PARTITION THEOREMS FOR SPACES
OF VARIABLE WORDS

VITALY BERGELSON, ANDREAS BLASS, *and* NEIL HINDMAN

ABSTRACT

Furstenberg and Katznelson applied methods of topological dynamics to Ramsey theory, obtaining a density version of the Hales–Jewett partition theorem. Inspired by their methods, but using spaces of ultrafilters instead of their metric spaces, we prove a generalization of a theorem of Carlson about variable words. We extend this result to partitions of finite or infinite sequences of variable words, and we apply these extensions to strengthen a partition theorem of Furstenberg and Katznelson about combinatorial subspaces of the set of words.

1. *Introduction*

Our topic is Ramsey theory, that is, theorems asserting the existence of large, structurally rich homogeneous sets for partitions, in the context of words and variable words over a finite alphabet Σ. (A set is called homogeneous for a partition if it is a subset of one of the pieces of the partition. A variable word is like a word except that an additional symbol $v \notin \Sigma$, called a variable, must occur at least once.) Carlson [5] obtained some strong results of this sort. In particular, Theorem 2 of [5] implies as corollaries many standard results of Ramsey theory. Although Theorem 2 is expressed in terms of the Ellentuck topology, which we do not consider in this paper, its finitary Ramsey-theoretic corollaries all follow from a single Ramsey-theoretic statement, Lemma 5.9 of [5], the main lemma in the proof of Theorem 2.

We present a new proof of Carlson's Lemma 5.9, as a consequence of a stronger result, which

(a) deals with variable words and (ordinary non-variable) words together,

(b) produces homogeneous sets with additional strong combinatorial properties, related to the concept of central subset of $\mathbb{N}^+$ (a concept introduced in [7] and defined just after Corollary 2.7 in § 2 below),

(c) deals with located words, that is, words that may have gaps between the letters and before the first letter, and

(d) establishes homogeneity for concatenations of non-consecutive words in a string.

The technique we use to obtain this and other results is inspired by Furstenberg's and Katznelson's approach [8] to Ramsey theory via topological dynamics. As in [2] and [5], we use ultrafilters and algebraic operations on them to represent dynamical concepts in a relatively simple form.

After this introductory section, we shall devote two sections to developing the necessary ultrafilter machinery. Then, in § 4, we prove the result described above and derive several consequences, including Carlson's Lemma 5.9. In § 5, we extend the result of § 4 to deal with finite sequences of (variable) words instead of single (variable) words, and in § 6 we extend it further to deal with infinite

sequences provided the partitions are reasonably well behaved. Finally, in § 7, we apply some of the preceding results to partitions of the finite-dimensional combinatorial subspaces of an infinite-dimensional space of words. Because we use an extensive amount of terminology, we include an index at the end of the paper.

In the remainder of this introduction, we shall explain some of the notation and terminology that we shall use, and we shall state Carlson's Lemma 5.9 and a related result involving concatenation of non-consecutive words in a sequence.

The set $\mathbb{N}^+$ is the set of positive integers; $\mathbb{N} = \mathbb{N}^+ \cup \{0\}$ is the set of natural numbers. Infinite sequences have their terms indexed by $\mathbb{N}$, that is, they are functions with domain $\mathbb{N}$. Finite sequences have terms indexed by a proper initial segment of $\mathbb{N}$.

As mentioned above, Σ is a finite, non-empty alphabet; we write σ for its cardinality, and we identify it with $\{1, 2, \ldots, \sigma\}$ whenever convenient. A *word* over Σ is a non-empty, finite sequence of elements of Σ, and $W(\Sigma)$ is the set of these words. The set $W(\Sigma)$ with the operation of concatenation (written $^\frown$ when necessary but usually simply indicated by juxtaposition) is a semigroup, the free semigroup on Σ. A *variable* can be any object $v \notin \Sigma$; usually, we shall deal with a single variable and denote it by v, but occasionally we shall need a list of distinct variables, and we shall denote them by $v_0, v_1, v_2, \ldots$. A *variable word* over Σ is a word over $\Sigma \cup \{v\}$ in which v actually appears, so the set of variable words is

$$W(\Sigma; v) = W(\Sigma \cup \{v\}) \backslash W(\Sigma).$$

A *d-variable word* over Σ (for $d \in \mathbb{N}$) is a word over $\Sigma \cup \{v_0, v_1, \ldots, v_{d-1}\}$ in which each of $v_0, v_1, \ldots, v_{d-1}$ actually occurs and all occurrences of v_i precede all occurrences of v_{i+1} (for $i = 0, 1, \ldots, d-2$). We write $W(\Sigma; v_0, v_1, \ldots, v_{d-1})$ for the set of d-variable words over Σ.

When only one variable v and only one alphabet Σ are under consideration, we shall refer to words over $\Sigma \cup \{v\}$ simply as *words*, and we shall refer to words over Σ as *constant words*. Thus, the words that are not constant words are exactly the variable words.

If w is a variable word and $a \in \Sigma \cup \{v\}$ then we write $w(a)$ for the result of replacing v by a in w. Thus $w(v) = w$ and, for $a \in \Sigma$, $w(a) \in W(\Sigma)$. Any $w(a)$ is called an *instance* of w, a *constant* instance if $a \in \Sigma$.

If $\vec{s} = (s_0(v), s_1(v), \ldots)$ is an infinite sequence of variable words, then a *reduction* of $\vec{s}$ is any infinite sequence of words or variable words obtained from $\vec{s}$ by replacing each $s_i(v)$ by one of its instances, dividing the resulting sequence (of words over $\Sigma \cup \{v\}$) into (infinitely many) finite blocks of consecutive words, and concatenating the members of each block. More formally, $\vec{t}$ is a reduction of $\vec{s}$ if there are $a_0, a_1, a_2, \ldots \in \Sigma \cup \{v\}$ and integers $n_0 < n_1 < n_2 < \ldots$ with $n_0 = 0$ such that, for all i,

$$t_i = s_{n_i}(a_{n_i})^\frown s_{n_i+1}(a_{n_i+1})^\frown \cdots ^\frown s_{n_{i+1}-1}(a_{n_{i+1}-1}).$$

For example, if $\vec{s}$ were $(avbv, v, bav, abvb, vavv, \ldots)$ then $\vec{t}$ might be $(aabavbab, ababbabb, \ldots)$. A reduction is called a variable reduction or a constant reduction if all of its terms are variable words or constant words, respectively; there are of course also mixed reductions, like the example above, where some but not all of the terms are variable words.

A reduction $\vec{t}$ of $\vec{s}$ can be specified by listing the elements $a_i \in \Sigma \cup \{v\}$ that are substituted for v in the various components s_i and specifying which of the resulting

words are to be concatenated. A convenient way to do this is by means of a sequence $\vec{w}$ of words, in which the nth component lists the a_i used in forming the nth component of $\vec{t}$. Thus, in the notation of the definition above,

$$w_i = a_{n_i} a_{n_i+1} \ldots a_{n_{i+1}-1},$$

and in the example following the definition, $\vec{w} = (avb, ab, \ldots)$. We express this relationship between $\vec{s}$, $\vec{t}$ and $\vec{w}$ by writing $\vec{t} = \vec{s}[\vec{w}]$.

By a *reduced word* of the sequence $\vec{s}$, we mean any word over the alphabet $\Sigma \cup \{v\}$ obtainable by taking a finite initial segment of $\vec{s}$, replacing each of its terms by an instance, and concatenating the results; that is, a reduced word of $\vec{s}$ has the form $s_0(a_0)\frown s_1(a_1)\frown \ldots \frown s_k(a_k)$ for some k and some $a_0, a_1, \ldots, a_k \in \Sigma \cup \{v\}$. In this situation, we denote the word $s_0(a_0)\frown s_1(a_1)\frown \ldots \frown s_k(a_k)$ by $\vec{s}[w]$ where w is the word $a_0 a_1 \ldots a_k$. Notice that $\vec{s}[w]$ is a variable word if and only if w is, that is, if and only if at least one of the a_i is v; in this case we call $\vec{s}[w]$ a *variable reduced word* of $\vec{s}$, and otherwise we call it a *constant reduced word*. Notice that a word is a reduced word of $\vec{s}$ if and only if it occurs as the first term in some reduction of $\vec{s}$.

THEOREM 1.1 (Lemma 5.9 of [5]). *Let $W(\Sigma; v)$ be partitioned into finitely many pieces, and let $\vec{s}$ be an infinite sequence from $W(\Sigma; v)$. Then $\vec{s}$ has a variable reduction $\vec{t}$ such that all variable reduced words of $\vec{t}$ lie in the same piece of the partition.*

By an *extraction* of $\vec{s}$, we mean a reduction of some subsequence of $\vec{s}$. Similarly, by an *extracted word* of $\vec{s}$, we mean a reduced word of some subsequence of $\vec{s}$. Thus our description of reduced words above becomes a description of extracted words if we replace 'initial segment' with 'subsequence'. Similarly, our description of how to obtain a reduction becomes a description of extractions if we replace 'dividing the resulting sequence into finite blocks of consecutive words' with 'choosing, in the resulting sequence, an infinite sequence of blocks of not necessarily consecutive terms, such that each of the chosen blocks ends before the next one begins'.

The general result in §4, from which we shall deduce Theorem 1.1, also implies the following 'extraction' analogue.

THEOREM 1.2. *Let $W(\Sigma; v)$ be partitioned into finitely many pieces, and let $\vec{s}$ be an infinite sequence from $W(\Sigma; v)$. Then $\vec{s}$ has a variable extraction $\vec{t}$ such that all variable extracted words of $\vec{t}$ lie in the same piece of the partition.*

Theorem 1.2 produces a $\vec{t}$ with a stronger homogeneity property than in Theorem 1.1, but to obtain such a $\vec{t}$ we must take an extraction of $\vec{s}$; a reduction will not do. (To see why it will not do, consider $\vec{s} = (av, bv, bv, bv, \ldots)$ and partition the variable words according to whether the first letter is a or not.)

To specify extracted words and extractions of $\vec{s}$ by means of words w and sequences of words $\vec{w}$, analogously to our specification of reduced words and reductions above, we must supplement each letter in w or in any w_i with an indication of which $s_j(v)$ it is to be substituted into, since some of the s_j may be skipped. Such a w, with an increasing sequence of integers j attached to its letters, can be viewed as a located word in the sense of the following definition.

A *located word* over Σ is a function from a finite non-empty subset of $\mathbb{N}$ into Σ. We define $L(\Sigma)$ to be the set of located words over Σ. The set of *variable located words* over Σ is

$$L(\Sigma; v) = L(\Sigma \cup \{v\})\backslash L(\Sigma).$$

Notice that a word, being a function on a finite initial segment of $\mathbb{N}$, is also a located word. As with words, we adopt for located words the convention that, when only one variable v and only one alphabet Σ are under discussion, then a *located word* is a located word over $\Sigma \cup \{v\}$, and a *constant located word* is a located word over Σ. We also adopt the following (mnemonic) notation for the sets of constant, variable, and all located words:

$$C = L(\Sigma),$$

$$V = L(\Sigma; v) = L(\Sigma \cup \{v\})\backslash L(\Sigma),$$

and

$$A = L(\Sigma \cup \{v\}) = C \cup V.$$

If $\vec{s}$ is a sequence of variable words and w is a located word, then we denote by $\vec{s}[w]$ the extracted word of $\vec{s}$ obtained by taking the terms $s_j(v)$ in $\vec{s}$ for $j \in \mathrm{Domain}(w)$, instantiating by the values of w to produce $s_j(w(j))$, and concatenating the results in order of increasing j. All extracted words of $\vec{s}$ arise in this way. If $\mathrm{Domain}(w)$ is an initial segment of $\mathbb{N}$, so w is a word, then this agrees with our previous use of the notation $\vec{s}[w]$ for reduced words. If $\vec{w}$ is an *orderly* sequence of located words, that is, a sequence of located words such that all elements of the domain of w_i precede all elements of the domain of w_{i+1} for all i, then we can analogously define an extraction $\vec{s}[\vec{w}]$ of $\vec{s}$, and all extractions of $\vec{s}$ are of this form.

The concepts of reduced word, reduction, extracted word, and extraction, which we applied above to sequences $\vec{s}$ of variable words, can also be applied to orderly sequences of variable located words. The only change needed is that concatenation of words is replaced by union of located words, where located words, being functions, are regarded as sets of ordered pairs so that union makes sense. Notice that the orderliness of $\vec{s}$ implies that its components have disjoint domains, so these unions are again functions and therefore located words. We use the previously introduced notation $w(a), \vec{s}[w]$, and $\vec{s}[\vec{w}]$ for located words as well.

Our main partition theorems will be about located words. They can be used to prove theorems about ordinary words by 'forgetting the locations', that is, by replacing the domain of a located word with the initial segment of $\mathbb{N}$ of the same length.

Located words relate nicely to positive integers. If we identify Σ with $\{1, 2, \ldots, \sigma\}$, then a located word w can be viewed as the base $\sigma + 1$ expansion of a positive integer (with places not in $\mathrm{Domain}(w)$ being filled with zeros),

$$\sum_{n \in \mathrm{Domain}(w)} w(n) \cdot (\sigma + 1)^n.$$

This connection with $\mathbb{N}^+$ will allow us to relate our partition theorems to concepts defined in the context of $\mathbb{N}^+$, particularly the concept of centrality.

