

The Density Hales-Jewett Theorem in a Measure Preserving Framework

Daniel Glasscock, April 2013

The purpose of this note is to explain in detail the reduction of the combinatorial density Hales-Jewett theorem (DHJ) to an equivalent measure-theoretic dynamical theorem. This approach was used by Furstenberg and Katznelson in the first proof [3] of DJH in 1991. The approach taken here more closely resembles their earlier work [2] on DHJ₃.

1. Introduction

Let $k \in \mathbb{N}$, and write $[k]$ for $\{1, 2, \dots, k\}$. The set of all words with letters from $[k]$ of length n is then $[k]^n$. Let $[k]^{<\omega}$ denote the free semigroup of all words with letters from $[k]$ of positive, finite length with concatenation.

A *variable word* is a word with letters from $[k] \cup \{t\}$ in which t appears. Variable words should be considered non-constant functions $[k] \rightarrow [k]^{<\omega}$ defined by substituting letters from $[k]$ in place of t . A *combinatorial line* in $[k]^{<\omega}$ is the image of a variable word.

Theorem (DHJc): For all $k \in \mathbb{N}$ and $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and $A \subseteq [k]^n$ with $|A| > \epsilon k^n$, the set A contains a combinatorial line.

A $[k]^{<\omega}$ -system is a probability space (X, μ) together with k infinite sequences of invertible, measure preserving transformations $\{T_j^{(i)}\}_{j \in \mathbb{N}}^{i \in [k]}$ and a map $T : [k]^{<\omega} \rightarrow \langle \{T_j^{(i)}\}_{j \in \mathbb{N}}^{i \in [k]} \rangle$ defined by

$$T(w) = T_1^{(w_1)} \circ T_2^{(w_2)} \circ \dots \circ T_n^{(w_n)}$$

where $w = w_1 w_2 \dots w_n$, $w_i \in [k]$.

Theorem (DHJm): For all $k \in \mathbb{N}$, $[k]^{<\omega}$ -systems (X, μ, T) , and $A \subseteq X$ of positive measure, there exists a combinatorial line L such that

$$\mu \left(\bigcap_{w \in L} T(w)^{-1} A \right) > 0.$$

It is in this form that Furstenberg and Katznelson [3] proved the density Hales-Jewett theorem. The goal here is to prove the equivalence between these two forms of the theorem.

2. Equivalent forms

The first step is to move **DHJc** into a “static” measure-theoretic form. A $[k]^{<\omega}$ -process in a probability space (X, μ) is a collection of measurable sets $\{B_w\}_{w \in [k]^{<\omega}}$ indexed by $[k]^{<\omega}$. A $[k]^n$ -process is defined similarly. We will show by combinatorial arguments that **DHJc** is equivalent to the following static form of the density Hales-Jewett theorem.

Theorem (DHJf): Let (X, μ) be a probability space. For all $k \in \mathbb{N}$ and $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and $[k]^n$ -processes $\{B_w\}$ in X for which $\mu(B_w) > \epsilon$ for all $w \in [k]^n$, there exists a combinatorial line L in $[k]^n$ such that $\mu(\cap_{w \in L} B_w) > 0$.

The following equivalent infinitary version will be more useful in establishing the dynamical connection. The equivalence is left as an easy exercise.

Theorem (DHJi): Let (X, μ) be a probability space. For all $k \in \mathbb{N}$, $\epsilon > 0$, and $[k]^{<\omega}$ -processes $\{B_w\}$ in X for which $\mu(B_w) > \epsilon$ for all $w \in [k]^{<\omega}$, there exists a combinatorial line L such that $\mu(\cap_{w \in L} B_w) > 0$.

