TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 311, Number 2, February 1989

INFIX CONGRUENCES ON A FREE MONOID

C. M. REIS

ABSTRACT. A congruence ρ on a free monoid X^* is said to be infix if each class C of ρ satisfies $u \in C$ and $xuy \in C$ imply $xy = 1$.

The main purpose of this paper is a characterization of commutative maximal infix congruences. These turn out to be congruences induced by homomorphisms τ from X^* to $\mathbf{N}^0$, the monoid of nonnegative integers under addition, with $\tau^{-1}(0) = 1$.

1. INTRODUCTION

Let X be a finite alphabet and X^* the free monoid on X. A subset T of X^* is an *infix code* if u and xuy in T imply $xy = 1$. A congruence ρ on X^* is said to be *infix* (resp. *f-disjunctive*) if each ρ-class is an infix code (resp. a finite infix code). f-disjunctive congruences, which form a subset of the set of infix congruences, were introduced in [6] and it is, in part, the purpose of this paper to further study these congruences within the broader context of infix congruences. In addition, results analogous to those obtained in [6] for f-disjunctive congruences will be proved for infix congruences.

In §2 we prove some general results on infix congruences. In particular we show that the class of infix congruences is strictly larger than the class of f-disjunctive congruences. We also prove that commutative infix congruences are in fact f-disjunctive and that commutative maximal infix congruences are cancellative.

A congruence ρ on $X^* = \{a_1, a_2, \ldots, a_n\}^*$ is said to be *p-linear* if there exist positive integers $l_1, l_2, \ldots, l_n$ such that $u \equiv v(\rho)$ if and only if $\sum_{i=1}^n l_i|u|_{a_i} = \sum_{i=1}^n l_i|v|_{a_i}$, where $|w|_{a_i}$ denotes the number of occurrences of the letter a_i in the word w. §3 is devoted to the characterization of commutative maximal infix congruences. These turn out to be precisely the p-linear congruences.

Throughout this paper, $\mathbf{R}$ will denote the real numbers, $\mathbf{Z}$ the integers and $\mathbf{N}$ the natural numbers. The length of a word $w \in X^*$ will be denoted by $|w|$ while the syntactic congruence of a language L over X^* will be denoted by P_L. Recall that P_L is defined by $u \equiv v(P_L)$ if, for all $x, y \in X^*$, $xuy \in L$ if and only if $xvy \in L$. A congruence ρ will be said to be *principal* if $\rho = P_L$

Received by the editors December 1, 1987.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 20M05.
Key words and phrases. Infix congruences, free monoid, commutative maximal infix congruences.

for some language L. If $f: M \to T$ is a monoid homomorphism, then $\ker f$ is the congruence defined by $u \equiv v(\ker f)$ if $f(u) = f(v)$.

As a general reference we recommend G. Lallement's book [5].

2. General results on infix congruences

In [6], it was shown that, given an f-disjunctive congruence ρ, there exists a principal f-disjunctive congruence P_L with $\rho \leq P_L$. Here we prove a somewhat weaker result for infix congruences. Although the proof is similar, we include it for the sake of completeness. We begin with a lemma.

Lemma 2.1. *Let ρ be an infix congruence on X^*. Then any nontrivial submonoid M of X^* meets infinitely many ρ-classes.*

Proof. Suppose M meets only finitely many ρ-classes, say $\bar{w}_1, \bar{w}_2, \dots, \bar{w}_n$ where $\bar{w}_i$ is the ρ-class of the word $w_i \in M$. Then $\{\bar{w}_1, \bar{w}_2, \dots, \bar{w}_n\}$ is a submonoid of X^*/ρ. Hence, $\bar{w}_j$, say, is idempotent, whence $w_j^2 \equiv w_j(\rho)$. Since ρ is infix, $w_j = 1$. It follows that for each $\bar{w}_i$, $\bar{w}_i^k = 1$ for some k. Hence $w_i = 1$ for all i. Thus M is the trivial monoid, a contradiction. □

Definition 2.2. A language $L \subset X^*$ is said to be *dense* if $X^*wX^* \cap L \neq \varnothing$ for all $w \in X^*$.

Theorem 2.3. *Let ρ be an infix congruence. Then there exists a dense language L such that $\rho \leq P_L$ and the restriction of ρ to L is P_L.*

Proof. Let $C_1, C_2, \dots$ be the classes of ρ, so numbered that $m(C_i) \leq m(C_j)$ if $i < j$ where $m(C_i) = \min\{|w| \mid w \in C_i\}$. Let $w_1 \leq w_2 \leq \cdots$ be a total ordering of X^+, the free semigroup on X, and choose a certain subset of the set of all ρ-classes as follows:

Choose $D_1 = C_{i_1}$ such that $X^*w_1X^* \cap C_{i_1} \neq \varnothing$ and let $\alpha_1 = u_1w_1v_1 \in D_1$. Choose $D_2 = C_{i_2}$ such that $X^*\alpha_1w_2X^* \cap C_{i_2} \neq \varnothing$ and let $\alpha_2 = u_2\alpha_1w_2v_2 \in D_2$. Now, by Lemma 2.1, $X^*\alpha_2w_3X^*$ meets infinitely many ρ-classes; hence we may choose $D_3 = C_{i_3}$ with $X^*\alpha_2w_3X^* \cap C_{i_3} \neq \varnothing$ and $|u_2\alpha_2w_2v_2| < m(D_3)$. Let $\alpha_3 = u_3\alpha_2w_3v_3 \in D_3$. Having chosen D_j, $3 \leq j \leq n$ with $\alpha_j \in D_j$, $\alpha_j = u_j\alpha_{j-1}w_jv_j \in D_j$ and $|u_{j-1}u_{j-2}\cdots u_2\alpha_{j-1}w_2v_2\cdots w_{j-1}v_{j-1}| < m(D_j)$, choose $D_{n+1} = C_{i_{n+1}}$ with $X^*\alpha_nw_{n+1}X^* \cap C_{i_{n+1}} \neq \varnothing$ and $|u_nu_{n-1}\cdots u_2\alpha_nw_2v_2\cdots w_nv_n| < m(D_{n+1})$. Again, this can be done since $X^*\alpha_nw_{n+1}X^*$ meets infinitely many ρ-classes by Lemma 2.1. Let $\alpha_{n+1} = u_{n+1}\alpha_nw_{n+1}v_{n+1} \in D_{n+1}$. Now set $L = \bigcup_{j=1}^{\infty} D_j$. Clearly $\rho \leq P_L$ since ρ saturates L and P_L is the coarsest congruence saturating L. We now show that $D_r \not\equiv D_s(P_L)$ if $r \neq s$. Assuming $r < s$, we have

$$\alpha_s = u_su_{s-1}\cdots u_{r+1}\alpha_rw_{r+1}v_{r+1}\cdots w_sv_s \in L$$

and thus $(u_su_{s-1}\cdots u_{r+1}, w_{r+1}v_{r+1}\cdots w_sv_s)$ is a context of α_r. Now

$$\gamma = u_su_{s-1}\cdots u_{r+1}\alpha_sw_{r+1}v_{r+1}\cdots w_sv_s \notin D_j$$

if $j \leq s$ since α_j is a proper factor of γ for $j \leq s$ and each D_j is infix. Moreover, $|\gamma| \leq |u_s u_{s-1} \cdots u_2 \alpha_s w_2 v_2 \ldots w_s v_s| < m(D_t)$ for all $t > s$ by our choice of the D_j's. Hence $\gamma \notin L$ and $D_r \not\equiv D_s(P_L)$ proving that the D_j's are P_L-classes. By our construction, it is clear that L is dense. □

Remark. This construction of L does not always yield an infix congruence P_L.

Corollary 2.4. *Let ρ be an infix, cancellative congruence on X^*. Then ρ is principal.*

Proof. Let L be as in Theorem 2.3 and let $u \equiv v(P_L)$. Since L is dense, there exist $x, y \in X^*$ with $xuy \in L$. But $xuy \equiv xuy(\rho)$ since the restriction of ρ to L is P_L. Since ρ is cancellative, $u \equiv v(\rho)$, whence $\rho = P_L$. □

It is tempting to conjecture that every infix congruence is f-disjunctive. The following shows that this is not the case.

Example 2.5. Let $w \in X^+$, $w = a_1^{m_1} a_2^{m_2} \cdots a_s^{m_s}$ where, for all i, $m_i > 0$, $a_i \in X$ and $a_i \neq a_{i+1}$. Then the *skeleton* of w, denoted $\mathrm{sk}(w)$, is the word $a_1 a_2 \cdots a_s$. Define $\mathrm{sk}(1) = 1$. For w as above, let $I(w)$ denote $a_1^{m_1}$ and let $F(w)$ denote $a_s^{m_s}$. Set $I(1) = F(1) = 1$. Now define a congruence ρ on X^* as follows:

$$u \equiv v(\rho) \quad \text{if } I(u) = I(v),\ F(u) = F(v) \text{ and } \mathrm{sk}(u) = \mathrm{sk}(v).$$

We prove that each ρ-class is infix. Let $u \equiv v(\rho)$. If $u = 1$, $\mathrm{sk}(v) = 1$ whence $v = 1$ and thus $[1]_\beta = \{1\}$. If $|\mathrm{sk}(u)| = 1$, then $|\mathrm{sk}(v)| = 1$ and since $I(u) = I(v)$, it follows that $u = v$. If $|\mathrm{sk}(u)| = 2$, since $I(u) = I(v)$ and $F(u) = F(v)$, we have $u = v$. Thus assume $|\mathrm{sk}(u)| > 2$. Then

$$u = b_1^{s_1} b_2^{s_2} \cdots b_n^{s_n}, \qquad v = b_1^{s_1} b_2^{t_2} \cdots b_{n-1}^{t_{n-1}} b_n^{s_n},$$

where all the exponents are positive, the b's are letters of X and $b_i \neq b_{i+1}$ for all i. Suppose $u = pvq$. If p contains a letter other than b_1, the skeleton of v is changed, which is not possible. Hence p is a power of b_1. But p cannot be a positive power of b_1 since $I(u) = I(v)$. Thus $p = 1$. Similarly, $q = 1$. Clearly ρ has infinitely many infinite classes and is therefore infix but not f-disjunctive.

As a particular example of the above, let $X = \{a, b\}$. Then the ρ-classes are as follows:

(1) Singleton classes: $[1]_\rho$, $[a^s]_\rho$, $[b^s]_\rho$, $[a^s b^t]_\rho$, $[b^t a^s]_\rho$ where s and t are positive integers. (2) Classes of the following four types, all infinite:

(i) $[a^i (ba)^s b^j]_\rho$, i, j, s positive;

(ii) $[b^i (ab)^s a^j]_\rho$, i, j, s positive;

(iii) $[a^i (ba)^s ba^j]_\rho$, i, j positive; s nonnegative;

(iv) $[b^i (ab)^s ab^j]_\rho$, i, j positive; s nonnegative.

We now consider maximal infix congruences and set the stage for a characterization in §3 of commutative maximal infix congruences.

