ISRAEL JOURNAL OF MATHEMATICS, Vol. 68, No. 3, 1989

IDEMPOTENTS IN COMPACT SEMIGROUPS AND RAMSEY THEORY

BY

H. FURSTENBERG[a] AND Y. KATZNELSON[b,†]

[a]*The E. Landau Center for Research in Mathematical Analysis, Department of Mathematics, The Hebrew University of Jerusalem, Jerusalem, Israel*; *and* [b]*Department of Mathematics, Stanford University, Stanford, CA 94305, USA*

ABSTRACT

We prove a theorem about idempotents in compact semigroups. This theorem gives a new proof of van der Waerden's theorem on arithmetic progressions as well as the Hales–Jewett theorem. It also gives an infinitary version of the Hales–Jewett theorem which includes results of T. J. Carlson and S. G. Simpson.

We shall show here that van der Waerden's theorem on arithmetic progressions and its variants, the Hales–Jewett theorem ([HJ]), the theorem of T. J. Carlson and S. G. Simpson ([CS]), and the theorem of T. J. Carlson (theorem 2 of [C]) can be obtained by an analysis of idempotents in compact semigroups. We obtain in this way a sharpening of Carlson's result (which provides an "infinite" version of a theorem of Graham and Rothschild on k-parameter sets ([GR])) as well as an extension of Hindman's theorem and the related theorem of K. Milliken and A. Taylor (cf. [H]). This extended result, which is the principal novel result of this paper, plays a crucial role in a density version of the Hales–Jewett theorem which we present elsewhere (see [FK]).

Our use of idempotents is modelled to some extent on Ellis' theory of enveloping semigroups that arise in topological dynamics ([E]). The relevance of topological dynamics for combinatorial theory has already been pointed out in [FW]. In that paper we showed that Hindman's theorem is a consequence of

[†] Research supported by the National Science Foundation under Grant No. DMS86-05098.
Received July 23, 1989

the existence of (non-trivial) idempotents in the "enveloping semigroup" of a group of continuous transformations of a compact space to itself. In the present paper we make use of a natural order relation among the idempotents in a semigroup.

It is possible to avoid the formalism of topological dynamics by making use of the Stone–Čech compactification of the structures we shall study. This approach has been adopted by Bergelson and Hindman ([BH]). We prefer the former approach which, we believe, lends itself to a more elementary exposition. The Stone–Čech compactification approach invokes the theory of ultrafilters as does Carlson and in fact this approach has a good deal in common with Carlson's exposition.

Our formulation of the Hales–Jewett theorem and its extensions takes place in an arbitrary semigroup Γ, and the machinery available from the theory of compact semigroups is applied to appropriate compactifications of Γ and Cartesian powers thereof. In this algebraic setting the significance of arithmetic progressions stems from the fact that the set of vectors $(a, a+b, \ldots, a+(k-1)b)$, $a \in \mathbf{Z}$, $b \geqq 0$ form a subsemigroup of $\mathbf{Z}^k$ and the non-diagonal elements ($b > 0$) form an "ideal" in this semigroup. For the Hales–Jewett theorem, $\mathbf{Z}$ is replaced by $W(k)$, the free semigroup of words in k-letters. The Hales–Jewett k-tuples then form the elements of an ideal in a semigroup of $W(k)^k$.

If we think of words of a given alphabet as forming the points of an infinite dimensional "combinatorial space", then the Hales–Jewett k-tuples form combinatorial lines in this space. Carlson's theorem can be regarded as a theorem about coloring the space of combinatorial lines. Our result extends this to the space of d-planes in combinatorial space: If one colors all combinatorial d-planes in infinite dimensional combinatorial space (for a fixed alphabet and a finite number of colors) then we can find an infinite dimensional subspace so that all its d-planes have the same color.

The universal nature of the free semigroup explains the significance of the Hales–Jewett theorem as being the prototype of all the results that can be obtained using this algebraic approach. The same can be said for the infinitary versions of Hales–Jewett which we also treat.

We develop the algebraic preliminaries in §1 and, in §2, we present the setting in which we obtain the van der Waerden and Hales–Jewett theorems (Theorem 2.2) as well as the other results mentioned above. Finally, in §3 we point out how the results involving infinite configurations carry over to the case of "compact coloring", which is the case we meet so often in applications.

1. Compact semigroups and idempotents

Let S be a semigroup endowed with a topology with respect to which S is a compact Hausdorff space. The following assumption will be made:

(*) For each $s \in S$, the map $t \to ts$, $t \in S$ is continuous.

We shall call S simply a compact semigroup, bearing in mind the asymmetry of the condition (*).

An *idempotent* $t \in S$ is an element satisfying $t^2 = t$. The following result of Ellis extending a standard theorem for finite semigroups is fundamental.

THEOREM 1.1. *A compact semigroup S contains an idempotent.*

PROOF. Let $\mathcal{M}$ be the family of compact subsemigroups $M \subset S$ ordered by inclusion. On account of compactness $\mathcal{M}$ contains a minimal member, say P. Now if $p \in P$ then, by (*), Pp is closed, and so compact, and it is clearly a semigroup. Hence

$$Pp = P.$$

This means that $Q = \{x \in P \mid xp = p\}$ is non-empty. But Q is again closed by (*) and it is also a semigroup $\subset P$. Hence $Q = P$; so $p^2 = p$. QED

We shall make use of the notion of an *ideal* in S. Say that $J \subset S$ is a *left ideal* if it is closed and $SJ \subset J$. Similarly for right ideals. Note the following difference between left and right ideals. If J is a left ideal (and so compact) then Jx is again a left ideal for any $x \in S$. Not so for right ideals (see the example at the end of this section). We also note that if J is minimal among left ideals then each Jx is also a minimal left ideal, and if $x \in J$, then $Jx = J$. We leave this as an exercise.

