

## Chapter 6

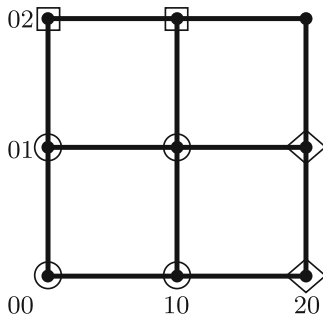
# Canonical Partitions

Originally, Ramsey theory investigates the behavior of structures with respect to colorings of substructures into a fixed number of classes, typically into two classes. Probably the most well-known example is the pigeon hole principle, saying that for every 2-coloring of  $\omega$  there exists an infinite subset  $F \subseteq \omega$  which is monochromatic. Of course, if we allow colorings with an unbounded number of colors then it is clear that the conclusion of the pigeon hole principle does not have to hold. For example, we could take  $\Delta(n) = n$  for every  $n < \omega$ . However, in this case we have an infinite set which meets each color in at most one element. Now it is an easy observation that one of these two possibilities must always occur. For every coloring  $\Delta : \omega \rightarrow \omega$  there exists an infinite set  $F \subseteq \omega$  such that either  $\Delta \upharpoonright F$  is monochromatic or  $\Delta \upharpoonright F$  is one-to-one, i.e., any two elements of  $F$  have different colors. This is the most elementary example of a *canonical partition theorem*, first introduced by Erdős and Rado (1950) studying unbounded colorings of finite sets. A coloring  $\Delta : [\omega]^k \rightarrow \omega$  of the  $k$ -subsets of the nonnegative integers is canonical if there exists a  $J \subseteq k$  such that  $\Delta(X) = \Delta(Y)$  if and only if  $X \upharpoonright J = Y \upharpoonright J$  for every pair  $X, Y \in [\omega]^k$ . The Erdős-Rado canonization Theorem 1.4 then asserts that for every coloring  $\Delta : [\omega]^k \rightarrow \omega$  there exists  $F \in [\omega]^\omega$  such that  $\Delta \upharpoonright [F]^k$  is canonical.

In this chapter unrestricted colorings of parameter words are investigated and their canonical patterns are determined. As applications we derive a canonizing version of van der Waerden's theorem from the corresponding result for zero-parameter words and the finite form of the Erdős-Rado canonization theorem from a canonizing version of the Graham-Rothschild theorem. In fact, throughout this chapter we will consider only parameter words over the trivial group.

A final remark concerns our notation. To indicate that we consider unbounded colorings we will always choose  $\omega$  as their range, although it will quite often happen that only finitely many colors can actually be used.

**Fig. 6.1** The canonical pattern on  $3^2$



### 6.1 Canonizing Hales-Jewett’s Theorem

In studying unbounded colorings of zero-parameter words we meet completely different pattern of ‘canonical colorings’ than for finite sets. Consider, e.g., the alphabet  $3 = \{0, 1, 2\}$  and the equivalence relation  $\approx$  on 3 having 0 and 1 in the same class, 2 in another one. Define an (unbounded) coloring  $\Delta_\approx : [3]_{(0)}^n \rightarrow \omega$  by  $\Delta_\approx(g) = g/\approx$ , where  $g/\approx \in [\{0, 2\}]_{(0)}^n$  is the  $\approx$ -quotient of  $g$ , i.e.,  $g/\approx(i) = 0$  if  $g(i) \in \{0, 1\}$ ,  $g/\approx(i) = 2$  otherwise. Observe that  $\Delta_\approx$  obeys a kind of uniform description. Any two  $m$ -parameter words inherit the same pattern from  $\Delta_\approx$ . In case  $m = 2$ , i.e. of planes, this pattern can be visualized as in Fig. 6.1.

Of course, every equivalence relation on the alphabet  $\{0, 1, 2\}$  leaves such a hereditary pattern. More general, let  $A$  be any finite alphabet and let  $\approx$  be an equivalence relation on  $A$ . Then every coloring  $\Delta_\approx : [A]_{(0)}^n \rightarrow \omega$  satisfying

$$\Delta_\approx(g) = \Delta_\approx(h) \quad \text{if and only if} \quad g/\approx = h/\approx \tag{6.1}$$

is hereditary in the sense that for every  $m$  and every  $f \in [A]_{(m)}^n$  the restriction  $\Delta_\approx \upharpoonright f \cdot A^m$  again satisfies (6.1). The following theorem shows that these are all ‘canonical colorings’.

**Theorem 6.1 (Canonical Hales-Jewett theorem).** *Let  $A$  be a finite alphabet and  $m$  be a positive integer. Then there exists a positive integer  $n = CHJ(|A|, m)$  such that for every unbounded coloring  $\Delta : [A]_{(0)}^n \rightarrow \omega$  there exists  $f \in [A]_{(m)}^n$  and there exists an equivalence relation  $\approx$  on  $A$  such that for all  $g, h \in [A]_{(0)}^n$  it follows that*

$$\begin{aligned} \Delta(f \cdot g) = \Delta(f \cdot h) \quad \text{if and only if} \quad g/\approx = h/\approx, \\ \text{i.e., } g(i) \approx h(i) \text{ for every } i < m. \end{aligned}$$

Observe that considering unbounded colorings we are only interested in the pattern of these colorings but not in the actual colors. This is taken into account by considering equivalence relations, thus abstracting from the actual colors.

A set  $\mathcal{E}$  of equivalence relations on  $[A]_k^{(m)}$  is called a *canonical set of equivalence relations* if  $\mathcal{E}$  is minimal (with respect to cardinality) such that there exists  $n$  so that for every unbounded coloring  $\Delta : [A]_k^{(n)} \rightarrow \omega$  there exists  $f \in [A]_k^{(n)}$  and an equivalence relation  $\approx$  in  $\mathcal{E}$  satisfying  $\Delta(f \cdot g) = \Delta(f \cdot h)$  if and only if  $h \approx g$ , i.e., the equivalence relation induced by  $\Delta$  coincides on  $f$  with  $\approx$ .

Theorem 6.1 together with the hereditary property of each of these equivalence relations imply that the set of all equivalence relations on  $[A]_0^{(m)}$  which are induced by equivalence relations on  $A$ , form a canonical set of equivalence relations on  $[A]_0^{(m)}$ . In fact, this is the unique canonical set of equivalence relations on  $[A]_0^{(m)}$ . Hence, it is justified to call a coloring  $\Delta : [A]_0^{(m)} \rightarrow \omega$  satisfying  $\Delta(g) = \Delta(h)$  if and only if  $g/\approx = h/\approx$  for some equivalence relation  $\approx$  on  $A$ , a *canonical coloring of zero-parameter words*.

*Proof of Theorem 6.1.* Assume that  $\Delta : [A]_0^{(n)} \rightarrow \omega$  is given. Consider the colorings that a line  $g \in [A]_1^{(n)}$  induces:  $\langle \Delta(g \cdot a) \mid a \in A \rangle$ . In the following we are not interested in the actual coloring of the line, but only in its *pattern*, i.e., for which  $a$ 's in  $A$  we get the same color and for which different ones. We can thus describe the pattern of a line by an equivalence relation on the alphabet  $A$ . Let  $r_a$  denote the number of equivalence relations on  $A$ . We just convinced ourselves that every coloring  $\Delta : [A]_0^{(n)} \rightarrow \omega$  gives rise to a coloring  $\Delta^* : [A]_1^{(n)} \rightarrow r_a$  which assigns to each line the equivalence relation on  $A$  that corresponds to the pattern induced by  $\Delta$  on this line. Observe that the Graham-Rothschild theorem implies that for  $n = GR(|A|, 1, M, r_a)$ , where  $M$  is yet to be determined, there exists an  $f \in [A]_M^{(n)}$  that is monochromatic with respect to  $\Delta^*$ .

