Languages and monoids with disjunctive identity*

Lila Kari and Gabriel Thierrin
Department of Mathematics, University of Western Ontario
London, Ontario, N6A 5B7 Canada

Abstract

We show that the syntactic monoids of insertion-closed, deletion- closed and dipolar-closed languages are the groups. If the languages are insertion-closed and congruence simple, then their syntactic monoids are the monoids with disjunctive identity. We conclude with some properties of dipolar-closed languages.

1 Introduction

Let M be a monoid with identity 1. If $L \subseteq M$ is a subset of M and if $u \in M$, then:

$$u^{-1}L = \{x \in M | ux \in L\},\ Lu^{-1} = \{x \in M | xu \in L\},$$

$$L..u = \{(x, y) | x, y \in M, xuy \in L\}.$$

The relations R_L and ${}_LR$ defined by:

$$u \equiv v\ (R_L) \Leftrightarrow\ u^{-1}L = v^{-1}L,\ \ u \equiv v\ ({}_LR)\ \Leftrightarrow\ Lu^{-1} = Lv^{-1},$$

are respectively a right congruence and a left congruence of M, called the *right principal* and the *left principal congruence* determined by L. These congruences were first considered by Dubreil in [3] as a way to extend to semigroups the construction of right congruences on groups (see also [1]).

The relation P_L defined by $u \equiv v\ (P_L)\ \Leftrightarrow\ L..u = L..v$ is a congruence of M called the *principal congruence* determined by L. This congruence was first considered in semigroups for describing their homomorphisms and a systematic study of their properties was given by Croisot in [2].

A subset $L \subseteq M$ is called *disjunctive* if the principal congruence P_L is the identity relation on M (see for example [10], [12]). Given any subset T of M, it is easy to see that the set of classes representing the elements of T is a disjunctive

*This research was supported by Grant OGP0007877 of the Natural Sciences and Engineering Research Council of Canada

set of the quotient monoid M/P_T. If $u \in M$ and if the set $\{u\}$ is disjunctive, u will be called a *disjunctive element* of M. In particular, it is possible for the identity 1 of a monoid M to be a disjunctive element. If this is the case, then the monoid is *simple* or *0-simple* (see [5]). Recall that (see [1]) a monoid M is simple if for any $u, v \in M$ there exist x, y such that $xuy = v$. M is 0-simple if it has a zero element, and if the preceding condition holds for any two nonzero elements u, v of M. The groups and the bicyclic monoid are examples of simple monoids.

Since in this paper, monoids with disjunctive identity will be associated with special classes of languages, we recall a few definitions related to formal languages.

Let X^* be the free monoid generated by the alphabet X where the identity 1 of X^* is the empty word. Elements and subsets of X^* are called respectively *words* and *languages* over X. The congruences R_L and P_L determined by a language $L \subseteq X^*$ are called respectively the *syntactic right congruence* and the *syntactic congruence* of L. The quotient monoid X^*/P_L is called the *syntactic monoid* of L. The syntactic monoid plays an important role for the characterization of several interesting classes of languages (see for example [10], [12]).

In this paper, we are interested in the syntactic monoid of some classes of languages related to the operations of insertion and deletion (see [6], [8]). The following three classes of languages L are considered:

(i) insertion-closed (or ins-closed): $u_1u_2 \in L$, $v \in L$ imply $u_1vu_2 \in L$;

(ii) deletion-closed (or del-closed): $u_1vu_2 \in L$, $v \in L$ imply $u_1u_2 \in L$;

(iii) dipolar-closed (or dip-closed): $u_1u_2 \in L$, $u_1vu_2 \in L$ imply $v \in L$.

We show that the syntactic monoids of insertion-closed, deletion-closed and dipolar-closed languages are the groups. If the languages are insertion-closed and congruence-simple, then their syntactic monoids are the monoids with disjunctive identity. Properties of insertion-closed or deletion-closed languages have been considered in [4]. Properties of dipolar-closed languages are given in the last section of the paper. Results associated with similar concepts in relation with codes can be found in [5].