It is significant that minimal left ideals always exist (by Zorn's lemma).

Since an ideal is also a compact semigroup, every ideal contains idempotents. Typically S will not be a group; so it has proper ideals. This implies that S contains idempotents distinct from the identity.

Now introduce a (partial) order among idempotents:

$$(**) \qquad \theta < \varphi \Leftrightarrow \theta\varphi = \varphi\theta = \theta.$$

THEOREM 1.2. *If φ is an idempotent in J, J a minimal left ideal, and if θ is an idempotent with $\theta < \varphi$, then $\varphi = \theta$.*

Thus, idempotents which belong to minimal left ideals are minimal, and distinct idempotents in the same minimal ideal are not comparable.

PROOF. Since $\theta < \varphi$, $\theta\varphi = \theta$. This implies $\theta \in J$. Now $J\theta \subset J$ is also a (closed) left ideal and by minimality, $J\theta = J$. Hence $\varphi = \psi\theta$ for some $\psi \in J$, which implies $\varphi\theta = \varphi$. But $\theta < \varphi$ implies $\varphi\theta = \theta$, whence $\theta = \varphi$. QED

In particular the partially-ordered set I of idempotents in S contains minimal elements.

THEOREM 1.3. *If φ is any idempotent and J is a left ideal, then $J\varphi$ contains an idempotent θ with $\theta < \varphi$.*

PROOF. Since $J\varphi$ is an ideal, and therefore a semigroup, it contains an idempotent $\omega = \psi\varphi$. Set $\theta = \varphi\omega = \varphi\psi\varphi$. Then

$$\theta^2 = \varphi\psi\varphi \cdot \varphi\psi\varphi = \varphi\psi\varphi\psi\varphi = \varphi\omega^2 = \varphi\omega = \theta.$$

So θ is an idempotent with $\theta \in J\varphi$. Moreover $\varphi\theta = \theta\varphi = \theta$. QED

THEOREM 1.4. *An idempotent is minimal iff it belongs to a minimal left ideal.*

PROOF. We have already noticed that, by Theorem 1.2, the idempotents in minimal left ideals are minimal. Conversely suppose φ is a minimal idempotent. Let J be any minimal left ideal. Then $J\varphi$ is a minimal ideal containing an idempotent $\theta < \varphi$ which implies $\theta = \varphi$. So φ is in a minimal ideal. QED

Note that every left ideal meets every right ideal: if J' is a right ideal and J'' a left ideal then $J'J'' \subset J' \cap J'$. Hence if J is a two-sided ideal then J meets, and therefore contains, every minimal left ideal. In particular every two-sided ideal contains all the minimal elements of I. This, combined with Theorem 1.3, proves:

THEOREM 1.5. *If φ is an idempotent and J is a two-sided ideal, then J contains an idempotent θ with $\theta < \varphi$.*

EXAMPLE 1.6. Let X be a compact metric space. X^X will denote the set of *all* transformations of $X \to X$, continuous or not. X^X is a semigroup (under iteration) and we topologize it by taking the product topology which, by Tychonof's theorem, is compact. A basis for open sets for this topology is given by the neighborhoods:

$$V_{x_1, x_2, \dots, x_p; \varepsilon}(f) = \{g \in X^X \mid d(f(x_1), g(x_1)) < \varepsilon, \dots, d(f(x_p), g(x_p)) < \varepsilon\}.$$

The map $f \mapsto f \circ g$ is continuous for each $g \in X^X$, as required by (*). On the other hand $f \mapsto g \circ f$ is continuous if, and only if, g is continuous. Thus X^X is an example of a compact semigroup as described in §1. The topology in X^X is not a metric topology but if $F \subset X^X$ and $g \in \bar{F}$, then for any *finite* set of points $x_1, \ldots, x_p \in X$ we can find a sequence $\{f_n\} \subset F$ with $f_n(x_j) \to g(x_j)$, $j = 1, \ldots, p$.

An idempotent $\varphi \in X^X$ is a map which equals the identity on its range, and for idempotents φ, θ, the order relation $\theta < \varphi$, i.e., $\varphi\theta = \theta\varphi = \theta$, implies that range $\varphi \supset$ range θ. In particular, the minimal idempotents are none other than the constant maps.

One obtains examples of right ideals by imposing "range conditions", e.g., "the set of all $\psi \in X^X$ whose range has at most m points in a given closed set $G \subset X^X$".

Examples of left ideals are described by restrictions on the size of the range on subsets in the domain, e.g., "the set of all $\psi \in X^X$ which are constant on each atom of a given partition of X".

The two-sided ideals are determined by a single integer m, and consist of all the transformations whose range has no more than m points.

X^X has a unique minimal left ideal, which is in fact two-sided (but is not minimal as right ideal), namely the set of all constant maps, i.e., all minimal idempotents.