We now repeat the above argument for  $m$ -spaces instead of lines. Every  $m$ -space  $g \in [A]_m^{(M)}$  induces a *pattern* with respect to the colors  $\langle \Delta((f \cdot g) \cdot h) \mid h \in [A]_0^{(m)} \rangle$  – and thus an equivalence relation on  $[A]_0^{(m)}$ . Let  $\hat{r}_a$  denote the number of equivalence relations on  $[A]_0^{(m)}$  and let  $\Delta^{**} : [A]_m^{(M)} \rightarrow \hat{r}_a$  denote the coloring that assigns to every  $m$ -space the equivalence relation on  $[A]_0^{(m)}$  that corresponds to the pattern induced by  $\Delta$  on this  $m$ -space. Applying the Graham-Rothschild theorem again implies that for  $M = GR(|A|, m, m + 1, \hat{r}_a)$  there exists a  $f' \in [A]_{m+1}^{(M)}$  that is monochromatic with respect to  $\Delta^{**}$ .

Observe that  $f, f'$  induce a coloring  $\hat{\Delta} : [A]_0^{(m+1)} \rightarrow \omega$ , defined by

$$\hat{\Delta}(\hat{f}) = \Delta((f \cdot f') \cdot \hat{f}) \quad \text{for every } \hat{f} \in [A]_0^{(m+1)}.$$

By construction we also have

- (1) The pattern which  $\hat{\Delta}$  leaves to lines are all the same, i.e.

$$\hat{\Delta}(\eta \cdot a) = \hat{\Delta}(\eta \cdot b) \quad \text{if and only if} \quad \hat{\Delta}(\eta' \cdot a) = \hat{\Delta}(\eta' \cdot b)$$

for all  $\eta, \eta' \in [A] \binom{m+1}{1}$  and all  $a, b \in A$ ,

and, additionally,

- (2) The pattern which  $\hat{\Delta}$  leaves to  $m$ -spaces are all the same, i.e.

$$\hat{\Delta}(\xi \cdot g) = \hat{\Delta}(\xi \cdot h) \quad \text{if and only if} \quad \hat{\Delta}(\xi' \cdot g) = \hat{\Delta}(\xi' \cdot h)$$

for all  $\xi, \xi' \in [A] \binom{m+1}{m}$  and all  $g, h \in [A] \binom{m}{0}$ .

(We remark in passing that by repeating the above argument multiple times we could even ensure that the pattern which  $\hat{\Delta}$  leaves to  $i$ -spaces are all the same – for all  $1 \leq i \leq m$ . However, in the following we do not need this generalization.)

In the following we use the notation  $\hat{\triangleq}$  to abbreviate facts (1) and (2). More precisely, for  $a, b \in A$  we write  $a \hat{\triangleq} b$  if  $\hat{\Delta}(\eta \cdot a) = \hat{\Delta}(\eta \cdot b)$  for some (and hence for all)  $\eta \in [A] \binom{m+1}{1}$ . Similarly, for  $g, h \in [A] \binom{m}{0}$  we write  $g \hat{\triangleq} h$  if  $\hat{\Delta}(\xi \cdot g) = \hat{\Delta}(\xi \cdot h)$  for some (and hence again for all)  $\xi \in [A] \binom{m+1}{m}$ .

We define the relation  $\approx$  on  $A$  as follows:  $a \approx b$  if and only if  $a \hat{\triangleq} b$ . The idea now is to show that  $g \hat{\triangleq} h$  if and only if  $g/\approx = h/\approx$ . Observe that in this case an  $m$ -parameter word  $f \cdot f' \cdot \xi \in [A] \binom{n}{m}$ , where  $\xi \in [A] \binom{m+1}{m}$  is an arbitrarily chosen  $m$ -parameter word, together with  $\approx$  satisfy the theorem.

First consider  $g, h \in [A] \binom{m}{0}$  such that  $g/\approx = h/\approx$ . We prove by induction that

$$(g_0, g_1, \dots, g_{m-1}) \hat{\triangleq} (h_0, h_1, \dots, h_{i-1}, g_i, \dots, g_{m-1})$$

for all  $i \leq m$ . This is trivially satisfied for  $i = 0$ . Assume it holds for some  $i < m$ , and consider the line  $\eta = (h_0, h_1, \dots, h_{i-1}, \lambda_0, g_{i+1}, \dots, g_{m-1}) \in [A] \binom{m}{1}$  and an arbitrary  $m$ -parameter word  $\xi \in [A] \binom{m+1}{m}$ . Observe that  $g_i \approx h_i$  implies  $\hat{\Delta}((\xi \cdot \eta) \cdot g_i) = \hat{\Delta}((\xi \cdot \eta) \cdot h_i)$ , and thus  $\eta \cdot g_i \hat{\triangleq} \eta \cdot h_i$ . As

$$\eta \cdot g_i = (h_0, h_1, \dots, h_{i-1}, g_i, \dots, g_{m-1}) \quad \text{and}$$

$$\eta \cdot h_i = (h_0, h_1, \dots, h_{i-1}, h_i, g_{i+1}, \dots, g_{m-1}),$$

we deduce that the induction hypothesis also holds for  $i + 1$ . Note that for  $i = m$  we get  $g \hat{\triangleq} h$ , as desired.

Let us now assume that  $g, h \in [A] \binom{m}{0}$  are such that  $g/\approx \neq h/\approx$ . Choose  $i \in m$  with  $g_i \not\approx h_i$  and consider  $\eta = (g_0, \dots, g_{i-1}, g_i, \lambda_0, g_{i+1}, \dots, g_{m-1}) \in [A] \binom{m+1}{1}$ . Then  $g_i \not\approx h_i$  implies that

$$\begin{aligned} \hat{\Delta}(g_0, \dots, g_{i-1}, g_i, g_i, g_{i+1}, \dots, g_{m-1}) &= \hat{\Delta}(\eta \cdot g_i) \\ &\neq \hat{\Delta}(\eta \cdot h_i) = \hat{\Delta}(g_0, \dots, g_{i-1}, g_i, h_i, g_{i+1}, \dots, g_{m-1}). \end{aligned} \quad (6.2)$$

In order derive a contradiction assume that  $g \stackrel{\Delta}{=} h$  and consider  $m$ -parameter words

$$\begin{aligned} \xi &= (\lambda_0, \dots, \lambda_{i-1}, \lambda_i, \lambda_i, \lambda_{i+1}, \dots, \lambda_{m-1}), \\ \xi' &= (\lambda_0, \dots, \lambda_{i-1}, \lambda_i, h_i, \lambda_{i+1}, \dots, \lambda_{m-1}) \in [A]^{\binom{m+1}{m}}. \end{aligned}$$

Then  $g \stackrel{\Delta}{=} h$  implies  $\hat{\Delta}(\xi \cdot g) = \hat{\Delta}(\xi \cdot h)$  and  $\hat{\Delta}(\xi' \cdot g) = \hat{\Delta}(\xi' \cdot h)$ . Closer inspection of the words  $\xi$  and  $\xi'$  yields that  $\xi \cdot h = \xi' \cdot h$ , thus

$$\begin{aligned} \hat{\Delta}(g_0, \dots, g_{i-1}, g_i, g_i, g_{i+1}, \dots, g_{m-1}) &= \hat{\Delta}(\xi \cdot g) \\ &= \hat{\Delta}(\xi' \cdot g) = \hat{\Delta}(g_0, \dots, g_{i-1}, g_i, h_i, g_{i+1}, \dots, g_{m-1}). \end{aligned}$$

which contradicts (6.2). Hence  $g \not\stackrel{\Delta}{=} h$ , as desired. This completes the proof of Theorem 6.1.

Schmerl (1993) applies this result to show that for every countable non-standard model  $\mathcal{M}$  of Peano arithmetic and every positive integer  $k \geq 2$  there exists a cofinal extension  $\mathcal{N}$  of  $\mathcal{M}$  such that the lattice  $\mathcal{L}(\mathcal{N}/\mathcal{M})$  of intermediate models is isomorphic to  $\Pi(k)$ , the lattice of equivalence relations of a  $k$ -element set (cf. also Schmerl 1985).