The probability space (X, μ) plays an unimportant role here. What really matters is the relative position of the sets amongst themselves. In order to highlight this, we will focus on the following object. The (joint) distribution of a $[k]^{<\omega}$ -process $\{B_w\}$ in X is the function d which for all $n \in \mathbb{N}$ is defined on subsets $W \subseteq [k]^n$ by

$$d(W) = \mu \left(\bigcap_{w \in W} B_w \right).$$

Since many different $[k]^{<\omega}$ -processes may have the same distribution, it is worthwhile to consider distributions without making reference to the processes from which they come. If $Z_k = \cup_n \mathcal{P}([k]^n)$ is the disjoint union of the power sets of each $[k]^n$, then distributions of $[k]^{<\omega}$ -processes (as functions on Z_k with values in $[0, 1]$) are points in the compact metric space $[0, 1]^{Z_k}$. (An example of a metric: $\sum_n 2^{-k^n - n} \sum_{W \subseteq [k]^n} d(W)$.) A point $\varphi \in [0, 1]^{Z_k}$ is called a (joint) distribution if it is the distribution of a $[k]^{<\omega}$ -process in some probability space (such a process is called a *parent* process for φ).

The equivalence between the previous theorem and the next is left as an easy exercise.

Theorem (DHJd): For all $k \in \mathbb{N}$, $\epsilon > 0$, and distributions $d \in [0, 1]^{Z_k}$ for which $d(\{w\}) > \epsilon$ for all $w \in [k]^{<\omega}$, there exists a combinatorial line L such that $d(L) > 0$.

To motivate the next definition, consider the $[k]^{<\omega}$ -process $\{T(w)^{-1}A\}$ associated with a $[k]^{<\omega}$ -system (X, μ, T) and a set $A \subseteq X$ of positive measure. This process is “stationary” in the following sense: if w_1 and w_2 are of the same length, then for all $v \in [k]^{<\omega}$,

$$\mu(T(w_1v)^{-1}A \cap T(w_2v)^{-1}A) = \mu(T(w_1)^{-1}A \cap T(w_2)^{-1}A).$$

A $[k]^{<\omega}$ -process $\{B_w\}$ in (X, μ) is *stationary* if for all $n \in \mathbb{N}$, $W \subseteq [k]^n$, and $v \in [k]^{<\omega}$,

$$\mu \left(\bigcap_{w \in W} B_{wv} \right) = \mu \left(\bigcap_{w \in W} B_w \right).$$

A distribution d is called *stationary* if for all $n \in \mathbb{N}$, $W \subseteq [k]^n$, and $v \in [k]^{<\omega}$, $d(Wv) = d(W)$. It follows from the definitions that a distribution is stationary if and only its parent processes are stationary.

It will be shown later that an arbitrary distribution has subdistributions which approximate stationary ones. This fact will allow us to show the equivalence of **DHJd** and the following stationary form.

Theorem (DHJs): For all $k \in \mathbb{N}$, $\epsilon > 0$, and stationary distributions $d \in [0, 1]^{Z_k}$ for which $d(\{w\}) > \epsilon$ for all $w \in [k]^{<\omega}$, there exists a combinatorial line L such that $d(L) > 0$.

The final step will be to show the equivalence of **DHJs** and **DHJm**. This is accomplished by realizing any stationary distribution as a distribution of a $[k]^{<\omega}$ -orbit of some set of positive measure.

In summary, we will prove

$$\mathbf{DHJc} \iff \mathbf{DHJf} \iff \mathbf{DHJi} \iff \mathbf{DHJd} \iff \mathbf{DHJs} \iff \mathbf{DHJm}.$$

3. Equivalence of **DHJc** and **DHJf**

Both of the following proofs are directly from [3]. First, **DHJc** implies **DHJf**.

Proof. Let $k \in \mathbb{N}$, $\epsilon > 0$. Let N be as given in **DHJc**. Let $n \geq N$ and $\{B_w\}$ be a $[k]^n$ -process in X for which $\mu(B_w) > \epsilon$ for all $w \in [k]^n$. For all $x \in X$, let $A(x) = \{w \in [k]^n \mid x \in B_w\}$. Then $|A(x)| = \sum_{w \in [k]^n} \chi_{B_w}$, so

$$\int_X |A(x)| d\mu > \epsilon k^n.$$

Thus there exists a $C \subseteq X$ of positive measure such that for all $x \in C$, $|A(x)| > \epsilon k^n$. By **DHJc**, for all $x \in C$, the set $A(x) \subseteq [k]^n$ contains a combinatorial line $L(x)$. Since there are only finitely many combinatorial lines in $[k]^n$, there exists a $C' \subseteq C$ of positive measure such that $L(x) = L$ is constant on C' . This means that $C' \subseteq \cap_{w \in L} B_w$, and we are done since C' has positive measure. \square

Conversely, **DHJf** implies **DHJc**.