2. *Partial semigroups and ultrafilters*

We wish to apply, in the context of located words, certain concepts and constructions that have been developed in the theory of discrete semigroups and their Stone–Čech compactifications. Unfortunately, the natural binary operation on located words, namely to regard them as sets of ordered pairs and to take their union, is well defined only when applied to located words with disjoint domains; otherwise the union may fail to be a function. (See [17] for a treatment of this situation in a very general context.) In fact, we shall want to restrict further the domain of our operation on located words, so that it behaves well with respect to the operation of concatenation on ordinary words, that is, so that forgetting the locations is a homomorphism.

Although located words, with this restricted operation of union, do not constitute a semigroup, we shall see that much of the theory of Stone–Čech compactifications of discrete semigroups can be carried over to this context and similar ones. Since we shall have to deal with several such contexts and with connections between them, it is worthwhile to set up this machinery in some generality, and that is the purpose of the present section.

A *partial semigroup* is a set S together with an operation $*$ that maps a subset D of $S \times S$ into S and satisfies the associative law

$$(x * y) * z = x * (y * z)$$

in the sense that if either side is defined then so is the other and they are equal.

Any partial semigroup S can be made into a semigroup by adjoining a new element $\perp$ (to be thought of as 'undefined') and extending the operation $*$ by giving it the value $\perp$ wherever it was not previously defined. This construction lets us reduce some questions about partial semigroups to questions about semigroups, but it seems not to behave well in connection with some of the concepts (for example, minimal ultrafilters) that we shall need to use.

Our primary example of a partial semigroup is the set $L(\Sigma)$ of located words, defined in § 1 as functions into Σ with non-empty finite domains which are subsets of $\mathbb{N}$. Its operation $*$ is defined by

$$u * w = u \cup w \quad \text{if } \max(\text{Domain}(u)) < \min(\text{Domain}(w)),$$

but with $u * w$ undefined otherwise, where the functions u and w are regarded as sets of ordered pairs in forming $u \cup w$.

Had we written 'Domain$(u) \cap$ Domain$(w) = \varnothing$' instead of the stronger condition $\max(\text{Domain}(u)) < \min(\text{Domain}(w))$, we would have another partial semigroup (an oid in the sense of [17]), but it would not behave as well in connection with concatenation of (unlocated) words (see Example 2.5 below).

Of course, any semigroup is also a partial semigroup; we shall have occasion to use the semigroups $(\mathbb{N}^+, +)$ and $(W(\Sigma), \frown)$. The partial semigroup obtained from $(\mathbb{N}^+, +)$ by restricting addition to pairs of numbers that can be added in base $\sigma + 1$ without carrying has played at least an implicit role in previous work on the Ramsey theory of $(\mathbb{N}^+, +)$.

A partial semigroup $(S, *)$ is *(left) adequate* if, for any finitely many elements $x_1, x_2, \ldots, x_k \in S$, there exists $y \in S$ such that all of the products $x_1 * y, x_2 * y, \ldots, x_k * y$ are defined. All the examples of partial semigroups mentioned above are adequate.

A *homomorphism* from one partial semigroup $(S, *)$ to another $(S', *')$ is a function $f\colon S \to S'$ such that, whenever $x * y$ is defined in S, then $f(x) *' f(y)$ is defined in S' and equals $f(x * y)$.

EXAMPLES. (2.1) If $S = S'$ and $*$ is the restriction of $*'$ to part of its domain, then the identity map is a homomorphism from $(S, *)$ to $(S', *')$ (but not in the other direction).

(2.2) If $\Sigma \subseteq \Sigma'$ then the inclusion $L(\Sigma) \hookrightarrow L(\Sigma')$ is a homomorphism.

(2.3) If v is a variable (that is, $v \notin \Sigma$) then for each $a \in \Sigma$ the substitution function $\hat{a}\colon L(\Sigma \cup \{v\}) \to L(\Sigma)$ sending $w(v)$ to $w(a)$ is a homomorphism.

(2.4) Reading located words over $\{1, 2, ..., \sigma\}$ as base $\sigma + 1$ expansions of numbers yields a homomorphism

$$(L(\Sigma), *) \to (\mathbb{N}^+, +)\colon w \mapsto \sum_{n \in \text{Domain}(w)} w(n) \cdot (\sigma + 1)^n,$$

even if we restrict addition in $\mathbb{N}^+$ to the non-carrying situation.

(2.5) Let c be the 'compression of domains' or 'forgetting locations' function from located words to (unlocated) words. Formally,

$$c(w)(i) = w(\text{the } i\text{th element of Domain}(w))$$

(where we start counting with the 0th element). Then c is a homomorphism from $(L(\Sigma), *)$ to $(W(\Sigma), \hat{})$. (This would not be true if we had used 'Domain$(u) \cap$ Domain$(w) = \varnothing$' instead of 'max(Domain(u)) < min(Domain(w))' in defining $*$.)

We use the usual notation βS for the set of ultrafilters on S, identifying the points of S with the principal ultrafilters. If $(S, *)$ is a partial semigroup, we write δS (or $\delta(S, *)$ if necessary) for the set of those ultrafilters $\mathcal{U} \in \beta S$ such that, for every element s of S, $\{t \in S \mid s * t \text{ is defined}\} \in \mathcal{U}$. The standard existence proof for ultrafilters (which is given in [11] as Theorem 6.5) shows that δS is non-empty if and only if the family of sets $\{t \in S \mid s * t \text{ is defined}\}$, as s varies over S, has the finite intersection property, that is, if and only if $(S, *)$ is adequate. This was the motivation for the definition of adequate.

Before proceeding with a discussion of ultrafilters on partial semigroups, we introduce some notation that will facilitate this discussion (and is also useful in other contexts involving ultrafilters). To any ultrafilter $\mathcal{U}$ on S, we associate the quantifier 'for $\mathcal{U}$-almost all $s \in S$', which we abbreviate as $(\mathcal{U}s)$. That is, for any predicate P defined on S,

$$(\mathcal{U}s) P(s) \quad \Leftrightarrow \quad \{s \in S \mid P(s)\} \in \mathcal{U}$$
$$\Leftrightarrow \quad (\exists X \in \mathcal{U})(\forall s \in X) P(s).$$

Thus, if one thinks of $\mathcal{U}$ as a (finitely additive, two-valued) measure, then $(\mathcal{U}s) P(s)$ means that $P(s)$ holds for all but a measure-zero set of s. That $\mathcal{U}$ is an ultrafilter means that this quantifier $(\mathcal{U}s)$ commutes with all propositional connectives (for example, not, and, or, implies). The quantifier $(\mathcal{U}s)$ completely determines the ultrafilter $\mathcal{U}$, since

$$X \in \mathcal{U} \quad \Leftrightarrow \quad (\mathcal{U}s) s \in X,$$

and every quantifier that commutes with all connectives is of the form $(\mathcal{U}s)$. (By a quantifier Q over a set S, we mean any function from the power set of S into

{true, false}, that is, transforming propositional functions $P(s)$ into truth values $(Qs)P(s) = Q(\{s|\ P(s)\}) \in \{\text{true, false}\}$.) Thus, one can define an ultrafilter $\mathcal{U}$ by specifying the associated quantifier and verifying that it commutes with connectives.

For example, if $\mathcal{U}$ is an ultrafilter on S and if $f: S \to S'$, then we define the ultrafilter $f(\mathcal{U})$ on S' by

$$(f(\mathcal{U})s')\, P(s') \quad \Leftrightarrow \quad (\mathcal{U}s)\, P(f(s)),$$

which commutes with connectives because $(\mathcal{U}s)$ and substitution do. Of course, this definition in terms of quantifiers is equivalent to the familiar definition

$$f(\mathcal{U}) = \{X \subseteq S'|\ f^{-1}(X) \in \mathcal{U}\}$$

of the continuous extension of $f: S \to S'$ to $f: \beta S \to \beta S'$, but the quantifier form will be more convenient in calculations.

Returning to the discussion of ultrafilters on adequate partial semigroups, we note that δS consists of the ultrafilters $\mathcal{U}$ such that $(\forall s)(\mathcal{U}t)$ ($s * t$ is defined). That is, for each s, $s * t$ is almost always defined.

The partial semigroup operation $*$ on S induces a (total) operation, which we denote with the same symbol $*$, on δS, defined in terms of the associated quantifiers by

$$((\mathcal{U} * \mathcal{V})s)\, P(s) \quad \Leftrightarrow \quad (\mathcal{U}s)(\mathcal{V}t)\, [s * t \text{ is defined and } P(s * t)].$$

(An equivalent formulation not using ultrafilter quantifiers is that a subset A of S belongs to $\mathcal{U} * \mathcal{V}$ if and only if $\{s \in S|\ -s * A \in \mathcal{V}\} \in \mathcal{U}$, where $-s * A = \{t \in S|\ s * t \text{ is defined and is in } A\}$.) To see that this quantifier defines an ultrafilter $\mathcal{U} * \mathcal{V}$ on S, we must check that it commutes with negation and conjunction (as the other connectives are definable from these). The case of conjunction is easy, for arbitrary ultrafilters $\mathcal{U}$ and $\mathcal{V}$. For negation, we need to use that $\mathcal{V} \in \delta S$ (though $\mathcal{U}$ can be arbitrary). We have

$\neg(\mathcal{V}t)\, [s * t \text{ is defined and } P(s * t)]$

$\qquad \Leftrightarrow \quad (\mathcal{V}t)\,\neg[s * t \text{ is defined and } P(s * t)]$

$\qquad \Leftrightarrow \quad (\mathcal{V}t)\, ([s * t \text{ is defined}] \text{ and } \neg[s * t \text{ is defined and } P(s * t)])$

$\qquad \Leftrightarrow \quad (\mathcal{V}t)\, [s * t \text{ is defined and } \neg P(s * t)],$

where the second step used $(\mathcal{V}t)\, [s * t \text{ is defined}]$, that is $\mathcal{V} \in \delta S$, and the first two steps used that $(\mathcal{V}t)$ commutes with negation and conjunction. Now using that $(\mathcal{U}s)$ commutes with negation, we find that $((\mathcal{U} * \mathcal{V})s)$ does also.

PROPOSITION 2.6. *The operation $*$ defined above makes δS a compact right-topological semigroup, under the topology inherited from βS.*

The term 'right-topological' means that, for each fixed $\mathcal{V}$, $\mathcal{U} * \mathcal{V}$ is a continuous function of $\mathcal{U}$. (Some people, including one of the authors, would call this left-topological, but not in this paper. Also, some people, including another of the authors, would define $\mathcal{U} * \mathcal{V}$ with the order of the quantifiers reversed, but not in this paper. Still another of the authors would not define $\mathcal{U} * \mathcal{V}$ in this way at all but rather as a convolution of finitely additive measures, but again not in this paper.)

Proof. First we check that $\mathcal{U} * \mathcal{V}$ is in δS if $\mathcal{U}$ and $\mathcal{V}$ are. We have, for any $r \in S$,

$$((\mathcal{U} * \mathcal{V})s)\,[r * s \text{ is defined}] \quad \Leftrightarrow \quad (\mathcal{U}s)(\mathcal{V}t)\,[s * t \text{ and } r * (s * t) \text{ are defined}]$$
$$\Leftrightarrow \quad (\mathcal{U}s)(\mathcal{V}t)\,[s * t \text{ and } (r * s) * t \text{ are defined}],$$

where the last equivalence uses the associativity of $*$. But the last formula in this chain of equivalences is true because $\mathcal{U}, \mathcal{V} \in \delta S$; therefore so is the first, which means that $\mathcal{U} * \mathcal{V} \in \delta S$.

The associativity and right-continuity of $*$ on δS are proved exactly as for βS when S is a (non-partial) semigroup; see [11]. (In fact, the 'adjoin $\perp$' construction reduces the partial semigroup case to the semigroup case.) Finally, δS is compact because it is closed in the compact space βS, being the intersection of the closures of the sets $\{t \mid s * t \text{ is defined}\}$ for all $s \in S$.

The following corollary summarizes some facts from the general theory of compact, right-topological semigroups [2, 3, 8] that we shall need for δS. A *left ideal* of δS is a subset I such that $\mathcal{U} * \mathcal{V} \in I$ for all $\mathcal{U} \in \delta S$ and $\mathcal{V} \in I$. Right ideals and two-sided ideals are defined similarly. We define a partial ordering $\leqslant$ on the set of idempotents in δS by

$$\mathcal{U} \leqslant \mathcal{V} \quad \Leftrightarrow \quad \mathcal{U} * \mathcal{V} = \mathcal{V} * \mathcal{U} = \mathcal{U}.$$

COROLLARY 2.7. (1) *Every closed (equivalently, compact) sub-semigroup of δS (in particular δS itself) contains an idempotent.*

(2) *Every left ideal includes a minimal left ideal.*

(3) *Every minimal left ideal is closed.*

(4) *An idempotent belongs to some minimal left ideal if and only if it is minimal under the ordering $\leqslant$. (Such an idempotent is called minimal.)*

(5) *For every idempotent $\mathcal{U}$, there is a minimal idempotent $\mathcal{V} \leqslant \mathcal{U}$.*

(6) *Every minimal left ideal is included in every two-sided ideal.*

We call a subset of S *central* if it belongs to some minimal idempotent in δS. For (non-partial) semigroups, S, this agrees with the definition of centrality in [2]. See [7] for information about uses of centrality (in $\mathbb{N}^+$) in Ramsey theory. Our partition theorems will produce central homogeneous sets of located words (that is, central sets included in a single piece of a given partition). To relate these results to centrality in the more familiar context of $\mathbb{N}^+$, we need information about the behaviour of central sets and minimal idempotents under homomorphisms, particularly the 'base $\sigma + 1$' homomorphism of Example 2.4.