The special case  $|A| = 2$  of Theorem 6.1 admits the following formulation.

**Corollary 6.2.** *Let  $m$  be a positive integer. Then there exists a positive integer  $n = CHJ(2, m)$  such that for every coloring  $\Delta : \mathcal{B}(n) \rightarrow \omega$  of the points of the  $n$ -dimensional Boolean lattice  $\mathcal{B}(n)$  there exists a  $\mathcal{B}(m)$ -sublattice  $\mathcal{L} \subseteq \mathcal{B}(n)$  such that either  $\Delta \upharpoonright \mathcal{L}$  is constant or  $\Delta \upharpoonright \mathcal{L}$  is one-to-one.  $\square$*

Here we have the same kind of result as for the unbounded pigeon hole principle: the substructure we are looking for must either be colored monochromatically or one-to-one. Nešetřil and Rödl (1978b, 1979) call this phenomenon *selectivity*. We will meet this phenomenon several times in the sequel, e.g., in the next section in connection with van der Waerden's theorem.

Recall that every finite poset can be embedded (order-preservingly) into some Boolean lattice  $\mathcal{B}(n)$ , cf. Sect. 4.2.5. Hence, we get immediately

**Corollary 6.3.** *Let  $Q$  be a finite poset. Then there exists a finite poset  $P$  such that for every coloring  $\Delta : P \rightarrow \omega$  of the points of  $P$  there exists a  $Q$ -subposet  $Q' \subseteq P$  such that either  $\Delta \upharpoonright Q'$  is monochromatic or  $\Delta \upharpoonright Q'$  is one-to-one.  $\square$*

## 6.2 Canonizing van der Waerden's Theorem

As indicated in Sect. 4.2, van der Waerden's theorem on arithmetic progressions is one of the most prominent applications of Hales-Jewett's theorem. The aim of this section is to show how a canonical version of van der Waerden's theorem can be obtained using the canonical Hales-Jewett theorem.

**Theorem 6.4 (Canonical van der Waerden theorem).** *Let  $t$  be a positive integer. Then there exists a positive integer  $n = EG(t)$  such that for every coloring  $\Delta : n \rightarrow \omega$  there exists a  $t$ -term arithmetic progression  $X \subseteq n$  such that either  $\Delta \upharpoonright X$  is constant or  $\Delta \upharpoonright X$  is one-to-one.*

At the first glance it may look somewhat astonishing that the canonical Hales-Jewett theorem which allows every pattern on the lines can be used in order to obtain a selectivity result for arithmetic progressions. The original proof of Erdős and Graham (1980) used Szemerédi's density result for arithmetic progressions. Later, an 'elementary' proof was obtained by Prömel and Rödl (1986). The proof given here is based on ideas from (Prömel and Rothschild 1987) which can also be used to prove a slightly stronger result, viz. a restricted version of the canonical van der Waerden theorem.

*Proof of Theorem 6.4.* Let  $\ell = (t - 1)^2 + 1$ . It is easy to see that the first  $\ell$  nonnegative integers have the following property:

- (1) Let  $\mu < \nu < t$  be arbitrary and let  $\approx$  be an equivalence relation on  $\ell$  such that every arithmetic progression of length  $t$  in  $\ell$  has its  $\mu$ th and its  $\nu$ th term in the same equivalence class. Then there is a  $t$ -term arithmetic progression in  $\ell$  which is completely contained in one equivalence class, e.g., the progression  $\mu + (\nu - \mu) \cdot j, j < t$ .

Let  $(X_i)_{i < z}$  be an enumeration of all arithmetic progressions of length  $t$  in  $\ell$  and assume  $X_i = \{x_{i,0}, \dots, x_{i,t-1}\}$  for every  $i < z$  is in ascending order.

Choose  $n = CHJ(\ell, z)$  according to the canonical Hales-Jewett theorem and let  $\Delta : (\ell - 1)n + 1 \rightarrow \omega$  be a coloring. Consider the coloring  $\Delta^* : [\ell] \binom{n}{0} \rightarrow \omega$  which is defined by

$$\Delta^*(g_0, \dots, g_{n-1}) = \Delta(\sum_{i < n} g_i).$$

By choice of  $n$  there exists  $f \in [\ell] \binom{n}{z}$  and an equivalence relation  $\approx$  on  $\ell$  such that  $\Delta^* \upharpoonright f$  is canonical, meaning that for all  $g, h \in [\ell] \binom{z}{0}$  we have:

$$\Delta^*(f \cdot g) = \Delta^*(f \cdot h) \quad \text{if and only if} \quad g_i \approx h_i \text{ for every } i < z.$$

Let  $F = \sum \{f_i \mid f_i \in \ell\}$  and put  $\zeta^j = (x_{0,j}, x_{1,j}, \dots, x_{z-1,j})$  for  $j < t$  and consider  $\{f \cdot \zeta^j \mid j < t\}$ . Observe that  $\{F + \sum_{i < z} x_{ij} \mid j < t\}$  forms a  $t$ -term arithmetic progression.

First assume that  $\Delta^*\{f \cdot \xi^j \mid j < t\}$  is one-to-one. Then, clearly,  $\Delta\{F + \sum_{i < z} x_{ij} \mid j < t\}$  is also one-to-one and we are done.

So assume that there exists  $\mu, \nu < t$  such that

$$\Delta^*(f \cdot \xi^\mu) = \Delta^*(f \cdot \xi^\nu).$$

But then  $x_{j,\mu} \approx x_{j,\nu}$  for every  $j < z$ . So by (1) there exists an arithmetic progression  $X_i$  such that  $x_{i,0} \approx x_{i,1} \approx \dots \approx x_{i,t-1}$ . Let

$$\xi^j = \underbrace{(0, \dots, 0)}_{z-1}, x_{ij},$$

for every  $j < t$ . Then  $\Delta^*\{f \cdot \xi^j \mid j < t\}$  is constant and hence, by definition, also  $\Delta\{F + x_{ij} \mid j < t\}$ . Observing that  $\{F + x_{ij} \mid j < t\}$  forms a  $t$ -term arithmetic progression completes the proof of Theorem 6.4.  $\square$

Concerning more than one dimension a canonical version of Gallai-Witt's theorem was proved by Deuber et al. (1983) for finite subsets of the integer lattice grid and by Spencer (1983) for arbitrary finite subsets of the Euclidean space, both based on Fürstenberg-Katznelson's density version of the Gallai-Witt result. Simplified proofs are given in Prömel and Rödl (1986) and Prömel and Rothschild (1987). Although the method used to prove the canonical van der Waerden theorem can easily be adopted to derive a canonical version of Gallai-Witt's theorem there exist additional canonical patterns in this higher dimensional case. We omit the result.

### 6.3 Canonizing Graham-Rothschild's Theorem

Next we consider an extension of the canonizing version of Hales-Jewett's theorem to higher dimensions. Here, the canonical colorings occurring in the Erdős-Rado canonization theorem and those from the canonical Hales-Jewett theorem come together, finding a kind of common generalization.

Consider the surjective mapping  $\phi : [A]_{\binom{m}{k}} \rightarrow [m]^k$  given by  $\phi(f) = \{\min f^{-1}(\lambda_i) \mid i < k\}$ , cf. Sect. 3.1.2. This mapping shows that every canonical coloring  $\Delta_J : [m]^k \rightarrow \omega$ , where  $J \subseteq k$  and  $\Delta_J(X) = X : J$ , gives rise to a canonical coloring

$$\Delta : [A]_{\binom{m}{k}} \rightarrow \omega \quad \text{via} \quad \Delta(f) = (\phi(f)) : J.$$

On the other hand, every equivalence relation  $\approx$  on  $A \cup \{\lambda_0, \dots, \lambda_{k-1}\}$  allows to color according to the  $\approx$ -quotient of the  $k$ -parameter words in  $[A]_{\binom{m}{k}}$ .