These descriptions depend on the fact that we are dealing with the full X^X. For subsemigroups of X^X, which are the only semigroups in this paper, the story may be much less obvious and much more pregnant with applications.

2. Finite coloring (the theorems of van der Waerden, Hales–Jewett, Carlson–Simpson etc.)

We apply the results of the previous section in the following context:

Let Γ be a discrete semigroup, and C a finite set. Denote $X = C^\Gamma$, the space of all mappings from Γ to C. C is often referred to as the space of colors, and the elements of X as C-colorings of Γ. Without loss of generality we assume that Γ has a unit element which we denote by e. (We can always add one if it is not there, and extend the coloring as we please.)

Γ acts on X by "translation", namely for $\gamma \in \Gamma$ and $x \in X$ we write

$$(\sigma(\gamma)x)(\gamma') = x(\gamma'\gamma). \tag{2.1}$$

Notice that $\sigma(\gamma_1\gamma_2) = \sigma(\gamma_1)\sigma(\gamma_2)$.

We denote by $E(\Gamma)$ the closure of $\sigma(\Gamma)$ in X^X.

For a fixed integer k, we write $Y = X^k$, and extend the action (2.1) of Γ to an action of Γ^k on Y, writing

$$\sigma(\gamma_1, \ldots, \gamma_k) = (\sigma(\gamma_1), \ldots, \sigma(\gamma_k)). \tag{2.2}$$

For a subsemigroup $G \subset \Gamma^k$ we denote by $E(G)$ the *enveloping semigroup* of G, that is the closure of $\sigma(G)$ in Y^Y. We extend the notation to arbitrary subsets; thus for $A \subset \Gamma^k$, $E(A)$ denotes the closure of $\sigma(A)$ in Y^Y. Because of the fact that the elements of $\sigma(G)$ act componentwise, we have $E(G) \subset (E(\Gamma))^k$. Notice that if $G \supset \Delta_k(\Gamma)$, the diagonal of Γ^k, then $E(G) \supset \Delta_k(E(\Gamma))$, the diagonal of $(E(\Gamma))^k$.

THEOREM 2.1. *Assume that $G \supset \Delta_k(\Gamma)$ and let I be a two-sided ideal in G. Then for any minimal idempotent $\theta \in E(\Gamma)$, $\tilde{\theta} = (\theta, \ldots, \theta) \in E(I)$.*

Before proving Theorem 2.1 we state the following corollary thereof:

THEOREM 2.2. *Let $G \supset \Delta_k(\Gamma)$ and let I be a two-sided ideal in G. Then for any finite coloring of Γ, I contains monochromatic elements.*

EXAMPLE 2.3. $\Gamma = \mathbf{Z}^+$, the non-negative integers, the operation being addition. Let G be the set of sequences of the form $(a, a+b, \ldots, a+(k-1)b) \in \Gamma^k$, $a, b \in \mathbf{Z}^+$ (non-decreasing arithmetic progressions). Let I be the ideal in G defined by the condition $b > 0$. Theorem 2.2 in this context is the theorem of van der Waerden.

EXAMPLE 2.4. $\Gamma = W(k)$ is the free semigroup generated by the k symbols $\{1, 2, \ldots, k\}$, i.e. the set of all finite words $w = w(1)w(2)\cdots w(n)$ with each $w(j) \in \{1, 2, \ldots, k\}$, and the operation is concatenation. We take for G the span of $\Delta_k(\Gamma)$ and the element $(1, 2, \ldots, k)$ and $I = G \setminus \Delta_k(\Gamma)$ (that is, the set of Hales–Jewett sequences). Theorem 2.2 for this context is the Hales–Jewett theorem.

PROOF OF THEOREM 2.1. The first observation is that $E(I)$ is a two-sided ideal in $E(G)$. This follows from the fact that if A and B are subsets of Γ^k, then

$$E(A)E(B) \subset E(AB). \tag{2.3}$$

To see that, we take $\phi \in E(A)$, $\psi \in E(B)$ and check that $\phi\psi \in E(AB)$ by taking an arbitrary finite set $Y_1 \subset Y$, an arbitrary finite set $\Gamma_1 \subset \Gamma^k$ and producing $a \in A$ and $b \in B$ such that $\sigma(ab)y(\tilde{\gamma}) = \phi\psi(y)(\tilde{\gamma})$ for all $y \in Y_1$, $\tilde{\gamma} \in \Gamma_1$.

Begin by finding $a \in A$ such that $\sigma(a)$ is close enough to ϕ to guarantee

$$\sigma(a)\psi y(\tilde{\gamma}) = \phi\psi y(\tilde{\gamma}) \qquad \text{for all } y \in Y_1, \quad \tilde{\gamma} \in \Gamma_1.$$

Now $\sigma(a)\psi y(\tilde{\gamma}) = \psi y(\tilde{\gamma}a)$, so that if $b \in B$ is such that $\sigma(b)y(\tilde{\gamma}a) = \psi y(\tilde{\gamma}a)$ (which exist since $\psi \in E(B)$) then $\sigma(ab)y(\tilde{\gamma}) = \sigma(b)y(\tilde{\gamma}a) = \psi y(\tilde{\gamma}a) = \phi\psi y(\tilde{\gamma})$ for all $y \in Y_1$, $\tilde{\gamma} \in \Gamma_1$.