Theorem 2.6. *Let ρ be an infix congruence of X^*. Then there exists a congruence μ containing ρ such that μ is maximal subject to being infix.*

Proof. Let $\mathscr{F} = \{\tau \mid \tau$ is an infix congruence, $\tau \supset \rho\}$ and let $\{\tau_i\}$ be a chain in $\mathscr{F}$. Then $\bigcup \tau_i$ is a congruence and if $u \equiv puq(\bigcup \tau_i)$, then $u \equiv puq(\tau_i)$ for some i, whence $p = q = 1$. Hence $\bigcup \tau_i \in \mathscr{F}$. By Zorn, $\mathscr{F}$ has a maximal element μ. □

Example 2.7. Referring back to Example 2.5 in the case $X = \{a, b\}$ it is easy to show, though tedious, that the congruence ρ is maximal infix. For, suppose that μ is an infix congruence with $\mu \supsetneq \rho$. There are various cases to consider. We check here only two cases to give the flavour of the other computations necessary to establish the maximality of ρ.

Case (i). Suppose $a^i(ba)^s b^j \equiv a^u(ba)^t b^v(\mu)$, i, j and s positive. Then

$$ba(ba)^s ba \equiv ba(ba)^t ba(\mu).$$

Hence $(ba)^{s+2} \equiv (ba)^{t+2}(\mu)$. Since μ is infix, $s = t$ and we have

$$a^i(ba)^s b^j \equiv a^u(ba)^s b^v(\mu).$$

Hence $(ba)^{s+1}b^j \equiv (ba)^{s+1}b^v(\mu)$. Again, since μ is infix, $j = v$. Similarly $i = u$.

Case (ii). Suppose $a^i(ba)^s ba^j \equiv b^u(ab)^t ab^v(\mu)$ where i, j, u and v are positive, s and t nonnegative.

Then $aba(ba)^s baba \equiv ab(ab)^t aba(\mu)$. Hence $(ab)^{s+3}a \equiv (ab)^{t+2}a(\mu)$. Since μ is infix, $s + 3 = t + 2$. Thus $a^i(ba)^s ba^j \equiv b^u(ab)^{s+1}ab^v(\mu)$. Therefore

$$ba(ba)^s ba^j \equiv b^{u+1}(ab)^{s+1}ab^v(\mu)$$

and thus

$$b(ab)^{s+1}a^j \equiv b^{u+1}(ab)^{s+1}ab^v(\mu).$$

Hence

$$b(ab)^{s+1}ab \equiv b^{u+1}(ab)^{s+1}ab^{v+1}(\mu).$$

This contradicts the fact that μ is infix since $u+1$ and $v+1$ are at least 2. □

We remark that the congruence ρ above is not cancellative. For example

$$ab^2a \equiv aba(\rho) \quad \text{but } ba \not\equiv a(\rho).$$

This is in sharp contrast to the situation when the congruence is commutative and maximal subject to being infix.

Theorem 2.8. *Let μ be a commutative congruence maximal subject to being infix. Then μ is cancellative.*

Proof. Define a congruence $\widehat{\mu}$ on X^* by $u \equiv v(\widehat{\mu})$ if there exists $w \in X^*$ with $wu \equiv wv(\mu)$. Clearly $\mu \leq \widehat{\mu}$ and $\widehat{\mu}$ is a congruence which is cancellative. Suppose now that $u \equiv xuy(\widehat{\mu})$. Then there exists $w \in X^*$ with $wu \equiv wxuy(\mu)$.

Since μ is commutative, $uw \equiv wuxy(\mu)$, thus proving that $xy = 1$ since μ is infix. Hence $\widehat{\mu}$ is infix, whence $\mu = \widehat{\mu}$. □

We now give an example of a class of commutative congruences which are maximal infix. It will be seen in the next section that this class constitutes all commutative maximal infix congruences.

Example 2.9. Let $X = \{a_1, a_2, \ldots, a_n\}$ and let $l_1, l_2, \ldots, l_n$ be positive integers. Let $|u|_{a_i}$ denote the number of occurrences of a_i in the word u. Define a congruence ρ by $u \equiv v(\rho)$ if $\sum l_i|u|_{a_i} = \sum l_i|v|_{a_i}$. ρ is in fact f-disjunctive since the l_i are all positive. It is clearly commutative and cancellative. Maximality will be proved in §3.

Definition 2.10. A congruence of the type described above will be called *p-linear.*

Given two words u and v in X^* we may ask under what conditions there exists an infix congruence ρ such that $u \equiv v(\rho)$. Clearly a necessary condition is that neither word be a factor of the other. That this is not sufficient is shown by the following example.

Example 2.11. Let $X = \{a, b\}$, $u = ababa$, $v = baba^2ba$ and let ρ be any congruence with $u \equiv v(\rho)$. Thus

$$\begin{aligned} v^2 &= baba\underline{ababa}baaba \\ &\equiv babababa\underline{ababa}aba(\rho) \\ &\equiv bab\underline{ababababaaba}aba(\rho) \\ &\equiv babv^2aba(\rho) \end{aligned}$$

showing that ρ is not infix.

At this writing, the author does not know of a necessary and sufficient condition for two words u and v to be congruent modulo an infix congruence. There is however a simple sufficient condition which we prove in Theorem 2.15 below.

Definition 2.12. On X^* define the partial order $\leq$ by $u \leq v$ if $u = u_1u_2\cdots u_n$, $v = v_1u_1v_2u_2\cdots u_nv_{n+1}$, $u_i, v_i \in X^*$.

This partial order was studied in [3] by Haines. In particular, he proved that any collection of elements in X^* which are incomparable with respect to this partial order must be finite.