It turns out that all colorings which are relevant for canonizing the Graham-Rothschild theorem can be produced by combining these two types of colorings appropriately.

Let  $J \subseteq k$  be any subset of  $k$  and put  $J^+ = J \cup \{k\}$ . For  $i \in k$  let  $pre(i) := \max\{j \in J^+ \mid j < i\}$  (and  $pre(i) = -1$  if there doesn't exist such element in  $J$ ), and  $suc(i) := \min\{j \in J^+ \mid j > i\}$ . Consider a family of equivalence relations  $\{\approx_i\}_{i \in J^+}$ , where  $\approx_i$  is defined on  $A \cup \{\lambda_0, \dots, \lambda_{i-1}\}$ . We associate to the pair  $\Pi = (J, (\approx_i)_{i \in J^+})$  an equivalence relation  $\approx_\Pi$  on  $[A]_k^n$  by putting

$$g \approx_\Pi h \quad \text{if and only if} \quad \text{for every } i \in J^+$$

$$(1) \quad \min g^{-1}(\lambda_i) = \min h^{-1}(\lambda_i),$$

$$(2) \quad g(v) \approx_i h(v) \quad \forall \min g^{-1}(\lambda_{pre(i)}) < v < \min g^{-1}(\lambda_i),$$

where we tacitly agree that  $\min g^{-1}(\lambda_{-1}) = -1$  and  $\min g^{-1}(\lambda_k) = m$ .

Note that the definition of  $\approx_\Pi$  does not depend on the dimension of the parameter words on which it is imposed. The pair  $\Pi = (J, (\approx_i)_{i \in J^+})$  is called an  $(A, k)$ -canonical pair, if and only if

- (3) For every  $j \in J$  we have  $\alpha \approx_j \beta$  implies  $\alpha \approx_{suc(j)} \beta$  for all  $\alpha, \beta \in A \cup \{\lambda_0, \dots, \lambda_{j-1}\}$ , i.e., the family of equivalence relations is getting coarser, and
- (4) For every  $j \in \{0, \dots, k-1\} \setminus J$  there exists  $\alpha \in A \cup \{\lambda_0, \dots, \lambda_{j-1}\}$  such that  $\alpha \approx_{suc(j)} \lambda_j$ .

Observe that condition (3) assures that the associated equivalence relation  $\approx_\Pi$  is hereditary, meaning that for every  $f \in [A]_m^n$  the restriction of  $\approx_\Pi$  to  $f$  yields the same equivalence relation, i.e.,  $f \cdot g \approx_\Pi f \cdot h$  if and only if  $g \approx_\Pi h$ . We prove now that any two equivalence relations which are associated to distinct canonical pairs are essentially different and then we show that the set of equivalence relations which come from  $(A, k)$ -canonical pairs indeed forms a canonical set of equivalence relations on  $[A]_k^n$ .

**Proposition 6.5.** *Let  $\Pi_0 = (J_0, (\approx_i^0)_{i \in J_0^+})$  and  $\Pi_1 = (J_1, (\approx_i^1)_{i \in J_1^+})$  be distinct  $(A, k)$ -canonical pairs. Then for every  $f \in [A]_m^n$  the restrictions of  $\approx_{\Pi_0}$  and  $\approx_{\Pi_1}$  to  $f$  are distinct.*

*Proof.* Fix some  $f \in [A]_m^n$ . First assume that  $J_0 \neq J_1$ . Without loss of generality we can assume that there exists  $j \in J_0$  such that  $j \notin J_1$ . By (4) we know that there exists  $\alpha \in A \cup \{\lambda_0, \dots, \lambda_{j-1}\}$  so that  $\alpha \approx_i^1 \lambda_j$  (where  $i > j$  is minimal so that  $i \in J_1^+$ ). Consider

$$g = (\lambda_0, \dots, \lambda_{j-1}, \lambda_j, \lambda_j, \lambda_{j+1}, \dots, \lambda_{k-1}, \lambda_0, \dots, \lambda_0) \in [A]_k^m$$

and

$$h = (\lambda_0, \dots, \lambda_{j-1}, \alpha, \lambda_j, \lambda_{j+1}, \dots, \lambda_{k-1}, \lambda_0, \dots, \lambda_0) \in [A]_k^m.$$



Then,

$$f \cdot g \not\approx_{\Pi_0} f \cdot h, \text{ as } \min(f \cdot g)^{-1}(\lambda_j) \neq \min(f \cdot h)^{-1}(\lambda_j), \text{ but}$$

$$f \cdot g \approx_{\Pi_1} f \cdot h, \text{ as } \alpha \approx_i^1 \lambda_j \text{ implies by (3) that } \alpha \approx_\ell^1 \lambda_j \text{ for every } i \leq \ell \leq k.$$

Now assume that  $J_0 = J_1$ , but there exist  $i \in J_0^+$  and  $\alpha, \beta \in A \cup \{\lambda_0, \dots, \lambda_{i-1}\}$  so that  $\alpha \not\approx_i^0 \beta$ , but  $\alpha \approx_i^1 \beta$ . Put

$$g = (\lambda_0, \dots, \lambda_{i-1}, \alpha, \lambda_i, \lambda_{i+1}, \dots, \lambda_{k-1}, \lambda_0, \dots, \lambda_0) \in [A] \binom{m}{k}$$

and

$$h = (\lambda_0, \dots, \lambda_{i-1}, \beta, \lambda_i, \lambda_{i+1}, \dots, \lambda_{k-1}, \lambda_0, \dots, \lambda_0) \in [A] \binom{m}{k}.$$

Then, obviously,  $f \cdot g \not\approx_{\Pi_0} f \cdot h$ , but  $f \cdot g \approx_{\Pi_1} f \cdot h$ , as above, completing the proof of Proposition 6.5.  $\square$

**Theorem 6.6 (Canonical Graham-Rothschild theorem).** *Let  $A$  be a finite alphabet and  $k, m$  be positive integers. Then there exists  $n = PV(|A|, k, m)$  such that for every coloring  $\Delta : [A] \binom{n}{k} \rightarrow \omega$  there exists  $f \in [A] \binom{n}{m}$  and there exists an  $(A, k)$ -canonical pair  $\Pi = (J, (\approx_i)_{i \in J^+})$  such that for all  $g, h \in [A] \binom{m}{k}$  we have*

$$\Delta(f \cdot g) = \Delta(f \cdot h) \quad \text{if and only if} \quad g \approx_\Pi h.$$

*Proof.* Proceeding as in the proof of the canonical Hales-Jewett theorem we observe that by using the (classical) Graham-Rothschild theorem twice we may assume that there exists  $\hat{f} \in [A] \binom{n}{m+1}$  such that  $\hat{\Delta} : [A] \binom{m+1}{k} \rightarrow \omega$ ,

$$\hat{\Delta}(g) := \Delta(\hat{f} \cdot g) \quad \text{for} \quad g \in [A] \binom{m+1}{k},$$

satisfies:

(1a) The pattern which  $\hat{\Delta}$  leaves to the  $(k+1)$ -parameter subwords are all the same, i.e.,

$$\hat{\Delta}(\eta \cdot a) = \hat{\Delta}(\eta \cdot b) \quad \text{if and only if} \quad \hat{\Delta}(\eta' \cdot a) = \hat{\Delta}(\eta' \cdot b)$$

for all  $\eta, \eta' \in [A] \binom{m+1}{k+1}$  and all  $a, b \in [A] \binom{k+1}{k}$ , and additionally,