2 Insertion and deletion closed languages

Insertion and deletion have been introduced in [6] and studied for example in [6], [7], [8], [9], as operations generalizing the catenation respectively left/right quotient of languages. For two words $u, v \in X^*$, the *insertion* of v into u is defined as

$$u \longleftarrow v = \{u_1vu_2 |\ u = u_1u_2\},$$

and the *deletion* of v from u as

$$u \longrightarrow v = \{x \in X^* |\ u = x_1vx_2, x = x_1x_2\}.$$

As noticed above, instead of adding (erasing) the word v to the right (from the left/right) extremity of u, the new operation inserts (deletes) it into (from) an arbitrary position of u. The result is usually a set with cardinality greater than one, which contains the catenation (left/right quotient) of the words as one of its elements. The study of insertion and deletion has triggered the consideration in [4] of two related notions. To the language $L \subseteq X^*$, one can associate the two languages *ins(L)* and *del(L)* defined by:

(i) $ins(L) = \{x \in X^* | \forall u \in L, u = u_1u_2 \ \Rightarrow \ u_1xu_2 \in L\}$;

(ii) $del(L) = \{x \in Sub(L) | \forall u \in L, u = u_1xu_2 \ \Rightarrow \ u_1u_2 \in L\}$, where *Sub(L)* is the set of subwords of the words in L.

The set $ins(L)$ consists of all words with the following property: their insertion into any word of L yields words belonging to L. Analogously, $del(L)$ consists of all words x with the following property: x is a subword of at least one word of L, and the deletion of x from any word of L is included in L. The condition that $x \in Sub(L)$ has been added because otherwise $del(L)$ would contain irrelevant elements, such as words which are not subwords of any word of L.

A language L such that $L \subseteq ins(L)$ is called *insertion-closed*. It is immediate that L is insertion-closed iff $u = u_1u_2 \in L$, $v \in L$ imply $u_1vu_2 \in L$.

A language L is called *deletion-closed* if $v \in L$ and $u_1vu_2 \in L$ imply $u_1u_2 \in L$. For example, let $X = \{a, b\}$. Then X^* and $L_{ab} = \{x \in X^* | \ |x|_a = |x|_b\}$ are insertion-closed languages that are also deletion-closed.

A language L such that L is a class of its syntactic congruence P_L is called a *congruence simple* or shortly a *c-simple language.* It is easy to see that L is c-simple iff $xLy \cap L \neq \emptyset$ implies $xLy \subseteq L$. Remark that, if L is a c-simple language and if $1 \in L$, then L is a submonoid of X^*.

Proposition 2.1 *Let L be a language that is insertion-closed and deletion-closed. Then L is c-simple.*

Proof. Suppose that $u, xuy \in L$. Since L is del-closed, $xy \in L$. Let $v \in L$. Since L is ins-closed, this implies $xvy \in L$. Hence $xLy \subseteq L$. □

Lemma 2.1 *If L is a c-simple language over the alphabet X and if $1 \in L$, then syn(L) is a monoid with a disjunctive identity. Conversely, if M is a monoid with a disjunctive identity, then there exists a c-simple language L over an alphabet X with $1 \in L$ and such that syn(L) is isomorphic to M.*

Proof. Let $e = [L]$ be the class of L modulo P_L. Since $1 \in L$, L is a submonoid of X^* and e is the identity of the monoid $syn(L)$. The element e is a disjunctive element of $syn(L)$ because L is a class of P_L.

Conversely, let X be a set of generators of M, let e be the identity of M and let X^* be the free monoid generated by X. Let $\phi : X^* \to M$ be the canonical mapping of X^* onto M defined by $\phi(x) = e$ if $x = 1$ and

$$\phi(x) = \phi(x_1)\phi(x_2)\cdots\phi(x_n) = x_1x_2\cdots x_n \in M$$

if $x = x_1 x_2 \cdots x_n \in X^+$ with $x_i \in X$. Clearly ϕ is a morphism of X^* onto M and θ, defined by $u \equiv v(\theta)$ iff $\phi(u) = \phi(v)$, is a congruence of X^* such that X^*/θ is isomorphic to M. Let $L = \phi^{-1}(e)$. Since e is a disjunctive element of M, $\theta = P_L$ is the syntactic congruence of L, L is a class of P_L and $syn(L)$ is isomorphic to M. □

