Ramsey Theory and Semigroup Colorings

Mathieu Rundstrom

McGill University

May, 2020

Introduction

Loosely speaking, Ramsey Theory is the study of objects for which, at a large enough scale, contain a certain amount of order. Starting at the basis of Ramsey Theory, we build a machinery up in the goal of understanding and proving Hindman's, and Hales and Jewett's theorems using the theory of ultrafilters. The latter theorem can be informally understood as for high enough dimension, a high dimensional generalization of tic-tac-toe cannot end in a draw. The structure and content of this paper follows the first chapters of Stevo Todorcevic's *Introduction to Ramsey Spaces* [3] quite closely. Finally, it cannot be overstated that this project was only made possible through the help and guidance of Prof. Marcin Sabok.

Table of Contents

Ultrafilters

It turns out that many Ramsey theoretic results can be deduce using set theory, specifically using a set of tools related to *ultrafilters*. The purpose of this section is to introduce ultrafilters, and give some intuition behind them. We note that the notion of an ultrafilter can be built up from the notion of either coideals or from filters. For the sake of conciseness, we only do the latter. Below we consider some set S with partial order $\leq$.

Definition 2.1 (Filter). *$\mathcal{F} \subseteq S$ is a* filter *if the following hold:*

- *$\mathcal{F}$ is nonempty*
- *For every $x, y \in \mathcal{F}$, there exists some $z \in \mathcal{F}$ such that $z \leq x$ and $z \leq y$.*
- *For every $x \in \mathcal{F}$ and $y \in S$, if $x \leq y$, then y is also in $\mathcal{F}$.*

We say a filter is *proper* if it is not equal to the whole set. Together with a notion of maximal, we obtain the definition of an ultrafilter, namely

Definition 2.2 (Ultrafilter). *$\mathcal{U} \subsetneq S$ is an* ultrafilter *if $\mathcal{U}$ is a filter and if $\mathcal{U}$ is not a proper subset of any other filter $\mathcal{F} \subsetneq S$.*

We note that an ultrafilter $\mathcal{U}$ on S may be considered a finitely additive measure on S, where every subset of S either has measure 0 or 1 depending on whether it belongs to $\mathcal{U}$ or not. Conversely, a non-trivial $0, 1$-valued finitely additive measure on the power set $P(S)$ of S induces an ultrafilter on S. A simple example of ultrafilters are so called *principal* ultrafilters. These ultrafilters are of form $\{X \subseteq I : k \in X\}$ for some $k \in I$, where

I is some index set, for example, $I = \mathbb{N}$. However, most ultrafilters that we are interested in are *nonprincipal*, i.e. are not of this form, and cannot even be described constructively. To understand this we give the proposition below. Note that for any set of sets, inclusion is a partial order, and recall the definition of a *chain*, a totally ordered subset of a partially ordered set. We use the following lemma due to Kuratowski and Zorn, equivalent to the axiom of choice [2]:

Lemma 2.3 (Kuratowski-Zorn)**.** *If P (together with $\leq$) is a partially ordered set such that every chain in P has an upper bound, then P has a maximal element.*

Proposition 2.4. *For every infinite set I there exists a nonprincipal ultrafilter $\mathcal{U} \subset P(I)$.*

Proof. We use the filter $\mathcal{F}$ of all subsets of I whose complement is finite, i.e. $\mathcal{F} = \{X \subseteq I : I \setminus X \text{ is finite}\}$. It is clear that this is a filter (it is often referred to as the Fréchet filter on I). Consider the set of filters $G = \{\mathcal{G} \subseteq P(I) : \mathcal{G} \text{ is a filter}, \mathcal{F} \subseteq \mathcal{G}\}$. This is a partially ordered set with respect to inclusion. Moreover, noting that the union of a chain of proper filters is a proper filter, any such chain has this union as an upper bound. Thus, we can apply the Kuratowski-Zorn lemma (2.3), and obtain a maximal filter $\mathcal{U}$ in G, thus $\mathcal{U}$ is an ultrafilter. □

A key property is that a given ultrafilter $\mathcal{U}$ can be associated with a quantifier $(\mathcal{U}x)$, where x ranges over elements in S. We mean by this that given a property $\varphi(x)$ of elements of S, we write $(\mathcal{U}x)\varphi(x)$ to signify $\{x \in S : \varphi(x)\} \in \mathcal{U}$. We make extensive use of this in our section concerning semigroup colorings. We now give some nice properties of such quantifiers, stating that these quantifiers interact with $\wedge, \vee$, and $\neg$ as one might wish.

Proposition 2.5.

(1) $(\mathcal{U}x)\varphi_0(x) \wedge (\mathcal{U}x)\varphi_1(x)$ is equivalent to $(\mathcal{U}x)(\varphi_0(x) \wedge \varphi_1(x))$.