(1b) The pattern which  $\hat{\Delta}$  leaves to the  $m$ -parameter subwords are all the same, i.e.,

$$\hat{\Delta}(\xi \cdot a) = \hat{\Delta}(\xi \cdot b) \quad \text{if and only if} \quad \hat{\Delta}(\xi' \cdot a) = \hat{\Delta}(\xi' \cdot b)$$

for all  $\xi, \xi' \in [A] \binom{m+1}{m}$  and all  $a, b \in [A] \binom{m}{k}$

We define the relation  $\triangleq$  similarly as in the proof of the canonical Hales-Jewett theorem: for  $t \in \{k+1, m\}$  and  $a, b \in [A]_k^{(t)}$ , we write  $a \triangleq b$  if  $\hat{\Delta}(f' \cdot a) = \hat{\Delta}(f' \cdot b)$  for some (and hence for all)  $f' \in [A]_t^{(m+1)}$ . We also extend this notation to other values  $t \in \{k+1, \dots, m+1\}$  as follows: for  $a, b \in [A]_k^{(t)}$ , we write  $a \triangleq b$  if  $\hat{\Delta}(f' \cdot a) = \hat{\Delta}(f' \cdot b)$  for all  $f' \in [A]_t^{(m+1)}$ . We will repeatedly make use of the following simple fact that shows that the relation  $\triangleq$  can be extended upwards:

(1c) If  $a \triangleq b$  for some  $a, b \in [A]_k^{(t)}$ , then  $f'' \cdot a \triangleq f'' \cdot b$  for every  $f'' \in [A]_{t'}^{(t)}$ ,  $t' \in \{t, \dots, m+1\}$ .

To see this fix some  $f'' \in [A]_{t'}^{(t)}$  and consider an arbitrary  $f''' \in [A]_{t'}^{(m+1)}$ ; then  $f''' \cdot f'' \in [A]_t^{(m+1)}$  and  $a \triangleq b$  thus implies that  $\hat{\Delta}(f''' \cdot f'' \cdot a) = \hat{\Delta}(f''' \cdot f'' \cdot b)$ .

It remains to find an  $(A, k)$ -canonical pair  $\Pi$  such that

$$g \triangleq h \quad \text{if and only if} \quad g \approx_{\Pi} h,$$

for every pair  $g, h \in [A]_k^{(m)}$ . Note that then  $\Pi$  together with an  $m$ -parameter word  $\hat{f} \cdot f$ , where  $f \in [A]_k^{(m+1)}$  is chosen arbitrarily, satisfies the theorem.

First we define equivalence relations  $\approx_i^*$  for all  $i \leq k$ . These equivalence relations will later be used to obtain a set  $J \subseteq k$  and a family of equivalence relations  $\approx_i$ ,  $i \in J^+$ , which form an  $(A, k)$ -canonical pair. Let  $\approx_i^*$  be defined on  $A \cup \{\lambda_0, \dots, \lambda_i\}$ , by  $\alpha \approx_i^* \beta$  if and only if  $\lambda^i(\alpha) \triangleq \lambda^i(\beta)$ , where

$$\lambda^i(x) = (\lambda_0, \dots, \lambda_{i-1}, x, \lambda_i, \dots, \lambda_{k-1}).$$

In order to later define the desired set  $J$ , we first exhibit three properties of the relations  $\approx_i^*$ :

- (2a)  $\alpha \approx_i^* \beta$  implies  $\alpha \approx_{i+1}^* \beta$ , thus  $\approx_{i+1}^*$  is coarser than  $\approx_i^*$ , for every  $i < k$ .  
(2b) Let  $\alpha \approx_i^* \lambda_i$  for some  $\alpha \in A \cup \{\lambda_0, \dots, \lambda_{i-1}\}$ . Then  $\approx_{i+1}^* [A \cup \{\lambda_0, \dots, \lambda_i\}] = \approx_i^*$ .

Every parameter word  $g \in [A]_k^{(m)}$  is naturally divided into  $k+1$  (possibly empty) pieces between the minimal occurrences of its  $k$  parameters. We denote by  $p(g, i) \subseteq m$  the positions of the  $i$ th of these  $k+1$  pieces. More formally,

$$p(g, i) = \{j < m \mid \min g^{-1}(\lambda_{i-1}) < j < \min g^{-1}(\lambda_i)\}$$

for  $i < k$ , where we assume that  $\min g^{-1}(\lambda_{-1}) = -1$  and  $\min g^{-1}(\lambda_k) = m$ .

- (2c) Let  $g \in [A]_k^{(m)}$  and  $\ell \in m$  such that  $\ell \in p(g, i)$  for some  $i \in k+1$ . Then for any  $\alpha \in A \cup \{\lambda_0, \dots, \lambda_i\}$  such that  $g_{\ell} \approx_i^* \alpha$  and

$$g' = (g_0, \dots, g_{\ell-1}, \alpha, g_{\ell+1}, \dots, g_{m-1}) \in [A]_k^{(m)}$$

we have  $g \triangleq g'$ .

*Proof of (2a):* Assume that  $\alpha \approx_i^* \beta$ . Applying (1c) with

$$\begin{aligned}\eta &= (\lambda_0, \dots, \lambda_i, \lambda_{i+1}, \lambda_i, \lambda_{i+2}, \dots, \lambda_k), \\ \eta' &= (\lambda_0, \dots, \lambda_i, \lambda_{i+1}, \alpha, \lambda_{i+2}, \dots, \lambda_k) \in [A]_{(k+1)}^{(k+2)}\end{aligned}$$

on  $\lambda^i(\alpha) \triangleq \lambda^i(\beta)$ , we get

$$\begin{aligned}\eta \cdot \lambda^i(\alpha) &= (\lambda_0, \dots, \lambda_{i-1}, \alpha, \lambda_i, \alpha, \lambda_{i+1}, \dots, \lambda_{k-1}) \\ &\triangleq (\lambda_0, \dots, \lambda_{i-1}, \beta, \lambda_i, \beta, \lambda_{i+1}, \dots, \lambda_{k-1}) = \eta \cdot \lambda^i(\beta),\end{aligned}$$

and

$$\begin{aligned}\eta' \cdot \lambda^i(\alpha) &= (\lambda_0, \dots, \lambda_{i-1}, \alpha, \lambda_i, \alpha, \lambda_{i+1}, \dots, \lambda_{k-1}) \\ &\triangleq (\lambda_0, \dots, \lambda_{i-1}, \beta, \lambda_i, \alpha, \lambda_{i+1}, \dots, \lambda_{k-1}) = \eta' \cdot \lambda^i(\beta).\end{aligned}$$

Thus, by transitivity,

$$(\lambda_0, \dots, \lambda_{i-1}, \beta, \lambda_i, \beta, \lambda_{i+1}, \dots, \lambda_{k-1}) \triangleq (\lambda_0, \dots, \lambda_{i-1}, \beta, \lambda_i, \alpha, \lambda_{i+1}, \dots, \lambda_{k-1}).$$

Now consider  $\eta'' = (\lambda_0, \dots, \lambda_{i-1}, \beta, \lambda_i, \dots, \lambda_k) \in [A]_{(k+1)}^{(k+2)}$  and observe that the equality above implies

$$\begin{aligned}\eta'' \cdot \lambda^{i+1}(\alpha) &= (\lambda_0, \dots, \lambda_{i-1}, \beta, \lambda_i, \alpha, \lambda_{i+1}, \dots, \lambda_{k-1}) \\ &\triangleq (\lambda_0, \dots, \lambda_{i-1}, \beta, \lambda_i, \beta, \lambda_{i+1}, \dots, \lambda_{k-1}) \\ &= \eta'' \cdot \lambda^{i+1}(\beta).\end{aligned}$$

Therefore, for any  $f \in [A]_{(k+2)}^{(m+1)}$  we have  $\hat{\Delta}((f \cdot \eta'') \cdot \lambda^{i+1}(\alpha)) = \hat{\Delta}((f \cdot \eta'') \cdot \lambda^{i+1}(\beta))$ , hence from (1a) we deduce  $\lambda^{i+1}(\alpha) \triangleq \lambda^{i+1}(\beta)$  which by definition implies  $\alpha \approx_{i+1}^* \beta$ , proving (2a).

