Coloring Equation.**

Assume

$$a + b + c = 2^{n+1} - 1 + \lambda.$$

We show that $\lambda = 0$.

By the definition of a'

$$a' + b + c = 2^{n+1} - 1.$$

Hence

$$a' + a + b + c - a = 2^{n+1} - 1 \implies$$

$$a' - a = 2^{n+1} - 1 - (a + b + c) = -\lambda \implies$$

$$a = a' + \lambda$$

By the same reasoning

$$b = b' + \lambda$$

Hence we can rewrite The Coloring Equation as

$$\text{COL}(a, b) = \text{COL}(a + \lambda, b) = \text{COL}(a, b + \lambda).$$

Since COL is a proper coloring, $\lambda = 0$.
End of proof of Claim 2.

□

So we have $d(f_3) \le \lg(\chi(2^n)) + O(1)$. To show this is $\le O(\sqrt{n})$ we take a detour into upper bounding χ.

0.4 An Upper Bound on χ and Hence on $d(f_3)$

We will need the following definition from Ramsey Theory.

Definition 0.4.

(1) A 3-AP is an arithmetic progression of length 3.
(2) Let $\chi^*(T)$ be the minimum number of colors needed to color $[T]$ such that there are no monochromatic 3-AP's.
(3) A set $A \subseteq [T]$ is 3-*free* if there do not exist any 3-AP's in A.
(4) Let $\mathrm{r}_3(T)$ be the size of the largest 3-free subset of $[T]$.

Lemma 0.1.

(1) $\chi(T) \le \chi^*(3T)$.
(2) $\chi^*(T) \le O(\frac{T \log T}{\mathrm{r}_3(T)})$.
(3) $\chi(T) \le O(\frac{T \log T}{\mathrm{r}_3(T)})$. *(This follows from 1 and 2.)*

Proof. 1) Let $c = \chi(3T)$. Let COL′ be a c-coloring of $[3T]$ with no monochromatic 3-AP's. Let COL be the following c-coloring of $[T] \times [T]$.

$$\mathrm{COL}(a,b) = \mathrm{COL}'(a+2b).$$

Assume, by way of contradiction, that COL is not a proper c-coloring. Hence there exist $a, b \in [T]$ and $\lambda \neq 0$ such that

$$\mathrm{COL}(a,b) = \mathrm{COL}(a+\lambda, b) = \mathrm{COL}(a, b+\lambda).$$

By the definition of COL the following are equal.

$$\mathrm{COL}'(a+2b) = \mathrm{COL}'(a+\lambda+2b) = \mathrm{COL}'(a+2\lambda+2b)$$

Hence $a+2b, a+2b+\lambda, a+2b+2\lambda$ form a monochromatic 3-AP, which yields a contradiction.

2) Let $A \subseteq [T]$ be a set of size $r_3(T)$ with no 3-AP's. We use A to obtain a 3-free coloring of $[T]$. The main idea is that we use randomly chosen translations of A to cover all of $[T]$.

Let $a \in [T]$. Pick a translation of A by picking $t \in [T]$. The probability that $a \in A+t$ is $\frac{|A|}{T}$. Hence the probability that $a \notin A+t$ is $1 - \frac{|A|}{T}$. 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

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

We need to pick s such that this quantity is < 1 We take $s = 2\frac{T \ln T}{|A|}$ which yields

$$Te^{-s\frac{|A|}{T}} = Te^{-2\ln T} = 1/T < 1.$$

We color T by coloring each of the s translates a different color. If a number is in two translates then we color it by one of them arbitrarily. Clearly this coloring has no monochromatic 3-APs. Note that it uses $\frac{T \ln T}{|A|} = O(\frac{T \log T}{\mathrm{r}_3(T)})$ colors. □

To get an upper bound on χ we need a lower bound on $\mathrm{r}_3(T)$. Behrend ([Behrend (1946)]) and Moser ([Moser (1953)]) both proved Part 1 of the next theorem. Behrend proved it first and with a smaller (hence better) value of c, but his proof was nonconstructive (i.e, the proof does not indicate how to actually find such a set). Moser's proof was constructive.