(2) $(\mathcal{U}x)\varphi_0(x) \vee (\mathcal{U}x)\varphi_1(x)$ is equivalent to $(\mathcal{U}x)(\varphi_0(x) \vee \varphi_1(x))$.

(3) $\neg(\mathcal{U}x)\varphi(x)$ is equivalent to $(\mathcal{U}x)\neg\varphi(x)$.

An interesting concept is the power of an ultrafilter. Let $\mathcal{U}$ be an ultrafilter on $\mathbb{N}$. For a set S we denote by $S^{[k]}$ the set of k-element subsets, borrowing notation from Todorcevic, as we will for most notation hereinafter. For some positive integer k, $\mathcal{U}^k$ is an ultrafilter on $\mathbb{N}^{[k]}$, defined as follows

$$A \in \mathcal{U}^k \text{ iff } (\mathcal{U}_{x_0})(\mathcal{U}_{x_1})\dots(\mathcal{U}_{x_{k-1}})\{x_0, x_1, \dots, x_n\} \in A.$$

Ramsey Theory

A simple and well known example, often seen in a first encounter to Ramsey theory is the following:

Fact. *In any group of six people, either at least three people pairwise know each other, or at least three people pairwise do not know each other.*

A more general question here would be: given integers m, n what is the smallest integer k for which at least m people pairwise know each other, or at least n people do not? In other words: what is the smallest integer k such that any 2-coloring of the complete graph on K vertices K_k either contains an m-clique of color 1, or an n-clique of color 2? These numbers are usually denoted $R(m, n)$ and are known as Ramsey numbers. $R(m, n)$ is well defined for all natural numbers. Generalizations of this idea exists to hypergraphs and to forcing other subgraphs than cliques. We now give a generalization of the above number to a finite number of colors instead of just two, and the result for an infinite complete graph. The original proof of which was a corollary of the following infinite version of this, due to Frank Plumpton Ramsey:

Theorem 3.1 (Ramsey)**.** *Color the edges of a complete graph on an infinite number of vertices using a finite number of colors. Then there exists an infinite subset of vertices in which all edges between vertices in this subset have the same color.*

Proof. Note first that a partition of an infinite set into a finite number of sets must contain a infinite set (this is the statement of Ramsey's theorem for 1-graphs). Let the graph described have vertex set V_0, and consider an arbitrary vertex $v_0 \in V_0$. Clearly, some color c_0 is the color of infinitely many edges incident to v_0. Denote by V_1 the set of vertices joined by these edges. We can iterate the process above to obtain a vertex $v_1 \in V_1$, a color c_1 and another set of vertices V_2. Repeating this, we obtain sets an infinite sequence of sets $V_0, V_1, V_2, \ldots$, observing that $V_0 \supset V_1 \supset V_2 \supset \ldots$, with corresponding v_i and c_i (with colors possibly repeated). By construction, for $i < j$, an edge (v_1, v_j) is colored c_i. In the sequence of colors $c_1, c_2, \ldots$ there is at least one color, say c, repeated infinitely often. Define now $C = \{v_i : c_i = c\}$. We are done by applying the previous observation and noting that C is infinite. $\square$

Constructing an analogous sequence of sets for hypergraphs yields a more general result we state soon. Since the main point of view of this paper is that of set theory, we note that edges are just pairs of vertices, i.e. 2-element subsets; with the analogous observation for r-graphs, edges are r-element subsets. Now, theorem (3.1) more generally and in this new language

Theorem 3.2 (Ramsey). *For every positive integer k and every finite coloring of $\mathbb{N}^{[k]}$, there is an infinite subset M of $\mathbb{N}$ such that $M^{[k]}$ is monochromatic, i.e. all sets in M are colored the same color.*

This result can be applied to obtain some results on relations, inlcuding the Erdós-Rado theorem (3.5) we see soon. First, a definition:

Definition 3.3. *Fix k a positive integer. For a sequence $\rho \in \{<, =, >\}^{k \times k}$, we define a relation $R_\rho \subseteq N^k$by*

$$R_\rho(x) \text{ iff } (\forall (i,j) \in k \times k) x_i \rho(i,j) x_j.$$

Relations on $\mathbb{N}$ formed from disjunctions of relations of the above form are called *canonical k-ary relations*. With this in mind, Ramsey arrived at the following theorem:

Theorem 3.4. *For every positive integer k and every relation $S \subseteq \mathbb{N}^k$ there is an infinite subset M of $\mathbb{N}$ and a canonical k-ary relation R on $\mathbb{N}$ such that $S \cap M^k = R \cap M^k$.*