Proof. Let $k \in \mathbb{N}$. First, we claim that **DHJc** holds with $N = 2$ if $\epsilon > 1 - 1/k$. Let $n \geq N$ and $A \subseteq [k]^n$ with $|A| > \epsilon k^n$. For $i \in [k]$, let $A_i = \{w \in A \mid w_1 = i\}$ and $A'_i \subseteq [k]^{n-1}$ be the projection of A_i onto the last $n-1$ letters. There exists a $v \in \cap_{i \in [k]} A'_i$ since otherwise $|A| \leq k^n - k^{n-1}$, a contradiction. Then $\{iv \mid i \in [k]\}$ is a combinatorial line in A .

Let ϵ_0 be the infimum over all ϵ for which the conclusion of **DHJc** is valid. If $\epsilon_0 = 0$, we are done. Otherwise, $0 < \epsilon_0 \leq 1 - 1/k$. Let m be large enough to satisfy the conclusions of **DHJf** for processes of measure greater than $\epsilon_0/2$. Let $\epsilon_1 = \epsilon_0(1 - k^{-m-2})$ so that

$$\epsilon_2 = \epsilon_1 + \frac{\epsilon_0}{2} k^{-m} > \epsilon_0.$$

Since $\epsilon_2 > \epsilon_0$, there exists an M such that the conclusion of **DHJc** is valid for $n \geq M$ and sets of density greater than ϵ_2 . We claim that the conclusion of **DHJc** is valid for sets of density greater than ϵ_1 and $n \geq m + M$, which, since $\epsilon_1 < \epsilon_0$, contradicts the definition of ϵ_0 .

Let $n \geq m + M$ and $A \subseteq [k]^n$ have density greater than ϵ_1 . Consider the $[k]^m$ -process $\{A'_w\}$ in $[k]^{n-m}$ defined by $A'_w = \{u \in [k]^{n-m} \mid wu \in A\}$. There are two cases to consider. Case 1: each A'_w has density greater than $\epsilon_0/2$. By **DHJf**, there exists a combinatorial line $L \subseteq [k]^m$ such that $\cap_{w \in L} A'_w$ has positive measure (density). In particular, the intersection is non-empty, so there exists a $v \in \cap_{w \in L} A'_w$. Now Lv is a combinatorial line in A .

Case 2: there exists a $w \in [k]^m$ such that A'_w has density at most $\epsilon_0/2$. By double counting A , the average of the densities of the A'_w 's is the density of A . Since A has density greater than ϵ_1 , by the definition of ϵ_2 , there exists a $v \in [k]^m$ such that $A'_v \subseteq [k]^{n-m}$ has density greater than ϵ_2 . Since $n - m \geq M$, there exists a combinatorial line L in A'_v . Then vL is a combinatorial line in A . \square

4. Subspaces, subprocesses, and subdistributions

Notions of subspaces of $[k]^{<\omega}$, subprocesses, and subdistributions are essential to proving these equivalences. A *variable sentence* Σ is a concatenation $w_1(t_1)w_2(t_2)\cdots \in ([k] \cup \{t_i\}_{i \in \mathbb{N}})^{\mathbb{N}}$ of infinitely many variable words, each with a distinct variable. Variable sentences may be thought of as functions $\Sigma : [k]^{<\omega} \rightarrow [k]^{<\omega}$ by defining

$$\Sigma v = w_1(v_1)\cdots w_n(v_n)$$

where $v = v_1 \cdots v_n$, $v_i \in [k]$. A (*infinite dimensional*) *subspace* of $[k]^{<\omega}$ is a subset of $[k]^{<\omega}$ of the form $\Sigma[k]^{<\omega}$ for some variable sentence Σ . Thus, elements of subspaces of $[k]^{<\omega}$ are indexed by $[k]^{<\omega}$. Note that subspaces of $[k]^{<\omega}$ and variable sentences are in 1-to-1 correspondence.