Theorem 0.3.

(1)

$$r_3(T) \geq \Omega(T^{1-\frac{c}{\sqrt{T}}})$$

(This is due to Behrend and Moser as noted above.)

(2)

$$\chi(T) \leq O(T^{\frac{c}{\sqrt{\lg T}}} \log(T))$$

(This follows from Lemma 0.1 and Part 1.)

(3)

$$\chi(2^n) \leq 2^{O(\sqrt{n})}.$$

(This follows from Part 2.)

(4)

$$\lg(\chi(2^n) \leq O(\sqrt{n}).$$

(This follows from Part 3.)

(5)

$$d(f_3) \leq O(\sqrt{n}).$$

(This follows from Part 4 and Theorem 0.2.)

0.5 Can the Upper Bound on $d(f_3)$ be Improved?

One way to obtain a smaller upper bound on $d(f_3)$ is to improve Theorem 0.3.1 by finding a better larger 3-free sets and hence a larger lower bound on $r_3(T)$. The largest known 3-free set from 1946 until 2008 were those from Theorem 0.3.1. Elkin (published [Elkin (2011)], simplified by Green and Wolf ([Green and Wolf (2010)])) obtained the following result:

$$r_3(T) \geq (\log T)^{1/4} \frac{T}{2^{2\sqrt{2}\sqrt{\log_2(T)}}}.$$

If we had used this bound instead of Theorem 0.3.1 we would still have $d(f_3) \leq O(\sqrt{n})$ though perhaps with a better constant.

Is there a mathematical obstacle to improving the lower bound on $r_3(T)$? In 2016 [Bloom (2016)] showed

$$r_3(T) \leq \frac{T(\ln \ln T)^4}{\ln T}$$

Hence the gap between the upper and lower bound on $r_3(T)$ is still large. Which is believed to be closer to the truth? I posed this question to a pretension of professors. I got several responses which are best summarized by one of them, Jacob Fox:

The good money is on the following two statements:

(1) The lower bound on $\mathrm{r}_3(T)$ by Elkin is close to the truth, and
(2) current methods are not up going to improve this.

If the consensus is correct then we cannot improve the upper bound on $d(f_3)$, by using 3-free sets. In this case to improve the upper bound on $d(f_3)$ will require some other technique. Having said that, I believe:

(1) If the consensus is correct then $d(f_3) = \Theta(\sqrt{n})$.
(2) The consensus is correct.

Hence I am forced to believe $d(f_3) = \Theta(\sqrt{n})$.

0.6 A Lower Bound on $d(f_3)$

We use the following lemma from communication complexity.

Lemma 0.2. *Let f be any function from $\{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n$ to $\{0,1\}$. Let P be a multiparty protocol for f.*

(1) Let TRAN *be a possible transcript of the protocol P. There exists $A_1, A_2, A_3 \subseteq \{0,1\}^n$ such that, for all $a_1, a_2, a_3 \in \{0,1\}^n$ the following holds: The protocol P on input (a_1, a_2, a_3) produces transcript* TRAN *iff $(a_1, a_2, a_3) \in A_1 \times A_2 \times A_3$.*
(2) Let $a_1, a_2, a_3 \in \{0,1\}^n$, $\sigma_1, \sigma_2, \sigma_3 \in \{\{0,1\}^n\}$, TRAN *be a transcript. Assume that σ_1 has a_1 as its first element, σ_2 has a_2 as its second element, σ_3 has a_3 as its third element. (In symbols, if $*$ means we don't care about the element, then*

$$\sigma_1 = (a_1, *, *)$$
$$\sigma_2 = (*, a_2, *)$$
$$\sigma_3 = (*, *, a_3).$$

Further assume that $\sigma_1, \sigma_2, \sigma_3$ all produces transcript TRAN. *Then (a_1, a_2, a_3) produces transcript* TRAN.

Theorem 0.4. $d(f_3) \geq \lg(\chi(2^n)) + \Omega(1)$.

Proof. Let P be a protocol for f. Let d be the maximum number of bits communicated. Note that the number of transcripts is bounded by 2^d. We use this protocol to create a proper 2^d-coloring of $[2^n] \times [2^n]$.

