Partition Regularity for Linear Equations over $\mathbb{N}$

Daniel Glasscock, November 2010

This exposition grew from a talk I gave at Central European University in November, 2010. I've attempted to reach the broadest mathematical audience by giving examples and motivating the proofs. Besides van der Waerden's theorem and Ramsey's theorem, this note is self contained. A good general reference is [2].

Introduction

Partition regularity is a topic in Ramsey theory. The basic idea is to take a set, partition it into (usually finitely many) pieces, and try to say as much as possible about the pieces of the partition. Structure in the pieces is usually described using structure from the original set.

Let $\mathbb{N} = \{1, 2, 3, \ldots\}$ be the set of natural numbers. A partition of $\mathbb{N}$ into finitely many pieces will be called *a finite coloring of* $\mathbb{N}$ and will be described by a (coloring) function $c : \mathbb{N} \to \{1, 2, \ldots, r\}$. The utility of this is twofold: coloring emphasizes partitions (no number gets two colors), and the word *monochromatic* allows us to talk about a single color (piece of the partition) without specifying which one.

One of the earliest results in Ramsey theory was given by German mathematician Issai Schur in 1916:

Theorem (Schur, 1916). *[7] In any finite coloring of $\mathbb{N}$ there exists a monochromatic solution to the equation $x + y = z$.*

In terms of the general framework laid out in the first paragraph, Schur tells us that despite the complexity of the finite partition of $\mathbb{N}$, one of the pieces of the partition must contain a solution to the equation $x + y = z$. To put it another way, solutions to $x + y = z$ in $\mathbb{N}$ are *resilient* in the sense that they cannot be separated by any finite partition.

Another celebrated result in the history of Ramsey theory was that of Dutch mathematician Bartel Leendert van der Waerden in 1927:

Theorem (van der Waerden, 1927). *[5] In any finite coloring of $\mathbb{N}$ there exist arbitrarily long monochromatic arithmetic progressions.*

Van der Waerden gives us another type of resilient structure in $\mathbb{N}$, namely finite arithmetic progressions. If we color $\mathbb{N}$ with finitely many colors and fix some finite pattern (or *constellation*) in $\mathbb{N}$, then his result gives us that the pattern must arise in a single color.

We now have two results concerning the structure of at least one of the pieces of a finite partition of $\mathbb{N}$. We might see one as describing a resilient algebraic structure (solutions to $x + y = z$) and the other as describing a resilient geometric structure (finite arithmetic progressions). Our aim will be to generalize Schur's result in order to say more about when solutions to certain linear equations are resilient.

We take a moment to prove Schur's theorem. In search of a monochromatic solution to $x+y = z$, we will translate a finite coloring on $\mathbb{N}$ into a finite edge coloring on the complete graph on n vertices, K_n. We aim to do this in such a way that the existence of a monochromatic triangle will give us a monochromatic solution to $x + y = z$. From here we will be able to apply one of the fundamental results in Ramsey theory: if n is large enough, we are guaranteed to have a monochromatic triangle in the colored K_n.

Proof of Schur's theorem. Suppose $\mathbb{N}$ is colored with r colors via $c : \mathbb{N} \to \{1, \ldots, r\}$. For any $n \geq 1$, we identify the vertices of K_n with the integers $\{1, 2, \ldots, n\}$ and color the edge $\{i, j\}$ for $i \neq j$ with

$c(|i-j|)$. Ramsey's theorem gives that if n is large enough (larger than $R_r(3)$), then K_n is guaranteed to have a monochromatic triangle. Fix such a sufficiently large n.

A monochromatic triangle in K_n is exactly $1 \le i < j < k \le n$ such that the edges $\{i,j\}$, $\{i,k\}$, and $\{j,k\}$ have the same color. This means that $x = k-j$, $y = j-i$, and $z = k-i$ are three positive integers, all with the same color. We see immediately that $x+y=z$, as desired. $\square$

Though we will rely on van der Waerden's result later on, we will not prove it here. For a nice exposition, see *[6]*.