Let Σ_1, Σ_2 be variable sentences. It is left to the reader to check that $\Sigma_1[k]^{<\omega} \subseteq \Sigma_2[k]^{<\omega}$ if and only if there exists a variable sentence Σ_3 for which $\Sigma_1 = \Sigma_2 \Sigma_3$. In case either (both) hold, $\Sigma_1[k]^{<\omega}$ is called a *subspace* of $\Sigma_2[k]^{<\omega}$. The first characterization is more apparent, while the second means that

subspaces of subspaces of $[k]^{<\omega}$ are subspaces of $[k]^{<\omega}$. Subsets of $[k]^{<\omega}$ of the form $\Sigma[k]^n$ for a variable sentence Σ are called *n-dimensional subspaces* of $[k]^{<\omega}$. Note that a combinatorial line is exactly a 1-dimensional subspace, and so a combinatorial line in a subspace of $[k]^{<\omega}$ is a combinatorial line in $[k]^{<\omega}$.

Let $\{B_w\}$ be a $[k]^{<\omega}$ -process in (X, μ) . A *subprocess* of $\{B_w\}$ is a $[k]^{<\omega}$ -process of the form $\{B_{\Sigma w}\}$ for some variable sentence Σ . Subprocesses of $\{B_w\}$ are also $\Sigma[k]^{<\omega}$ -processes, but it will be advantageous to consider all processes and subprocesses as indexed by $[k]^{<\omega}$. Similarly, a *subdistribution* of a distribution d is a distribution of the form $d \circ \Sigma$ for some variable sentence Σ . (By a slight abuse of notation, Σ acts element-wise to take subsets of $[k]^{<\omega}$ to subsets of $[k]^{<\omega}$.)

5. The regular process and space of distributions

The next lemma shows that the probability space (X, μ) plays a secondary role. Let m be the Lebesgue measure on $[0, 1]$, and for each $n \in \mathbb{N}$, fix an ordering of $\mathcal{P}([k]^n)$.

Lemma 1: *For all $k \in \mathbb{N}$ and distributions $d \in [0, 1]^{Z_k}$, there exists a parent process in $([0, 1], m)$.*

Proof. Let $k \in \mathbb{N}$, $d \in [0, 1]^{Z_k}$ be a distribution, and $\{B_w\}$ be a parent $[k]^{<\omega}$ -process in (X, μ) . We will describe a $[k]^{<\omega}$ -process in $([0, 1], m)$ with distribution d . Let $n \in \mathbb{N}$, and for each $W \subseteq [k]^n$, let $p(W) = \mu(\cap_{w \in W} B_w \setminus \cup_{w \notin W} B_w)$. That is, p measures the cells of the partition of X created by $\{B_w\}_{w \in [k]^n}$. For each $W \subseteq [k]^n$, in order, let I_W be a closed interval of length $p(W)$ in $[0, 1]$ disjoint from the previous ones with left endpoint situated as close to 0 as possible. This is possible since $\sum_{W \subseteq [k]^n} p(I_W) = 1$. For each $w \in [k]^n$, let $C_w = \cup_{W \ni w} I_W$. It is straightforward to check that the resulting $[k]^{<\omega}$ -process $\{C_w\}$ in $[0, 1]$ has distribution d . \square

Given a distribution d , the parent process $\{C_w\}$ created above will be called the *regular process* for d .