We define $\mathrm{COL}(a,b)$ as follows. First find c such that $a+b+c = 2^{n+1}-1$. Then run the protocol on (a,b,c). The color is the transcript produced.

Claim: COL is a proper coloring of $[2^n] \times [2^n]$.
Proof: Let $\lambda \in [2^n]$ be such that

$$\mathrm{COL}(a,b) = \mathrm{COL}(a+\lambda,b) = \mathrm{COL}(a,b+\lambda).$$

We denote this value TRAN (for Transcript). We show that $\lambda = 0$.

Let c be such that

$$a + b + c = 2^{n+1} - 1.$$

We know that the following tuples produce the same transcript TRAN, which ends with a YES.

- (a, b, c).
- $(a + \lambda, b, c - \lambda)$.
- $(a, b + \lambda, c - \lambda)$.

By Lemma 0.2.2 the tuple $(a, b, c-\lambda)$ also goes to TRAN and ends with a YES. Hence $a + b + c - \lambda = 2^{n+1} - 1$. Since $a + b + c = 2^{n+1} - 1 + \lambda$, so $\lambda = 0$.
End of Proof of Claim □

Theorem 0.5.

(1) $\chi(T) \geq \Omega(\log\log T)$. *This is due to Graham and Solymosi ([Graham and Solymosi (2006)]).*
(2) $\chi(2^n) \geq \Omega(\log n)$ *(This follows from Part 1.)*
(3) $\lg(\chi(2^n)) \geq \Omega(\log\log n)$ *(This follows from Part 2.)*
(4) $d(f_3) \geq \Omega(\log\log n)$ *(This follows from Part 3 and Theorem 0.4.)*

0.7 What About Alice, Bob, Carol, Donna, . . ., Zelda?

Consider the following extension of our problem. There are k people $A_1, \ldots, A_k$. Person A_i has n-bit number a_i on her forehead. They want to know if $\sum_{i=1}^{k} a_i = 2^{n+1} - 1$. How many bits do they need to exchange? Formally:

Definition 0.5. Let $n \in \mathbb{N}$. $f_k : (\{0,1\}^n)^k \to \{\mathrm{YES}, \mathrm{NO}\}$ is defined by:

$$f_k(a_1, \ldots, a_k) = \begin{cases} \mathrm{YES} & \text{if } \sum_{i=1}^{k} a_i 2^{n+1} - 1 \\ \mathrm{NO} & \text{if } \sum_{i=1}^{k} a_i \neq 2^{n+1} - 1 \end{cases} \quad (0.2)$$

Let $d(f_k)$ be the number of bits the parties must exchange so that everyone knows $f_k(a_1, \ldots, a_k)$.

We state the lemmas needed to upper bound $d(f_k)$. Unless otherwise noted the proofs are similar to those for upper bounding $d(f_3)$ (which is $d(f_3)$).

Definition 0.6. Let $k, T \in \mathbb{N}$. A *k-dimensional widget of* $[T]^k$ is a set of the form:

0) $(a_1, \ldots, a_k)$
1) $(a_1 + \lambda, a_2, a_3, \ldots, a_k)$
2) $(a_1, a_2 + \lambda, a_3, \ldots, a_k)$
3) $(a_1, a_2, a_3 + \lambda, \ldots, a_k)$

$\vdots$

k) $(a_1, a_2, a_3, \ldots, a_{k-1}, a_k + \lambda)$

When k and T are understood we just call them *widgets*.

Definition 0.7. Let $c, T \in \mathbb{N}$.

(1) A *proper c-coloring of* $[T]^k$ is a function COL : $[T]^k \to [c]$ such that there are no monochromatic widgets.
(2) Let $\chi(T)$ be the least c such that there is a proper c-coloring of $[T]^k$.

Theorem 0.6. *Let $n \in \mathbb{N}$. Then $d(f_k) \le \lg(\chi(2^n)) + O(k)$. (We think of k as a constant so the last term is really $O(1)$; however, there are applications where k is a function of n.)*

Definition 0.8.