PROPOSITION 2.8. *Let $f: (S, *) \rightarrow (S', *')$ be a surjective homomorphism of adequate partial semigroups. Then the canonical (continuous) extension of f to $f: \beta S \rightarrow \beta S'$ maps δS into $\delta S'$, and its restriction to δS is a homomorphism $f: (\delta S, *) \rightarrow (\delta S', *')$ of semigroups.*

We abuse language by writing f for all three of the functions that would properly be called f, βf, and δf. The abuse will cause no confusion as the functions agree wherever two of them are defined.

Proof. If $\mathcal{U} \in \delta S$ and $s' \in S'$, then, using surjectivity of f to write $s' = f(s)$ with $s \in S$, we have $(\mathcal{U}t)\,[s * t \text{ is defined}]$ and therefore, as f is a homomorphism,

$(\mathscr{U}t)\,[s'*'f(t)$ is defined], which means that $(f(\mathscr{U})t')\,[s'*'t'$ is defined]. So $f(\mathscr{U})\in\delta S$. To show that $f(\mathscr{U}*\mathscr{V})=f(\mathscr{U})*'f(\mathscr{V})$ for $\mathscr{U},\mathscr{V}\in\delta S$, we compute

$$
\begin{aligned}
(f(\mathscr{U}*\mathscr{V})s')\,P(s') \quad &\Leftrightarrow \quad ((\mathscr{U}*\mathscr{V})s)\,P(f(s)) \\
&\Leftrightarrow \quad (\mathscr{U}s)(\mathscr{V}t)\,[s*t \text{ is defined and } P(f(s*t))] \\
&\Rightarrow \quad (\mathscr{U}s)(\mathscr{V}t)\,[f(s)*'f(t) \text{ is defined and } P(f(s)*'f(t))] \\
&\Leftrightarrow \quad (f(\mathscr{U})s')(f(\mathscr{V})t')\,[s'*'t' \text{ is defined and } P(s'*'t')] \\
&\Leftrightarrow \quad ((f(\mathscr{U})*'f(\mathscr{V}))s')\,P(s').
\end{aligned}
$$

This shows that $f(\mathscr{U}*\mathscr{V})\subseteq f(\mathscr{U})*'f(\mathscr{V})$. As no ultrafilter is properly included in another, the inclusion must be an equality.

PROPOSITION 2.9. *Let* $b\colon L(\Sigma)\to\mathbb{N}^+$ *be defined (as in Example* 2.4*) by*

$$
b(w)=\sum_{n\in\mathrm{Domain}(w)}w(n)\cdot(\sigma+1)^n
$$

and write b *also for the continuous extension mapping* $\beta L(\Sigma)$ *into* $\beta\mathbb{N}^+$. *Then*

(1) *every idempotent in* $(\beta\mathbb{N}^+,+)$ *is* b *of some member of* $\delta L(\Sigma)$;

(2) *if* $\mathscr{U}$ *is a minimal idempotent in* $\delta L(\Sigma)$, *then* $b(\mathscr{U})$ *is a minimal idempotent in* $\beta\mathbb{N}^+$;

(3) *if* X *is a central subset of* $L(\Sigma)$, *then* $b[X]$ *is a central subset of* $\mathbb{N}^+$.

Proof. (1) Let $\mathscr{U}'$ be an idempotent in $\beta\mathbb{N}^+$. As b maps onto $\mathbb{N}^+$, its continuous extension maps onto $\beta\mathbb{N}^+$, so let $\mathscr{U}\in\beta L(\Sigma)$ satisfy $b(\mathscr{U})=\mathscr{U}'$. We shall show that $\mathscr{U}\in\delta L(\Sigma)$.

To this end, we fix an arbitrary $w\in L(\Sigma)$ and we seek to prove that $(\mathscr{U}s)\,[w*s$ is defined]. Let $k=\max(\mathrm{Domain}(w))+1$. By definition of $*$ in $L(\Sigma)$, what we must prove is that $(\mathscr{U}s)\,[\min(\mathrm{Domain}(s))\geqslant k]$, that is, $(\mathscr{U}s)\,[b(s)$ is divisible by $(\sigma+1)^k]$, or equivalently $(\mathscr{U}'n)\,[n$ is divisible by $(\sigma+1)^k]$, that is,

$$
\mathbb{N}^+\cdot(\sigma+1)^k\in\mathscr{U}'.
$$

But it is well known that, for all $m\in\mathbb{N}^+$, $\mathbb{N}^+\cdot m\in\mathscr{U}'$. (For some j, $\{n\in\mathbb{N}^+\mid n\equiv j\bmod m\}\in\mathscr{U}'$ and hence $\{n\in\mathbb{N}^+\mid n\equiv 2j\bmod m\}\in\mathscr{U}'+\mathscr{U}'=\mathscr{U}'$. So $j\equiv 0\bmod m$.)

(2) If $\mathscr{U}$ is a minimal idempotent of $\delta L(\Sigma)$, then, by Proposition 2.8 above and Lemma 1.2 of [13], $b(\mathscr{U})$ is a minimal idempotent of the image of the homomorphism $b\colon\delta L(\Sigma)\to\beta\mathbb{N}^+$. But, by (1) just proved, this image contains all the idempotents of $\beta\mathbb{N}^+$. It follows immediately that $b(\mathscr{U})$ is a minimal idempotent of $\beta\mathbb{N}^+$.

(3) If X is central in $L(\Sigma)$, then it belongs to a minimal idempotent $\mathscr{U}\in\delta L(\Sigma)$. Then by (2), $b(\mathscr{U})$ is a minimal idempotent of $\beta\mathbb{N}^+$ containing $b[X]$, so this set is central in $\mathbb{N}^+$.

3. *Minimal idempotent ultrafilters over words*

This section is devoted to obtaining the ultrafilters that will play a key role in the proofs of our partition theorems. We fix a non-empty finite alphabet Σ and a variable $v\notin\Sigma$. We use the notation introduced in §1:

$$
C=L(\Sigma),\quad V=L(\Sigma;v)=L(\Sigma\cup\{v\})\backslash L(\Sigma),
$$
$$
\text{and}\quad A=L(\Sigma\cup\{v\})=C\cup V.
$$

The operation $*$ on $L(\Sigma) = C$ defined in §2 makes C an adequate partial semigroup. Applying the same definition with $\Sigma \cup \{v\}$ in place of Σ, we see that A is an adequate partial semigroup, and so is V with the restriction of this operation. In fact, V is a two-sided ideal in A, that is, if $u * w$ is defined in A and at least one of u and w is in V, then $u * w \in V$.

By the results in §2, δC, δV, and δA are compact right-topological semigroups. We identify βC and βV with subsets of βA via the extensions to ultrafilters of the inclusion maps of C and V into A. Because δA consists of those $\mathcal{U} \in \beta A$ that contain $\{w \in A \mid \min(\text{Domain}(w)) > k\}$ for all k and because of similar descriptions of δC and δV, we have $\delta C = \beta C \cap \delta A$ and $\delta V = \beta V \cap \delta A$. Thus, δC and δV are subsemigroups of δA. We remark, for use in quantifier calculations, that if an inclusion map, say from C to A, sends $\mathcal{U}$ to $\mathcal{U}'$, then, for any predicate P on A,

$$(\mathcal{U}'x)\, P(x) \quad \Leftrightarrow \quad (\mathcal{U}x)\, P(x),$$

where the P on the right is regarded as a predicate on C by restriction. Because of this, it does no harm to use the same notation for $\mathcal{U}$ and $\mathcal{U}'$; in other words, our identifications will not cause confusion. Notice that an ultrafilter $\mathcal{U}$ on A is (the inclusion-image of) an ultrafilter on C if and only if $(\mathcal{U}x)\,(x \in C)$.

Because V is a two-sided ideal in A, δV is a two-sided ideal in δA. Indeed, if $\mathcal{U} \in \delta A$ and $\mathcal{V} \in \delta V$, then $(\mathcal{U}s)(\mathcal{V}t)\,[s * t \in V]$, which means that

$$((\mathcal{U} * \mathcal{V})s)\,[s \in V],$$

that is, $\mathcal{U} * \mathcal{V} \in \delta V$, and δV is a left ideal. The proof that it is a right ideal is similar.

Recall from Example 2.3 the homomorphisms

$$\hat{a}\colon A \to C\colon w(v) \mapsto w(a)$$

substituting an element $a \in \Sigma$ for v. As $\hat{a}$ is surjective (because it is the identity on $C \subseteq A$), it yields, by Proposition 2.8, a homomorphism $\hat{a}\colon \delta A \to \delta C$.

Applying Corollary 2.7(1), fix a minimal idempotent $\mathcal{U}$ in δC. It is, of course, also an idempotent in δA, but it need not be minimal there. (Indeed, we shall soon see that $\mathcal{U}$ cannot be minimal in δA.) By Corollary 2.7(5), fix a minimal idempotent $\mathcal{V} \in \delta A$ with $\mathcal{V} \leq \mathcal{U}$.

PROPOSITION 3.1. *The idempotent $\mathcal{V}$ is minimal in δV.*

Proof. Corollary 2.7(4) and (6) ensure that $\mathcal{V}$ (or any minimal idempotent of δA) is in the two-sided ideal δV. For minimality, notice that any counter-example to minimality of $\mathcal{V}$ (that is, any idempotent $\mathcal{W} < \mathcal{V}$) in δV would also be a counter-example in δA.

Notice that this proposition implies that $\mathcal{U}$ is not minimal in δA, for if it were minimal then we would have $\mathcal{V} = \mathcal{U} \in \delta C$, contrary to $\mathcal{V} \in \delta V$.

PROPOSITION 3.2. *For each $a \in \Sigma$, $\hat{a}(\mathcal{V}) = \mathcal{U}$.*

Proof. As $\hat{a}$ is a homomorphism and $\mathcal{V}$ is idempotent, $\hat{a}(\mathcal{V})$ is also idempotent. Also, as $\mathcal{V} \leq \mathcal{U}$ in δA, we have $\hat{a}(\mathcal{V}) \leq \hat{a}(\mathcal{U})$ in δC. But $\hat{a}$ is the identity on C, so

$\hat{a}(\mathcal{U}) = \mathcal{U}$. Thus, $\hat{a}(\mathcal{V})$ is an idempotent below $\mathcal{U}$ in δC. As $\mathcal{U}$ is minimal in δC, it follows that $\hat{a}(\mathcal{V}) = \mathcal{U}$.

4. Partition theorems for located words

We retain the notation $C = L(\Sigma)$, $V = L(\Sigma; v)$, and $A = C \cup V = L(\Sigma \cup \{v\})$ from § 3. As in § 1, we call a finite or infinite sequence $w_0, w_1, \ldots, w_n)$ from A orderly if $\max(\text{Domain}(w_i)) < \min(\text{Domain}(w_{i+1}))$ for all i $(<n)$, so $w_i * w_{i+1}$ is defined. Recall also from the end of § 1 the notions of an extraction $\vec{s}[\vec{w}]$ and an extracted word $\vec{s}[w]$ of an orderly sequence $\vec{s}$ from V.

THEOREM 4.1. *Let V and C each be partitioned into finitely many pieces. Then there exist pieces V_0 and C_0 in the given partitions of V and C, respectively, and there exists an infinite orderly sequence $\vec{w}$ from V such that all extracted words of $\vec{w}$ lie in V_0 or C_0, and such that V_0 and C_0 are central in V and C, respectively.*

The proof will show that if the partition of C has several central pieces, any one of them could be used as C_0.

Proof. By the results of § 3, find minimal idempotents $\mathcal{U} \in \delta C$ and $\mathcal{V} \in \delta V$ with $\mathcal{V} \leq \mathcal{U}$ (in δA) and with $\mathcal{U} = \hat{a}(\mathcal{V})$ for all $a \in \Sigma$. Let C_0 and V_0 be the pieces of the given partition that lie in $\mathcal{U}$ and $\mathcal{V}$, respectively. This ensures the centrality of C_0 and V_0. (If a central piece C_0 of C were prescribed, we could choose a minimal idempotent $\mathcal{U} \in \delta C$ containing it, then choose $\mathcal{V}$ as in § 3, and then proceed with the proof.)