Partition regularity

Motivated by Schur's result, we define an equation $a_1x_1 + \cdots + a_nx_n = 0$ with $a_i \in \mathbb{Z}$ to be *partition regular over* $\mathbb{N}$, denoted $PR_{\mathbb{N}}$, if there is a monochromatic solution in any finite coloring of $\mathbb{N}$.

Schur's theorem is then exactly that the equation $x_1+x_2-x_3=0$ is partition regular over $\mathbb{N}$. We called the solutions to $x_1+x_2-x_3=0$ resilient; if $a_1x_1+\ldots+a_nx_n=0$ is $PR_{\mathbb{N}}$ then its solutions are resilient in the same way, that is they cannot be separated by finite partitions.

The most natural direction now is to determine which equations are $PR_{\mathbb{N}}$. Richard Rado classified $PR_{\mathbb{N}}$ equations (and much more) in his thesis in 1933:

Theorem (Rado, 1933). *[4] The equation $a_1x_1+\cdots+a_nx_n=0$ is $\mathrm{PR}_{\mathbb{N}}$ if and only if some non-empty subsum of the coefficients $\{a_1,\ldots,a_n\}$ is zero.*

The nonempty subsums of a set $A \subset \mathbb{Z}$ are

$$subsum(A) = \left\{ \sum_{x \in X} x \mid X \subset A \text{ nonempty} \right\}.$$

With this, Rado's result is that $a_1x_1+\cdots+a_nx_n=0$ is $PR_{\mathbb{N}}$ exactly when $0 \in subsum(\{a_1,\ldots,a_n\})$.

With this result, we see immediately that $x_1+x_2-x_3=0$ is $PR_{\mathbb{N}}$ since $0 = 1-1$ is a subsum of the coefficients. The related equation $x_1+x_2-2x_3=0$ is also $PR_{\mathbb{N}}$ since $0=1+1-2$. Actually, solutions to $x_1+x_2-2x_3=0$ are exactly 3 term arithmetic progressions, so van der Waerden's theorem also gives that $x_1+x_2-2x_3=0$ is $PR_{\mathbb{N}}$.

What about $x_1+x_2-3x_3=0$? Is it $PR_{\mathbb{N}}$? Since $0 \notin subsum(\{1,1,-3\})$, Rado's theorem gives that there exits a finite coloring of $\mathbb{N}$ in which there is *not* a monochromatic solution to $x_1+x_2-3x_3=0$. What might such a coloring look like?

Consider the following coloring $c_5 : \mathbb{N} \to \{1,2,3,4\}$. If n is coprime to 5, then define $c_5(n) = n \bmod 5$. If n is divisible by 5 but not by 25, then define $c_5(n) = n/5 \bmod 5$. In general, if 5 divides n evenly k times but not $k+1$ times, then we define $c_5(n) = n/5^k \bmod 5$.

Note that this coloring has the property that $c_5(5n) = c_5(n)$ for all $n \in \mathbb{N}$. This means that if we have a monochromatic solution $s_1+s_2-3s_3=0$ and the s_i's are all divisible by 5, then $\frac{s_1}{5}+\frac{s_2}{5}-3\frac{s_3}{5}=0$ is also a monochromatic solution.

Let's assume for a contradiction that such a monochromatic solution exists, i.e. that there exists $s_1, s_2, s_3 \in \mathbb{N}$ monochromatic (with respect to c_5) such that $s_1+s_2-3s_3=0$. By dividing out by

5 as many times as we can, we may assume that at least one of the s_i's is coprime to 5.

Now consider the equation $s_1 + s_2 - 3s_3 \equiv 0$ modulo 5. The fact that s_1, s_2, and s_3 are monochromatic is exactly that each is either equal to 0 modulo 5 or equal to a common value $s \in \{1,2,3,4\}$ modulo 5 which corresponds to their color. (Since at least one of the s_i's is coprime to 5, at least one is equal to s modulo 5.) Since s is coprime to 5, we may divide it out, yielding an equation of the form $a \equiv 0$ modulo 5 for some $a \in \mathit{subsum}(\{1,1,-3\})$. Since no non-empty subsum of $\{1,1,-3\}$ is equal to 0 or divisible by 5, this is a contradiction, and no such monochromatic solution can exist.