Proposition 2.2 *If L is an insertion-closed language over the alphabet X, $1 \in L$, and if L is a c-simple language, then syn(L) is a monoid with a disjunctive identity. Conversely, if M is a monoid with a disjunctive identity, then there exists an insertion-closed and c-simple language L over an alphabet X with $1 \in L$ and such that syn(L) is isomorphic to M.*

Proof. The first part follows immediately from Lemma 2.1, because L being ins-closed with $1 \in L$, is a submonoid of X^*. For the converse note that, by the same lemma, there exists a c-simple language L such that $syn(L)$ is isomorphic to M. L is ins-closed, because $vw \in L, u \in L$ implies $[vw] = e = [u]$ and therefore:

$$[vuw] = [v][u][w] = [v][w] = [vw] = e.$$

Consequently, $vuw \in L$. □

A language L is called *dipolar-closed* or simply *dip-closed* if $u_1 u_2 \in L$, $u_1 x u_2 \in L$ imply $x \in L$. This notion is related to the operation of dipolar deletion (see [6], [9]). Recall that, for two words $u, v \in X^*$, the *dipolar deletion* of v from u is defined as

$$u \rightleftharpoons v = \{x \in X^* | \; u = v_1 x v_2, v = v_1 v_2\}.$$

In other words, the dipolar deletion erases, if possible, from u a prefix and a suffix whose catenation equals v. Remark that every nonempty dipolar-closed language L contains the empty word 1, because $u_1 u_2 \in L$ implies $u_1.1.u_2 \in L$ and hence $1 \in L$. Examples and properties of dipolar-closed languages are given in the last section.

Proposition 2.3 *If L is an insertion-closed, deletion-closed and dipolar-closed language over the alphabet X, then syn(L) is a group or a group with zero.*

Conversely, if G is a group or a group with zero, then there exists an insertion-closed, deletion-closed and dipolar-closed language L over an alphabet X such that $syn(L)$ is isomorphic to G.

Proof. By Proposition 2.1, L is a class of P_L. Let $e = [L]$ be the class of L modulo P_L. Then, by Lemma 2.1 and Proposition 2.2, e is the identity and a disjunctive element of $syn(L)$.

Every monoid with disjunctive identity is either simple or 0-simple (see [5]). Suppose first that $syn(L)$ is simple and let $[u]$ be the class of u modulo P_L. There exist $x, y \in X^*$ such that $xuy \in L$ and $xuyxuy \in L$. Since $x.uyx.uy$ in

L and L is dip-closed, we have $uyx \in L$. This implies $[u][yx] = e$. Similarly $xu.yxu.y$ and $xuy \in L$ imply $yxu \in L$, i.e. $[yx][u] = e$. Since every $[u]$ has a right and a left inverse, it follows that $syn(L)$ is a group. Suppose now that $syn(L)$ is 0-simple. If the class $[u]$ of u is $\neq 0$, then, because $syn(L)$ is 0-simple, we have $[x][u][y] = e$ for some $x, y \in X^*$, or equivalently $xuy \in L$. Since L is dip-closed, by a similar argument as before we deduce that $uyx \in L$ and $yxu \in L$. Therefore every $[u] \neq 0$ has an inverse in $syn(L)$. Let $T = syn(L)\backslash\{0\}$ and let $r, s \in T$. If $rs = 0$, then, since both r and s have inverses r^{-1} and s^{-1}, we have $e = rss^{-1}r^{-1} = 0$, a contradiction. Therefore $syn(L)\backslash\{0\}$ is a group and $syn(L)$ is a group with zero.

For the converse, let X be a set of generators of G, let e be the identity of G and let X^* be the free monoid generated by X. If $\phi : X^* \to G$ is the canonical mapping of X^* onto G, then, as above, it can be shown that ϕ is a morphism of X^* onto G. Moreover θ, defined as in Lemma 2.1, is a congruence of X^* such that X^*/θ is isomorphic to G. If $L = \phi^{-1}(e)$, then $\theta = P_L$ is the syntactic congruence of L and $syn(L)$ is isomorphic to G.

If $vw, u \in L$, then, since G is group or a group with 0, $e = [v][w] = [u]$. Consequently, $[vuw] = [v][u][w] = [v][w] = e$. Therefore $vuw \in L$ and L is ins-closed.

If $vuw\ , u \in L$, then $[v][u][w] = e = [u]$. Since $e = [u]$ is the identity of G, $e = [v][u][w] = [v][w]$. Hence $vw \in L$ and L is del-closed.