Definition 2.13 [7]. A subset T of X^* is a *hypercode* if T is a set of incomparable words relative to the partial order $\leq$.

Theorem 2.14. *Let ρ be a commutative infix congruence on X^*. Then each ρ-class is a hypercode, whence ρ is f-disjunctive.*

Proof. If $u_1u_2\cdots u_n \equiv v_1u_1v_2u_2\cdots u_nv_{n+1}(\rho)$, then

$$u_1u_2\cdots u_n \equiv u_1u_2\cdots u_nv_1v_2\cdots v_{n+1}(\rho).$$

Since ρ is infix, $v_1v_2\cdots v_{n+1} = 1$. □

Theorem 2.15. *Let u and v be words of X^+, $u \neq v$. Then there exists a commutative, cancellative infix congruence with $u \equiv v(\rho)$ if and only if either $|u| = |v|$ or there exist letters $a_s, a_t \in X$ with $|u|_{a_s} > |v|_{a_s}$ and $|u|_{a_t} < |v|_{a_t}$.*

Proof. Suppose ρ is a commutative, cancellative infix congruence with $u \equiv v(\rho)$ and suppose $|u| \neq |v|$ with $|u| > |v|$, say. Let $|u|_{a_i} = x_i$, $|v|_{a_i} = y_i$ for all i. Then

$$a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n} \equiv a_1^{y_1} a_2^{y_2} \cdots a_n^{y_n} (\rho).$$

Since $\sum x_i > \sum y_i$, there exists s with $x_s > y_s$. If $x_i \geq y_i$ for all i, by cancellativity we would have

$$a_1^{x_1 - y_1} a_2^{x_2 - y_2} \cdots a_s^{x_s - y_s} \cdots a_n^{x_n - y_n} \equiv 1(\rho),$$

contradicting infixity of ρ. Hence there exists t with $x_t < y_t$. Conversely, if $|u| = |v|$, then the length congruence λ defined by $w \equiv x(\lambda)$ if $|w| = |x|$ will do. Suppose now $|u| \neq |v|$ and $|u|_{a_s} > |v|_{a_s}$, $|u|_{a_t} \leq |v|_{a_t}$. Again, letting $|u|_{a_i} = x_i$, $|v|_{a_i} = y_i$, let $S = \{i \mid x_i - y_i \geq 0\}$ and let $T = \{i \mid x_i - y_i < 0\}$. By hypothesis $S \neq \varnothing \neq T$. Consider the nontrivial words $\prod_{i \in S} a_i^{x_i - y_i}$ and $\prod_{i \in T} a_i^{y_i - x_i}$ (the order in which the a_i appear is immaterial). Let $x_i - y_i = s_i$, for all $i \in S$, and let $y_i - x_i = t_i$ for all $i \in T$. Let $l_1, l_2, \ldots, l_n$ be positive integers satisfying $\sum_{i \in S} l_i s_i = \sum_{i \in T} l_i t_i$. Then for some permutation π of $1, 2, \ldots, n$ the p-linear congruence ρ defined by $w \equiv x(\rho)$ if $\sum l_{\pi(i)} |w|_{a_i} = \sum l_{\pi(i)} |x|_{a_i}$ is the required congruence identifying u and v. $\square$

3. A characterization of commutative, maximal infix congruences

It is the purpose of this section to prove that the commutative maximal infix congruences are precisely the p-linear congruences defined in 2.10. We begin by proving the easier half of this result.

Theorem 3.1. *Every p-linear congruence ρ of X^* is maximal infix.*

Proof. Let $X = \{a_1, a_2, \ldots, a_n\}$ and let $l_1, l_2, \ldots, l_n$ be n positive integers. Let ρ be the p-linear congruence defined by $u \equiv v(\rho)$ if $\sum l_i |u|_{a_i} = \sum l_i |v|_{a_i}$. Let μ be a maximal infix congruence with $\mu > \rho$. By Theorem 2.8, μ is cancellative. Since $\mu > \rho$, there exist $v, w \in X^*$ with $v \equiv w(\mu)$ but $v \not\equiv w(\rho)$. Let $v = a_1^{u_1} a_2^{u_2} \cdots a_n^{u_n}$, $w = a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n}$. Since $v \not\equiv w(\rho)$, $l_1 x_1 + l_2 x_2 + \cdots + l_n x_n \neq l_1 u_1 + l_2 u_2 + \cdots + l_n u_n$. Assume w.l.o.g. that $l_1 x_1 + l_2 x_2 + \cdots + l_n x_n > l_1 u_1 + \cdots + l_n u_n$. Since $a_1^{kl_2} \equiv a_2^{kl_1}(\mu)$ for all positive integers k,

$$a_1^{x_1 + kl_2} a_2^{x_2} \cdots a_n^{x_n} \equiv a_1^{u_1} a_2^{kl_1 + u_2} a_3^{u_3} \cdots a_n^{u_n} (\mu)$$

for all k. We may thus assume w.l.o.g. that $x_1 > u_1$. Similarly, since $a_2^{kl_3} \equiv a_3^{kl_2}(\mu)$ for all positive k,

$$a_1^{x_1} a_2^{x_2 + kl_3} a_3^{x_3} \cdots a_n^{x_n} \equiv a_1^{u_1} a_2^{u_2} a_3^{u_3 + kl_2} \cdots a_n^{u_n} (\mu)$$

for all k. Thus, again, we may assume that $x_1 > u_1$, $x_2 > u_2$. Continuing in this fashion, we may assume that $x_i > u_i$, $i = 1, 2, \dots, n-1$. If $x_n \geq u_n$,