Lemma 2: *For all $k \in \mathbb{N}$, the set of distributions in $[0, 1]^{Z_k}$ is closed.*

Proof. Let $(d_i)_{i \in \mathbb{N}}$ be a sequence of distributions converging to φ in $[0, 1]^{Z_k}$. For each $i \in \mathbb{N}$, let $\{C_{i,w}\}$ be the regular $[k]^{<\omega}$ -process for d_i . For each $w \in [k]^{<\omega}$, the sequence of endpoints of the intervals composing $C_{i,w}$ converge to the endpoints of a set of intervals which compose a C_w . It is easily verified that the resulting $[k]^{<\omega}$ -process $\{C_w\}$ in $[0, 1]$ has distribution φ . \square

6. Equivalence of DHJd and DHJs

That **DHJd** implies **DHJs** is clear. The reverse implication will follow by the compactness of $[0, 1]^{Z_k}$ and Lemma 2. More specifically, we will show that any distribution has subdistributions limiting to a stationary distribution. After applying **DHJs** to the corresponding stationary distribution, we may approximate it well enough to carry the conclusion back to a subdistribution of the original distribution.

In order to accomplish this, we need to formalize the notion of being close to stationary. If $\epsilon > 0$, a distribution d is ϵ -stationary if for all $n \leq \epsilon^{-1}$, $W \subseteq [k]^n$, and $v \in [k]^{<\omega}$,

$$|d(Wv) - d(W)| < \epsilon.$$

For a distribution $d \in [0, 1]^{Z_k}$, let $S(d)$ be the closure in $[0, 1]^{Z_k}$ of the set of subdistributions of d .

To prove the next lemma, we will make use of the following theorem due to Carlson and Simpson [1]. It can be thought of as having the same relationship to the Hales-Jewett theorem as Hindman's theorem has to Schur's.

Theorem (CS): For all $k \in \mathbb{N}$ and finite colorings of $[k]^{<\omega}$, there exists a monochromatic subspace.

Lemma 3: For all $k \in \mathbb{N}$ and distributions $d \in [0, 1]^{Z_k}$, $S(d)$ contains stationary distributions.

Proof. Suppose first that $S(d)$ contains ϵ -stationary distributions for arbitrary small ϵ . For each $i \in \mathbb{N}$, let d_i be an ϵ_i -stationary distribution in $S(d)$ where $\epsilon_i \searrow 0$. Since $[0, 1]^{Z_k}$ is compact, there exists a subsequence $(d_{n_i})_{i \in \mathbb{N}}$ which converges to a point $\varphi \in [0, 1]^{Z_k}$. Since $S(d)$ and the set of distributions are both closed, φ is a distribution in $S(d)$.

To see that φ is stationary, let $n \in \mathbb{N}$, $W \subseteq [k]^n$, and $v \in [k]^{<\omega}$. Let $\delta > 0$. Choose $i \in \mathbb{N}$ such that $\epsilon_{n_i} < \min(1/n, \delta/3)$ and such that the distance between d_{n_i} and φ in the metric on $[0, 1]^{Z_k}$ is so small that $|\varphi(W) - d_{n_i}(W)| < \delta/3$ and $|\varphi(Wv) - d_{n_i}(Wv)| < \delta/3$. Then

$$|\varphi(Wv) - \varphi(W)| \leq |\varphi(Wv) - d_{n_i}(Wv)| + |d_{n_i}(Wv) - d_{n_i}(W)| + |d_{n_i}(W) - \varphi(W)| < \delta.$$

Since δ was arbitrary, $\varphi(Wv) = \varphi(W)$, so φ is stationary.

Thus it suffices to show that for all $\epsilon > 0$, $S(d)$ contains an ϵ -stationary distribution. Let $\epsilon > 0$. We will prove the following by induction: for all $N \in \mathbb{N}$, there exists a variable word Σ such that for all $n \leq N$, $W \subseteq [k]^n$, and $v \in [k]^{<\omega}$, $|d(\Sigma W) - d(\Sigma Wv)| < \epsilon$.