Proof of Rado's Theorem

To prove one direction of Rado's theorem, we will need a generalization of the c_5 coloring to any prime p. Define $c_p : \mathbb{N} \to \{1, \ldots, p-1\}$ as $c_p(n) = n/p^k \bmod p$ when p divides n evenly k times but not $k+1$ times. In the same way as c_5, $c_p(pn) = c_p(n)$ for all $n \in \mathbb{N}$.

Now we can prove that if $a_1x_1 + \cdots + a_nx_n = 0$ is $PR_{\mathbb{N}}$, then some non-empty subsum of the coefficients is 0.

Proof of Rado's theorem ($\Rightarrow$). Suppose for a contradiction that the equation $E : a_1x_1 + \cdots + a_nx_n = 0$ with $a_i \in \mathbb{Z}$ is $PR_{\mathbb{N}}$ and that no non-empty subsum of the coefficients is zero. Let p be a prime which does not divide any element of $\mathit{subsum}(\{a_1, \ldots, a_n\})$, and color $\mathbb{N}$ with the c_p coloring described above.

Since E is $PR_{\mathbb{N}}$, it has a monochromatic solution $a_1s_1 + \cdots + a_ns_n = 0$. By dividing the s_i's by common factors of p (which does not change the color of the solution), we may assume that one of the s_i's is not divisible by p.

Now consider the equation $a_1s_1 + \cdots + a_ns_n \equiv 0$ modulo p. Since the s_i's are monochromatic, they are each either equal to 0 modulo p or equal to a common value $s \in \{1, \ldots, p-1\}$ modulo p. By dividing out by s, we are left with an equation of the form $a \equiv 0$ modulo p for some $a \in \mathit{subsum}(\{a_1, \ldots, a_n\})$. This is a contradiction since $0 \notin \mathit{subsum}(\{a_1, \ldots, a_n\})$ and p does not divide any element of $\mathit{subsum}(\{a_1, \ldots, a_n\})$. $\square$

The other direction in Rado's theorem is a bit more difficult. The following argument helps to simplify understanding the approach. Suppose that 0 is a subsum of coefficients of the equation $a_1x_1 + \cdots + a_nx_n = 0$; by relabeling, we may assume $a_1 + \cdots + a_j = 0$ for some $1 \leq j \leq n$.

Now suppose that we have a solution $a_1s_1 + \cdots + a_ns_n = 0$. We see that for any $d \in \mathbb{Z}$, $a_1(ds_1) + \cdots + a_n(ds_n) = 0$ is still a solution. Since $a_1 + \cdots + a_j = 0$, we have that for any $c \in \mathbb{Z}$, $a_1(c+s_1) + \cdots + a_j(c+s_j) + a_{j+1}s_{j+1} + \cdots + a_ns_n = 0$ is also a solution. Combining these facts, we have that one solution - the s_i's - has generated a whole family of solutions in $\mathbb{N}$:

$$a_1(c + ds_1) + \cdots + a_j(c + ds_j) + a_{j+1}(ds_{j+1}) + \cdots + a_n(ds_n) = 0 \quad \forall\, c, d \geq 1$$

To prove Rado's theorem, we need to find a *monochromatic* solution given an arbitrary finite coloring of $\mathbb{N}$. We will be able to guarantee, with the right tool, that the family of solutions above contains a monochromatic solution. The right tool in this case is a corollary to van der Waerden's theorem.

Corollary (to van der Waerden). *Let $k, s \geq 1$ be integers. In any finite coloring of $\mathbb{N}$ there exist integers $c, d \geq 1$ such that the set*

$$\{c, c+d, c+2d, \ldots, c+(k-1)d, sd\}$$

is monochromatic.

This corollary allows us to find a monochromatic k-term arithmetic progression along with an arbitrary multiple s of the progression's jump size d in the same color. The "shape" of this set is very similar to the one we need for solutions to the equation above.