The proof consists of bounding the number of equivalence classes a certain equivalence relation E corresponding to S has, and then applying (3.2). Using this result, Erdós and Rado proved the following:

Theorem 3.5. *For every equivalence relation E on $\mathbb{N}^{[k]}$ there is an infinite subset M of* N *and an index set $I \subseteq \{0, 1, \ldots, k-1\}$ such that $E|M^{[k]} = E_I|M^{[k]}$.*

Proof. We define a relation R_E as follows

$$R_E = \{(x_0, \ldots, x_{2k-1}) \in \mathbb{N}^{2k} : \{x_0, \ldots, x_{2k-1}\}E\{x_0, \ldots, x_{2k-1}\}\}.$$

By (3.4), we have M and infinite subset of $\mathbb{N}$ and a sequence of symbols $\Sigma \subseteq \{<, =, >\}^{2k \times 2k}$ such that R_E is equal to the disjunction of relations R_ρ with $\rho \in \Sigma$. Let

$$I = \{i < k : (\forall \rho \in \Sigma \rho(i, k+i) = =\}.$$

Let $N \subset M$ such that between every two integers of N there is at least one integer of M. We show that $E|N^{[k]} = E_I|N^{[k]}$. Suppose $s, t \in N^{[k]}$ in increasing order agree on indices from I. We show that s and t are equivalent with respect to E. We do this done by induction on the cardinality of the set

$$D(s, t) = i < k : s_i \neq t_i,$$

i.e. the number of positions where the two sets or sequences differ. Clearly, if $D(s, t)$, we are done. Suppose now that $D(s, t) = \{i\}$ for some i. Since $i \neq I$, we must have some $\rho \in \Sigma$ where the relation between i and $k + i$ is not equality. By the choice of N, we can find u s.t. sEu and tEu. By transitivity, we conclude sEt. The case of $|D(s, t)| \geq 2$ is reduced to a previous case by choosing a minimal i and replacing the ith member of t by s_i to obtain t'. By transitivity and the induction hypothesis, the rest follows from the base case. The converse of the statement has a clear proof with appropriate choice of ρ. □

Results on Semigroup Colorings

In this section we define semigroups and give some nice properties of these. We use this and the prior notion of ultrafilters to arrive at our final destination: the theorem of Hindman, and the theorem of Hales and Jewett. First, however, a small tangent into (the axiomatization of) set theory.

4.1 Semigroup foundations

A *semigroup* is an algebraic structure; it is a generalization of a group without requiring the existence of an identity element or inverses. Formally:

Definition 4.1 (Semigroup). *A* semigroup $(S, \cdot)$ *is a set* S *together with an associative binary operation "*$\cdot$*", i.e. the following holds:*

$$\forall x, y, z \in S, (a \cdot b) \cdot c = a \cdot (b \cdot c).$$

Recall the notion of a *Hausdorff topology*: for any two distinct points there exist neighbourhoods of each which are disjoint from each other. For example, any metric space is Hausdorff with respect to the metric. With this in mind, we have the following definition:

Definition 4.2 (Compact semigroup). *A* compact semigroup *is a non-empty semigroup together with a Hausdorff topology for which the map*

$$x \mapsto xs$$

is continuous for all $s \in S$.

Definition 4.3 (Idempotent)**.** *An element x of a (compact) semigroup S is* idempotent *if $x^2 = x$.*

As a consequence of the Kuratowski-Zorn lemma (2.3), we have the following key result:

Lemma 4.4 (Ellis)**.** *Every compact semigroup S has an idempotent.*

Proof. By compactness and Zorn's Lemma, there is a minimal compact subsemigroup of S, say R. As R is non-empty, take $s \in R$. We have that Rs is a compact semigroup. Since it is closed under the semigroup operation, and since R is minimal, $R = Rs$. Let $P = \{x \in R : xs = s\}$. Then, since $s \in Rs$, P is non-empty, so $P = R$. Hence, $s \in P$, moreover, $s^2 = s$. $\square$

This lemma has some important consequences, as we see now and also later on. Fix a compact semigroup S.

Definition 4.5 (Left-ideal)**.** *A* left-ideal *of S is a nonempty subset I of S such that $SI \subseteq I$.*

Similarly, when $IS \subseteq I$, we say I is a *right-ideal*. In the context of compact semigroups, left-ideals are of greater interest than their counterpart. For example, note that for every $x \in S$, Sx is a closed left-ideal, so every minimal left-ideal is closed. If a left-ideal is minimal among all closed left-ideals, then it is also minimal among all left- ideals. Now, every closed ideal of S is a compact subsemigroup of S, so by(4.4) it contains idempotents. As we will now see, idempotents belonging to minimal left-ideals are rather special. We now define a transitive and antisymmetric order $\leq$ on S

$$x \leq y \text{ iff } xy = yx = x.$$

Since $x \leq x$ occurs if and only if $x^2 = x$, $\leq$ is a partial order on the idempotents of S. Here is a consequence of Ellis's lemma that we will use to prove the Infinite Hales-Jewett theorem (4.19).