Base case: $N = 1$. Partition $[0, 1]$ into length ϵ subintervals $\alpha_1, \alpha_2, \dots, \alpha_P$, and color $[k]^{<\omega}$ with P^{2^k} colors in the following way. For each $v \in [k]^{<\omega}$, let $c(v) = (c_W)_{W \subseteq [k]}$ where $d(Wv) \in \alpha_{c_W}$. By **CS**, there exists a variable sentence $\Sigma' = w_1(t_1)w_2(t_2) \dots$ with monochromatic range. Let $\Sigma = w_0(t_0)w_2(t_2) \dots$ where $w_0(t_0) = t_0w_1(1)$. Now suppose $W \subseteq [k]$ and $v \in [k]^{<\omega}$. Then for $w \in W$, $\Sigma w = w\Sigma'1$ and $\Sigma wv = w\Sigma'1v$. Since $\Sigma'1$ and $\Sigma'1v$ have the same color,

$$|d(\Sigma W) - d(\Sigma Wv)| = |d(W\Sigma'1) - d(W\Sigma'1v)| < \epsilon.$$

Assume now that the induction hypothesis holds for some $N \in \mathbb{N}$. Since sub-subdistributions are sub-distributions, we may without loss of generality assume that d is such that for all $n \leq N$, $W \subseteq [k]^n$, and $v \in [k]^{<\omega}$, $|d(W) - d(Wv)| < \epsilon$.

Partition $[0, 1]$ into length ϵ subintervals $\alpha_1, \alpha_2, \dots, \alpha_P$, and color $[k]^{<\omega}$ with $P^{2^{k^n}}$ colors in the following way. For each $v \in [k]^{<\omega}$, let $c(v) = (c_W)_{W \subseteq [k]^n}$ where $d(Wv) \in \alpha_{c_W}$. By **CS**, there exists a variable sentence $\Sigma' = w_1(t_1)w_2(t_2) \dots$ with monochromatic range. Let $\Sigma = t_{-N} \dots t_{-1}w_0(t_0)w_2(t_2) \dots$ where $w_0(t_0) = t_0w_1(1)$.

Suppose $n \leq N$, $W \subseteq [k]^n$, and $v \in [k]^{<\omega}$. Then for $w \in W$, $\Sigma w = w$ and $\Sigma wv = w\tilde{v}$ where \tilde{v} is a word independent of w . Since $W \subseteq [k]^n$, by the induction hypothesis,

$$|d(\Sigma W) - d(\Sigma Wv)| = |d(W) - d(W\tilde{v})| < \epsilon.$$

Suppose $n = N + 1$, $W \subseteq [k]^n$, and $v \in [k]^{<\omega}$. Then for $w \in W$, $\Sigma w = w\Sigma'1$ and $\Sigma wv = w\Sigma'1v$. Since $\Sigma'1$ and $\Sigma'1v$ have the same color,

$$|d(\Sigma W) - d(\Sigma Wv)| = |d(W\Sigma'1) - d(W\Sigma'1v)| < \epsilon.$$

The induction yields an ϵ -stationary distribution in $S(d)$. □

We now have the tools to prove that **DHJs** implies **DHJd**.

Proof. Let $k \in \mathbb{N}$, $\epsilon > 0$, and $d \in [0, 1]^{Z_k}$ be a distribution for which $d(\{w\}) > \epsilon$ for all $w \in [k]^{<\omega}$. By Lemma 3, there exists a stationary distribution \tilde{d} in $S(d)$. Let $(\Sigma_i)_{i \in \mathbb{N}}$ be a sequence of variable sentences such that $d \circ \Sigma_i \rightarrow \tilde{d}$ in $[0, 1]^{Z_k}$.