We have noticed earlier that $E(G) \supset \Delta_k(E(\Gamma))$; thus $\tilde{\theta} \in E(G)$, and by 1.5, $E(I)$ contains an idempotent $\tilde{\phi} = (\phi_1, \dots, \phi_k) < \tilde{\theta}$. That implies that for all j, ϕ_j is an idempotent and $\phi_j < \theta$. By Theorem 1.2, $\phi_j = \theta$ for all j and $\tilde{\theta} \in E(I)$. QED

Proof of Theorem 2.2. Denote by C the range of the given coloring and define $X = C^{\Gamma}$, the space of all colorings of Γ with range contained in C. Let $x \in X$ denote the given coloring and let $\tilde{\theta} = (\theta, \dots, \theta) \in E(I)$. Write $x_0 = \theta x$, and let $\tilde{\gamma} = (\gamma_1, \dots, \gamma_k) \in I$ be such that $\sigma(\tilde{\gamma})$ and $\tilde{\theta}$ are close enough so that $\sigma(\tilde{\gamma})(x, x, \dots, x)$ and $\tilde{\theta}(x, x, \dots, x) = (x_0, \dots, x_0)$ assign the same color, coordinate by coordinate, to the identity e. But that means that $x(\gamma_j) = x_0(e)$ for all j. QED

The proof of Theorem 2.2 does not use the fact that $\tilde{\theta}$ is an idempotent, only the fact that $E(I)$ intersects the diagonal. The following theorem, which is a slightly veiled generalization of the Carlson–Simpson theorem, does use the idempotence.

Theorem 2.5. *As in Theorem 2.2, assume $G \supset \Delta_k(\Gamma)$ and I a two-sided ideal in G. Then for any finite coloring of Γ, there exists a sequence $\{\tilde{\gamma}_n\} \subset I$, $\tilde{\gamma}_n = (\gamma_{n,1}, \dots, \gamma_{n,k})$, such that the set $\tilde{\Gamma} = \tilde{\Gamma}(\{\tilde{\gamma}_n\})$ of all finite products $\{\gamma_{1,j_1} \cdots \gamma_{m,j_m}\}$ is monochromatic.*

Proof. Keeping the notation of the proof of Theorem 2.2, we shall show that one can choose $\{\tilde{\gamma}_n\} \subset I$ such that the color of every $\{\gamma_{1,j_1} \cdots \gamma_{m,j_m}\}$ is $x_0(e)$. Take $\tilde{\gamma}_1 \in I$ so that $\sigma(\tilde{\gamma}_1)$ and $\tilde{\theta}$ are close in their actions on $y = (x, x, \dots, x)$ and $y_0 = (x_0, \dots, x_0)$, enough so that corresponding coordinates of $\sigma(\tilde{\gamma}_1)(x, x, \dots, x)$ and $\tilde{\theta}(x, x, \dots, x) = (x_0, \dots, x_0)$ assign the same color to the identity e, and similarly for $\sigma(\tilde{\gamma}_1)(x_0, \dots, x_0)$ and $\tilde{\theta}(x_0, \dots, x_0) = (x_0, \dots, x_0)$. This means

$$(2.4) \qquad x(\gamma_{1,j}) = x_0(\gamma_{1,j}) = x_0(e), \qquad 1 \leqq j \leqq k.$$

Now choose $\tilde{\gamma}_2$ so that its action matches that of $\tilde{\theta}$ on y, y_0 as above, but closer, so that corresponding components assign the same color not only to e but to the points $\{\gamma_{1,j}\}$, $j = 1, \dots, k$ as well. This means

$$x(\gamma_{1,j_1}\gamma_{2,j_2}) = \sigma(\gamma_{2,j_2})x(\gamma_{1,j_1}) = \theta x(\gamma_{1,j_1}) = x_0(\gamma_{1,j_1}) = x_0(e).$$

Similarly

$$x_0(\gamma_{1,j_1}\gamma_{2,j_2}) = \sigma(\gamma_{2,j_2})x_0(\gamma_{1,j_1}) = \theta x_0(\gamma_{1,j_1}) = x_0(\gamma_{1,j_1}) = x_0(e).$$

Having chosen $\tilde{\gamma}_n$, $n \leqq m$, we choose $\tilde{\gamma}_{m+1}$ so that its action matches that of $\tilde{\theta}$ on y and on y_0 close enough so that corresponding components assign the same color to each of the points $\{\gamma_{1,j_1}\cdots\gamma_{m,j_m}\}$ as well as to e. As above

$$\begin{aligned} x(\gamma_{1,j_1}\cdots\gamma_{m+1,j_{m+1}}) &= \sigma(\gamma_{m+1,j_{m+1}})x(\gamma_{1,j_1}\cdots\gamma_{m,j_m}) \\ &= \theta x(\gamma_{1,j_1}\cdots\gamma_{m,j_m}) \\ &= x_0(\gamma_{1,j_1}\cdots\gamma_{m,j_m}) = x_0(e), \end{aligned} \tag{2.5}$$

the last equality by induction. The analogous check for x_0, as above, prepares the next stage of the induction. QED

REMARK. The Carlson–Simpson theorem is our Theorem 2.5 in the context of Example 2.4. A slight rewording of the proof shows that if we write $\gamma_{m,0} = e$ for all m, even the set $\tilde{\Gamma}_0$ of all finite products $\{\gamma_{1,j_1}\cdots\gamma_{m,j_m}\}$ with at least one index $j_m \neq 0$, is monochromatic. In this form our Theorem 2.5 is also an extension of Hindman's theorem.