$$a_1^{u_1} a_2^{u_2} \cdots a_n^{u_n} \cdot a_1^{x_1 - u_1} a_2^{x_2 - u_2} \cdots a_n^{x_n - u_n} \equiv a_1^{u_1} a_2^{u_2} \cdots a_n^{u_n} (\mu)$$

contradicting the fact that μ is infix. Thus assume that $x_n < u_n$. Since μ is cancellative we have

$$a_1^{x_1 - u_1} a_2^{x_2 - u_2} \cdots a_{n-1}^{x_{n-1} - u_{n-1}} \equiv a_n^{u_n - x_n} (\mu)$$

where $l_1(x_1 - u_1) + \cdots + l_{n-1}(x_{n-1} - u_{n-1}) > l_n(u_n - x_n)$. Let $s_i = x_i - u_i$, $i = 1, 2, \dots, n-1$, $s_n = u_n - x_n$. Then we have

$$a_1^{s_1} a_2^{s_2} \cdots a_{n-1}^{s_{n-1}} \equiv a^{s_n} (\mu) \quad \text{with } l_1 s_1 + \cdots + l_{n-1} s_{n-1} > l_n s_n.$$

Now

$$a_n^{l_i s_i} \equiv a_i^{l_n s_i} (\mu), \qquad i = 1, 2, \dots, n-1.$$

Thus

$$a_1^{l_n s_1} \cdot a_2^{l_n s_2} \cdots a_{n-1}^{l_n s_{n-1}} \equiv a_n^{l_1 s_1 + l_2 s_2 + \cdots + l_{n-1} s_{n-1}} (\mu).$$

But

$$a_1^{l_n s_1} a_2^{l_n s_2} \cdots a_{n-1}^{l_n s_{n-1}} \equiv a_n^{l_n s_n} (\mu),$$

whence

$$a_n^{l_1 s_1 + l_2 s_2 + \cdots + l_{n-1} s_{n-1}} \equiv a_n^{l_n s_n} (\mu).$$

But $l_1 s_1 + l_2 s_2 + \cdots + l_{n-1} s_{n-1} > l_n s_n$, contradicting the fact that μ is infix. Hence ρ is maximal infix. □

Remark. If ρ is commutative, cancellative and infix, then ρ is not necessarily maximal infix. For example, the congruence ρ defined by $u \equiv v(\rho)$ if $|u|_{a_i} = |v|_{a_i}$ for all $a_i \in X$ is clearly commutative, cancellative and infix but $\rho \lneq \lambda$, λ the length congruence.

To prove that every commutative maximal congruence is p-linear, we need several lemmas, some involving ideas from linear topological spaces.

We start with the following simple lemma.

Lemma 3.2. *Let ρ be a commutative cancellative congruence on X^*. Then ρ is infix if and only if $[1]_\rho = \{1\}$.*

Proof. The necessity is clear. Assume that $[1]_\rho = \{1\}$. If $u \equiv xuy(\rho)$, then, by commutativity, $u \equiv uxy(\rho)$, whence $xy \equiv 1(\rho)$ by cancellativity. Therefore $xy = 1$ and ρ is infix. □

We now start proving a series of results which properly belong to functional analysis. We refer the reader to [4] for a more general discussion.

Definition 3.3. (a) Let $\mathbf{R}$ denote the real numbers. A *cone* C is a subset of $\mathbf{R}^n$ such that (i) $C + C \subset C$; (ii) $aC \subset C$ for all nonnegative real numbers a.

(b) Let $\alpha \in \mathbf{R}^n$, $\alpha \neq 0$. The set of vectors γ such that $\alpha \cdot \gamma \geq 0$ is called a *half-space*. Here "$\cdot$" denotes the usual dot product.

Lemma 3.4 [4]. *Let C be a cone of $\mathbf{R}^n$, $C \neq \mathbf{R}^n$. Then C is contained in a half-space.*

Proof. We may assume w.l.o.g. that the interior C^0 of C is not empty. Choose $\xi \in C^0$. If $-\xi \in C$, then $0 = \xi + (-\xi) \in C^0$ since translation by $-\xi$ is a homeomorphism. Thus an open neighbourhood of 0 is contained in C. But since $aC \subset C$ for all nonnegative real numbers a, it follows that $C = \mathbf{R}^n$, a contradiction. Therefore $-\xi \notin C$. Applying Zorn's Lemma, there exists a cone M maximal subject to $-\xi \notin M$ and $\xi \in M^0$. Now let $\alpha \notin M$. Then by the maximality of M, $-\xi \in M + \mathrm{sp}^+\{\alpha\}$ where $\mathrm{sp}^+\{T\}$ denotes the set of all linear combinations of vectors of T with nonnegative coefficients. Hence $-\xi = \mu + b\alpha$, $\mu \in M$, $b > 0$. Therefore $-\alpha = (1/b)(\xi + \mu)$. Since $\xi \in M^0$ and both translation by μ and multiplication by $1/b$ are homeomorphisms it follows that $-\alpha \in M^0$. Thus

$$\mathbf{R}^n = (M \cup -M^0) \cap (-M \cup M^0) = (M \cap (-M)) \cup M^0 \cup -M^0.$$