For all $w \in [k]^{<\omega}$, $d \circ \Sigma_i(\{w\}) \longrightarrow \tilde{d}(\{w\})$, whereby $\tilde{d}(\{w\}) > \epsilon/2$. Invoking **DHJs** for \tilde{d} , there exists a combinatorial line L such that $\tilde{d}(L) > 0$. Since $d \circ \Sigma_i(L) \longrightarrow \tilde{d}(L)$, there exists an $i \in \mathbb{N}$ for which $d \circ \Sigma_i(L) > 0$. Now $L' = \Sigma_i L$ is a combinatorial line for which $d(L') > 0$. \square

7. Equivalence of DHJs and DHJm

It remains to show that **DHJs** is equivalent to **DHJm**. Since the $[k]^{<\omega}$ -process $\{T(w)^{-1}A\}$ associated with a $[k]^{<\omega}$ -system (X, μ, T) and a set $A \subseteq X$ of positive measure has stationary distribution, it is clear that **DHJs** implies **DHJm**. The reverse implication will follow from the fact that any stationary distribution looks like the distribution of a $[k]^{<\omega}$ -orbit of a set of positive measure.

We will rely on the following lemma in order to craft an appropriate $[k]^{<\omega}$ -system on which to apply **DHJm**.

Lemma 4: *Let \mathcal{B}_0 and \mathcal{B}_1 be finite algebras of measurable sets in $([0, 1], m)$. Assume that there is a measure preserving isomorphism $\bar{\psi} : \mathcal{B}_0 \longrightarrow \mathcal{B}_1$. Then there exists an invertible measure preserving transformation ψ of $([0, 1], m)$ which induces $\bar{\psi}$, i.e. $\bar{\psi}(B) = \psi^{-1}(B)$.*

We conclude this note with the proof that **DHJm** implies **DHJs**.

Proof. Let $k \in \mathbb{N}$, $\epsilon > 0$, and $d \in [0, 1]^{\mathbb{Z}^k}$ be a stationary distribution for which $d(\{w\}) > \epsilon$ for all $w \in [k]^{<\omega}$. Using the subspace $\Sigma = 1t_1t_2 \cdots$ and passing to the subdistribution $d \circ \Sigma$, we may assume without loss of generality that d is constant on the set $\{\{1\}, \dots, \{k\}\}$. Let $\{C_w\}$ be the regular $[k]^{<\omega}$ -process for d .

For a finite set Y of measurable sets, denote by $\langle Y \rangle$ the finite algebra generated by Y . For all $v \in [k]$, $\langle \{C_v\} \rangle$ is isomorphic to $\langle \{C_1\} \rangle$ because $m(C_v) = m(C_1)$. By Lemma 4, for all $v \in [k]$, there exists an invertible, measure preserving transformation $T_1^{(v)}$ of $[0, 1]$ such that $T_1^{(v)}C_v = C_1$.

Let $n \geq 2$. Since d is stationary, for all $v \in [k]$, the obvious correspondence between $\langle \{C_w \mid w \in [k]^{n-1}\} \rangle$ and $\langle \{C_{wv} \mid w \in [k]^{n-1}\} \rangle$ is an isomorphism. By Lemma 4, for all $v \in [k]$, there exists an invertible, measure preserving transformation $T_n^{(v)}$ such that for all $w \in [k]^{n-1}$, $T_n^{(v)}C_{wv} = C_w$.

Now $([0, 1], m)$, the maps $\{T_j^{(i)}\}_{j \in \mathbb{N}}^{i \in [k]}$ defined above, and $T(w) = T_1^{(w_1)} \circ T_2^{(w_2)} \circ \dots \circ T_n^{(w_n)}$ yield a $[k]^{<\omega}$ -system. By the choice of the $T_j^{(i)}$'s, for all $w \in [k]^{<\omega}$, $T(w)C_w = C_1$. Applying **DHJm** with $A = C_1$, there exists a combinatorial line L such that $m(\cap_{w \in L} T(w)^{-1}C_1) > 0$. But $T(w)^{-1}C_1 = C_w$, and so $d(\cap_{w \in L} C_w) > 0$. \square

References

- [1] Timothy J. Carlson and Stephen G. Simpson. A dual form of Ramsey's theorem. *Adv. in Math.*, 53(3):265–290, 1984.
- [2] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett theorem for $k = 3$. *Discrete Math.*, 75(1-3):227–241, 1989. Graph theory and combinatorics (Cambridge, 1988).
- [3] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett theorem. *J. Anal. Math.*, 57:64–119, 1991.