*Proof of (2b):* Let us assume  $\alpha \approx_i^* \lambda_i$  for some  $\alpha \in A \cup \{\lambda_0, \dots, \lambda_{i-1}\}$ . From (2a) we already know that  $\approx_{i+1}^*$  is coarser than  $\approx_i^*$ . We need to show that  $\approx_{i+1}^*$  restricted to  $A \cup \{\lambda_0, \dots, \lambda_i\}$  is not strictly coarser than  $\approx_i^*$ . In other words, we need to show that for  $\beta, \gamma \in A \cup \{\lambda_0, \dots, \lambda_i\}$  with  $\beta \approx_{i+1}^* \gamma$  we also have  $\beta \approx_i^* \gamma$ . Observe that the assumption  $\alpha \approx_i^* \lambda_i$  implies that  $\alpha \approx_{i+1}^* \lambda_i$  (as  $\approx_{i+1}^*$  is coarser). That is, without loss of generality we may assume that neither  $\beta$  nor  $\gamma$  is equal to  $\lambda_i$ .

We proceed similarly as in the proof of (2a). Applying (1c) with

$$\begin{aligned}\eta &= (\lambda_0, \dots, \lambda_{i-1}, \lambda_i, \beta, \lambda_{i+1}, \dots, \lambda_k), \\ \eta' &= (\lambda_0, \dots, \lambda_{i-1}, \lambda_i, \gamma, \lambda_{i+1}, \dots, \lambda_k) \in [A]_{(k+1)}^{(k+2)}\end{aligned}$$

on  $\lambda^i(\alpha) \triangleq \lambda^i(\lambda_i)$ , we get

$$\begin{aligned} (\lambda_0, \dots, \lambda_{i-1}, \alpha, \beta, \lambda_i, \dots, \lambda_{k-1}) &\triangleq (\lambda_0, \dots, \lambda_{i-1}, \lambda_i, \beta, \lambda_i, \dots, \lambda_{k-1}), \\ (\lambda_0, \dots, \lambda_{i-1}, \alpha, \gamma, \lambda_i, \dots, \lambda_{k-1}) &\triangleq (\lambda_0, \dots, \lambda_{i-1}, \lambda_i, \gamma, \lambda_i, \dots, \lambda_{k-1}). \end{aligned}$$

Similarly, applying (1c) with

$$\eta'' = (\lambda_0, \dots, \lambda_{i-1}, \lambda_i, \lambda_{i+1}, \lambda_i, \lambda_{i+2}, \dots, \lambda_k) \in [A]_{(k+1)}^{(k+2)}$$

on  $\lambda^{i+1}(\beta) \triangleq \lambda^{i+1}(\gamma)$ , which follows from  $\beta \approx_{i+1}^* \gamma$ , implies

$$(\lambda_0, \dots, \lambda_i, \beta, \lambda_i, \lambda_{i+1}, \dots, \lambda_{k-1}) \triangleq (\lambda_0, \dots, \lambda_i, \gamma, \lambda_i, \lambda_{i+1}, \dots, \lambda_{k-1}).$$

Therefore, by transitivity we have

$$(\lambda_0, \dots, \lambda_{i-1}, \alpha, \beta, \lambda_i, \dots, \lambda_{k-1}) \triangleq \Delta(\lambda_0, \dots, \lambda_{i-1}, \alpha, \gamma, \lambda_i, \dots, \lambda_{k-1}).$$

Thus, applying  $\eta''' = (\lambda_0, \dots, \lambda_{i-1}, \alpha, \lambda_i, \dots, \lambda_k) \in [A]_{(k+1)}^{(k+2)}$  on the previous equality, we have

$$\begin{aligned} \eta''' \cdot \lambda^i(\beta) &= (\lambda_0, \dots, \lambda_{i-1}, \alpha, \beta, \lambda_i, \dots, \lambda_{k-1}) \\ &\triangleq (\lambda_0, \dots, \lambda_{i-1}, \alpha, \gamma, \lambda_i, \dots, \lambda_{k-1}) \\ &= \eta''' \cdot \lambda^i(\gamma). \end{aligned}$$

By the same argument as in the proof of (2a) we deduce that  $\beta \approx_i^* \gamma$ , thus proving (2b).

*Proof of (2c):* Let  $g \in [A]_{(k)}^{(m)}$  and  $\ell \in p(g, i)$  for some  $i \in k + 1$ , and consider any  $\alpha \in A \cup \{\lambda_0, \dots, \lambda_i\}$  such that  $\alpha \approx_i^* g_\ell$ . Then by applying (1c) with

$$\eta = (g_0, \dots, g_{\ell-1}, \lambda_i, g_{\ell+1}^*, \dots, g_{m-1}^*) \in [A]_{(k+1)}^{(m)},$$

where

$$g_v^* = \begin{cases} g_v & \text{if } g_v \in A \cup \{\lambda_0, \dots, \lambda_{i-1}\} \\ \lambda_{\mu+1} & \text{if } g_v = \lambda_\mu \text{ for } \mu \geq i, \end{cases}$$

on  $\lambda^i(\alpha) \triangleq \lambda^i(g_\ell)$ , we get

$$g = \eta \cdot \lambda^i(g_\ell) \triangleq \eta \cdot \lambda^i(\alpha) = g',$$

which completes the proof of (2c).

*Completing the proof:* With properties (2a), (2b) and (2c) at hand we complete the proof of the theorem as follows. Let

$$J := \{i < k \mid \alpha \not\approx_i^* \lambda_i \text{ for every } \alpha \in A \cup \{\lambda_0, \dots, \lambda_{i-1}\}\} \quad \text{and} \\ \approx_i := \approx_i^* \upharpoonright A \cup \{\lambda_0, \dots, \lambda_{i-1}\} \text{ for every } i \in J^+.$$

We now show that these  $J \subseteq k$  and  $\approx_i, i \in J^+$ , are such that  $\Pi = (J, (\approx_i)_{i \in J^+})$  is as required in the theorem.

By (2a) and the definition of  $J$  it is obvious that  $\Pi = (J, (\approx_i)_{i \in J^+})$  is an  $(A, k)$ -canonical pair. In the remainder of the proof we verify that  $g \stackrel{\Delta}{=} h$  if and only if  $g \approx_{\Pi} h$ , for all  $g, h \in [A]_k^m$ . In doing so we will repeatedly use the following observation which immediately follows from the definition of  $J$  and (2b):

(\*) If  $i < k, j \in J$  are such that  $\text{pre}(j) < i \leq j$ , then  $\pi_i^* = \pi_j \upharpoonright A \cup \{\lambda_0, \dots, \lambda_i\}$

First assume that  $g \approx_{\Pi} h$ . We show, by induction, that there exist  $k$ -parameter words  $g^0, \dots, g^k, h^0, \dots, h^k \in [A]_k^m$ , such that for each  $t \in k + 1$  the following holds:

(3a)  $\min(g^t)^{-1}(\lambda_i) = \min(h^t)^{-1}(\lambda_i)$  for  $i \in t$ , i.e., the first occurrences of each of the first  $t$  parameters are identical in  $g^t$  and  $h^t$ .