We intend to construct the required orderly sequence $\vec{w}$ by induction, choosing one of its terms w_n at each stage and simultaneously choosing sets V_n and C_n so as to satisfy the following requirements, in which all statements of the form '$w * u \in \ldots$' are to be construed as including the assertion that $w * u$ is defined:

(1) $V_n \in \mathcal{V}$;

(2) $C_n \in \mathcal{U}$;

(3) $w_n \in V_n$;

(4) $\hat{a}(w_n) \in C_n$ for all $a \in \Sigma$;

(5) $(\mathcal{V}w)[w_n * w \in V_n]$;

(6) $(\mathcal{V}w)[\hat{a}(w_n) * w \in V_n]$ for all $a \in \Sigma$;

(7) $(\mathcal{U}u)[w_n * u \in V_n]$;

(8) $(\mathcal{U}u)[\hat{a}(w_n) * u \in C_n]$ for all $a \in \Sigma$;

(9) $\max(\text{Domain}(w_{n-1})) < \min(\text{Domain}(w_n))$ if $n > 0$.

For the basis, $n = 0$, of the induction, we have already obtained V_0 and C_0 satisfying (1) and (2). Since (9) is not a condition for $n = 0$, we need to prove that some $w_0 \in V_0$ satisfies (3) through (8). In fact, $\mathcal{V}$-almost all w_0 will do, because $V_0 \in \mathcal{V}$ (for (3)), $C_0 \in \mathcal{U} = \hat{a}(\mathcal{V})$ (for (4)), $V_0 \in \mathcal{V} = \mathcal{V} * \mathcal{V}$ (for (5)), $V_0 \in \mathcal{V} = \mathcal{U} * \mathcal{V} = \hat{a}(\mathcal{V}) * \mathcal{V}$ (for (6)), $V_0 \in \mathcal{V} = \mathcal{V} * \mathcal{U}$ (for (7)), and $C_0 \in \mathcal{U} = \mathcal{U} * \mathcal{U} = \hat{a}(\mathcal{V}) * \mathcal{U}$ (for (8)). (We have treated (4), (6), and (8) as σ separate conditions each; this is legitimate because ultrafilter quantifiers commute with conjunction.)

For the induction step, suppose we have already obtained V_{n-1}, C_{n-1}, and w_{n-1}

satisfying (1) through (9). Then we define

$$V_n = V_{n-1} \cap \{w| \ w_{n-1} * w \in V_{n-1}\} \cap \bigcap_{a \in \Sigma} \{w| \ \hat{a}(w_{n-1}) * w \in V_{n-1}\}.$$

Recall our convention that statements of the form '$w * u \in \dots$' include the assertion that $w * u$ is defined; thus, (9) follows from (3) and can be ignored. Our definition of V_n satisfies (1) because each of the $\sigma + 2$ sets being intersected in the definition is in $\mathcal{V}$, by induction hypotheses (1), (5), and (6). Similarly, we define

$$C_n = C_{n-1} \cap \{u| \ w_{n-1} * u \in V_{n-1}\} \cap \bigcap_{a \in \Sigma} \{u| \ \hat{a}(w_{n-1}) * u \in C_{n-1}\}.$$

This satisfies (2) by induction hypotheses (2), (7), and (8). Finally, we choose w_n satisfying (3)–(8) (remember that (9) follows from (3) and our definition of V_n). That such a w_n exists, indeed that $\mathcal{V}$-almost all w_n satisfy (3)–(8), follows exactly as in the case $n = 0$. This completes the inductive construction of the sequence $\vec{w}$.

We have already pointed out that V_0 and C_0 are central, and (9) ensures that $\vec{w}$ is orderly, so it remains only to show that all extracted words of $\vec{w}$ are in V_0 or C_0. To this end, let $u_1 * u_2 * \dots * u_k$ be such an extracted word, each u_j being an instance of w_{i_j}, where $i_1 < i_2 < \dots < i_k$. We shall establish the following assertion for $0 \le r \le i_k$ by decreasing induction on r.

(∗) Let p be the smallest integer with $r \le i_p$. (So $1 \le p \le k$.) Then $u_p * \dots * u_k$ is in V_r or C_r (according to whether it is in V or C).

Notice that when $r = 0$ then $p = 1$ and (∗) is the desired result. So the proof will be complete once we establish (∗).

The basis of the decreasing induction is at $r = i_k$. So $p = k$. We must show that $u_k \in V_r \cup C_r$. But u_k is an instance of $w_{i_k} = w_r$. So (3) and (4) for $n = r$ give what we need.

For the induction step, suppose that (∗) holds for a certain $r > 0$, and we wish to prove it for $r - 1$. If $r - 1$ is not one of the i_j, then p is the same for $r - 1$ as for r, and the desired conclusion follows immediately from $V_r \subseteq V_{r-1}$ and $C_r \subseteq C_{r-1}$ (see the definitions of V_n and C_n). There remains the non-trivial case, where $r - 1 = i_j$. Then the p in (∗) is $j + 1$ in the induction hypothesis (for r) and j in the conclusion we want (for $r - 1$). Let $z = u_{j+1} * \dots * u_k$. By the induction hypothesis, $z \in V_r \cup C_r$. We must show that $u_j * z \in V_{r-1} \cup C_{r-1}$. Here u_j is an instance of $w_{i_j} = w_{r-1}$. We consider four cases, depending on which of u_j and z contain the variable v.

Case 1: $u_j = w_{r-1}$, $z \in V_r$. By definition of V_n, we have

$$z \in V_r \subseteq \{w| \ w_{r-1} * w \in V_{r-1}\},$$

so $u_j * z = w_{r-1} * z \in V_{r-1}$.

Case 2: $u_j = w_{r-1}$, $z \in C_r$. By definition of C_n, we have

$$z \in C_r \subseteq \{u| \ w_{r-1} * u \in V_{r-1}\},$$

so $u_j * z = w_{r-1} * z \in V_{r-1}$.

Case 3: $u_j = \hat{a}(w_{r-1})$ with $a \in \Sigma$ and $z \in V_r$. By definition of V_n, we have

$$z \in V_r \subseteq \{w| \ \hat{a}(w_{r-1}) * w \in V_{r-1}\},$$

so $u_j * z = \hat{a}(w_{r-1}) * z \in V_{r-1}$.

Case 4: $u_j = \hat{a}(w_{r-1})$ *with* $a \in \Sigma$ *and* $z \in C_r$. By definition of C_n, we have

$$z \in C_r \subseteq \{u| \ \hat{a}(w_{r-1}) * u \in C_{r-1}\},$$

so $u_j * z = \hat{a}(w_{r-1}) * z \in C_{r-1}$.

This completes the proof of (*) and thus of the theorem.

The following extension of Theorem 4.1, obtained by almost the same proof, will be useful in § 5.

THEOREM 4.2. *Let* $\mathcal{U}$ *and* $\mathcal{V}$ *be minimal idempotents in* δC *and* δV, *respectively, with* $\mathcal{V} \leqslant \mathcal{U}$ *in* δA, *and let* $V'_n \in \mathcal{V}$ *and* $U'_n \in \mathcal{U}$ *for all* $n \in \mathbb{N}$. *Then there is an infinite orderly sequence* $\vec{w}$ *from* V *such that, if* $z = u_1 * u_2 * \ldots * u_q$ *is an extracted word of* $\vec{w}$, u_j *being an instance of* w_{i_j} *with* $i_1 < i_2 < \ldots < i_q$, *then* $z \in V'_k$ *or* U'_k *(according to whether* $z \in V$ *or* C*), where* $k = \max(\text{Domain}(w_{i_1-1}))$ *if* $i_1 > 0$ *and* $k = 0$ *if* $i_1 = 0$.

Proof. Proceed as in the proof of Theorem 4.1 except that $V_0 = V'_0$ and $C_0 = U'_0$ and, at subsequent stages of the induction, V_n and C_n are chosen to be subsets of V'_k and U'_k, respectively, where $k = \max(\text{Domain}(w_{n-1}))$. By simply including V'_k or U'_k among the sets being intersected in the definitions of V_n and C_n, respectively, we obtain these inclusions without damaging (1) or (2), so the induction goes through as before. The proof of (*) also proceeds as before, but instead of taking $r = 0$ in (*) we take $r = i_1$. If $i_1 > 0$, this gives us (since $p = 1$)

$$u_1 * u_2 * \ldots * u_k \in V_{i_1} \cup C_{i_1} \subseteq V'_k \cup U'_k,$$

where $k = \max(\text{Domain}(w_{i_1-1}))$. If $i_1 = 0$, it gives us

$$u_1 * u_2 * \ldots * u_k \in V_0 \cup C_0 \subseteq V'_0 \cup U'_0.$$

The rest of this section is devoted to corollaries of Theorem 4.1, including the two theorems stated in the introduction.

COROLLARY 4.3. *Let* V *and* C *each be partitioned into finitely many pieces, and let* $\vec{s}$ *be an infinite orderly sequence from* V. *There are pieces* V_0 *and* C_0 *of the given partitions of* V *and* C, *and there is a variable extraction* $\vec{t}$ *of* $\vec{s}$ *such that all extracted words of* $\vec{t}$ *lie in* V_0 *or* C_0.

Theorem 4.1 is the special case of this corollary where s_n is the located word consisting of v at location n. The following proof shows how to reduce the general case of the corollary to this special case.

Proof. For any located word w, we write, as in § 1, $\vec{s}[w]$ for the located word extracted from $\vec{s}$ by forming, for each $j \in \text{Domain}(w)$, the located word $s_j(w_j)$ and then combining these located words, in order of increasing j, by the operation $*$. The map

$$g: A \to A: w \mapsto \vec{s}[w]$$

maps V into V and C into C. The given partitions Π_V of V and Π_C of C induce, via g, new partitions

$$\Pi'_V = \{g^{-1}(X)| \ X \in \Pi_V\}, \quad \Pi'_C = \{g^{-1}(X)| \ X \in \Pi_C\}.$$

Apply Theorem 4.1 to these new partitions, obtaining an infinite orderly sequence $\vec{w}$ from V such that all of its extracted words in V lie in one class $g^{-1}(V_0)$ of Π'_V, while all its extracted words in C lie in one class $g^{-1}(C_0)$ of Π'_C. (Here $V_0 \in \Pi_V$ and $C_0 \in \Pi_C$.) Let $\vec{t} = \vec{s}[\vec{w}]$, that is, $t_j = \vec{s}[w_j]$ for all j. Then $\vec{t}$ is a variable extraction of $\vec{s}$, and each of its extracted words has the form $\vec{s}[u] = g(u)$ for some extracted word u of $\vec{w}$ and therefore lies in V_0 or C_0.

The next corollary deals with unlocated words.

COROLLARY 4.4. *Let $W(\Sigma; v)$ and $W(\Sigma)$ be partitioned into finitely many pieces. There is an infinite sequence $\vec{w}$ from $W(\Sigma; v)$ such that all its extracted words in either of the sets $W(\Sigma; v)$ and $W(\Sigma)$ lie in one piece of the given partition of that set.*

Proof. The 'forgetting locations' homomorphism $c\colon L(\Sigma \cup \{v\}) \to W(\Sigma \cup \{v\})$ of Example 2.5 and the given partitions of $W(\Sigma; v)$ and $W(\Sigma)$ induce new partitions of V and C, as in the proof of Corollary 4.3; the pieces of the new partitions are c^{-1} of the pieces of the given partitions. Apply Theorem 4.1 to these new partitions, and let $\vec{w}'$ be an infinite orderly sequence of located variable words with homogeneity of extracted words as in Theorem 4.1. Define $\vec{w}$ by $w_j = c(w'_j)$. All extracted words of $\vec{w}$ are c-images of extracted words of $\vec{w}'$ and therefore lie in the same pieces of the two given partitions.

COROLLARY 4.5. *Let $W(\Sigma; v)$ and $W(\Sigma)$ be partitioned into finitely many pieces, and let $\vec{s}$ be an infinite sequence of variable words. There is a variable extraction $\vec{t}$ of $\vec{s}$ such that all extracted words of $\vec{t}$ in either of the sets $W(\Sigma; v)$ and $W(\Sigma)$ lie in one piece of the given partition of that set.*

Proof. The given partitions of $W(\Sigma; v)$ and $W(\Sigma)$ induce partitions of V and C by taking inverse images along

$$g\colon V \cup C = A \to W(\Sigma \cup \{v\}) = W(\Sigma; v) \cup W(\Sigma),$$
$$w \mapsto \vec{s}[w].$$

Let $\vec{w}$ be as in Theorem 4.1 for these partitions. Then $\vec{t} = \vec{s}[\vec{w}]$ is as desired, because every extracted word of $\vec{t}$ is $\vec{s}[u] = g(u)$ for some extracted word u of $\vec{w}$.

Notice that the part of Corollary 4.5 about $W(\Sigma; v)$ is Theorem 1.2. The next corollary similarly includes Theorem 1.1, Carlson's Lemma 5.9 [5].