If $vw, vuw \in L$, then $[v][w] = [v][u][w] = e$. If $[v]^{-1}$ and $[w]^{-1}$ are the inverses of $[v]$ and $[w]$, then: $e = [v]^{-1}[v][u][w][w]^{-1} = [u]$. Therefore $u \in L$ and L is dip-closed. □

A monoid with a disjunctive identity is either simple or 0-simple (see [5]). However such a monoid is not necessarily a group or a group with zero. For example, the bicyclic monoid B is simple and its identity 1 is disjunctive. However B is not a group.

Since the bicyclic monoid has a disjunctive identity, we can use Lemma 2.1 and Proposition 2.2 to construct a c-simple insertion-closed and deletion-closed language L_B, called the *bicyclic language,* having B as its syntactic monoid. Since B is finitely generated, the alphabet of the language L_B is also finite.

Recall that the bicyclic monoid B can be defined in the following way (see for example [1]). If N denotes the set of the non-negative integers, then $B = N \times N$ with the product defined by: $(m, n)(r, s) = (m+r-min(n, r), n+s-min(n, r))$. The element $(0, 0)$ is the identity element of B and B is generated by the pair $a = (1, 0)$ and $b = (0, 1)$. Let $X = \{a, b\}$, let X^* be the free monoid generated by X, let $e = (0, 0)$ and let ϕ be the canonical morphism of X^* onto B. Then the language $L_B = \phi^{-1}(e)$ is an ins-closed language such that $syn(L)$ is isomorphic to B.

The language L_B is del-closed. Suppose that $uwv, w \in L_B$. Then $\phi(w) = e = \phi(uwv) = \phi(u)\phi(w)\phi(v) = \phi(u)\phi(v) = \phi(uv)$. Consequently, $uv \in \phi^{-1}(e) = L_B$, and hence L_B is del-closed. The language L_B is not dip-closed. Indeed, it is easy to verify that $(0,1)(1,0) = (0,0)$ and $(0,1)(1,1)(1,0) = (0,0)$. If $c = \phi^{-1}((1,1))$, then in X^* we have $ba \in L_B$ and $bca \in L_B$. If L_B were dip-closed, we would have $c \in L_B$, i.e. $(1,1) = (0,0)$, which is impossible.

The next example is a language that is ins-closed, but not del-closed, not dip-closed and not c-simple. Let $X = \{a, b\}$ and $L = X^* \backslash \{a, a^2\}$. Clearly L is ins-closed. Since $a.a^3.a, a^3 \in L$, but $a^2 \notin L$, L is not del-closed. Since $a.a^2, a.a.a^2 \in L$, but $a \notin L$, it follows that L is not dip-closed. The language L is not c-simple. Indeed, we have $a.b.1 \in L$ with $b \in L$, hence $a.L.1 \cap L \neq \emptyset$. If L were c-simple, this would imply $a.L.1 \subseteq L$. Since $1 \in L$, we have $a = a.1.1 \in L$, a contradiction.

3 Dipolar-closed languages

Properties of insertion-closed and deletion-closed languages have been thoroughly studied in [4]. The aim of this section is to complete this investigation by studying properties of the related dipolar-closed languages. First we give some examples of dipolar-closed languages.

Examples. (1) Let $X = \{a, b\}$ and let m, n be two fixed positive integers. Let $L(a, m, b, n) = \{u \in X^* | \ \ |u|_a = km, |u|_b = kn\}$, where k is a positive integer. Then $L(a, m, b, n)$ is dip-closed, ins-closed and del-closed. Special case: $L_{ab} = L(a, 1, b, 1)$.

(2) Given a language L, $Sub(L)$ is a dip-closed language. Special case: $L = Sub(L)$. For example, the language $L = \{1, a, b, ab\}$ is dip-closed, del-closed but not ins-closed.

(3) Let L be an outfix code, i.e. $L \subseteq X^+$ and $u_1u_2, u_1xu_2 \in L$ implies $x = 1$. Then $L \cup \{1\}$ is dip-closed, but not ins-closed.