We prove that $M \cap (-M)$, M^0 and $-M^0$ are mutually disjoint and that $M \cap (-M)$ is a hyperplane, i.e., a subspace of dimension $n-1$. Let $\alpha \in M \cap -M^0$. Then $\alpha \in M$ and $-\alpha \in M^0$, showing that $0 \in M^0$, whence, as before $M = \mathbf{R}^n$, a contradiction. Therefore $M \cap -M^0 = \varnothing$ and consequently $-M \cap M^0 = \varnothing$. Hence the three sets $M \cap -M$, M^0 and $-M^0$ are disjoint. Clearly $M \cap -M$ is a subspace. We show that it is a hyperplane by proving that each $\alpha \in \mathbf{R}^n$ is a linear combination of ξ and some element of $M \cap (-M)$. Let $\alpha \in -M^0$. Then $\{t \mid t\alpha + (1-t)\xi \in M^0\}$ and $\{t \mid t\alpha + (1-t)\xi \in -M^0\}$ are nonintersecting sets, open in $[0, 1]$ and which do not cover $[0, 1]$ since $[0, 1]$ is connected. Hence there exists t_0 with $0 < t_0 < 1$ such that $t_0\alpha + (1-t_0)\xi \notin M^0 \cup -M^0$. Hence $t_0\alpha + (1-t_0)\xi \in M \cap (-M)$ since $\mathbf{R}^n = M \cap (-M) \cup M^0 \cup -M^0$. Therefore, α is a linear combination of the fixed vector ξ and a vector of $M \cap (-M)$. If $\alpha \in M^0$, $-\alpha \in -M^0$ and again $-\alpha$, and thus α, is a linear combination of ξ and an element of $M \cap -M$. Hence $M \cap -M$ is indeed a hyperplane and C lies in the half-space M. □

Definition 3.5. Let $\alpha_1, \alpha_2, \dots, \alpha_k$ be vectors of $\mathbf{R}^n$ and let S be a subset of $\mathbf{R}$ containing 0. We shall say that the set $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ of vectors has property P with respect to S if the following holds: $\sum c_i\alpha_i = 0$ and $c_i \in S$ for all i imply $c_i = 0$ for all i.

Corollary 3.6. *Let $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ be a set of vectors of $\mathbf{R}^n$ having property P with respect to $\mathbf{R}^+$, where $\mathbf{R}^+$ denotes the nonnegative real numbers. Then there exists a nonzero vector α such that $\alpha \cdot \alpha_i \geq 0$ for all i.*

Proof. $C = \mathrm{sp}^+\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ is a cone. If β and $-\beta \in C$, then $\beta = \sum b_i\alpha_i$, $-\beta = \sum b_i'\alpha_i$, with b_i and b_i' nonnegative. Thus $0 = \sum(b_i + b_i')\alpha_i$ whence $b_i + b_i' = 0$ for all i. Since b_i and b_i' are nonnegative for all i, it follows that $b_i = b_i' = 0$ for all i. Hence, for all $\gamma \in C$, $\gamma \neq 0$, $-\gamma \notin C$ showing that $C \neq \mathbf{R}^n$. By Lemma 3.4, C is contained in a half-space determined by the hyperplane $\alpha \cdot \xi = 0$. By choosing α with the appropriate sign, we may assume $\alpha \cdot \alpha_i \geq 0$ for all i. □

Corollary 3.7. *Let* $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ *be vectors of* $\mathbf{R}^n$ *with property* P *with respect to* $\mathbf{R}^+$. *Then there exists a vector* β *such that* $\beta \cdot \alpha_i > 0$ *for all* i.

Proof. By the previous result, there exists $\alpha \neq 0$ with $\alpha \cdot \alpha_i \geq 0$ for all i. Let β be a vector with $\beta \cdot \alpha_i > 0$ for the largest number of α_i's. Let S be the set of those α_i with $\beta \cdot \alpha_i > 0$ and T the set of those α_i with $\beta \cdot \alpha_i = 0$ (note that S could be $\varnothing$). Let $V = \operatorname{sp}\{\alpha_i \mid \alpha_i \in T\}$. Then $\dim V = t \leq n-1$. By the previous result, there exists a hyperplane H of V with T contained in the half-space determined by H. Now $H = \widehat{H} \cap V$ where $\widehat{H}$ is a hyperplane of $\mathbf{R}^k$ given by $\gamma \cdot \zeta = 0$, say. Clearly T lies on one side of $\widehat{H}$ and we may assume $\gamma \cdot \alpha_i \geq 0$ for all $\alpha_i \in T$. If $\gamma \cdot \alpha_i = 0$ for all $\alpha_i \in T$, $V = \operatorname{sp}\{\alpha_i \mid \alpha_i \in T\} \subset \widehat{H}$, whence $V = \widehat{H} \cap V = H$, a contradiction. Thus $\gamma \cdot \alpha_i > 0$ for some $\alpha_i \in T$. By multiplying γ by a suitable positive scalar we may assume $\|\gamma\| < \beta \cdot \alpha_i / \|\alpha_i\|$ for all $\alpha_i \in S$. (Here, $\| \ \|$ denotes the usual Euclidean norm). Consider now $\beta + \gamma$. If $\alpha_i \in T$, $(\beta + \gamma) \cdot \alpha_i = \beta \cdot \alpha_i + \gamma \cdot \alpha_i = \gamma \cdot \alpha_i$ and if $\alpha_i \in S$, $(\beta + \gamma) \cdot \alpha_i = \beta \cdot \alpha_i + \gamma \cdot \alpha_i$. But by the Cauchy-Schwarz inequality,

$$|\gamma \cdot \alpha_i| \leq \|\gamma\| \, \|\alpha_i\| \leq \beta \cdot \alpha_i / \|\alpha_i\| \cdot \|\alpha_i\| = \beta \cdot \alpha_i .$$

Hence if $\alpha_i \in S$, $(\beta + \gamma) \cdot \alpha_i = \beta \cdot \alpha_i + \gamma \cdot \alpha_i > 0$. But for at least one $\alpha_j \in T$, $\beta \cdot \alpha_j > 0$, whence $(\beta + \gamma)\alpha_j > 0$. The maximality of S is thus contradicted proving that $\beta \cdot \alpha_i > 0$ for all i. □

Corollary 3.8. *Given vectors* $\alpha_1, \alpha_2, \ldots, \alpha_k$ *in* $\mathbf{Z}^n$ *with property* P *with respect to* $\mathbf{N}^0$, *then there exists a vector* $\alpha \in \mathbf{Z}^n$ *such that* $\alpha \cdot \alpha_i > 0$ *for all* i.