We are now ready to finish the proof of Rado's theorem. We need to show that if 0 is a subsum of the coefficients of an equation, then it has a monochromatic solution in any finite coloring of $\mathbb{N}$.

Proof of Rado's theorem ($\Leftarrow$). Let $\mathbb{N}$ be finitely colored and let $E : a_1x_1 + \cdots + a_nx_n = 0$ with $a_i \in \mathbb{Z}$ be such that $0 \in \mathit{subsum}(\{a_1, \ldots, a_n\})$. By relabeling, we may assume that $a_1 + \cdots + a_j = 0$ for some $1 \leq j \leq n$.

If $j = n$, then $a_1(1) + \cdots + a_n(1) = 0$ is a monochromatic solution, and we are done. Similarly if $j < n$ and $a_{j+1} + \cdots + a_n = 0$, then $x_i = 1$ for all $1 \leq i \leq n$ is still a monochromatic solution, and we are done. We may assume that we are in neither of these two cases.

Define $A = \gcd(a_1, \ldots, a_j)$ and $B = a_{j+1} + \cdots + a_n$; we may assume $j < n$ and $B \neq 0$ by the previous paragraph. Since $-BA$ is an integer multiple of $\gcd(a_1, \ldots, a_j)$, basic number theory tells us that there are $\lambda_i \in \mathbb{Z}$, $1 \leq i \leq j$ such that $-BA = \lambda_1 a_1 + \cdots + \lambda_j a_j$. If $l \in \mathbb{Z}$, then

$$-BA = \lambda_1 a_1 + \cdots + \lambda_j a_j = (\lambda_1 + l)a_1 + \cdots + (\lambda_j + l)a_j$$

since $a_1 + \cdots + a_j = 0$. Thus by shifting the λ_i's if necessary, we may assume that $\lambda_i \in \mathbb{N}$ for all $1 \leq i \leq j$.

Now a solution to E in $\mathbb{N}$ comes from rewriting the equation $-BA + BA = 0$ with the work above:

$$a_1(\lambda_1) + \cdots + a_j(\lambda_j) + a_{j+1}(A) + \cdots + a_n(A) = 0.$$

Following the argument outlined above, we will now use the corollary to van der Waerden's theorem to find a monochromatic subset of $\mathbb{N}$ in which we can find a solution to E. Let $k = 1 + \max_i \lambda_i$ and $s = A$; both are positive integers. Then the corollary to van der Waerden gives that there are integers $c, d \geq 1$ such that the set $\{c, c + d, c + 2d, \ldots, c + (k-1)d, sd\}$ is monochromatic.

Since k is larger than all of the λ's, we see that $c + \lambda_i d$ is the same color for all $1 \leq i \leq j$. We have finally that

$$a_1(c + \lambda_1 d) + \cdots + a_j(c + \lambda_j d) + a_{j+1}(Ad) + \cdots + a_n(Ad) = 0$$

is a monochromatic solution to E in $\mathbb{N}$, as desired. $\square$

Open problems

Despite having a nice classification of partition regular homogeneous equations over $\mathbb{N}$, there still remain a few natural loose ends. We will touch on two open problems; see *[3]* for a more in-depth discussion.

Problem 1. In the proof of Rado's theorem, we showed that $0 \in \mathit{subsum}(\{a_1, \ldots, a_n\})$ for any $PR_{\mathbb{N}}$ equation $E : a_1x_1 + \cdots + a_nx_n = 0$ by considering the c_p coloring for a certain prime p. We used the fact that E was $PR_{\mathbb{N}}$ to guarantee that it had a monochromatic solution in that specific c_p coloring.

What this means is that we could have significantly relaxed our assumptions about the regularity of E and still reached the same conclusion. For example, instead of assuming that E is $PR_{\mathbb{N}}$, we could have assumed that it has a monochromatic solution in each c_p coloring. An even weaker assumption would have been that E has a monochromatic solution in a c_p coloring for *some* prime p not dividing a subsum of its coefficients.