Lemma 4.6. *If y is an idempotent and if I is a closed left-ideal, then the left-ideal Iy contains an idempotent x such that $x \leq y$.*

The proof below consist of simple algebraic manipulation and repeated application of the idempotent definition.

Proof. By (4.4), there is some idemoptent w in Iy. Let $v \in I$ satisfy $w = vy$, and let $x = yw$. Using the fact that y and w are idempotents, it follows that

$$x^2 = yvyyw = yvyw = yww = yw = x,$$

i.e. x is an idempotent. Observe also that

$$yx = yyw = yw = x,$$

and

$$xy = yvyy = yvy = x.$$

Hence, $x \leq y$ as desired. $\square$

From (4.6) we have two corollaries, noting that an idempotent belonging to a minimal left ideal is minimal with respect to $\leq$.

Corollary 4.7. *An idempotent is minimal if and only if it belongs to some minimal left-ideal.*

Corollary 4.8. *Any ideal which is both a left- and right-ideal of S contains all the minimal idempotents of S.*

4.2 A Connection to Ultrafilters: The Galvin-Glazer Theorem

The following theory can be stated mostly in terms of semigroups; however, it can also be stated in slightly more general terms without much increase in difficulty. Thus, we generalize the notion of a semigroup to the notion of a *partial semigroup*.

Definition 4.9 (Partial semigroup)**.** *A* partial semigroup *is a set S together with a partial map $*: S^2 \to S$ satisfying associativity, i.e. whenever one of $(x * y) * z$ and $x * (y * z)$ is defined, so is the other, and equality holds.*

We say that a partial semigroup is *directed* if for every finite sequence $(x_i)_{i=1}^n$ of elements of S there exists $y \in S$ such that y and x_i are distinct for all $i \in [n]$ and such that $x_i * y$ is defined for these i.

We now arrive at a connection to ultrafilters. Given a directed partial semigroup $(S, *)$, define γS to be the space of all ultrafilters cU on S satisfying

$$(\forall x \in S)\{y \in S : x \star y\} \in \mathcal{U}.$$

We can consider γS a nonempty closed subspace of the Čech-Stone compactification[1] βS, i.e. we consider it a compact Hausdorff space with the topology generated by the sets of form

$$\bar{A} = \{\mathcal{U} \in \gamma S : A \in \mathcal{U}\}$$

where $A \subseteq S$. We can extend the partial semigroup operation $*$ on S to a *total* operation $*$ on γS defined using ultrafilter quantifiers as

$$\mathcal{U} * \mathcal{V} = \{A \subseteq S : (\mathcal{U}x)(\mathcal{V}y)x * y \in A.$$

We have the following properties for $*$ on γS:

Lemma 4.10.

(1) $\mathcal{U} * \mathcal{V}$ whenever $\mathcal{U}, \mathcal{V} \in \gamma S$.

(2) $(\mathcal{U} * \mathcal{V}) * \mathcal{W} = \mathcal{U} * (\mathcal{V} * \mathcal{W})$

(3) For every $\mathcal{V} \in \gamma S$, the map $\mathcal{U} \mapsto \mathcal{U} * \mathcal{V}$ is continuous from γS into γS.

[1] Čech-Stone compactification is a property that holds in Tychonov spaces, however, the above is everything we need from this notion, and the theory behind it is beyond the scope of this paper, which is why we do not cover it in more detail.

The proof of these consist of applying definitions and manipulating ultrafilter quantifiers. Of importance to us are the following corollaries:

Corollary 4.11. *The space* $(\gamma S, *)$ *is a compact semigroup for every partial directed semigroup* $(S, *)$.

For clarity, we recall a definition from algebra:

Definition 4.12 (Left cancellative)**.** *An element* a *in a semigroup* $(S, \cdot)$ *is* left cancellative *if* $a \cdot b = a \cdot c$ *implies* $b = c$ *for all* $b, c \in S$. *If every element in* S *is left cancellative, we say* S *is left cancellative.*

With this terminology, we have the following:

Corollary 4.13. *For every directed partial semigroup*$(S, *)$ *that does not have idempotents itself or is left cancellative, there is a nonprincipal ultra filter* $\mathcal{U}$ *in* S *such that* $\mathcal{U} * \mathcal{U} = \mathcal{U}$.

We give a proof of the second corollary (4.13).