(4) Let L be an ideal of X^*, $L \neq X^*$. Then $L^c = X^* \backslash L$ is dip-closed, but in general not ins-closed or del-closed. Take for example $X = \{a, b, c\}$ and $L = X^*abX^*$. Then L^c is dip-closed, but not ins-closed since $a, b \in L^c$ with $ab \notin L^c$, and not del-closed since $acb, c \in L^c$ with $ab \notin L^c$.

(5) Let X such that $|X| \geq 2$ and let $Y \subseteq X$, $Y \neq X$, be a nonempty subalphabet of X. Then $L = Y^*$ is dip-closed, ins-closed and del-closed. In particular, a^* is dip-closed for all $a \in X$.

Proposition 3.1 *Let L be a dipolar-closed language. If L is c-simple, then L is insertion-closed and deletion-closed.*

Proof. Since L is dip-closed, $1 \in L$. If $uv, w \in L$, then $u.1.v \in L$ and since L is c-simple, this implies $uLv \subseteq L$ and $uwv \in L$.

Suppose that $w \in L$ and $uwv \in L$. Since L is c-simple, $uLv \subseteq L$. From $1 \in L$ follows $uv \in L$ and hence L is del-closed. □

A language that is dip-closed and del-closed is not in general ins-closed. For example, take $L = \{1, u\}$, $u \neq 1$.

It is easy to see that the family of dip-closed languages is closed under intersection and inverse homomorphism, but, as the next result shows, it is not closed under other basic operations of formal languages.

Proposition 3.2 *The family of dipolar-closed languages is not closed under union, complementation, catenation, catenation closure, homomorphism and intersection with regular languages.*

Proof. Let $X = \{a, b\}$.

Union: Let $L_1 = \{1, aba\}$ and $L_2 = \{1, a^2\}$. Then both L_1 and L_2 are dip-closed, but the union $L_3 = \{1, a^2, aba\}$ is not. Indeed $a.a \in L_3$, $a.b.a \in L_3$, but $b \notin L_3$.

Complementation: Let $L = a^*$. Then L is dip-closed. We have $b^2, bab \in L^c$, but $a \notin L^c$ and hence L^c is not dip-closed.

Catenation: Let $L = \{1, ab\}$. Then L is dip-closed and $L^2 = \{1, ab, abab\}$. We have $a.b, a.ba.b \in L^2$, but $ba \notin L^2$. Hence L^2 is not dip-closed.

Catenation closure: Let $L = \{1, ab\}$. Then a.b, $a.ba.b \in L^*$, but $ba \notin L^*$. Hence L^* is not dip-closed.

Homomorphism: Let $L = a^*, \phi(a) = ab$ and $\phi(b) = b$. Then $\phi(L) = (ab)^*$ that is not dip-closed, because $ab, a.ba.b \in \phi(L)$ but $ba \notin \phi(L)$.

Intersection with regular languages: Let $L = \{1, a, b, ab\}$ and $R = \{1, b, ab\}$. Then L is dip-closed, R is regular and $L \cap R = R$ is not dip-closed. □

Proposition 3.3 *Let $u, v \in X^+, u \neq v$. Then there exists a dipolar-closed language L such that:*

(i) $u \in L, v \notin L$;

(ii) if L' is a dipolar-closed language such that $L \subseteq L'$ and $v \notin L'$, then $L' = L$.

Proof. The language $L_u = \{1, u\}$ is dip-closed and $v \notin L_u$.

Let $DP(L) = \{L_i | i \in I\}$ be the family of dip-closed languages L_i containing u with $v \notin L_i$. Let $\cdots \subseteq L_j \subseteq \cdots$, $j \in I$, be a chain of languages L_j with $L_j \in DP(L)$ and let $U = \cup_{j \in I} L_j$. If $rs, rxs \in U$, then $rs \in L_i$ and $rxs \in L_j$ where L_i and L_j are in the chain. Hence there exists a language L_k in the chain such that $L_i, L_j \subseteq L_k$ and $rs, rxs \in L_k$. Therefore $x \in L_k \subseteq U$ and U is dip-closed.