COROLLARY 4.6. *Let $W(\Sigma; v)$ and $W(\Sigma)$ be partitioned into finitely many pieces, and let $\vec{s}$ be an infinite sequence of variable words. There is a variable reduction $\vec{t}$ of $\vec{s}$ such that all reduced words of $\vec{t}$ in either of the sets $W(\Sigma; v)$ and $W(\Sigma)$ lie in one piece of the given partition of that set.*

Proof. Induce partitions of V and C via the map

$$g\colon A \to W(\Sigma \cup \{v\})\colon w \mapsto \vec{s}[c(w)].$$

Let $\vec{w}$ be as in Theorem 4.1 for these new partitions. Define a 'partially compressed' version $\vec{w}'$ of $\vec{w}$ as follows. (Intuitively, we compress the sequence $\vec{w}$

as a whole, not its individual terms.) Let D be the union of the domains of all the w_j, and let the domain of w_j consist of the n_jth to $(n_{j+1} - 1)$th elements of D inclusive, so $0 = n_0 < n_1 < \ldots$. (We have used here that $\vec{w}$ is orderly.) Then w_j' has as its domain the interval $[n_j, n_{j+1})$ of $\mathbb{N}$ and has $c(w_j') = c(w_j)$. In other words, w_j' is obtained by moving w_j from its original location to $[n_j, n_{j+1})$. Now define $\vec{t}$ by $t_j = c(\vec{s}[w_j'])$. Since the domains of the w_j' are consecutive intervals in $\mathbb{N}$, $\vec{t}$ is a reduction of $\vec{s}$. Furthermore, any reduced word of $\vec{t}$ can be written (for suitable $a_0, a_1, \ldots, a_k \in \Sigma \cup \{v\}$) as

$$c(\vec{s}[w_0'(a_0)])\,^\frown c(\vec{s}[w_1'(a_1)])\,^\frown \ldots \,^\frown c(\vec{s}[w_k'(a_k)])$$
$$= \vec{s}[w_0'(a_0) * w_1'(a_1) * \ldots * w_k'(a_k)]$$
$$= \vec{s}[c(w_0(a_0) * w_1(a_1) * \ldots * w_k(a_k))]$$
$$= g(w_0(a_0) * w_1(a_1) * \ldots * w_k(a_k)),$$

so it is g of an extracted (in fact reduced) word of $\vec{w}$. Therefore, by the homogeneity of extracted words of $\vec{w}$, all constant reduced words of $\vec{t}$ lie in the same piece of the given partition of $W(\Sigma)$ and all its variable reduced words lie in the same piece of the given partition of $W(\Sigma; v)$.

Finally, we give the translation of Theorem 4.1 to the context of $\mathbb{N}^+$ via the map $b: L(\Sigma \cup \{v\}) \to \mathbb{N}^+$ given by 'base $\sigma + 2$ expansion', where we identify Σ with $\{1, 2, \ldots, \sigma\}$ and v with $\sigma + 1$. (See Example 2.4.) By an *instance* of a number $q \in \mathbb{N}^+$, we mean the result of replacing all occurrences of $\sigma + 1$, in the base $\sigma + 2$ expression for q, with a digit in $\{1, 2, \ldots, \sigma, \sigma + 1\}$, the same digit for all occurrences. (We allow $\sigma + 1$ as a replacement digit, so q is an instance of itself.)

COROLLARY 4.7. *Let $\mathbb{N}^+$ be partitioned into finitely many pieces. Then there exist pieces X and Y of this partition and there exists a sequence $(q_n)_{n \in \mathbb{N}}$ from $\mathbb{N}^+$ such that:*

(1) *q_n has at least one $\sigma + 1$ in its base $\sigma + 2$ expansion;*

(2) *q_{n+1} is divisible by a power of $\sigma + 2$ that is greater than q_n;*

(3) *all finite sums of instances of distinct q_n lie in X or Y according to whether or not they have $\sigma + 1$ as a digit;*

(4) *X is central in $\mathbb{N}^+$.*

We remark that, although Y need not be central, and indeed the relevant part

$$Y' = \{q \in Y \mid q \text{ does not have } \sigma + 1 \text{ as a digit}$$
$$\text{in its base } \sigma + 2 \text{ expansion}\}$$

cannot be central (cf. our remarks after the proof of Theorem 3.1 about $\mathcal{U}$ not being minimal in δA), the set Y'', obtained by reading the base $\sigma + 2$ expansions of all $q \in Y'$ as base $\sigma + 1$ expansions, is central. We refer the reader to [7] for additional combinatorial properties of central sets. (It is shown in [2] that the definition of 'central' used there and here is equivalent, for $\mathbb{N}^+$, to the definition in [7].) For example, Proposition 8.9 of [7] (see also § 5 of [2]) allows us to add to the conclusion of Corollary 4.7 that X contains arbitrarily long arithmetic progressions.

5. *Partition theorems for finite sequences of located words*

In this section, we shall extend Theorem 4.1 to obtain homogeneity for k-tuples of words. This extension is analogous to the extension by Milliken [15] and Taylor [19] of the finite sum theorem [10]. (See also [1] where the Milliken–Taylor result is obtained by methods similar to those of the present section.) With notation C, V, and A as in previous sections, we write $[A]^k$ for the set of orderly k-tuples $(w_0, w_1, ..., w_{k-1})$ of located words $w_j \in A = L(\Sigma \cup \{v\})$, and similarly for $[C]^k$ and $[V]^k$. Furthermore, for any set $J \subseteq \{0, 1, ..., k-1\}$, we set

$$[A]_J^k = \{(w_0, w_1, ..., w_{k-1}) \in [A]^k| \text{ for all } j, w_j \in V \Leftrightarrow j \in J\}.$$

Thus, in particular, $[A]_\varnothing^k = [C]^k$ and $[A]_{\{0,1,...,k-1\}}^k = [V]^k$.

THEOREM 5.1. *Let $[A]^k$ be partitioned into finitely many pieces. Then there is an orderly infinite sequence $\bar{w}$ from V, and there is, for each $J \subseteq \{0, 1, ..., k-1\}$, a piece P_J of the given partition such that, for each such J, all k-tuples in $[A]_J^k$ whose components are extracted words of $\bar{w}$ lie in P_J.*

Notice that the dependence of P_J on J is unavoidable, since the given partition may have its pieces included in single sets $[A]_J^k$ and since every $\bar{w}$ has k-tuples of extracted words in every $[A]_J^k$. It is also unavoidable that we obtain homogeneity only for orderly k-tuples; were we to partition the set of all k-tuples of words (or even just k-tuples whose components have disjoint domains), the partition class of a k-tuple could depend on how the domains of the components intermesh, and no homogeneity could be expected (see [15]).

The proof of the theorem is based upon two lemmas, the first of which is just the theorem for a single J, and the second of which extends the first just as Corollary 4.3 extended Theorem 4.1.

LEMMA 5.2. *Let $[A]^k$ be partitioned into finitely many pieces, and let $J \subseteq \{0, 1, ..., k-1\}$. Then there is an orderly infinite sequence $\bar{w}$ from V such that all k-tuples in $[A]_J^k$ whose components are extracted words of $\bar{w}$ lie in the same piece of the partition.*

Proof. Let $\mathcal{U}$ and $\mathcal{V}$ be ultrafilters with all the properties listed in § 3. For $0 \leq i < k$, let $\mathcal{W}_i$ be $\mathcal{V}$ if $i \in J$ and let $\mathcal{W}_i$ be $\mathcal{U}$ if $i \notin J$. Since every orderly k-tuple is in some piece of the given partition Π and since

$$(\mathcal{W}_0 w_0)(\mathcal{W}_1 w_1) ... (\mathcal{W}_{k-1} w_{k-1}) [(w_0, w_1, ..., w_{k-1}) \text{ is orderly}]$$

(because each $\mathcal{W}_i \in \delta A$), we have

$$(\mathcal{W}_0 w_0)(\mathcal{W}_1 w_1) ... (\mathcal{W}_{k-1} w_{k-1}) \bigvee_{P \in \Pi} [(w_0, w_1, ..., w_{k-1}) \in P],$$

where $\bigvee$ denotes a disjunction. Since ultrafilter quantifiers commute with finite disjunctions, we infer that

$$\bigvee_{P \in \Pi} (\mathcal{W}_0 w_0)(\mathcal{W}_1 w_1) ... (\mathcal{W}_{k-1} w_{k-1}) [(w_0, w_1, ..., w_{k-1}) \in P].$$

So fix a class $P \in \Pi$ such that

$$(\mathcal{W}_0 w_0)(\mathcal{W}_1 w_1) ... (\mathcal{W}_{k-1} w_{k-1}) [(w_0, w_1, ..., w_{k-1}) \in P].$$

By definition of the ultrafilter quantifier $(\mathcal{W}_0 w_0)$, find $X_\varnothing \in \mathcal{W}_0$ such that, for each $w_0 \in X_\varnothing$,

$$(\mathcal{W}_1 w_1) \dots (\mathcal{W}_{k-1} w_{k-1})\, [(w_0, w_1, \dots, w_{k-1}) \in P].$$

Then, by definition of $(\mathcal{W}_1 w_1)$, find for each $w_0 \in X_\varnothing$, a set $X_{w_0} \in \mathcal{W}_1$ such that for all $w_1 \in X_{w_0}$,

$$(\mathcal{W}_2 w_2) \dots (\mathcal{W}_{k-1} w_{k-1})\, [(w_0, w_1, \dots, w_{k-1}) \in P].$$

Then find $X_{w_0, w_1} \in \mathcal{W}_2$ for each $w_0 \in X_\varnothing$, $w_1 \in X_{w_0}$, and so forth, with analogous properties. To simplify notation, we also set $X_{w_0, w_1, \dots, w_j} = A$ if some $w_i \notin X_{w_0, w_1, \dots, w_{i-1}}$. Then we have X_s defined for all sequences s of length at most k with the following properties:

$X_s \in \mathcal{W}_i$ if s has length i;

if $w_0, w_1, \dots, w_{k-1}$ satisfy $w_i \in X_{w_0, w_1, \dots, w_{i-1}}$ for all $i = 0, 1, \dots, k-1$, then $(w_0, w_1, \dots, w_{k-1}) \in P$.

We say that a sequence s is *below* a number $n \in \mathbb{N}$ if all terms w in s have $\max(\text{Domain}(w)) \leq n$. Notice that, for each n, there are only finitely many sequences (of length at most k) below n. We define, for all $n \in \mathbb{N}$,

$$V'_n = \bigcap \{X_s \mid s \text{ is a sequence from } A \text{ of length in } J \text{ and } s \text{ is below } n\}$$

and

$$U'_n = \bigcap \{X_s \mid s \text{ is a sequence from } A \text{ of length in } \{0, 1, \dots, k-1\} \backslash J$$
$$\text{and } s \text{ is below } n\}.$$

Notice that, for s as in the definition of V'_n, we have $X_s \in \mathcal{W}_{\text{length}(s)} = \mathcal{V}$. As only finitely many of the X_s are being intersected, we have $V'_n \in \mathcal{V}$. Similarly $U'_n \in \mathcal{U}$.

Thus, we are in a position to apply Theorem 4.2. Let $\vec{w}$ be an infinite orderly sequence from V satisfying the conclusion of that theorem. We show that it also satisfies the conclusion of the present Lemma 5.2.

Consider a k-tuple $(z_0, z_1, \dots, z_{k-1}) \in [A]^k_J$, where each z_i is an extracted word of $\vec{w}$; we shall complete the proof by showing that this k-tuple lies in P. By the properties of the X_s, it suffices to show that $z_i \in X_{(z_0, z_1, \dots, z_{i-1})}$ for $i = 0, 1, \dots, k-1$. Consider a particular i in this range, and let $z_i = u_1 * u_2 * \dots * u_r$, where each u_p is an instance of some w_{q_p} with $q_1 < q_2 < \dots < q_r$. Let $n = \max(\text{Domain}(w_{q_1 - 1}))$ (or $n = 0$ if $q_1 = 0$) and notice that $(z_0, z_1, \dots, z_{i-1})$ is below n (since $(z_0, z_1, \dots, z_{k-1})$ is orderly).

If $i \in J$ then, as $(z_0, z_1, \dots, z_{k-1}) \in [A]^k_J$, z_i is a variable located word. Thus, by our choice of $\vec{w}$ in accordance with Theorem 4.2, $z_i \in V'_n$. As $(z_0, z_1, \dots, z_{i-1})$ is a sequence below n of length $i \in J$, $X_{(z_0, z_1, \dots, z_{i-1})}$ is one of the sets we intersected to define V'_n. So $z_i \in X_{(z_0, z_1, \dots, z_{i-1})}$, as desired.

If $i \in \{0, 1, \dots, k-1\} \backslash J$ then $z_i \in C$, so, by our choice of $\vec{w}$, $z_i \in U'_n$. But $X_{(z_0, z_1, \dots, z_{i-1})}$ is one of the sets we intersected to define U'_n. So $z_i \in X_{(z_0, z_1, \dots, z_{i-1})}$, as desired.

LEMMA 5.3. *Let $[A]^k$ be partitioned into finitely many pieces, let $J \subseteq \{0, 1, \dots, k-1\}$, and let $\vec{s}$ be an infinite orderly sequence from V. There is a variable extraction $\vec{t}$ of $\vec{s}$ such that all k-tuples in $[A]^k_J$ whose components are extracted words of $\vec{t}$ lie in one piece of the given partition.*

Proof. Proceed as in the proof of Corollary 4.3, using Lemma 5.2 where that proof used Theorem 4.1. Remark also that if $(w_0, w_1, \ldots, w_{k-1}) \in [A]_J^k$, then $(\vec{s}[w_0], \vec{s}[w_1], \ldots, \vec{s}[w_{k-1}]) \in [A]_J^k$ for the same J.