Proof. If $\mathcal{U} \in \gamma S$ idempotent ultrafilter that is principal, then $a * a = a$. Moreover, if $(S, *)$ is left cancellative, then $\gamma S \setminus S$ left ideal of γS. We can therefore apply Ellis's lemma (4.4) to obtain an idempotent. □

Definition 4.14 (Basic sequence)**.** *A finite or infinite sequence* $X = (x_n)$ *of elements of a directed partial semigroup* $(S, *)$ *is basic if its elements appear exactly once in the sequence and if* $x_{n_0} * x_{n_1} * \cdots * x_{n_k}$ *is defined for every* $n_0, \ldots, n_k$ *in the domain* $|X|$ *of* X.

Accordingly, if X is a basic sequence, we let

$$[x] = \{x_{n_0} * \cdots * x_{n_k} : k \in \mathbb{N}, n_0 < \cdots < n_k < |X|\}.$$

Observe that $([X], *)$ forms a directed partial semigroup whenever X is an infinite basic sequence of elements of S. With this observation, we are ready for the main theorem of this section, the Galvin-Glazer theorem.

Theorem 4.15 (Galvin-Glazer). *If $(S, *)$ is a partial semigroup that either has no idempotents or is left cancellative, then for every finite coloring of $(S, *)$, there is an infinite basic sequence $X = (x_n)$ of elements of S such that $[X]$ is monochromatic.*

Before proving (4.15), we give some of its direct consequences, i.e. Hindman's theorem, where the set $S = \mathbb{N}$ and the operation $*$ is addition. Basic in this context means simply that the sum (or sequence) has no repeated terms.

Theorem 4.16 (Hindman). *For every finite coloring of $\mathbb{N}$ there is an infinite sequence $X = (x_n)$ of elements of $\mathbb{N}$ such that the set of all finite nonrepeating sums $x_{n_0} + \cdots + x_{n_k}$ is monochromatic.*

Hindman's theorem (4.16) can be seen as a strong generalization of a theorem of Shur's, see (4.17). For the sake of curiosity, and for the sake of staying connected with the origins of Ramsey Theory, we state it and give a brief proof using Ramsey numbers.

Theorem 4.17 (Schur). *For every $k \geq 1$, there exists n so that in every k-coloring of $[n]$, one can find a monochromatic solution to $x + y = z$.*

Proof. Let $n = R_k(3, 3, \ldots, 3)$, where R_k denotes the Ramsey number on k colors. Let $G = K_n$ be the complete graph on $[n]$, i.e. $[n]^{[2]}$. For $x, y \in [n]$, color the edge (x, y) using the color $|x - y|$. By definition of Ramsey numbers, there exists $a, b, c \in [n]$ such that $|a - b|, |a - c|, |b - c|$ have the same color. We can assume w.l.o.g. that $a > b > c$. Now, let $x = a - b, y = b - c, z = a - c$, then $x + y = z$ is a monochromatic solution. $\square$

The above argument can be slightly generalized, for example, to sums of three terms, and using $R_k(4, 4, \ldots, 4)$. Regardless, we now return to the proof of (4.15).

Proof. Fix some coloring of S. Choose $\mathcal{U} \in \gamma S \setminus S$ an idempotent ultrafilter. Fix $P_0 \in \mathcal{U}$ a monochromatic set relative to the given coloring of S. By definition of $*$ on γS and the idempotence of $\mathcal{U}$

$$(\mathcal{U}x)(\mathcal{U}y) x * y \in P_0,$$

hence, we can choose $x_0 \in P_0$ such that

$$P_1 = \{y \in P_0 : x_0 * y \in P_0\} \in \mathcal{U}.$$

Repeating this procedure by choosing $x_1 \in P_1$ we obtain P_2. Further repeating this, we obtain new P_i with the respective x_i forming an infinite basic sequence $X = (x_n) \subseteq S$. We claim that X has the following property

Claim 4.18. *We have $x_{n_0} * x_{n_1} * \cdots * x_{n_k} \in P_{n_0}$ for every finite sequence of non-negative integers $n_0 < n_1 < \cdots < n_k$.*

Proof. We proceed by induction on k. The base case $x_{n_0} \in P_{n_0}$ holds by construction. Assume the claim holds for some k, and let $x = x_{n_0} * x_{n_1} * \cdots * x_{n_k}$. By the inductive hypothesis, $x \in P_{n_1}$, and since $n_1 \geq n_0 + 1$

$$x \in P_{n_1} \subseteq P_{n_0+1} = \{y \in P_{n_0} : x_{n_0} * y \in P_{n_0}.$$

Therefore $x_{n_0} * x \in P_{n_0}$. □

The above claim directly yields $[X] \subseteq P_0$, which concludes the proof. □

4.3 The Hales-Jewett Theorem(s)

We are now ready to tackle the Hales-Jewett theorem. Let $L = \cup_{n=0}^{\infty} L_n$ be an alphabet, which is decomposed into a chain of finite subsets L_n. Further, let v be a variable not present in L. Denote by W_L (or W) the set of *words*, i.e. finite strings over L, and denote by W_{Lv} the set of *variable-words* over L, i.e. words over $L \cup \{v\}$ containing V once at minimum. If $s = s[v] \in W(v)$ and $a \in L \cup \{v\}$, then we denote by $s[a]$ the (variable-)word obtained by replacing every occurrence of v in s by a. Observe that $s[a] \in W$ if $a \neq v$, and $s[a] \in W_{Lv}$ otherwise.