The set $\tilde{\Gamma}(\{\tilde{\gamma}_n\})$ is a combinatorial homomorphic image of $W(k)$ under the map

$$\tilde{\gamma} : (j_1, \ldots, j_m) \mapsto \gamma_{1,j_1}\cdots\gamma_{m,j_m} \tag{2.6}$$

and $\tilde{\Gamma}_0(\{\tilde{\gamma}_n\})$ is a generalisation of our notion of an IP-system in Γ, cf. [FK2]. Notice that neither $\tilde{\Gamma}(\{\tilde{\gamma}_n\})$ nor $\tilde{\Gamma}_0(\{\tilde{\gamma}_n\})$ need be subsemigroups of Γ.

Our final theorem of this section deals with a variant of Example 2.4.

EXAMPLE 2.6. $\Gamma = W(k, l)$ denotes the free semigroup generated by the $k + l$ symbols $\{1, 2, \ldots, k, t_1, \ldots, t_l\}$. The reason that we write $W(k, l)$ and not $W(k + l)$, even though the two are clearly the same, is the different role that we have in mind for the last l digits; namely, we want to regard these as variables. We denote by $\Gamma_0 = W(k)$ the span of the generators $\{1, 2, \ldots, k\}$. $G \subset \Gamma^{k+l}$ is the subsemigroup spanned by $\Delta_{k+l}(\Gamma_0)$ and the element $(1, 2, \ldots, k, t_1, \ldots, t_l)$. Finally $I = G \setminus \Delta_{k+l}(\Gamma_0)$. Note that the $(k + l)$-tuples of G are Hales–Jewett sequences whose common part is taken from $\{1, 2, \ldots, k\}$.

Denote by $\Xi(l)$ the space of all finite words on the l digits in $\Lambda = \{t_1, \ldots, t_l\}$, with no runs longer than one (i.e., no consecutive multiple appearance of any digit), and by $\Xi_N(l)$ the set of words in $\Xi(l)$ whose length is less than or equal to N.

Define $\pi_l : \Gamma \mapsto \Xi(l)$ to be the collapsing map which erases all the digits not in Λ, shortens the runs to singletons, and removes all the empty spaces. For example, for $l = 1$, $\Xi(1)$ is the set containing the empty word and the singleton t_1, and π_1 just distinguishes between words that contain the digit t_1 and those that do not. We extend the notation to sequences of the form $j_1, \ldots, j_m$ with $1 \leqq j \leqq k + l$ and write $\pi_l(j_1, \ldots, j_m)$ as if the sequence were a word in Γ.

Given a partition (coloring) of $\Xi(l)$, we obtain a coloring of Γ by lifting via π_l.

THEOREM 2.7. *Denote by I the ideal defined in Example 2.6. Given any finite coloring x of $\Gamma = W(k, l)$, and an integer N, there exists a sequence $\{\tilde{\gamma}_n\} \subset I$, $\tilde{\gamma}_n = (\gamma_{n,1}, \ldots, \gamma_{n,k+l})$, such that the restriction of x to the set of all finite products $\{\gamma_{1,j_1} \cdots \gamma_{m,j_m}\}$ for which $\pi_l(j_1, \ldots, j_m) \in \Xi_N(l)$, is a lift via π_l of a coloring of $\Xi_N(l)$.*

PROOF. Let θ be a minimal idempotent in $E(\Gamma_0)$ and denote, as above $\tilde{\theta} = (\theta, \ldots, \theta) \in E(G)$. By 1.5, $E(I)$ contains an idempotent $\tilde{\phi} = (\phi_1, \ldots, \phi_{k+l}) < \tilde{\theta}$. That implies that for every j, ϕ_j is an idempotent which is minimal in the enveloping semigroup of the action on the jth coordinate, i.e. $E(\Gamma_0)$ for $j \leqq k$, and $E(\Gamma_j)$ where Γ_j denotes the span of Γ_0 and t_{j-k}, for $k < j \leqq k + l$. It follows that

$$\tilde{\phi} = (\theta, \ldots, \theta, \phi_{k+1}, \ldots, \phi_{k+l}),$$

and $\phi_j < \theta$.

Write $\phi(j_1, \ldots, j_m) = \phi_{j_1} \cdots \phi_{j_m}$. The idempotence of ϕ_j and the order relations imply that $\phi(j_1, \ldots, j_m) = \phi(\pi_l(j_1, \ldots, j_m))$. We shall show that we can choose the sequence $\{\tilde{\gamma}_n\} \subset I$, $\tilde{\gamma}_n = (\gamma_{n,1}, \ldots, \gamma_{n,k+l})$, such that

$$x(\gamma_{1,j_1} \cdots \gamma_{m,j_m}) = \phi(j_1, \ldots, j_m)x(e). \tag{2.7}$$

Since $\phi(j_1, \ldots, j_m)$ only depends on $\pi_l(j_1, \ldots, j_m)$, this will prove our theorem. In fact, it will be more convenient for the induction to claim the (formally) stronger claim, namely

$$\phi(r_1, \ldots, r_s)x(\gamma_{1,j_1} \cdots \gamma_{m,j_m}) = \phi(j_1, \ldots, j_m, r_1, \ldots, r_s)x(e), \tag{2.8}$$

provided $\pi(j_1, \ldots, j_m, r_1, \ldots, r_s) \in \Xi_N(l)$.