(1) A k-AP is an arithmetic progression of length k.
(2) A set $A \subseteq [T]$ is *k-free* if there do not exist any k-AP's in A.
(3) Let $\mathrm{r}_k(T)$ be the size of the largest k-free subset of $[T]$.

Lemma 0.3. $\chi(T) \le O(\frac{T \log(T)}{\mathrm{r}_k(T)})$.

The following was first proven by Rankin [Rankin (1960—1961)] but later rediscovered (and is unfortunately behind paywalls) by Laba and Lacey [Laba and Lacey (2001)].

Lemma 0.4.

$$\mathrm{r}_k(T) \ge \frac{T}{e^{\Omega((\log T)^{1/(\log_2(k-1)+1)})}}$$

Using Lemma 0.3 and 0.4 we obtain:

Theorem 0.7. $d_k(f) \leq O(n^{1/(\log_2(k-1)+1})$.

What about lower bounds? The following is a corollary of the Gallai-Witt theorem, also called the multi-dimensional van der Warden theorem.

Lemma 0.5. *For all k, for all c, there exists T such that for all c-colorings of $[T]^k$ there exists a monochromatic widget.*

There are essentially two proofs of Theorem 0.5

(1) The classical proof gives Ackerman-type bounds fot T. It was discovered independently by Gallai ([Rado (1933)],[Rado (1943)]), where the theorem is stated and credited to him, and Witt (see [Witt (1951)]).
(2) The theorem is an easy corollary of the Hales-Jewitt Theorem. If you use the bounds from the proof by Hales-Jewitt (see [Hales and Jewett (1963)]) then you still get Ackerman-type bounds for T. But if you use the proof of Hales-Jewitt due to Shelah (see [Shelah (1988)]) then you get primitive recursive bound on T. However, they are still quite large: bigger than TOWER or WOWER.

There is no good notation for the bounds nor do we want to invent one. Hence all we can deduce from Lemma 0.5 is the following.

Theorem 0.8. $d(f_k) \geq \Omega(1)$.

Jacob Fox and David Conlon, two prominent Ramsey theorists, both told me that improving the $\Omega(1)$ to something reasonable is hard.

0.8 Whose Problem Is This? Whose Bounds are They?

In 1983 Chandra et al. [Chandra *et al.* (1983)] introduced multiparty communication complexity in order to get better bounds on constant width branching programs. They wanted to get non-polynomial lower bounds; however, they got *non-linear* lower bounds. A later paper by Ajtai et al [Ajtai *et al.* (1986)] got $\Omega(n \log n)$ bounds. However, Barrington [Barrington (1989)] showed that you could always get polynomial upper bounds. Hence the goal of Chandra et al. was not possible.

Chandra et al. proved:

(1) $\lg(\chi(2^n) + \Omega(1) \leq d(f_k) \leq 2\lg(\chi(2^n) + O(1)$
(2) $d(f_3) \leq \sqrt{\log N}$

(3) for all $k \geq 3$, $d(f_k) \geq \Omega(1)$.

In 2006 [Beigel *et al.* (2006)] proved:

(1) $d(f_3) \geq \Omega(\log \log n)$.
(2) $d(f_k) \leq O(n^{1/(\log_2(k-1)+1)})$.

0.9 Favorite Paper

Shortly after we first proved the results in [Beigel *et al.* (2006)], I we had the following conversation with Clyde:

Begin Conversation

Bill: I have two results!

Clyde: Yeah!

Bill: But all I did was combine two known theorems from Ramsey Theory (Lemma 0.4 and Theorem 0.5.a) to a theorem in complexity theorem (Theorem 0.1).

Clyde: Who else knew all three of those theorems?

Bill: I would guess fewer than five people. Actually it might just be me.

Clyde: You took A and applied it to B. There is a term for that. Its called Research.

End Conversation

This is one of my favorite papers! Why? Because I was able to take some Ramsey Theory (my favorite field of mathematics) and apply it to complexity theory (my favorite field of computer science) and, without having to be clever, get out some results. I am the only person who could have written it (downside: I am also the only person who cares).