For a sequence $X = (x_0, x_1, \dots)$ over W_{Lv}, we let X_L, respectively X_{Lv}, denote the partial subsemigroup W_L, respectively W_{Lv}, generated by X defined in the following way:

$$[X]_L = \{x_{n_0}[\lambda_0]^\frown \dots ^\frown x_{n_k}[\lambda_k] \in W_L : n_0 < \dots < n_k, \lambda_i \in L_{n_i}, i \leq k\}$$

$$[X]_{Lv} = \{x_{n_0}[\lambda_0]^\frown \dots ^\frown x_{n_k}[\lambda_k] \in W_{Lv} : n_0 < \dots < n_k, \lambda_i \in L_{n_i} \cup \{v\}, i \leq k\}$$

We now state the Infinite Hales-Jewett Theorem; the finite version will be a clear corollary.

Theorem 4.19 (Infinite Hales-Jewett Theorem)**.** *For every finite coloring of $W_L \cup W_{Lv}$, there is an infinite sequence $X = (x_n)$ of elements of W_{Lv} such that the partial subsemigroups $[X]_L$ and $[X]_{Lv}$ are both monochromatic.*

The proof uses a similar idea to the one of Galvin-Glazer, noting that $S = W_L \cup W_{Lv}$ is a semigroup and considering the compactification $(\beta S, ^\frown)$.

Proof. We have the following subsemigroup of S^*

$$S_L^* = \{\mathcal{U} \in S^* : W_L \in \mathcal{U}\},$$

and a two-sided ideal

$$S^*_{Lv} = \{\mathcal{U} \in S^* : W_{Lv} \in \mathcal{U}\}.$$

Applying (4.6), we take $\mathcal{W}$ to be a minimal idempotent in S^*_L, and $\mathcal{V}$ a minimal idempotent in S^*_{Lv} satisfying $\mathcal{V} \leq \mathcal{W}$. Now, for each letter $\lambda \in L$, we have the corresponding substitution map $x \mapsto x[\lambda]$ from $W_L \cup W_{Lv}$ to W_L, which is the identity when restricted to W_L. This map extends to a map, a continuous homomorphism, $\mathcal{U} \mapsto \mathcal{U}[\lambda]$ from $S^*_L \cup S^*_{Lv}$ to S^*_L, which similarly is the identity on S^*_L.

Claim 4.20. *$\mathcal{V}[\lambda] = \mathcal{W}$ for all $\lambda \in L$.*

Proof. As noted above, $\mathcal{U} \mapsto \mathcal{U}[\lambda]$ is a homomorphism. Thus, $\mathcal{V}[\lambda]$ is an idempotent of S^*_L and $\mathcal{V}[\lambda] \leq \mathcal{W}[\lambda] = \mathcal{W}$. By definition, $\mathcal{W}$ is minimal in S^*_L, implying $\mathcal{V}[\lambda] = \mathcal{W}$. $\square$

Now, let P_v be the color of the given coloring that belongs to $\mathcal{V}$, and let P_W be the color which belongs to $\mathcal{W}$. Recursing on n, we now construct $X = (x_k)$, an infinite sequence of variable-words, as well as $\{P^n_W\}$ and $\{P^n_v\}$ infinite decreasing sequences of elements of $\mathcal{W}$ and $\mathcal{V}$, respectively, satisfying for all n the following four properties

$$(a)_n \quad x_n \in P^n_v,$$

$$(b)_n \quad \forall\lambda \in L_n \forall x \in P^n_v \; x[\lambda] \in P^n_W,$$

$$(c)_n \quad (\mathcal{V}_y)(\forall\lambda \in L_n \cup \{v\}) \; x_n[\lambda]^\frown y \in P^n_v,$$

$$(d)_n \quad (\mathcal{W}t) x_n{}^\frown t \in P^n_v.$$

First, let $P^0_W = P_W \cap W_L$, and $P^9_v = \{x \in P_v \cap W_{Lv} : \forall\lambda \in L_0 x[\lambda] \in P^0_W\}$. By (4.20), noting that $P^0_W \in \mathcal{W}$, P^0_v is a finite intersection of elements of $\mathcal{V}$, whence $P^0_v \in cV$. Thus

$$(\forall y \in L_0 \cup \{v\}) P^0_v \in \mathcal{V}[\lambda]^\frown \mathcal{V},$$

or equivalently

$$(\mathcal{V}x)(\mathcal{V}y)(\forall\lambda \in L_0 \cup \{v\}) \; x[\lambda]^\frown y \in P^0_v.$$