Denote by Y_N the set $\{x\} \cup \{\phi(j_1, \ldots, j_m)x : \pi(j_1, \ldots, j_m) \in \Xi_N(l)\}$, and by $\tilde{Y}_N$ the set in X^{k+l} of all the elements of the form $(y, \ldots, y) : y \in Y_N$.

Take $\tilde{\gamma}_1 \in I$ so that $\sigma(\tilde{\gamma}_1)$ and $\tilde{\phi}$ are close in their actions on every $z \in \tilde{Y}_N$, enough so that corresponding coordinates of $\sigma(\tilde{\gamma}_1)z$ and $\tilde{\phi}z$ assign the same color to the identity e. This gives (2.8) for $m = 1$. Checking for $z = x$ we obtain (2.7) for $m = 1$.

We proceed by induction. Assume that we have chosen $\{\tilde{\gamma}_m\}$, $m < n$, and that (2.8) is valid for $m < n$ and all "vectors" $\{(j_1, \ldots, j_m, r_1, \ldots, r_s)$ such that $\pi(j_1, \ldots, j_m, r_1, \ldots, r_s) \in \Xi_N(l)\}$.

Take $\tilde{\gamma}_n \in I$ so that $\sigma(\tilde{\gamma}_n)$ and $\tilde{\phi}$ are close in their actions on every $z \in \tilde{Y}_N$, enough so that corresponding coordinates of $\sigma(\tilde{\gamma}_n)z$ and $\tilde{\phi}z$ assign the same color to the identity e as well as to each $\gamma_{1,j_1} \cdots \gamma_{m,j_m}$, $1 \leqq m \leqq n-1$.

We have

$$\begin{aligned} \phi(r_1, \ldots, r_s)x(\gamma_{1,j_1} \cdots \gamma_{n,j_n}) &= \sigma(\gamma_{n,j_n})\phi(r_1, \ldots, r_s)x(\gamma_{1,j_1} \cdots \gamma_{n-1,j_{n-1}}) \\ &= \phi_{j_n}\phi(r_1, \ldots, r_s)x(\gamma_{1,j_1} \cdots \gamma_{n-1,j_{n-1}}) \qquad (2.9) \\ &= \phi(j_1, \ldots, j_n, r_1, \ldots, r_s)x(e) \end{aligned}$$

provided $\pi_l(j_1, \ldots, j_m, r_1, \ldots, r_s) \in \Xi_N(l)$. QED

REMARK. The case $N = 1$ of Theorem 2.7 was proved by Carlson, [C]. The case $k = 1$, $l = 1$ is Hindman's theorem. The case $k = 1$, l arbitrary, and the products $\gamma_{1,j_1} \cdots \gamma_{m,j_m}$ for which $\pi_l(j_1 \cdots j_m) = (t_1 \cdots t_l)$ (being monochromatic) was proved independently by Milliken, and by Taylor; see [H].

The need to specify N in the statement of Theorem 2.7 is apparent in the proof. The order in which the $\sigma(\gamma)$'s act is such that $\sigma(\gamma_{1,j_1})$ acts last while we have to choose $\tilde{\gamma}_1$ first. It is not hard to show that the statement of Theorem 2.7 without the limitation given by N is false. Consider for example the following coloring of $W(k, l)$: denote by $b(w)$ the index of the first occurrence of one of the digits $t_1, \ldots, t_l$ in w; denote by $p(w)$ the length of the word $\pi_l(w) \in \Xi(l)$. Color w white if $b(w) \leqq p(w)$, black otherwise. It is clear that if we restrict ourselves to words with $p(w) \leqq N$ we can obtain the sequence $\{\tilde{\gamma}_n\}$ such that $b(\gamma_{1,j_1} \cdots \gamma_{n,j_n})$ is always bigger than N (just by choosing $\tilde{\gamma}_1$ properly), so that the words generated are all black. On the other hand if $p(w)$ is unrestricted, then for any choice of $\{\tilde{\gamma}_n\} \subset I$, we will have both $b(w) \leqq p(w)$ and $p(w) < b(w)$ for w of the form $\gamma_{1,j_1} \cdots \gamma_{n,j_n}$. In some sense this is the only "counter-example". For any choice of a sequence $\{N(q)\}$ (even choosing $N(q+1)$ after $\tilde{\gamma}_q$) add to equation (2.8) the condition

$$\phi(r_1, \ldots, r_s)x(\gamma_{1,j_1}\cdots\gamma_{m,j_m}) = \phi(j_{q+1}, \ldots, j_m, r_1, \ldots, r_s)x(j_1, \ldots, j_q) \tag{2.8'}$$
$$\text{provided } \pi(j_{q+1}, \ldots, j_m, r_1, \ldots, r_s) \in \Xi_{N(q)}(l);$$

and our proof above gives

THEOREM 2.8. *Denote by I the ideal defined in Example* 2.6. *Given any finite coloring* x *of* $\Gamma = W(k, l)$, *there exists a sequence* $\{\tilde{\gamma}_n\} \subset I$, $\tilde{\gamma}_n = (\gamma_{n,1}, \ldots, \gamma_{n,k+l})$, *such that*

$$x(\gamma_{1,j_1}\cdots\gamma_{m,j_m}) = \phi(j_{q+1}\cdots j_m)x(\gamma_{1,j_1}\cdots\gamma_{q,j_q}) \tag{2.10}$$

provided $\pi(j_{q+1}, \ldots, j_m) \in \Xi_{N(q)}(l)$.