Proof. We first need to show that property P with respect to $\mathbf{N}^0$ implies property P with respect to $\mathbf{R}^+$. Suppose by way of contradiction that this is not the case. We may assume w.l.o.g. that there exist real numbers r_i, $i = 1, 2, \ldots, k$, all nonzero, such that $\sum r_i \alpha_i = 0$. Define a linear transformation $T: \mathbf{R}^k \to \operatorname{sp}\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ by $T: (c_1, c_2, \ldots, c_k) \to \sum_{i=1}^k c_i \alpha_i$. The kernel of T has a basis $\beta_1, \beta_2, \ldots, \beta_s$ where $\beta_i \in \mathbf{Z}^k$. Let $(r_1, r_2, \ldots, r_k) = \sum b_i \beta_i$ and let $N((r_1, r_2, \ldots, r_k); \delta)$ be a neighbourhood of $(r_i, r_2, \ldots, r_k)$ of radius $\delta > 0$ such that $N((r_1, r_2, \ldots, r_k); \delta) \subset \{(x_1, x_2, \ldots, x_k) \mid x_i > 0\}$. Choose rationals p_i/q_i so that $|p_i/q_i - b_i| < \delta / \sum \|\beta_j\|$ for all i. Then

$$\begin{aligned} \left\| (r_1, r_2, \ldots, r_n) - \sum (p_i/q_i)\beta_i \right\| &= \left\| \sum (b_i - p_i/q_i)\beta_i \right\| \\ &\leq \sum |b_i - p_i/q_i| \, \|\beta_i\| < \delta . \end{aligned}$$

Hence

$$\gamma = \sum (p_i/q_i)\beta_i \in N((r_1, r_2, \ldots, r_k); \delta) \subset \{(x_1, x_2, \ldots, x_k) \mid x_i > 0\}.$$

But $\gamma \in \ker T$ and γ has positive rational coordinates. Therefore some positive multiple of γ, say $(c_1, c_2, \dots, c_k)$, has positive integral coordinates. But $\sum c_i\alpha_i = 0$, a contradiction. Therefore $\mathrm{sp}^+\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ has the property in the hypothesis of Corollary 3.7. Hence, by Corollary 3.7, there exists $\beta \in \mathbf{R}^n$ with $\beta \cdot \alpha_i > 0$ for all i. But, by continuity of the dot product, there exists a vector α' with rational coordinates such that $\alpha' \cdot \alpha_i > 0$ for all i. Multiplying α' by a suitable positive integer yields a vector $\alpha \in \mathbf{Z}^n$ with $\alpha \cdot \alpha_i > 0$ for all i. $\square$

The next theorem is the cornerstone of the proof that each commutative maximal infix congruence is p-linear.

Theorem 3.9. *Let $M = \langle m_1, m_2, \dots, m_k\rangle$ be a finitely generated cancellative, commutative monoid with trivial group of units. Then there exists a homomorphism $\chi: M \to \mathbf{N}^0$ given by*

$$\chi: \prod(m_i^{c_i}) \to \sum_{i=1}^{k} l_i c_i, \qquad l_i \textit{ positive integers}.$$

Proof. Let G be the group of fractions of M. Then G is generated, qua group, by $m_1, m_2, \dots, m_k$. Thus $G \overset{\varphi}{\approx} T \oplus \mathbf{Z}^n$ where $k \geq n$, T is a finite abelian group and $n \geq 1$ since M is aperiodic. Let $\varphi(m_i) = (t_i, \alpha_i)$, $i = 1, 2, \dots, k$. Suppose that $\sum_{i=1}^k c_i\alpha_i = 0$, $c_i \in \mathbf{N}^0$ and let $p \in \mathbf{N}$ with $pT = (0)$. Then $\varphi(\prod m_i^{c_i p}) = (0, \sum pc_i\alpha_i) = (0, 0)$. Since φ is injective, it follows that $\prod m_i^{c_i p} = 1$. But since the group of units of M is trivial, we have $c_i p = 0$ for all i and thus $c_i = 0$ for all i. Letting π denote the projection of G onto $\mathbf{Z}^n$, the mapping $\psi = \pi\varphi\,|_M$ is a monoid homomorphism of M to $\mathbf{Z}^n$. Let $V = \{\sum_{i=1}^k c_i\alpha_i \mid c_i \in \mathbf{N}^0\}$. We now prove that there exists a monoid homomorphism $\tau: V \to \mathbf{N}^0$ given by $\tau(\sum_{i=1}^k c_i\alpha_i) = \sum_{i=1}^k l_i c_i$ where the l_i are positive integers. Since $\{\sum_{i=1}^k r_i\alpha_i \mid r_i \in \mathbf{R}\} = \mathbf{R}^n$, we may assume w.l.o.g. that $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ forms a basis. Let $\alpha_j = \sum_{i=1}^n b_{ji}\alpha_i$, for all j, $j \geq n+1$ where the b_{ji} are rational since the α's are all in $\mathbf{Z}^n$. Let $\beta_j = (b_{j1}, b_{j2}, \dots, b_{jn})$, $j \geq n+1$. We now show that the system of inequalities $\beta_j \cdot \xi > 0$, $j \geq n+1$, has a solultion $t = (p_1, p_2, \dots, p_n)$ where each p_i is a positive integer. Let $\mathscr{E}_i$, $i = 1, 2, \dots, n$, be the standard basis of $\mathbf{R}^n$ and consider the set $\{\mathscr{E}_1, \mathscr{E}_2, \dots, \mathscr{E}_n, \beta_{n+1}, \dots, \beta_k\}$. Let $c_1\mathscr{E}_1 + c_2\mathscr{E}_2 + \cdots + c_n\mathscr{E}_n + c_{n+1}\beta_{n+1} + \cdots + c_k\beta_k = 0$, $c_i \in \mathbf{N}^0$. Then $c_i + \sum_{j=n+1}^k c_j b_{ji} = 0$ for $i = 1, 2, \dots, n$. Now