E has a monochromatic solution in *any*
finite coloring of $\mathbb{N}$ (E is $PR_{\mathbb{N}}$)
$\Downarrow$
E has a monochromatic solution in
the c_p coloring for every prime p
$\Downarrow$
E has a monochromatic solution in the c_p coloring for a prime
p not dividing any element of $subsum(\{a_1, \ldots, a_n\})$
$\Downarrow$
$0 \in subsum(\{a_1, \ldots, a_n\})$
$\Downarrow$
E is $PR_{\mathbb{N}}$

The last implication (by Rado's theorem) gives that the 4 statements above are equivalent. In some way, this means that partition regularity in the c_p colorings is enough to guarantee partition regularity in an arbitrary finite coloring.

This leads us to ask the following questions: *Is there a more direct proof that partition regularity in the c_p colorings yields general partition regularity? In what sense, if any, do the c_p colorings generate an arbitrary finite coloring?*

Problem 2 (Rado's boundedness conjecture). Take an equation $F : a_1x_1 + \cdots + a_nx_n = 0$ for which $0 \notin subsum(\{a_1, \ldots, a_n\})$, and let p be a prime not dividing any element of $subsum(\{a_1, \ldots, a_n\})$. We know of a $(p-1)$-coloring in which F does not have a monochromatic solution, namely c_p.

Note, however, that the number of colors we use to "split" the solutions of F $(p-1)$ defined in this way is not bounded above in terms of n, the number of variables of F. For example, let F_k be the equation $k!x_1 - 2x_2 + x_3 = 0$ for $k \in \mathbb{N}$. Rado's theorem gives that F_k is not $PR_{\mathbb{N}}$ for any $k \geq 3$. Each F_k has $n = 3$ variables but the corresponding prime p_k (one which does not divide any element of $subsum(\{k!, -2, 1\})$) is necessarily greater than k (which is independent of n).

What we'd like to be able to say is that n^2 or 2^n or even n^{n^n} colors is enough to be able to split the solutions of F. In other words, we'd like to find a bound in terms of n on the number of colors needed to split the solutions of an equation in n variables which is not $PR_{\mathbb{N}}$.

Rado put forward the following conjecture in 1933:

Rado's boundedness conjecture. *For any integer $n \geq 1$, there exists an integer $B(n) \geq 1$ such that if $a_1x_1 + \cdots + a_nx_n = 0$ ($a_i \in \mathbb{Z}$) is not* $\mathrm{PR}_{\mathbb{N}}$*, then there exists an r-coloring of $\mathbb{N}$, $r \leq B(n)$, in which F has no monochromatic solutions.*

The $n = 3$ case of this problem was recently solved by Jacob Fox and Daniel Kleitman. They show that if an equation $a_1x_1 + a_2x_2 + a_3x_3 = 0$ is not $PR_{\mathbb{N}}$, then there is a coloring of $\mathbb{N}$ with 24 colors or fewer in which the equation does not have a monochromatic solution. See *[1]* for the details.

The $n \geq 4$ case is open as of the time of writing this note.

References

[1] J. Fox, D. Kleitman, *Abstract On Rados Boundedness Conjecture* ,
http://www.math.ucsd.edu/ sbutler/seminar/FoxKleitman2.pdf

[2] R. Graham, B. Rothschild, J. Spencer, *Ramsey Theory*, Wiley, New York, 1990.

[3] N. Hindman, I. Leader, D. Strauss, *Open problems in partition regularity*, Comb. Prob. and Comp. **12** (2003), 571583.

[4] R. Rado, *Studien zur Kombinatorik*, Math. Z. **36** (1933), no. 1, 424-470.

[5] I. Schur, *Uber die Kongruenz* $x^m + y^m \equiv z^m \pmod{p}$, Jahresber Deutsch. Math. Verein. **25** (1916), 114-117.

[6] R. Swan, *van der Waerden's theorem on arithmetic progressions*, http://math.uchicago.edu/ swan/expo/vdW.pdf

[7] B. L. van der Waerden, *Beweis einer Baudet'schen Vermutung*, Nieuw Arch. Wiskunde **15** (1927), 212-216.