3. Combinatorial spaces and subspaces

If $k = p^\nu$ is a prime power, we can identify the words of length d of $W(k)$ with points of a d-dimensional vector space over the field with k elements. With this identification the Hales–Jewett theorem is used to prove that if we color the points of a sufficiently high dimensional vector space over a fixed finite field with a fixed number of colors, then there will be a monochromatic r-plane, where r is preassigned. Based on this analogy we refer to $W(k)$ as an infinite dimensional *combinatorial space*.

Consider now the word $t_1t_2\cdots t_l$ in $\Xi(l)$ and let $W^*(k, l) = W(k, l) \cap \pi_l^{-1}(t_1t_2\cdots t_l)$. The words of $W^*(k, l)$ contain all the variables $t_1, t_2, \ldots, t_l$ and in order. If v is a word in $W^*(k, l)$, we can think of it as a variable point in $W(k)$ and letting the variables $t_1, \ldots, t_l$ undergo all the substitutions $t_j \mapsto a_j \in \{1, \ldots, k\}$ we obtain the analogue of an l-dimensional subspace. We take this to be the definition of a *combinatorial l-dimensional subspace* of $W(k)$, namely it is the range of some $v \in W^*(k, l)$ as $(t_1, \ldots, t_l)$ ranges over $\{1, \ldots, k\}^l$. Alternatively we could write $v = w_1(t_1)w_2(t_2)\cdots w_l(t_l)$ where $w_j(t)$ are monomials in one variable, i.e., elements of $W(k, 1)$. The points of the subspace in question will then be

$$\{w_1(j_1)\cdots w_l(j_l) \mid j_1, \ldots, j_l \in \{1, 2, \ldots, k\}\}.$$

In the same way we can define infinite dimensional subspaces of $W(k)$ by choosing an infinite sequence of words $\{w_j(t)\} \subset W(k, 1) \setminus W(k)$ (so that t actually occurs in each $w_j(t)$). We then set

$$\tilde{w}(j_1, \ldots, j_n) = w_1(j_1)\cdots w_n(j_n), \qquad j_i \in \{1, \ldots, k\}$$

and refer to $\tilde{w}(W(k))$ as the *infinite dimensional combinatorial subspace* of $W(k)$ defined by the sequence $w_1, w_2, \ldots$.

It is not hard to see that if L is a d-dimensional subspace of $W(k)$, then $\tilde{w}(L)$ is again a d-dimensional subspace contained in $\tilde{w}(W(k))$. Moreover, all combinatorial subspaces of $\tilde{w}(W(k))$ are obtained in this way.

The following result may now be deduced from Theorem 2.7:

THEOREM 3.1. *Given an arbitrary finite coloring of the set of all l-dimensional combinatorial subspaces in $W(k)$, there exists an infinite dimensional combinatorial subspace $\tilde{w}(W(k))$ all of whose l-dimensional combinatorial subspaces are colored by the same color.*

PROOF. A coloring of the l-dimensional combinatorial subspaces in $W(k)$ is a coloring of $W(k, l) \cap \pi_l^{-1}(t_1 t_2 \cdots t_l)$. Extend this arbitrarily to $W(k, l)$. By Theorem 2.7 we find $\{\tilde{\gamma}_n\} \subset I \subset$ the set of $(k+l)$-tuples of words, so that the set of all products $\{\gamma_{1,j_1} \cdots \gamma_{m,j_m}\}$ with $\pi_l(j_1, \ldots, j_m) = t_1 \cdots t_l$ have the same color. Since $\tilde{\gamma}_n \in I$, we can write

$$\tilde{\gamma}_n = (w_n(1), w_n(2), \ldots, w_n(k), w_n(t_1), \ldots, w_n(t_l))$$

for some monomial w_n. It is not hard to see that the products $\gamma_{1,j_1} \cdots \gamma_{m,j_m}$ with $j_1, \ldots, j_m \in \{1, \ldots, k\}$ form an infinite dimensional combinatorial subspace whose l-dimensional subspaces correspond to $j_1, \ldots, j_m \in \{1, 2, \ldots, k, t_1, \ldots, t_l\}$ for which $\pi_l(j_1, \ldots, j_m) = t_1 t_2 \cdots t_l$. Since this is monochromatic, the theorem follows. QED

4. Compact coloring — uniform continuity

We stated Theorem 2.2, 2.5, 2.7 and 2.8 for the classical case of a finite coloring. If the space of colors, C, is an infinite compact metric space, we can cover it by a finite number of ε-balls, and replace "monochromatic" in the statements of the theorems by "contained in an ε-ball". This is all we can do for Theorem 2.2, but for Theorems 2.5 and 2.7 we can improve the results by decreasing ε as the induction proceeds and, defining an appropriate metric on the set of finite products appearing in the statements, obtain that with respect to that metric the coloring is uniformly continuous on the set of finite products. Our interest in compact coloring is motivated by applications in the context of $W(k)$ and $W(k, l)$, and we confine our discussion to that context (noting though that since any finitely generated semigroup is the homomorphic image of $W(k)$, there is no real loss of generality in doing so).