Similarly, $P_v^0 \in \mathcal{V} = \mathcal{V}^\frown\mathcal{W}$, i.e.

$$(\mathcal{V}x)(\mathcal{W}t)x\,^\frown t \in P_v^0.$$

Then, we can choose $x_0 \in P_v^0$ satisfying the above four properties for $n = 0$. Suppose now the four properties hold up to some n. Consider the $(n+1)$th step; by $(d)_n$

$$P_W^{n+1} = \{t \in P_W^n : x_n\,^\frown\, t \in P_v^n\} \in \mathcal{W}.$$

By $(c)_n$

$$Q_v^n = \{y \in P_v^n : (\forall \lambda \in L_v \cup \{v\})x_n[\lambda]^\frown y \in P_v^n\} \in \mathcal{V}.$$

Applying (4.20), we have

$$P_v^{n+1} = \{x \in Q_v^n : (\forall \lambda \in L_{n+1} x[\lambda] \in P_w^{n+1}\}$$

is in $\mathcal{V}$. Again, for $\mathcal{V}$-almost all choices x_{n+1} from P_v^{n+1} satisfy $(c)_{n+1}$ and $(d)_{n+1}$.

Claim 4.21.

(1) $x_{n_0}[\lambda_0]^\frown \ldots ^\frown x_{n_k-1}[\lambda_{k-1}]^\frown y \in P_v^{n_0}$ for every $k > 0, n_0 < \ldots, n_k, \lambda_i \in L_{n_i} \cup \{v\}, i < k, y \in P_v^{n_k}$.

(2) $x_{n_0}\,^\frown x_{n_1}[\lambda_1]^\frown \ldots ^\frown x_{n_k-1}[\lambda_{k-1}] \in P_v^{n_0}$ for every $k \geq 0, n_0 < \ldots, n_k, \lambda_i \in L_{n_i}, o < i \leq k$.

(3) $x_{n_0}[\lambda_0]^\frown \ldots ^\frown x_{n_k}[\lambda_k] \in P_W^{n_0}$ for every $k \geq 0, n_0 < \ldots, n_k, \lambda_i \in L_{n_i}, i \leq k$.

Proof. We proceed by induction on k. When $k = 1$, (1) follows from the previous claim, namely when going from n_0 to $n_0 + 1$, we have that $P_v^{n_1} \subseteq P_v^{n_0+1} \subseteq Q_v^{n_0}$. Now, assume (1) holds for k. In this spirit, let $n_0 < \cdots < n_{k+1}, \lambda_i \in L_{n_i} \cup \{v\}, i < k$, and $y \in P_v^{n_{k+1}}$. Define

$$y' = x_{n_1}[\lambda_1]^\frown \ldots ^\frown x_{n_k}[\lambda_k]^\frown y.$$

By the induction hypothesis, $y' \in P_v^{n_1}$, and by the previous claim, $y' \in P_v^{n_1} \subseteq P_v^{n_0+1} \subseteq Q_v^{n_0}$. Thus, $y = x_{n_0}[\lambda_0]^\frown y' \in P_v^{n_0}$. Letting $y = x_{n_k}$, we obtain (3) from (1). We have yet to prove

(2), so let $k \geq 0, n_0 < \ldots, n_k, \lambda_i \in L_{n_i}, o < i \leq k$, as in its statement. Further, let

$$t' = x_{n_1}[\lambda_1]^\frown \ldots ^\frown x_{n_k}[\lambda_k].$$

Applying (3), we have $t' \in P_W^{n_1}$, and again according to the stop going from n_0 to $n_0 + 1$, we have

$$t' \in P_W^{n_1} \subseteq P_W^{n_0+1} = \{t \in P_W^{n_0} : x_{n_0}{}^\frown t \in P_v^{n_0}\}.$$

We conclude (2), since $x_{n_0}{}^\frown t' \in P_v^{n_0}$ □

Having established the above properties, it remains to show that if $X = (x_n)$ is the infinite sequence produced above, then $[X]_L \subseteq P_W$ and $[X]_{Lv} \subseteq P_v$. The former follows from (3) just proven. To prove the latter, consider

$$x = x_{n_0}[\lambda_0]^\frown \ldots ^\frown x_{n_k}[\lambda_k],$$

with $n_0 < \ldots, n_k, \lambda_i \in L_{n_i} \cup \{v\}, i < k$, where at for at least one i, $\lambda_i = v$. Define l to be the maximum i less than or equal to k for which this is the case. By (2)

$$y = x_{n_l}[\lambda_l]^\frown x_{n_{l+1}}[\lambda_{l+1}]^\frown \ldots ^\frown x_{n_k}[\lambda_k] \in P_v^{n_l}.$$