If $v \in U$, then $v \in L_j$ for some $j \in I$, a contradiction. Since the union of languages from any chain in $DP(L)$ is also an language in $DP(L)$, we can apply the Zorn's lemma. Therefore there exists a maximal dip-closed language, say L, such that $u \in L, v \notin L$ and this implies *(ii)*. □

Let $L \subseteq X^*$ and let $M(L) = \{x \in X^* | \ \exists u = x_1 v x_2 \in L, v \in X^*, x = x_1 x_2\}$. In other words, $M(L)$ contains words which are the catenation of a prefix and suffix of the same word in L. To the language L one can associate the set $dip(L)$ consisting of all words $x \in X^*$ with the following property: x is in $M(L)$ and the dipolar deletion of x from any word of L yields words belonging to L. (The condition $x \in M(L)$ has been added so that $dip(L)$ does not contain irrelevant words, such as words that cannot be deleted from any word of L.) Formally, $dip(L)$ is defined by:

$$dip(L) = \{x \in M(L)|\ \ u \in L, u = x_1 v x_2, x = x_1 x_2 \Longrightarrow v \in L\}.$$

Examples. Let $X = \{a, b\}$. Then $dip(X^*) = X^*$ and
– $dip(L_{ab}) = L_{ab}$.
– if $L = \{a^n b^n|\ n \geq 0\}$ then $dip(L) = L$.
– if $L = b^* a b^*$ then $dip(L) = b^*$.

Remark that, if $L \subseteq X^*$ then $x, y \in dip(L)$ and $xy \in M(L)$ imply $xy \in dip(L)$. In particular, if $M(L)$ is a submonoid of X^*, then $dip(L)$ is a submonoid of X^*.

In the following we show how, for a given language L, the set $dip(L)$ can be constructed. The construction involves the deletion operation which is, in some sense, inverse to the dipolar deletion operation.

Proposition 3.4 *Let $L \subseteq X^*$. Then $dip(L) = (L \longrightarrow L^c)^c \cap M(L)$.*

Proof. Let $x \in dip(L)$. From the definition of $dip(L)$ it follows that $x \in M(L)$. Assume now that $x \notin (L \longrightarrow L^c)^c$. This means there exist $u \in L$ such that $u = x_1 v x_2$, $x = x_1 x_2$ and $v \in L^c$. We arrived at a contradiction as $x \in dip(L)$, $x_1 v x_2 \in L$, $x = x_1 x_2$ but $v \notin L$.

For the other inclusion, let $x \in (L \longrightarrow L^c)^c \cap M(L)$. As $x \in M(L)$, if $x \notin dip(L)$ there exist $x_1 u x_2 \in L$ such that $x = x_1 x_2$ but $u \notin L$. This further implies that $x \in (L \longrightarrow L^c)$ - a contradiction with the initial assumption about x. □

Corollary 3.1 *If L is regular then $dip(L)$ is regular and can be effectively constructed.*

Proof. It follows from the fact that the family of regular languages is closed under complementation, intersection and deletion, the proofs are constructive (see [11], [9]) and, moreover the set $M(L)$ can be effectively constructed. □

Notice that a language $L \subseteq X^*$ is dip-closed iff $L \rightleftharpoons L \subseteq L$.

Proposition 3.5 *Let $L \subseteq X^*$ be an insertion-closed language. Then L is dipolar-closed if and only if $L = (L \rightleftharpoons L)$.*

Proof. Since L is dip-closed, $L \rightleftharpoons L \subseteq L$. Now let $u \in L$. Since L is ins-closed, $uu \in L$. Therefore, $u \in (L \rightleftharpoons L)$. We can conclude that $L = (L \rightleftharpoons L)$. The other implication is obvious. □

If L is a nonempty language, then the intersection of all the dip-closed languages containing L is a dip-closed language called the *dip-closure* of L. The dip-closure of L is the smallest dip-closed language containing L.

We will now define a sequence of languages whose union is the dipolar-closure of a given language L. Let $D_0(L) = L \cup \{1\}$, $D_{k+1}(L) = D_k(L) \rightleftharpoons D_k(L), k \geq 0$. Clearly $D_k(L) \subseteq D_{k+1}(L)$. Let

$$D(L) = \bigcup_{k \geq 0} D_k(L).$$

Proposition 3.6 *$D(L)$ is the dipolar-closure of the language L.*

Proof. Clearly $L \subseteq D(L)$. Let now $u_1u_2 \in D(L)$ and $u_1vu_2 \in D(L)$. Then $u_1u_2 \in D_i(L)$ and $u_1vu_2 \in D_j(L)$ for some integers $i, j \geq 0$. If $k = max\{i, j\}$, then $u_1u_2 \in D_k(L)$ and $u_1vu_2 \in D_k(L)$. This implies $v \in D_{k+1}(L) \subseteq D(L)$. Therefore $D(L)$ is a dip-closed language containing L.