$$\begin{aligned}\sum_{i=1}^{k} c_i\alpha_i &= \sum_{i=1}^{n} c_i\alpha_i + \sum_{j=n+1}^{k}\sum_{i=1}^{n} c_j b_{ji}\alpha_i \\ &= \sum_{i=1}^{n}\left(c_i + \sum_{j=n+1}^{k} c_j b_{ji}\right)\alpha_i = 0.\end{aligned}$$

Hence $c_i = 0$ for all i, $1 \le i \le k$, since $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ has property P with respect to $\mathbf{N}^0$. Thus the set of vectors $\{\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_n, \beta_{n+1}, \ldots, \beta_k\}$ has property P with respect to $\mathbf{N}^0$. By Corollary 3.8, there exists a vector $\mathscr{H} = (q_1, q_2, \ldots, q_n) \in \mathbf{Z}^n$ such that $\mathscr{H} \cdot \mathscr{E}_i > 0$ for all i, $1 \le i \le k$, and $\mathscr{H} \cdot \beta_j > 0$ for all j, $n+1 \le j \le k$. Since $\mathscr{H} \cdot \mathscr{E}_i = q_i$, it follows that all the q_i are positive integers. Since $\mathscr{H} \cdot \beta_j$ is rational for all j, an appropriate positive integral multiple, say $(p_1, p_2, \ldots, p_n)$, of $\mathscr{H}$ is such that $(p_1, p_2, \ldots, p_n) \cdot \beta_j$ is a positive integer for all j. Now define a linear transformation $T: \mathbf{R}^n \to \mathbf{R}$ by setting $T(\alpha_i) = p_i$, $i = 1, 2, \ldots, n$. Then $T(\alpha_j) = (p_1, p_2, \ldots, p_n) \cdot \beta_j \in \mathbf{N}$ for $j \ge n+1$. Hence the restriction τ of T to V is a monoid homomorphism from V to $\mathbf{N}^0$. Let $\chi = \tau\psi$. Then

$$\chi(\pi m_i^{c_i}) = \sum_{i=1}^{n} p_i c_i + \sum_{j=n+1}^{k} [(p_1, p_2, \ldots, p_n) \cdot \beta_j] c_j.$$

Hence, setting $l_i = p_i$, $i = 1, 2, \ldots, n$, and $l_j = (p_1, \ldots, p_n) \cdot \beta_j$, $j \ge n+1$, we have that the l_i are positive integers and $\chi(\prod m_i^{c_i}) = \sum_{i=1}^{k} l_i c_i$. □

We are now able to prove the result we have been aiming for.

Theorem 3.10. *Let μ be a commutative maximal infix congruence on $X^* = \{a_1, a_2, \ldots, a_n\}^*$. Then μ is p-linear.*

Proof. By Theorem 2.8, X^*/μ is cancellative. Also, since μ is infix, the group of units is trivial. By the preceding theorem, there exists a homomorphism $\chi: X^*/\mu \to \mathbf{N}^0$ given by $\chi(\prod \overline{a}_i^{c_i}) = \sum l_i c_i$ where the l_i are positive integers and $\overline{a}_i$ is the μ-class of a_i. Let $\nu: X^* \to X^*/\mu$ be the quotient homomorphism. Then $\ker \chi\nu \ge \mu$ and $\ker \chi\nu$ is cancellative. But since $[1]_{\ker \chi\nu} = \{1\}$, by Lemma 3.2, $\ker \chi\nu$ is infix. Hence $\ker \chi\nu = \mu$ by maximality of μ. Thus $a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n} \equiv a_1^{u_1} a_2^{u_2} \cdots a_n^{u_n} (\mu)$ if and only if

$$\chi\nu(a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n}) = \chi\nu(a_1^{u_1} a_2^{u_2} \cdots a_n^{u_n})$$

i.e., if and only if

$$\sum l_i x_i = \sum l_i u_i. \quad \square$$

References

1. S. Eilenberg, *Automata, languages and machines*, vol. A, Academic Press, 1974.
2. Y. Q. Guo, H. J. Shyr and G. Thierrin, *f-disjunctive languages*, Internat. J. Comput. Math. **18** (1986), 219–237.
3. L. H. Haines, *On free monoids partially ordered by embedding*, J. Combin. Theory **6** (1969), 94–98.
4. J. L. Kelley and I. Namioka, *Linear topological spaces*, Van Nostrand, Princeton, N.J., 1963.
5. G. Lallement, *Semigroups and combinatorial applications*, Wiley, 1979.
6. C. M. Reis, *A note on f-disjunctive languages*, Semigroup Forum **36** (1987), 159–165.
7. G. Thierrin and H. J. Shyr, *Hypercodes*, Inform. and Control **24** (1974), 45–54.

Department of Mathematics, University of Western Ontario, London, Ontario, Canada