(3b)  $g^t \approx_{\Pi} h^t$ , and

(3c)  $g^{t-1} \stackrel{\Delta}{=} g^t$  and  $h^{t-1} \stackrel{\Delta}{=} h^t$ ,

where  $g^{-1} = g$  and  $h^{-1} = h$ . For  $t = 0$ , all three properties are trivially satisfied for  $g^0 = g$  and  $h^0 = h$ . Assume now that the claim holds for some  $t \in k$ . If  $\min(g^t)^{-1}(\lambda_i) = \min(h^t)^{-1}(\lambda_i)$ , then  $g^{t+1} = g^t$  and  $h^{t+1} = h^t$  satisfies the claim for  $t + 1$ . Otherwise, without loss of generality we assume that  $\min(g^t)^{-1}(\lambda_i) > \min(h^t)^{-1}(\lambda_i)$ . Note that this implies  $t \notin J$  (as  $g^t \approx_{\Pi} h^t$ ) and  $\ell = \min(h^t)^{-1}(\lambda_i) \in p(g^t, t)$ . From (\*) and  $g^t \approx_{\Pi} h^t$  it thus follows that  $g_{\ell}^t \approx_i^* h_{\ell}^t$ . Thus by applying (2c) with  $\alpha = h_{\ell}^t = \lambda_i$  on  $g^t$  we get  $g^{t+1} \in [A]_k^m$ ,

$$g^{t+1} = (g_0^t, \dots, g_{\ell-1}^t, \lambda_i, g_{\ell+1}^t, \dots, g_{m-1}^t),$$

such that  $g^t \stackrel{\Delta}{=} g^{t+1}$ . It is easy now to see that  $g^{t+1}$ , together with  $h^{t+1} = h^t$ , satisfies all three properties of the claim. For  $t = k$ , (3a) implies that  $g^k$  and  $h^k$  agree on all first occurrences of parameters. Thus for each  $\ell \in m$  such that  $g_{\ell}^k \neq h_{\ell}^k$  we have  $\ell \in p(g^k, i)$ , for some  $i \in k + 1$ . Since  $g^k \approx_{\Pi} h^k$ , we can apply (2c) together with (\*) for  $\alpha = h_{\ell}^k$  on  $g^k$ , hence completely matching  $g^k$  and  $h^k$ . Therefore  $g^k \stackrel{\Delta}{=} h^k$ , and from (3c) we conclude  $g \stackrel{\Delta}{=} h$ .

Let us now assume that  $g \not\approx_{\Pi} h$ . First we show that we may assume without loss of generality that  $g$  and  $h$  are such that there exists a position  $\ell$  and an index  $i < k$  such that the following three properties are satisfied:

(4a)  $g_{\ell} \not\approx_i^* h_{\ell}$ ,

(4b) For all  $i' < i$  we have  $\min g^{-1}(\lambda_{i'}) = \min h^{-1}(\lambda_{i'}) < \ell$ ,

$$(4c) \quad \ell \leq \min g^{-1}(\lambda_i) \leq \min h^{-1}(\lambda_i).$$

If the first occurrences of the parameters  $\lambda_j$  for  $j < k$  are all identical, then  $g \not\approx_{\Pi} h$  together with (\*) easily implies that there exist indices  $i$  and  $\ell$  that satisfy (4a)–(4c). Otherwise choose  $i < k$  minimal such that  $\min g^{-1}(\lambda_i) \neq \min h^{-1}(\lambda_i)$ . We may assume without loss of generality (rename  $g$  and  $h$  if necessary) that  $\ell := \min g^{-1}(\lambda_i) < \min h^{-1}(\lambda_i)$ . If  $g_{\ell} \not\approx_i^* h_{\ell}$  then we have found  $\ell$  and  $i$  that satisfy (4a)–(4c). So assume that  $g_{\ell} \approx_i^* h_{\ell}$ . Apply (2c) to deduce that  $h' = (h_1, h_{\ell-1}, g_{\ell}, h_{\ell+1}, \dots, h_{m-1})$  satisfies  $h \triangleq h'$ . Note that (\*) implies that we also have that  $h' \approx_{\Pi} h$ . We may thus assume without loss of generality that  $h = h'$ . Repeating this process we see that we either find the desired  $\ell$  and  $i$  or we end up with  $g$  and  $h$  such that for all  $i < k$  we have  $\min g^{-1}(\lambda_i) = \min h^{-1}(\lambda_i)$ , which is the case that we already handled.

So assume now that  $\ell$  and  $i$  are such that (4a)–(4c) hold. Consider the  $(k + 1)$ -parameter word  $\eta \in [A]_{(k+1)}^{(m+1)}$ ,

$$\eta = (g_0, \dots, g_{\ell-1}, \lambda_i, g_{\ell}^*, \dots, g_{m-1}^*),$$

where

$$g_v^* = \begin{cases} g_v & \text{if } g_v \in A \cup \{\lambda_0, \dots, \lambda_{i-1}\} \\ \lambda_{j+1} & \text{if } g_v = \lambda_j \text{ for } j \geq i. \end{cases}$$

Note that  $g_{\ell} \not\approx_i^* h_{\ell}$  implies, by definition,  $\lambda^i(g_{\ell}) \triangleq \lambda^i(h_{\ell})$ . Then from (1a) we also have  $\eta \cdot \lambda^i(g_{\ell}) \triangleq \eta \cdot \lambda^i(h_{\ell})$ , thus

$$\begin{aligned} \eta \cdot \lambda^i(g_{\ell}) &= (g_0, \dots, g_{\ell-1}, g_{\ell}, g_{\ell}, \dots, g_{m-1}) \\ &\triangleq (g_0, \dots, g_{\ell-1}, h_{\ell}, g_{\ell}, \dots, g_{m-1}) = \eta \cdot \lambda^i(h_{\ell}). \end{aligned} \tag{6.3}$$

For a contradiction, let us assume  $g \triangleq h$ . Then applying (1c) with

$$\begin{aligned} \xi &= (\lambda_0, \dots, \lambda_{\ell-1}, \lambda_{\ell}, \lambda_{\ell}, \lambda_{\ell+1}, \dots, \lambda_{m-1}), \\ \xi' &= (\lambda_0, \dots, \lambda_{\ell-1}, h^*, \lambda_{\ell}, \lambda_{\ell+1}, \dots, \lambda_{m-1}) \in [A]_{(m)}^{(m+1)}, \end{aligned}$$

where  $h^* = h_{\ell}$  if  $h_{\ell} \in A$  and  $h^* = \lambda_{\min h^{-1}(\lambda_j)}$  if  $h_{\ell} = \lambda_j$ , we get

$$\begin{aligned} \xi \cdot g &= (g_0, \dots, g_{\ell-1}, g_{\ell}, g_{\ell}, \dots, g_{m-1}) \\ &\triangleq (h_0, \dots, h_{\ell-1}, h_{\ell}, h_{\ell}, \dots, h_{m-1}) = \xi \cdot h, \end{aligned}$$

and

$$\xi' \cdot g = (g_0, \dots, g_{\ell-1}, h_{\ell}, g_{\ell}, \dots, g_{m-1})$$

$$\triangleq (h_0, \dots, h_{\ell-1}, h_\ell, h_\ell, \dots, h_{m-1}) = \xi' \cdot h.$$

Note that  $\xi' \cdot g = (g_0, \dots, g_{\ell-1}, h_\ell, g_\ell, \dots, g_{m-1})$  comes from the fact that  $\min g^{-1}(\lambda_j) = \min h^{-1}(\lambda_j)$ , in case  $h_\ell = \lambda_j$ . Therefore, by transitivity we have

$$(g_0, \dots, g_{\ell-1}, g_\ell, g_\ell, \dots, g_{m-1}) \triangleq (g_0, \dots, g_{\ell-1}, h_\ell, g_\ell, \dots, g_{m-1}),$$

which contradicts (6.3). Hence  $g \not\triangleq h$ , which completes the proof of Theorem 6.6.  $\square$

This result was proved in Prömel and Voigt (1983), cf. also Prömel and Voigt (1986).

## 6.4 Applications

Every result which can be proved using the Graham-Rothschild theorem admits some kind of canonization using the canonizing version of Graham-Rothschild's theorem instead. Here we will only discuss three examples where applying the canonizing Graham-Rothschild theorem easily gives a canonical set of equivalence relations.