Let T be a dip-closed language such that $L = D_0(L) \subseteq T$. Since T is dip-closed, if $D_k(L) \subseteq T$ then $D_{k+1}(L) \subseteq T$. By an induction argument, it follows that $D(L) \subseteq T$. □

Since, by [9], the family of regular languages is closed under dipolar deletion, it follows that if L is regular, then the languages $D_k(L)$, $k \geq 0$, are also regular. The following result shows that $D(L)$ is regular for any regular language L.

Recall that, when the principal congruence P_L of a language L has a finite index (finite number of classes), the language L is regular.

Proposition 3.7 *If $L \subseteq X^*$ is regular then its dipolar closure is regular.*

Proof. We show that if $u \equiv v(P_{D_k(L)})$ then $u \equiv v(P_{D_{k+1}(L)})$. Let $u \equiv v(P_{D_k(L)})$ and let $xuy \in D_{k+1}(L)$. Then, there exists a word $\alpha_1xuy\alpha_2 \in D_k(L)$ such that $\alpha_1\alpha_2 \in D_k(L)$. From the fact that $u \equiv v(P_{D_k(L)})$ and that $P_{D_k(L)}$ is a congruence, we deduce that $\alpha_1xuy\alpha_2 \equiv \alpha_1xvy\alpha_2(P_{D_k}(L))$. Since $D_k(L)$ is a union of classes of $P_{D_k(L)}$, it follows then that $\alpha_1xvy\alpha_2 \in D_k(L)$. This further implies that $xvy \in D_{k+1}(L)$. In the same way, $xvy \in D_{k+1}(L)$ implies $xuy \in D_{k+1}(L)$. Consequently, $u \equiv v(P_{D_{k+1}(L)})$ holds. This means that the number of congruence classes of $P_{D_{k+1}(L)}$ is smaller or equal to that of $P_{D_k(L)}$. Therefore, since the index of $P_{D_k}(L)$ is finite, there exists an integer t such that $P_{D_t}(L) = P_{D_{t+k}}(L), k \geq 1$. For every $i \geq 0$, $D_i(L) \subseteq D_{i+1}(L)$ and $D_i(L)$ is a union of classes of $P_{D_i}(L)$. Therefore $D(L) = D_t(L)$ and consequently, $D(L)$ is regular. □

References

[1] A.H. Clifford and G.B. Preston. The Algebraic Theory of Semigroups. *Math. Surveys Series, Vol. 1 and 2, Amer. Math. Soc.*, Providence (1961, 1967).

[2] R. Croisot. Equivalences principales bilatères définies dans un demi-groupe. *J. Math. Pures Appl.* (9), 36, 1957, pp.373-417.

[3] P. Dubreil. Contribution à la théorie des demi-groupes. *Mém. Acd. Sci. Inst. France.* (2) 63, 1941, 52 pp.

[4] M. Ito, L. Kari and G. Thierrin. Insertion and deletion closure of languages. To appear in *TCS*.

[5] H. Jurgensen, H.J. H.J. Shyr and G. Thierrin. Monoids with disjunctive identity. *Acta Mathematica Hungarica* 47(1986), pp.299-312.

[6] L. Kari. On insertion and deletion in formal languages. *Ph.D. Thesis*, University of Turku (1991).

[7] L.Kari. On language equations with invertible operations. *Theoretical Computer Science*, 132(1994), pp.129-150.

[8] L. Kari. Insertion and deletion of words: determinism and reversibility. *Lecture Notes in Computer Science* 629(1992), pp.315-327.

[9] L. Kari. Deletion operations: closure properties. *Intern.J.Computer Math.*, 52(1994), 23-42.

[10] G. Lallement. *Semigroups and Combinatorial Applications*. John Wiley, New York 1979.

[11] A. Salomaa. *Formal languages*. Academic Press, New York, 1973.

[12] H.J. Shyr. *Free monoids and languages.* Lecture Notes, Institute of Applied Mathematics, National Chung-Hsing University, Taichung, Taiwan (1991).