Proof of Theorem 5.1. List the 2^k subsets J of $\{0, 1, \ldots, k-1\}$ in a sequence $J_0, J_1, \ldots, J_{2^k-1}$. Apply Lemma 5.2 with J_0 as J, and call the sequence $\vec{w}$ that you obtain $\vec{s}^{(0)}$. All k-tuples in $[A]_{J_0}^k$ of extracted words of $\vec{s}^{(0)}$ lie in one piece of the given partition; call that piece P_{J_0}. Notice that any extraction of $\vec{s}^{(0)}$ enjoys the same property.

Reasoning inductively, after $\vec{s}^{(i-1)}$ has been chosen, an infinite orderly sequence from V, apply Lemma 5.3 with $\vec{s}^{(i-1)}$ as $\vec{s}$ and with J_i as J. Call the $\vec{t}$ that you obtain $\vec{s}^{(i)}$. All k-tuples in $[A]_{J_i}^k$ of extracted words of $\vec{s}^{(i)}$ lie in one piece, called P_{J_i}, of the given partition. Notice that, for $j < i$, $\vec{s}^{(i)}$ is an extraction of $\vec{s}^{(j)}$, so $\vec{s}^{(i)}$ inherits the homogeneity property of $\vec{s}^{(j)}$: all k-tuples in $[A]_{J_j}^k$ of extracted words of $\vec{s}^{(i)}$ lie in P_{J_j}. Therefore, $\vec{s}^{(2^k-1)}$ has these homogeneity properties for all J_j, that is, for all $J \subseteq \{0, 1, \ldots, k-1\}$. It is therefore a $\vec{w}$ of the sort required in Theorem 5.1.

Having extended Theorem 4.1 to deal with k-tuples instead of single words, we could similarly extend all its corollaries. One of these extensions will be needed in § 7, so we record it here. We omit the rest of the corollaries and their routine proofs.

We write $W(\Sigma \cup \{v\})^k$ for the set of k-tuples of (unlocated) words over $\Sigma \cup \{v\}$. For $J \subseteq \{0, 1, \ldots, k-1\}$, we write $W(\Sigma \cup \{v\})_J^k$ for the set of those $(w_0, w_1, \ldots, w_{k-1}) \in W(\Sigma \cup \{v\})^k$ such that $\{i \mid v \text{ occurs in } w_i\} = J$. If $\vec{s}$ is an infinite sequence of variable words, then we call a k-tuple $\vec{t}$ of extracted words of $\vec{s}$ *orderly* if the terms in $\vec{s}$ used to produce t_i precede (in $\vec{s}$) those used to produce t_{i+1}, for each meaningful i.

COROLLARY 5.4. *Let $W(\Sigma \cup \{v\})^k$ be partitioned into finitely many pieces. There is an infinite sequence $\vec{w}$ of variable words over Σ, and there is, for each $J \subseteq \{0, 1, \ldots, k-1\}$, a piece P_J of the given partition, such that all orderly k-tuples of extracted words of $\vec{w}$ that belong to $W(\Sigma \cup \{v\})_J^k$ also belong to P_J.*

This corollary, with the additional information that $\vec{w}$ can be taken to be an extraction of any prescribed infinite sequence from $W(\Sigma; v)$, can be proved by the same method that proved Corollary 4.4.

At the referee's suggestion, we briefly discuss the connection between the results proved here and partition theory for k-parameter sets. Parts of that theory can be deduced from our work. As a typical example, we deduce the following result of Prömel and Voigt [16, Theorem F]. Recall that a d-variable word over Σ is a word over $\Sigma \cup \{v_0, \ldots, v_{d-1}\}$ in which all the variables v_i actually occur and all occurrences of v_i precede all occurrences of v_{i+1}. Similarly, we define an ∞-variable word to be an infinite sequence from $\Sigma \cup \{v_0, v_1, \ldots\}$ in which all the (infinitely many) variables occur and all occurrences of v_i precede all occurrences of v_{i+1} for all i. Instances of multi-variable words are obtained by replacing each variable by either a variable or a member of Σ.

COROLLARY 5.5 [16]. *Let the set of d-variable words over Σ be partitioned into*

finitely many pieces. Then there is an ∞-variable word q such that one piece of the given partition contains all those d-variable words that are instances of initial segments of q that end just before the first occurrence of some variable.

Proof. Let g be the function which to any $d + 1$-tuple of variable words

$$(u_0(v), u_1(v), ..., u_d(v)) \in W(\Sigma \cup \{v\})_J^{d+1}$$

where $J = \{0, 1, ..., d\}$, associates the d-variable word

$$u_0(v_0)^\frown u_1(v_1)^\frown ... ^\frown u_{d-1}(v_{d-1})^\frown \bar{u}_d,$$

where $\bar{u}_d$ means the initial segment of u_d up to but not including the first occurrence of the variable. Given any partition of the set of d-variable words, we can pull it back along g, obtaining a partition of $W(\Sigma \cup \{v\})_J^{d+1}$, and we can apply Corollary 5.4 to the resulting partition, obtaining a homogeneous infinite sequence $\vec{w}$ as there. (Technically, Corollary 5.4 requires a partition of all of $W(\Sigma \cup \{v\})^{d+1}$, not just the part $W(\Sigma \cup \{v\})_J^{d+1}$ that we are interested in; partition the rest arbitrarily.) Now let q be the ∞-variable word

$$q = w_0(v_0)^\frown w_1(v_1)^\frown ... .$$

The desired homogeneity property of q follows from that of $\vec{w}$ once we observe that any d-variable word obtainable by instantiating the initial segment of q ending just before some variable's first occurrence can also be obtained as follows. Take an initial segment of $\vec{w}$ long enough to match the part of q actually used. Partition it into $d + 1$ segments, such that for $0 \le i < d$ the ith segment corresponds to the part of q where v_i appeared in the instantiation, and such that the last segment contains the place where q was truncated just before some variable's first occurrence. Finally apply g to this $d + 1$-tuple of variable words.

Notice that the proof of Corollary 5.5 used only a small part of the strength of Corollary 5.4. For one thing, we needed only the case $J = \{0, 1, ..., d\}$. For another, we applied the homogeneity of $\vec{w}$ only to those orderly $d + 1$-tuples of extracted words that are obtained by concatenation of consecutive segments from the beginning of $\vec{w}$, that is, those $d + 1$-tuples whose concatenation is a reduced (rather than merely extracted) word of $\vec{w}$.

Prömel and Voigt give a similar result [16, Theorem A] for partitions of words that are more general than d-variable words in that only the first (rather than every) occurrence of v_i is required to precede the first occurrence of v_{i+1}. That is, the occurrences of the variables may be intermeshed with each other. There seems to be little chance of obtaining such a result as a consequence of ours, for the following reason. It is easy to check that, to obtain this sort of intermeshed homogeneity, one must allow intermeshed variables in q as well. (Consider a partition into two classes, according to whether or not meshing occurs in a word.) But our methods always yield unmeshed (orderly) homogeneous sequences. And even if they were modified to avoid this, we could always extract, from any of our homogeneous sequences, an unmeshed subsequence, which would retain the homogeneity properties because these are expressed in terms of extracted words.

It might seem that the inapplicability of our results to situations with intermeshed variables can be traced to our convention that composition in $L(\Sigma)$ is defined only when the domain of one located word entirely precedes the domain

of the other. Might one obtain better results by extending the composition operation to be defined whenever the domains of the two words are disjoint? Unfortunately, the answer is 'No' because this change in the definition of $L(\Sigma)$ would have no effect at all on $\delta L(\Sigma)$.

6. *Partition theorems for infinite sequences of located words*

We wish to do for infinite orderly sequences what we did in §5 for orderly k-tuples. Extending the notation from the preceding section, let $[A]^\infty$ be the set of infinite orderly sequences of located words over $\Sigma \cup \{v\}$ and, for $J \subseteq \mathbb{N}$, let

$$[A]^\infty_J = \{\vec{s} \in [A]^\infty| \text{ for all } j \in \mathbb{N}, s_j \in V \Leftrightarrow j \in J\}.$$

Given a partition Π of $[A]^\infty$, we wish to find $\vec{w} \in [V]^\infty$ such that all extractions of $\vec{w}$ in $[A]^\infty_J$ lie in the same piece.

There are two obstacles to this. The first is the general conflict between infinitary partition theorems and the axiom of choice, a conflict familiar from the simpler context of Ramsey's Theorem where the analogue of Ramsey's Theorem for partitions of the set of infinite subsets of $\mathbb{N}$ contradicts the axiom of choice (see [6, § 12]). In our present context, it is easy to construct, using the axiom of choice and a continuum-length transfinite induction, a partition of $[A]^\infty$ into two (or even continuum many) pieces such that every $\vec{w} \in [V]^\infty$ has extractions in $[A]^\infty_J$ (for every J) in every piece of the partition. In the context of Ramsey's Theorem, there is a well-known way around this obstacle: require the partitions to be reasonably well-behaved, for example, Borel [9] or analytic [18] in a suitable topology. We shall see that the same idea works in the present context.

The second obstacle is that there are too many sets J. In Theorem 5.1, there were 2^k sets J and we handled them in an induction of length 2^k. Now there are continuum many J, but the induction cannot be continued transfinitely. We can handle finitely many J without difficulty, and we can handle countably many with a slight weakening of the desired theorem, but we see no way to handle all the J at once. Our first theorem will deal with a single J, but before stating it we must define the well-behaved partitions to which it will apply.

We put the discrete topology on A, and we topologize the set A^∞ of infinite sequences from A as the (Tychonoff) product of countably many copies of the discrete space A. The set $[A]^\infty$ of orderly infinite sequences is a subset of A^∞, and we topologize it as a subspace. (An equivalent definition of the topology is that it is induced by the following metric: if $\vec{w}$ and $\vec{w}'$ are distinct and j is the first index with $w_j \neq w'_j$, then the distance from $\vec{w}$ to $\vec{w}'$ is 2^{-j}.) We let $\mathscr{C}$ be the smallest σ-algebra of subsets of $[A]^\infty$ that includes all open subsets and is closed under Souslin's operation $\mathscr{A}$. (See [14, p. 409].) In particular, $\mathscr{C}$ includes the algebra of Borel sets (the smallest σ-algebra containing the open sets) and the class of analytic sets (sets obtained by applying $\mathscr{A}$ to Borel sets). The members of $\mathscr{C}$ are known as C-sets in classical descriptive set theory [14, p. 455].

THEOREM 6.1. *Let $J \subseteq \mathbb{N}$ and let $[A]^\infty$ be partitioned into finitely many C-sets. Then there is $\vec{w} \in [V]^\infty$ such that all extractions $\vec{t} \in [A]^\infty_J$ of $\vec{w}$ lie in one piece of the partition.*

Proof. Let J and a partition Π be given as in the hypothesis. Let $\mathscr{U}$ and $\mathscr{V}$ be

as in § 3. By a *J-tree*, we mean a set T of finite sequences of members of A with the following properties:

(1) T is non-empty and closed under initial segments (and therefore contains the empty sequence $\varnothing$);

(2) for every $k \in J$ and every $(w_0, w_1, \ldots, w_{k-1}) \in T$, we have that

$$(\mathcal{V}w)\,[(w_0, w_1, \ldots, w_{k-1}, w) \in T];$$

(3) for every $k \notin J$ and every $(w_0, w_1, \ldots, w_{k-1}) \in T$, we have that

$$(\mathcal{U}w)\,[(w_0, w_1, \ldots, w_{k-1}, w) \in T].$$

We observe that the intersection of finitely many J-trees is again a J-tree, because $\mathcal{V}$ and $\mathcal{U}$ are closed under finite intersections. We also observe that

$$T_J = \{(w_0, w_1, \ldots, w_{k-1}) \mid \ k \in \mathbb{N} \text{ and } (w_0, w_1, \ldots, w_{k-1}) \text{ is an orderly}$$
$$k\text{-tuple from } A \text{ and, for all } i < k,\, w_i \in V \Leftrightarrow i \in J\}$$

$$= \bigcup_{k \in \mathbb{N}} [A]^k_{J \cap \{0,1,\ldots,k-1\}}$$

is a J-tree. By a *path* through a J-tree T, we mean an infinite sequence $\vec{w}$ whose initial segments are in T. If $T \subseteq T_J$, then all its paths are in $[A]^\infty_J$. Note that a J-tree cannot have terminal nodes (by (2) and (3)), so every J-tree has paths through it.

LEMMA 6.2. *For each piece P of the given partition Π, there is a J-tree $T(P)$ such that all or none of the paths through it lie in P.*

Proof. We use Theorem 2, Theorem 3, and Corollary 3.1 of [4]. (Actually, all these results have $\mathbb{N}$ where we have A, but the results can be transferred along any bijection between $\mathbb{N}$ and A.) Specifically, Theorem 3(a) asserts that a certain class of sets, there called **C**, is a Boolean σ-algebra, and Corollary 3.1 asserts that **C** is closed under Souslin's operation. So **C** contains all C-sets (in the notation of the present paper), in particular P. The definition of **C** in [4] then shows that certain games associated to P are determined (i.e., that one or the other player has a winning strategy). Now we apply Theorem 2, with $\mathcal{X} = P$, $\mathcal{Q}_k = \mathcal{V}$ if $k \in J$, and $\mathcal{Q}_k = \mathcal{U}$ if $k \notin J$, and observe that, because $\mathcal{U}$ and $\mathcal{V}$ are ultrafilters, $\check{\mathcal{Q}}_k = \mathcal{Q}_k$. Then, of the three statements whose equivalence Theorem 2 asserts, (b) is the determinacy just derived from Theorem 3 and Corollary 3.1, while (c) is the existence of the desired J-tree.