This raises the question of: *Which of your papers is your favorite?* What should the criteria be? Rather than discuss it here, at the end of a long chapter where it won't be read, I will discuss it in a later chapter.

Bibliography

Ajtai, M., babai, L., Hajnal, P., Kolmos, J., Pudlak, P., Rodl, V., Szemeredi, E., and Turán, G. (1986). Two lower bounds on branching programs, in *Proceedings of the Eighteenth Annual ACM Symposium on the Theory of Computing,* Berkeley CA, pp. 30–38.

Barrington, D. (1989). Bounded-width polynomial-size branching programs recognize exactly those languages in NC^1, *Journal of Computer and System Sciences* **38**, pp. 150–164.

Behrend, F. (1946). On set of integers which contain no three in arithmetic progression, *Proc. of the National Academy of Science (USA)* **23**, pp. 331–332, http://www.pnas.org/content/32/12/331.full.pdf+html?sid=aeac9fed-6205-4b25-9f16-7a34a17e6042.

Beigel, R., Gasarch, W., and Glenn, J. (2006). The multiparty communication complexity of exact-t: improved bounds and new problems, in *Proceedings of the 31th International Symposium on Mathematical Foundations of Computer Science 2001,* Stara Lesna, Slovakia, pp. 146–156.

Bloom, T. (2016). A quantitative improvement for Roth's theorem on arithmetic progressions, *Journal of the London Mathematical Society* **93**, pp. 643–663, an earlier version from 2014 was posted on arXiv: https://arxiv.org/abs/1405.5800.

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

Elkin, M. (2011). An improved construction of progression-free sets, *Israel Journal of Mathematics* An earlier version was posted on arxiv in 2008: https://arxiv.org/abs/0801.4310.

Graham, R. and Solymosi, J. (2006). Monochromatic equilateral right triangles on the integer grid, *Topics in Discrete Mathematics, Algorithms and Combinatorics* **26**, www.math.ucsd.edu/~/ron/06_03_righttriangles.pdf or www.cs.umd.edu/~/vdw/graham-solymosi.pdf.

Green, B. and Wolf, J. (2010). A note on Elkin's improvement of Behrend's construction, in D. Chudnovsky and G. Chudnovsky (eds.), *Additive Number*

Theory: Festschift in honor of the 60th birthday of Melvyn B. Nathanson (Springer Verlag, New York, Heidelberg, Berlin), pp. 141–144.

Hales, A. and Jewett, R. (1963). Regularity and positional games, *Transactions of the American Math Society* **106**.

Laba, I. and Lacey, M. T. (2001). On sets of integers not containing long arithmetic progressions, `arxiv.org/pdf/math.CO/0108155` or `www.math.ubc.ca/~ilaba/preprints` or `http://www.cs.umd.edu/~gasarch/vdw/vdw.html`.

Moser, L. (1953). On non-averaging sets of integers, *Canadian Journal of Mathematics* **5**, pp. 245–252.

Rado, R. (1933). Studien zur Kombinatorik, *Mathematische Zeitschrift* **36**, pp. 424–480, `http://www.cs.umd.edu/~gasarch/vdw/vdw.html`. Includes Gallai's theorem and credits him.

Rado, R. (1943). Notes on combinatorial analysis, *Proceedings of the London Mathematical Society* **48**, pp. 122–160, `http://www.cs.umd.edu/~gasarch/vdw/vdw.html`. Includes Gallai's theorem and credits him.

Rankin, R. (1960—1961). Sets of integers containing not more than a given number of terms in an arithmetic progression, *Proceedings of the Royal Society of Edinburgh Sect. A 65* , pp. 332–344.

Shelah, S. (1988). Primitive recursive bounds for van der Waerden numbers, *Journal of the American Mathematical Society* **1**, pp. 683–697, `http://www.jstor.org/view/08940347/di963031/96p0024f/0`.

Witt, E. (1951). Ein kombinatorischer satz de elementargeometrie, *Mathematische Nachrichten* **6**, pp. 261–262, `http://www.cs.umd.edu/~gasarch/vdw/vdw.html`. Contains Gallai-Witt Theorem, though Gallai had it first so it is now called Gallai's theorem.