From which it follows that

$$x = x_{n_0}^\frown \ldots ^\frown x_{n_{l-1}}[\lambda_{l-1}]^\frown y$$

satisfies the assumptions in (1). We finally conclude that $x \in P_v^{n_0} \subseteq P_v$, completing the proof. □

As a corollary, we have the finite version of the Heales-Jewett theorem.

Theorem 4.22 (Finite Hales-Jewett Theorem)**.** *For every finite alphabet L and positive integer k, there exists a positive integer n such that for every k-coloring of $W_L(n)$ of all words over L having length n, there is a variable-word x of length n such that the set $\{x[\lambda] : \lambda \in L\}$ is* *monochromatic.*

We include, for interest of the reader, an alternate proof of (4.22), without using the machinery developed in this paper. This proof was presented during a graduate course in combinatorics by Prof. Sergey Norin [1], we do not give all details for the sake of brevity, but the idea should still be clear. Below we restate the theorem in the language seen there.

Theorem 4.23. *For every r and t, there exists $d = HJ(t, r)$ such that if A is an alphabet with $|A| = t$ and A^d is colored in r colors, then there exists a monochromatic combinatorial line.*

We now see some new definitions that have previously seen analogs.

Definition 4.24 (Root). *A root τ is a word of length d in the alphabet $A \cup \{\star\}$, where $\star$ is a symbol not in A, containing at least one $\star$. The word $\tau(a)$ by substitution (analogously to the previously seen $x[\lambda]$ notation).*

Definition 4.25 (Combinatorial line). *A combinatorial line in A^d is a set $L_\tau = \{\tau(a) : a \in A\}$ where τ is a root of length d.*

Note that the notion of a combinatorial line can be seen in (4.22). We are now ready to give the proof.

Proof. Let $n = HJ(t, r)$. We will prove its existence by induction on t for fixed r. First, $HJ(1, r) = 1$, since combinatorial lines of length 1 are certainly monochromatic. Now, assume $n = HJ(t-1, r)$. Let $N_1 = r^{t^n}$ and define iteratively $N_i = r^{t^n + \sum^{i-1} N_j}$. We will show that $HJ(r, t) \leq N$, i.e. if χ is a coloring of A^N using r colors then χ contains a monochromatic combinatorial line.

We say $a, b \in A^n$ are neighbors if there exists some $i \leq n$ such that

$$a = a_1 a_2 \dots a_{i-1} 0 a_{i+1} \dots a_n,$$

and

$$b = a_1 a_2 \dots a_{i-1} 1 a_{i+1} \dots a_n.$$

Let τ be a root of length N such that $\tau = \tau_1 \dots \tau_n$, and τ_i of length N_i. Define

$$\tau(a) = \tau_1(a_1) \dots \tau_n(a_n).$$

Define χ_τ by $\chi_\tau(a) = \chi(\tau(a))$. We are not guaranteed that we have monochromatic combinatorial lines in A^n , but we will try to com- press to A^{n-1} where we do have such lines. Now, there exists roots $\tau_1, \dots, \tau_n$ with length as above and for neighbours a, b we have $\chi_\tau(a) = \chi_\tau(b)$, which can be seen by a clever application of the pigeonhole principle, made posible by our choices of N_i. Restrict χ_τ to $(A - \{0\})^n$ and denote this by χ'_τ. By induction hypothesis, we have a monochromatic combinatorial line, i.e. there exists $v = v_1 \dots v_n$ with $v_i \in (A - \{0\}) \cup \{\star\}$ such that $v(1), \dots, v(t-1)$ have the same color. We now want that the corresponding $\tau(v)$ is monochromatic in the original coloring. It remains to show that $\chi_\tau(v(0)) = \chi_\tau(v(1))$. In each position, $v(0)$ and $v(1)$ have either a 0 or a 1. Thus, after potentially changing positions iteratively, using the fact that the coloring acts identically on neighbours, we see they have the same coloring. □

The underlying connections between the ultrafilter techniques seen here and Ramsey theory are captivating, and are worth exploring further.

Bibliography

[1] HANSON, E. Lecture notes based on lectures by prof. sergey norin., April 2016.

[2] SUN, Y. R. Lecture notes based on lectures by prof. marcin sabok, December 2015.

[3] TODORCEVIC, S. *Introduction to Ramsey Spaces*. Annals of mathematics studies. Princeton University Press, 2010.