### 6.4.1 Finite Unions and Finite Sums

The first application of the canonical Graham-Rothschild theorem is a canonizing version of the finite union theorem (cf. Sect. 5.2.4). Recall that every nonempty subset of  $n$  can be interpreted as an element of  $[1] \binom{n}{1}$ . Observing that there are precisely three  $(\{0\}, 1)$ -canonical pairs, viz.  $(\emptyset, (\{0, \lambda\})_{\approx_1})$ ,  $(\{0\}, (\{0\}, \{\lambda\})_{\approx_1})$  and  $(\{0\}, (\{0, \lambda\})_{\approx_1})$ , we obtain

**Theorem 6.7.** *Let  $m$  be a positive integer. Then there exists  $n = n(m)$  such that for every coloring  $\Delta : \mathcal{B}(n) \rightarrow \omega$  there exist  $m$  mutually disjoint and non empty subsets  $X_0, \dots, X_{m-1} \in \mathcal{B}(n)$  such that one of the following three cases is valid for all nonempty  $I, J \subseteq m$ :*

- (1)  $\Delta(\bigcup_{i \in I} X_i) = \Delta(\bigcup_{j \in J} X_j)$
- (2)  $\Delta(\bigcup_{i \in I} X_i) = \Delta(\bigcup_{j \in J} X_j)$  if and only if  $I = J$
- (3)  $\Delta(\bigcup_{i \in I} X_i) = \Delta(\bigcup_{j \in J} X_j)$  if and only if  $\min I = \min J$ .  $\square$

Using again the bijection  $\kappa : \mathcal{B}(n) \rightarrow 2^n$  given by  $\kappa(B) = \sum_{i \in B} 2^i$  for every  $B \subseteq n$  we obtain a canonical Rado-Folkman-Sanders theorem, viz.

**Theorem 6.8.** *Let  $m$  be a positive integer. Then there exists  $n = n(m)$  such that for every coloring  $\Delta : n \rightarrow \omega$  there exist mutually distinct positive integers  $a_0, \dots, a_{m-1}$  such that one of the following three cases is valid for all nonempty  $I, J \subseteq m$ :*

- (1)  $\Delta(\sum_{i \in I} a_i) = \Delta(\sum_{j \in J} a_j)$
- (2)  $\Delta(\sum_{i \in I} a_i) = \Delta(\sum_{j \in J} a_j)$  if and only if  $I = J$
- (3)  $\Delta(\sum_{i \in I} a_i) = \Delta(\sum_{j \in J} a_j)$  if and only if  $\min I = \min J$  □

It is interesting to note that if finite subsets of  $\omega$  are partitioned, instead of subsets of some finite  $n$ , respectively  $\omega$  instead of  $n$ , and we ask for the canonical patterns on finite unions, respectively finite sums, then it turns out that three patterns are no longer sufficient (Taylor 1976).

## 6.4.2 Boolean Lattices

From the canonical Hales-Jewett theorem we obtained that coloring the points (i.e.,  $\mathcal{B}(0)$ -sublattices) of a sufficiently large Boolean lattice always yields a  $\mathcal{B}(m)$ -sublattice which is either colored monochromatically or one-to-one (Corollary 6.2). Clearly, these two patterns do not longer suffice if we color  $\mathcal{B}(1)$ -sublattices, i.e., 2-element chains.

Every 2-element chain in a Boolean lattice is given by a pair  $(X_0, X_0 \cup X_1)$ , where  $X_1 \neq \emptyset$  and  $X_0 \cap X_1 = \emptyset$ . On the other hand every such chain can be interpreted as a one parameter word over the alphabet  $\{0, 1\}$ . Using this interpretation, the canonizing Graham-Rothschild theorem gives a canonical set of equivalence relations as follows: on the left hand side as  $(2, 1)$ -canonical pairs, on the right hand side in terms of 2-element chains saying that  $(X_0, X_0 \cup X_1)$  is equivalent to  $(Y_0, Y_0 \cup Y_1)$  if and only if the equation(s) in the second column is (are) fulfilled:

$$\begin{array}{ll}
 J = \emptyset \text{ and} & \\
 \{\{0, \lambda\}, \{1\}\}_{\approx_1} & X_0 = Y_0 \\
 \{\{0\}, \{1, \lambda\}\}_{\approx_1} & X_0 \cup X_1 = Y_0 \cup Y_1 \\
 \{\{0, 1, \lambda\}\}_{\approx_1} & \text{always} \\
 \\
 J = \{0\} \text{ and} & \\
 \{\{0\}, \{1\}\}_{\approx_0}, \{\{0\}, \{1\}, \{\lambda\}\}_{\approx_1} & X_0 = Y_0 \quad \text{and} \quad X_1 = Y_1 \\
 \{\{0\}, \{1\}\}_{\approx_0}, \{\{0, 1\}, \{\lambda\}\}_{\approx_1} & \{x \in X_0 \mid x < \min X_1\} = \{y \in Y_0 \mid y < \min Y_1\} \\
 & \text{and} \quad X_1 = Y_1 \\
 \{\{0\}, \{1\}\}_{\approx_0}, \{\{0, \lambda\}, \{1\}\}_{\approx_1} & X_0 = Y_0 \quad \text{and} \quad \min X_1 = \min Y_1 \\
 \{\{0\}, \{1\}\}_{\approx_0}, \{\{0\}, \{1, \lambda\}\}_{\approx_1} & X_0 \cup X_1 = Y_0 \cup Y_1 \quad \text{and} \quad \min X_1 = \min Y_1 \\
 \{\{0\}, \{1\}\}_{\approx_0}, \{\{0, 1, \lambda\}\}_{\approx_1} & \{x \in X_0 \mid x < \min X_1\} = \{y \in Y_0 \mid y < \min Y_1\} \\
 & \text{and} \quad \min X_1 = \min Y_1 \\
 \{\{0, 1\}\}_{\approx_0}, \{\{0, 1\}, \{\lambda\}\}_{\approx_1} & X_1 = Y_1 \\
 \{\{0, 1\}\}_{\approx_0}, \{\{0, 1, \lambda\}\}_{\approx_1} & \min X_1 = \min Y_1
 \end{array}$$

In general, coloring  $\mathcal{B}(k)$ -lattices one obtains a canonizing version in the same way interpreting the  $(2, k)$ -canonical pairs in terms of sets. For sublattices of



Boolean lattices, i.e., for arbitrary distributive lattices, the situation gets slightly more complicated. The interested reader will find a discussion of this in Prömel and Voigt (1982).

### 6.4.3 Finite Sets

The last application of the canonical Graham-Rothschild theorem we mention in this section is another proof of a finite version of the Erdős-Rado canonization theorem.

**Theorem 6.9.** *Let  $k$  and  $m$  be positive integers. Then there exists a positive integer  $n = ER(k, m)$  such that for every coloring  $\Delta : [n]^k \rightarrow \omega$  there exists an  $m$ -subset  $M \in [n]^m$  and there exists a (possibly empty) set  $J \subseteq k$  such that*

$$\Delta(X) = \Delta(Y) \quad \text{if and only if} \quad X : J = Y : J$$

for every pair  $X, Y \in [M]^k$ .

*Proof.* Let  $n$  be according to Theorem 6.6 with respect to  $A = \{0\}$ ,  $k$  and  $m$ . Let  $\Delta : [n]^k \rightarrow \omega$  be a coloring. Define  $\Delta' : [\{0\}]_k^n \rightarrow \omega$  by  $\Delta'(g) = \Delta(\phi \cdot g)$ . Then there exist a  $(\{0\}, k)$ -canonical pair  $\Pi = (J, (\approx_i)_{i \in J^+})$  and an  $f \in [\{0\}]_m^n$  satisfying Theorem 6.6. Observe that by definition of  $\Delta'$ , every  $\approx_i$  can only have one equivalence class. But this implies immediately that  $M = \{f^{-1}(\lambda_i) \mid i < m\} \in [n]^m$  and  $J \subseteq k$  satisfy Theorem 6.9.  $\square$