Recall that the elements of $W(k)$ are finite words $w = w(1)w(2)\cdots w(n)$ with each $w(j) \in \{1, 2, \ldots, k\}$, and the operation is concatenation. We introduce a metric on $W(k)$ by

$$\rho(w_1, w_2) = \inf\{2^{-q} \mid w_1(j) = w_2(j) \quad \text{for } 1 \leqq j \leqq q\}. \tag{4.1}$$

Thus, two distinct words are close if they are both long and they match for a long time. The completion of $W(k)$ in this metric is (can be identified with) $W(k) \cup \Omega(k)$ where $\Omega(k)$ is the space of all infinite words on $\{1, 2, \ldots, k\}$.

For $W(k, l)$ we do not simply take the metric induced by that of $W(k + l)$; instead we take

$$\rho'(w_1, w_2) = \inf\{2^{-q} \mid w_1(j) = w_2(j) \in \{1, 2, \ldots, k\} \quad \text{for } 1 \leqq j \leqq q\}. \tag{4.2}$$

The completion of $W(k, l)$ relative to this metric is $W(k, l) \cup \Omega(k)$, its infinite words using only the digits $\{1, 2, \ldots, k\}$.

Recall that a combinatorial subspace $\tilde{W}(\{\tilde{w}_n\})$ of $W(k)$ (resp. $W(k, l)$) is defined by means of a sequence[†] $\{\tilde{w}_n\} \subset I \subset W(k)^k$ (resp. $W(k, l)^{k+l}$), I being the ideal described in Example 2.4 (resp. 2.6), as the combinatorial homomorphic image of $W(k)$ (resp. $W(k, l)$) under the map

$$\tilde{w} : (j_1, \ldots, j_m) \mapsto w_{1, j_1} \cdots w_{m, j_m} \tag{4.3}$$

THEOREM 4.1. *Let C be a compact metric space, and $f \colon W(k) \to C$ an arbitrary function. Then there exists a combinatorial subspace $\tilde{W}(\{\tilde{w}_n\}) \subset W(k)$ such that the restriction of f to it is uniformly continuous.*

PROOF. Repeat the proof of Theorem 2.5, using ε-balls as suggested in the observation made at the beginning of the section. QED

The completion of the combinatorial subspace $\tilde{W}(\{\tilde{w}_n\})$ is obtained by adding to it the "points at infinity" which are the infinite words of the form $w_{1, j_1} \cdots w_{m, j_m} \cdots$, that is the image $\tilde{\Omega}(\{\tilde{w}_n\})$ of $\Omega(k)$ under the obvious extension of $\tilde{w}$. The uniform continuity of f on $\tilde{W}(\{\tilde{w}_n\})$ permits extension by continuity to its completion.

THEOREM 4.2. *Let C be a compact metric space, and $f \colon W(k, l) \to C$ an arbitrary function. Then there exists a combinatorial subspace $\tilde{W}(\{\tilde{w}_n\}) \subset$*

[†] Our notation in (2.6) used the letter γ and we replace it by w since now $\Gamma = W(k)$.

$W(k, l)$ such that the restriction of f to $\tilde{W}(\{\tilde{w}_n\}) \cap \pi_l^{-1}\xi$ is uniformly continuous for every $\xi \in \Xi(l)$.

PROOF. This is an immediate consequence of Theorem 2.8 and the "diminishing ε-balls observation". QED

REMARK. The completion of $\tilde{W}(\{\tilde{w}_n\}) \cap \pi_l^{-1}\xi$ is again its union with $\tilde{\Omega}(\{\tilde{w}_n\})$ defined above. Thus for every $\xi \in \Xi(l)$ we obtain an extension f_ξ of f to $\tilde{\Omega}(\{\tilde{w}_n\})$.

REFERENCES

[BH] V. Bergelson and N. Hindman, *Non-metrizable topological dynamics and Ramsey theory*, preprint.

[C] T. J. Carlson, *Some unifying principles in Ramsey theory*, Discrete Math. **68** (1988), 117–169.

[CS] T. J. Carlson and S. G. Simpson, *A dual form of Ramsey's theorem*, Advances in Math. **53** (1984), 265–290.

[E] R. Ellis, *Introduction to Topological Dynamics*, W. A. Benjamin, New York, 1969.

[FK] H. Furstenberg and Y. Katznelson, *A density version of the Hales–Jewett theorem for* $k = 3$, Discrete Math. **75** (1989), 227–242.

[FK1] H. Furstenberg and Y. Katznelson, *A density version of the Hales–Jewett theorem*, to appear.

[FK2] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. Analyse Math. **45** (1985), 117–168.

[GR] R. Graham and B. Rothschild, *Ramsey's theorem for n-parameter sets*, Trans. Am. Math. Soc. **159** (1971), 257–292.

[HJ] A. W. Hales and A. I. Jewett, *Regularity and positional games*, Trans. Am. Math. Soc. **106** (1963), 222–229.

[H] N. Hindman, *Ultrafilters and combinatorial number theory*, in *Number Theory Carbondale 1979*, Lecture Notes in Math. #751, Springer-Verlag, Berlin, 1979, pp. 119–184.