We claim that, for some piece $P \in \Pi$, all paths through $T(P)$ lie in P. To see this, suppose not, so no paths through $T(P)$ lie in P, for any $P \in \Pi$. Then $\bigcap_{P \in \Pi} T(P) \cap T_J$ is a J-tree whose paths lie in no $P \in \Pi$; this is absurd as the pieces P cover $[A]^\infty$.

So fix $P \in \Pi$ such that all paths through the J-tree $T(P)$ are in P. Intersecting $T(P)$ with T_J, we assume henceforth that $T(P) \subseteq T_J$.

For any $(w_0, w_1, \ldots, w_{k-1}) \in T(P)$, let

$$B_{(w_0, w_1, \ldots, w_{k-1})} = \{w \mid (w_0, w_1, \ldots, w_{k-1}, w) \in T(P)\}.$$

Thus, if the length k of $s = (w_0, w_1, \ldots, w_{k-1})$ is in J, then $B_s \in \mathcal{V}$ and if $k \notin J$ then $B_s \in \mathcal{U}$. For $n \in \mathbb{N}$, we set

$$V'_n = \bigcap \{B_s \mid s \in T(P) \text{ and length}(s) \in J \text{ and } s \text{ is below } n\}$$

and

$$U'_n = \bigcap \{B_s \mid s \in T(P) \text{ and length}(s) \notin J \text{ and } s \text{ is below } n\}.$$

(As in the proof of Lemma 5.3, $(w_0, w_1, \ldots, w_{k-1})$ is said to be below n if $\max(\text{Domain}(w_{k-1})) \leq n$.) Thus, $V'_n \in \mathcal{V}$ and $U'_n \in \mathcal{U}$ for all $n \in \mathbb{N}$.

By Theorem 4.2, let $\vec{w}$ be an infinite orderly sequence from V such that, if $z = u_1 * u_2 * \ldots * u_q$ is an extracted word of $\vec{w}$, u_j being an instance of w_{i_j} with $i_1 < i_2 < \ldots < i_q$, then $z \in V'_k$ or U'_k (according to whether $z \in V$ or C), where $k = \max(\text{Domain}(w_{i_1-1}))$ if $i_1 > 0$ and $k = 0$ if $i_1 = 0$. We shall show that $\vec{w}$ is as required in the conclusion of Theorem 6.1. So let $\vec{t} \in [A]^\infty_J$ be an extraction of $\vec{w}$. We shall show that $\vec{t} \in P$ by showing that $\vec{t}$ is a path through $T(P)$, that is, by showing that $(t_0, t_1, \ldots, t_{n-1}) \in T(P)$ for all $n \in \mathbb{N}$. We proceed by induction on n, the basis being trivial since $\varnothing$ is in every J-tree. For the induction step, suppose that $(t_0, t_1, \ldots, t_{n-1}) \in T(P)$. Let $k = \max(\text{Domain}(t_{n-1}))$ if $n > 0$ and $k = 0$ if $n = 0$. Then, by our choice of $\vec{w}$, we have $t_n \in V'_k$ or U'_k according to whether $t_n \in V$ or C. We consider two cases, according to whether or not $n \in J$.

If $n \in J$ then, as $\vec{t} \in [A]^\infty_J$, we have $t_n \in V$ and therefore $t_n \in V'_k$. But, in the definition of V'_k, one of the sets B_s being intersected has $s = (t_0, t_1, \ldots, t_{n-1})$ because this s is in $T(P)$ by the induction hypothesis, its length n is in J by the case hypothesis, and s is below k by choice of k. So $t_n \in B_{(t_0, t_1, \ldots, t_{n-1})}$ and therefore $(t_0, t_1, \ldots, t_{n-1}, t_n) \in T(P)$ as desired.

If $n \notin J$ then, similarly to the preceding case, we have $t_n \in C$ and therefore $t_n \in U'_k$. But one of the B_s being intersected in the definition of U'_k has $s = (t_0, t_1, \ldots, t_{n-1})$. So $t_n \in B_{(t_0, t_1, \ldots, t_{n-1})}$ and therefore $(t_0, t_1, \ldots, t_{n-1}, t_n) \in T(P)$ as desired.

COROLLARY 6.3. *Let* $J \subseteq \mathbb{N}$, *let* $[A]^\infty$ *be partitioned into finitely many C-sets, and let* $\vec{s} \in [V]^\infty$. *Then* $\vec{s}$ *has an extraction* $\vec{t} \in [V]^\infty$ *such that all extractions* $\vec{z} \in [A]^\infty_J$ *of* $\vec{t}$ *lie in one piece of the partition.*

Proof. Proceed as in the proof of Corollary 4.3, pulling the given partition back along the map g sending $\vec{w}$ to $\vec{s}[\vec{w}]$, and apply Theorem 6.1 to this new partition to obtain $\vec{w}$. Then $\vec{s}[\vec{w}]$ is the desired $\vec{t}$. The only new ingredient of the proof is the observation that g is continuous, so pre-images of C-sets are C-sets.

COROLLARY 6.4. *Let* $J_1, J_2, \ldots, J_r \subseteq \mathbb{N}$, *let* $[A]^\infty$ *be partitioned into finitely many C-sets, and let* $\vec{s} \in [V]^\infty$. *There is an extraction* $\vec{t} \in [V]^\infty$ *of* $\vec{s}$ *and there are pieces* $P_1, P_2, \ldots, P_r$ *of the given partition such that, for each* $i = 1, 2, \ldots, r$, *all extractions* $\vec{z} \in [A]^\infty_{J_i}$ *of* $\vec{t}$ *lie in* P_i.

Proof. Apply Corollary 6.3 repeatedly, to each J_i in turn, as in the proof of Theorem 5.1.

COROLLARY 6.5. *Let* $J_1, J_2, \ldots$ *be subsets of* $\mathbb{N}$, *let* $[A]^\infty$ *be partitioned into finitely many C-sets, and let* $\vec{s} \in [V]^\infty$. *There is an extraction* $\vec{t} \in [V]^\infty$ *of* $\vec{s}$ *and there are pieces* $P_1, P_2, \ldots$ *of the given partition such that, for each* $i \in \mathbb{N}$, *all extractions* $\vec{z} \in [A]^\infty_{J_i}$ *of* $(t_i, t_{i+1}, \ldots)$ *lie in* P_i.

Proof. Again, we apply Corollary 6.3 repeatedly, to each J_i in turn, but (to

prevent the sequence from disappearing entirely after infinitely many steps) we do not change the first i components after stage i. More precisely, we let $\vec{s}^{(0)} = \vec{s}$ and, after $\vec{s}^{(n)}$ is defined, we define $\vec{s}^{(n+1)}$ to have the same 0th to $(n-1)$th components as $\vec{s}^{(n)}$, while the remaining components of $\vec{s}^{(n+1)}$ constitute an extraction from the remaining components of $\vec{s}^{(n)}$, chosen to satisfy the conclusion of Corollary 6.3 for J_n. Then we define $\vec{t}$ by taking $t_i = s_i^{(j)}$ for any $j > i$. (All such j give the same t_i.) Then $(t_i, t_{i+1}, \ldots)$ is an extraction of $(s_i^{(i+1)}, s_{i+1}^{(i+1)}, s_{i+2}^{(i+1)}, \ldots)$ and therefore inherits the latter's homogeneity with respect to J_i.

The method used to deduce Corollary 5.5 from Corollary 5.4 can be applied to Theorem 6.1 (with $J = \{0, 1, \ldots, k-1\}$) to obtain the following corollary, first proved in somewhat greater generality (with C-sets replaced by arbitrary sets having the Baire property) by Prömel and Voigt [16, Theorem E].

COROLLARY 6.6. *Let X be the set of all infinite sequences from $\Sigma \cup \{v_0, v_1, \ldots, v_{k-1}\}$ in which all the variables occur and every occurrence of v_i precedes every occurrence of v_{i+1}. Give $\Sigma \cup \{v_0, v_1, \ldots, v_{k-1}\}$ the discrete topology, and topologize X as a subspace of the product of infinitely many copies of $\Sigma \cup \{v_0, v_1, \ldots, v_{k-1}\}$. If X is partitioned into finitely many C-sets, then there is an ∞-variable word over Σ all of whose instances in X belong to the same piece of the partition.*

7. Combinatorial subspaces

Furstenberg and Katznelson [8] introduced the notions of finite- and infinite-dimensional subspaces of $W(\Sigma)$ and proved a Ramsey theorem for these spaces. We introduce here a more general notion of infinite-dimensional subspaces and prove the corresponding Ramsey theorem, obtaining the result from [8] as a corollary.

To avoid trivialities, we assume throughout this section that $|\Sigma| \geq 2$.

For $d \in \mathbb{N}^+$, call a set $L \subseteq W(\Sigma)$ a *d-dimensional subspace* of $W(\Sigma)$ if there are $w_0, w_1, \ldots, w_{d-1} \in W(\Sigma; v)$ such that

$$L = \{w_0(a_0)w_1(a_1) \ldots w_{d-1}(a_{d-1}) | \ a_0, a_1, \ldots, a_{d-1} \in \Sigma\}.$$

Equivalently, L consists of all the instances in $W(\Sigma)$ of the d-variable word $w_0(v_0)w_1(v_1) \ldots w_{d-1}(v_{d-1})$.

We call a set $L \subseteq W(\Sigma)$ a *weak infinite-dimensional subspace* of $W(\Sigma)$ if there is an infinite sequence $\vec{w}$ from $w(\Sigma; v)$ such that

$$L = \{w_0(a_0)w_1(a_1) \ldots w_{k-1}(a_{k-1}) | \ k \in \mathbb{N}^+ \text{ and } a_0, a_1, \ldots, a_{k-1} \in \Sigma\}.$$

In other words, the members of L are the constant reduced words of $\vec{w}$. These weak infinite-dimensional subspaces were simply called subspaces in [8]; we add the adjective 'weak' to distinguish this notion from the following one.

We call a set $L \subseteq W(\Sigma)$ a *strong infinite-dimensional subspace* of $W(\Sigma)$ if there is an infinite sequence $\vec{w}$ from $W(\Sigma; v)$ such that

$$L = \{w_{i_0}(a_0)w_{i_1}(a_1) \ldots w_{i_{k-1}}(a_{k-1}) | \ k \in \mathbb{N}^+ \text{ and } i_0 < i_1 < \ldots < i_{k-1}$$
$$\text{and } a_0, a_1, \ldots, a_{k-1} \in \Sigma\}.$$

In other words, the members of L are the constant extracted words of $\vec{w}$.

We write Weak($\vec{w}$) and Strong($\vec{w}$) for the weak and strong infinite-dimensional subspaces determined by $\vec{w}$ as above. Clearly, Weak($\vec{w}$) $\subseteq$ Strong($\vec{w}$). Similarly, we write Space($w_0, w_1, \ldots, w_{d-1}$) for the d-dimensional subspace determined by $w_0, w_1, \ldots, w_{d-1}$ as above.

Let $\vec{w}$ be an infinite sequence from $W(\Sigma; v)$, and let $z_0, z_1, \ldots, z_{d-1}$ be variable extracted words of $\vec{w}$. If all the elements of $\vec{w}$ used in forming z_i precede all those used in forming z_{i+1}, for $i = 0, 1, \ldots, d-2$, then we call $(z_0, z_1, \ldots, z_{d-1})$ an *orderly d-tuple extracted from* $\vec{w}$. In this situation, clearly

$$\text{Space}(z_0, z_1, \ldots, z_{d-1}) \subseteq \text{Strong}(\vec{w}).$$

If, furthermore, the elements of $\vec{w}$ used in forming $z_0, z_1, \ldots, z_{d-1}$ together constitute an initial segment of $\vec{w}$, then Space($z_0, z_1, \ldots, z_{d-1}$) $\subseteq$ Weak($\vec{w}$). It is not hard to check that all d-dimensional subspaces of Weak($\vec{w}$) arise in this manner. (We shall make no use of this, so we omit the proof.) On the other hand, not all d-dimensional subspaces of Strong($\vec{w}$) arise in this manner; those that do will be called *natural d-dimensional subspaces* of Strong($\vec{w}$). Here is an example of an unnatural subspace. Let $\Sigma = \{1, 2\}$. Let $w_0 = v11$, $w_1 = 2v2$, and $w_n = 111v$ for all $n \geqslant 2$. Then the one-dimensional subspace Space($21v$) $=$ $\{211, 212\} = \{w_0(2), w_1(1)\}$ is a subspace of Strong($\vec{w}$) but clearly not a natural subspace.

THEOREM 7.1. *Let $d \in \mathbb{N}^+$ and let the collection of all d-dimensional subspaces of $W(\Sigma)$ be partitioned into finitely many pieces. There exists a strong infinite-dimensional subspace H of $W(\Sigma)$ whose d-dimensional subspaces all lie in the same piece of the partition.*

Proof. The proof proceeds in three steps. First, we find a strong infinite-dimensional subspace $H = \text{Strong}(\vec{w})$ whose *natural* d-dimensional subspaces all lie in the same piece of the partition. Second, we show that the sequence $\vec{w}$ can be chosen to enjoy an additional combinatorial property. Finally, we show that this additional combinatorial property ensures that *all* the d-dimensional subspaces of H are natural subspaces of H.

For the first step, we partition $W(\Sigma; v)^d$ into finitely many pieces, putting two d-tuples $(w_0, w_1, \ldots, w_{d-1})$ into the same piece if and only if the d-dimensional spaces Space($w_0, w_1, \ldots, w_{d-1}$) lie in the same piece of the given partition. By Corollary 5.4, we obtain an infinite sequence $\vec{w}$ from $W(\Sigma; v)$ such that all variable extracted words $(t_0, t_1, \ldots, t_{d-1})$ of $\vec{w}$ lie in the same piece of this new partition. (We have used Corollary 5.4 only for $J = \{0, 1, \ldots, k-1\}$.) But this means that all the d-dimensional spaces Space($t_0, t_1, \ldots, t_{d-1}$), for all such $(t_0, t_1, \ldots, t_{d-1})$, lie in the same piece of the original partition. These subspaces are exactly the natural d-dimensional subspaces of $H = \text{Strong}(\vec{w})$, so the first step of the proof is complete.

In the second step, we show that the sequence $\vec{w}$ in the first step could be chosen so that the lengths of its terms satisfy

$$(1) \qquad\qquad \text{length}(w_n) > \sum_{i<n} \text{length}(w_i).$$

To do this, we merely concatenate many successive components of $\vec{w}$. To be precise, we inductively define an increasing sequence of natural numbers $k(n)$ so that $k(0) = 0$ and the total length of all the w_i for $k(n) \leqslant i < k(n+1)$ exceeds the total length of all the w_i for $i < k(n)$. Then we let w'_n be the concatenation of the

w_i for $k(n) \leq i < k(n+1)$. This sequence $\vec{w}'$ clearly satisfies (1) and inherits the homogeneity of $\vec{w}$, so the second step of the proof is complete.

For the third and final step, we assume that (1) holds, and we prove that every d-dimensional subspace L of $H = \text{Strong}(\vec{w})$ is a natural subspace of H. Let $L = \text{Space}(u_0, u_1, \ldots, u_{d-1})$ and let u be the d-variable word $u_0(v_0)u_1(v_1) \ldots u_{d-1}(v_{d-1})$. So L is the set of all constant instances of u. All the words in L, being in H, have the form

(2) $$w_{i_1}(a_1)w_{i_2}(a_2) \ldots w_{i_k}(a_k)$$

for some k in $\mathbb{N}^+$, some $i_1 < i_2 < \ldots < i_k$ in $\mathbb{N}$, and some $a_1, a_2, \ldots, a_k$ in Σ.

We claim that $k, i_1, i_2, \ldots, i_k$ are the same for all words in L. This is a consequence of (1) and the fact that all the words in L have the same length, namely the length of u. Indeed, we show by induction on m that all words of length m of the form (2) have the same $i_1, i_2, \ldots, i_k$. Suppose that $w_{i'_1}(a'_1)w_{i'_2}(a'_2) \ldots w_{i'_k}(a'_k)$ also has length m. If $i_k < i'_{k'}$, then

$$m = \sum_{j=1}^{k} \text{length}(w_{i_j}) \leq \sum_{i=0}^{i_k} \text{length}(w_i) < \text{length}(w_{i'_{k'}}) \leq m,$$

where the strict inequality is from (1). This contradiction and the analogous contradiction from $i'_{k'} < i_k$ show that $i_k = i'_{k'}$. But then the induction hypothesis applied to $w_{i_1}(a_1)w_{i_2}(a_2) \ldots w_{i_{k-1}}(a_{k-1})$ and its primed analogue shows that the other values of i and i' also match. This proves the claim.

We write z for the k-variable word

$$w_{i_1}(v_0)w_{i_2}(v_1) \ldots w_{i_k}(v_{k-1})$$

of which all the words in L are instances by the claim just proved. That is, every constant instance of u is also an instance of z. It follows immediately that u and z have the same length, that any position occupied by an $a \in \Sigma$ in z is occupied by the same a in u (because otherwise u would have constant instances with some $b \in \Sigma \backslash \{a\}$ at that position, and these cannot be instances of z), that therefore all positions occupied by variables in u are also occupied by (possibly other) variables in z, and that two positions cannot be occupied by the same variable in z and different variables in u (because otherwise u would have constant instances with different elements of Σ at these positions, and these cannot be instances of z). These observations imply that u can be obtained from z by substituting variables and constants for variables: u is a d-variable instance of the k-variable word z. Furthermore, as the variables occur in proper order in u, all the variables in z that are replaced by v_i in forming u must precede all those that are replaced by v_{i+1}. This means that we can break $\{0, 1, \ldots, k-1\}$ into d intervals $I_0, I_1, \ldots, I_{d-1}$ such that, for all $i \in I_j$, v_i in z is replaced with v_j or a constant in forming u, and at least one such v_i is replaced with v_j. For each j, if $I_j = \{h, h+1, \ldots, h'-1\}$, consider the corresponding segment

$$w_{i_{h+1}}(v_h)w_{i_{h+2}}(v_{h+1}) \ldots w_{i_{h'}}(v_{h'-1})$$

of z and the segment $t_j(v_j)$ of u that it becomes under the substitution that transforms z into u. Then the d-tuple $(t_0(v), t_1(v), \ldots, t_{d-1}(v))$ is an orderly d-tuple extracted from $\vec{w}$, and therefore $\text{Space}(t_0(v), t_1(v), \ldots, t_{d-1}(v))$ is a natural d-dimensional subspace of $\text{Strong}(\vec{w}) = H$. On the other hand, this subspace consists precisely of all of the instances of $t_0(v_0)t_1(v_1) \ldots t_{d-1}(v_{d-1}) = u$,

so it coincides with L. This shows that L is a natural d-dimensional subspace of H and completes the final step of the proof.

This theorem immediately implies the corresponding result for weak infinite-dimensional subspaces instead of strong ones, simply because $\text{Weak}(\vec{w}) \subseteq \text{Strong}(\vec{w})$. This corollary is Theorem 3.1 of [8].

By using the results of § 6, specifically Theorem 6.1 with $J = \mathbb{N}$ and with words in place of located words (via the 'forgetting locations' homomorphism), we can obtain a partition theorem for infinite-dimensional subspaces in place of d-dimensional ones. Call $\text{Strong}(\vec{t})$ a natural infinite-dimensional subspace of $\text{Strong}(\vec{w})$ if $\vec{t}$ is an extraction of $\vec{w}$. In the following theorem, C-sets in $W(\Sigma; v)^\infty$ are understood with respect to the product of discrete topologies.

THEOREM 7.2. *Let the set of strong infinite-dimensional subspaces of $W(\Sigma)$ be partitioned into finitely many pieces such that, for each piece P, the set $\{\vec{w} \mid \text{Strong}(\vec{w}) \in P\}$ is a C-set in $W(\Sigma; v)^\infty$. Then there is a strong infinite-dimensional subspace H of $W(\Sigma)$ all of whose natural infinite-dimensional subspaces lie in the same piece of the given partition.*

We omit the proof because it is entirely analogous to the first step in the proof of Theorem 7.1. Notice that the analogue of the second step would also work here, but the analogue of the third step fails; the subscripts in (2), now an infinite sequence, can depend on the particular word being represented. That is why, in Theorem 7.2, we assert homogeneity only for the natural subspaces.

Index

instance (of a number): § 4 after Corollary 4.6.

instance (of a variable word): § 1 before Theorem 1.1.

J-tree: § 6 before Lemma 6.2.

$L(\Sigma)$: § 1 after Theorem 1.2.

$L(\Sigma; \{v\})$: § 1 after Theorem 1.2.

$\leqslant$(order of idempotents): § 2 before Corollary 2.7.

left adequate: § 2 before Example 2.1.

left ideal: § 2 after Proposition 2.6.

located word: § 1 after Theorem 1.2.

minimal idempotent: § 2 in Corollary 2.7.

$\mathbb{N}$: § 1 before Theorem 1.1.

$\mathbb{N}^+$: § 1 before Theorem 1.1.

natural d-dimensional subspace: § 7 before Theorem 7.1.

orderly d-tuple extracted from $\vec{w}$: § 7 before Theorem 7.1.

orderly (k-tuple of extracted words): § 5 after Lemma 5.3.

orderly (sequence of located words): § 1 after Theorem 1.2.

partial semigroup: § 2 before Example 2.1.

reduced word: § 1 before Theorem 1.1.

reduction: § 1 before Theorem 1.1.

right topological: § 2 after Proposition 2.6.

$\vec{s}[\vec{w}]$: § 1 after Theorem 1.2.

Space($w_0, w_1, \ldots, w_{d-1}$): § 7 before Theorem 7.1.

strong infinite-dimensional subspace: § 7 before Theorem 7.1.

Strong($\vec{w}$): § 7 before Theorem 7.1.

($\mathcal{U}s$) (quantifier): § 2 after Example 2.5.

$V (=L(\Sigma; \{v\}))$: § 1 after Theorem 1.2.

variable located word: § 1 after Theorem 1.2.

variable reduced word: § 1 before Theorem 1.1.

variable reduction: § 1 before Theorem 1.1.

variable word: § 1 before Theorem 1.1.

$W(\Sigma)$: § 1 before Theorem 1.1.

$W(\Sigma \cup \{v\})^k$: § 5 after Lemma 5.3.

$W(\Sigma \cup \{v\})_j^k$: § 5 after Lemma 5.3.

weak infinite-dimensional subspace: § 7 before Theorem 7.1.

Weak($\vec{w}$): § 7 before Theorem 7.1.

word: § 1 before Theorem 1.1.

References

1. V. BERGELSON and N. HINDMAN, 'Ultrafilters and multidimensional Ramsey theorems', *Combinatorica* 9 (1989) 1–7.
2. V. BERGELSON and N. HINDMAN, 'Nonmetrizable topological dynamics and Ramsey theory', *Trans. Amer. Math. Soc.* 320 (1990) 293–320.
3. J. BERGLUND, H. JUNGHENN, and P. MILNES, *Analysis on semigroups* (Wiley, New York, 1989).
4. A. BLASS, 'Selective ultrafilters and homogeneity', *Ann. Pure Appl. Logic* 38 (1988) 215–255.
5. T. CARLSON, 'Some unifying principles in Ramsey theory', *Discrete Math.* 68 (1988) 117–169.
6. P. ERDŐS, A. HAJNAL, A. MÁTÉ, and R. RADO, *Combinatorial set theory: partition relations for cardinals* (North-Holland, Amsterdam, 1984).
7. H. FURSTENBERG, *Recurrence in ergodic theory and combinatorial number theory* (Princeton University Press, 1981).
8. H. FURSTENBERG and Y. KATZNELSON, 'Idempotents in compact semigroups and Ramsey theory', *Israel J. Math.* 68 (1989) 257–270.

9. F. GALVIN and K. PRIKRY, 'Borel sets and Ramsey's theorem', *J. Symbolic Logic* 38 (1973) 193–198.

10. N. HINDMAN, 'Finite sums from sequences within cells of a partition of *N*', *J. Combin. Theory (A)* 17 (1974) 1–11.

11. N. HINDMAN, 'Ultrafilters and combinatorial number theory', *Number theory Carbondale 1979* (ed. M. Nathanson), Lecture Notes in Mathematics 751 (Springer, Berlin, 1979), pp. 119–184.

12. N. HINDMAN, 'The ideal structure of the space of κ-uniform ultrafilters on a discrete semigroup', *Rocky Mountain J. Math.* 16 (1986) 685–701.

13. N. HINDMAN and J. PYM, 'Free groups and semigroups in βN', *Semigroup Forum* 30 (1984) 177–193.

14. K. KURATOWSKI and A. MOSTOWSKI, *Set theory with an introduction to descriptive set theory* (North-Holland, Amsterdam, 1978).

15. K. MILLIKEN, 'Ramsey's theorem with sums or unions', *J. Combin. Theory (A)* 18 (1975) 276–290.

16. H. J. PRÖMEL and B. VOIGT, 'Baire sets of k-parameter words are Ramsey', *Trans. Amer. Math. Soc.* 291 (1985) 189–201.

17. J. PYM, 'Semigroup structures on Stone–Čech compactifications', *J. London Math. Soc.* (2) 36 (1987) 421–428.

18. J. SILVER, 'Every analytic set is Ramsey', *J. Symbolic Logic* 35 (1970) 60–64.

19. A. TAYLOR, 'A canonical partition theorem for finite subsets of ω', *J. Combin. Theory (A)* 21 (1976) 137–146.

Department of Mathematics
Ohio State University
Columbus
Ohio 43210
U.S.A.

Department of Mathematics
University of Michigan
Ann Arbor
Michigan 48109
U.S.A.

Department of Mathematics
Howard University
Washington
D.C. 20059
U.S.A.