*International Journal of Computer Mathematics*
Vol. 88, No. 8, May 2011, 1619–1628

# Sub-classes of the monoid of left cancellative languages

Cao Chunhua[a], Yang Di[b,c]* and Liu Yun[d]

[a] *Department of Mathematics, Yunnan University, Yunnan, Kunming 650091, China;* [b] *The Faculty of Materials and Metallurgical Engineering, Kunming University of Science and Technology, Yunnan, Kunming 650031, China;* [c] *Department of Computer Science, Yunnan University of Finance and Economics, Yunnan, Kunming 650237, China;* [d] *Department of Mathematics, Yuxi Normal University, Yunnan, Kunming 653100, China*

(*Received 30 November 2007; resubmitted 20 January 2009; resubmitted 19 October 2009; resubmitted 6 February 2010; revised version received 28 March 2010; second revision received 14 May 2010; third revision received 8 August 2010; fourth revision received 24 August 2010; accepted 15 September 2010*)

A language $A$ is left cancellative if from $AB = AC$, it follows that $B = C$, for any two languages $B$ and $C$. Semi-singular and inf-singular languages are two disjoint sub-sets of left cancellative languages and are introduced by Hsieh and Shyr [*Left cancellative elements in the monoid of languages*, Soochow J. Math. 4 (1978), pp. 7–15]. In this paper, we further study them. It is shown that all non-dense and all maximal left cancellative languages are semi-singular while all right dense left cancellative languages are inf-singular. Finally, a theorem shows that there is a left cancellative language which is neither semi-singular nor inf-singular.

**Keywords:** left cancellative language; semi-singular language; inf-singular language; dense language; maximal left cancellative language

*2000 AMS Subject Classifications*: 68Q70; 68Q45; 94A45

## 1. Introduction

Prefix codes are widely used in information theory and computer science, for example, in encoding and decoding, data compression and transmission, DES and Huffman's algorithms [4,5,7–9]. Left cancellative languages are a kind of generalization of prefix codes. For the properties of left cancellative languages, see [3,10–12]. Especially in [12], maximal left cancellative languages are studied. From [10], we know that left singular languages are a kind of left cancellative languages and they are studied in [2,6,10]. In this paper, we find that maximal left cancellative languages and left singular languages have some characteristics in common. They are all semi-singular languages. In fact, the notions of semi-singular and inf-singular languages are discussed and introduced in [3]. Based on [3], we make a further study on semi-singular and inf-singular languages.

The paper is organized as follows. Section 2 gives some definitions and properties used in the paper. To investigate semi-singular and inf-singular languages, we propose some properties of

*Corresponding author. Email: yangdi65@yahoo.cn

ISSN 0020-7160 print/ISSN 1029-0265 online

DOI: 10.1080/00207160.2010.526705
http://www.informaworld.com

semi-singular and inf-singular words in Section 3. In Section 4, some results on semi-singular and inf-singular languages are proved. First, we prove that the set of semi-singular languages is a strongly prefix sub-semi-group and the set of inf-singular languages is a left ideal of the monoid of left cancellative languages which generalize the result that the set of semi-singular languages is a sub-semi-group of left cancellative languages in [3]. Then, we prove that all non-dense and all maximal left cancellative languages are semi-singular while all right dense left cancellative languages are inf-singular and all inf-singular languages are dense. At last, a theorem is given to show that there is a left cancellative language which is neither semi-singular nor inf-singular. So the monoid of left cancellative languages is the union of three disjoint sub-classes of semi-singular languages, inf-singular languages and left cancellative languages which are neither semi-singular nor inf-singular.

## 2. Definitions and preliminaries

Let $X$ be a non-empty finite set of letters. Any finite string over $X$ is called *a word*. For example, $w = abab^2a$ is a word over $\{a, b\}$. The word that contains no letter is called *the empty word*, denoted by 1. The set of all words is denoted by $X^*$, which is a free monoid with concatenation. For example, the production of two words $x = ab^2$ and $y = ab^3a$ is the word $xy = ab^2ab^3a$. For any word $w$ in $X^*$, let $lg(w)$ be the number of letters that occur in $w$. Then $lg(w) = 6$ for the former $w = abab^2a$. Let $X^+ = X^* \setminus \{1\}$. Any non-empty sub-set of $X^+$ is called *a language*. The set $M = \{A \mid A \subseteq X^+$ or $A = \{1\}\}$ with concatenation is the monoid of languages. And $D(M) = \{A \in M \mid AB = AC$ implies $B = C$ for all $B, C \in M\}$ is the monoid of left cancellative languages. An element in $D(M)$ is called *a left cancellative language*.

Let $A$ be a language and $Z_A = A \setminus AX^+$. For example, let $X = \{a, b\}$ and $A = \{a, b, a^3, b^3, aba, bab\}$. Then $Z_A = \{a, b\}$. For every language $A$, we can see $A \subseteq Z_A \cup Z_A X^+$. A language $A$ is called *a prefix code* if $Z_A = A$ (see [1,11]). A prefix code $A$ is called *a maximal prefix code* if $A \cup \{x\}$ is not a prefix code for all $x \in X^+ \setminus A$ (see [1,11]).

In [3], semi-singular and inf-singular languages are defined as follows. Let $A$ be a language. For $v \in Z_A$ and $x \in X^+$, the word $vx$ is called *A-semi-singular* if $vxr = yz$ for some $y \in A$ and $r, z \in X^*$, then $v = y$. Let $S_A = \{x \in X^+ \mid vx$ is $A$-semi-singular for some $v \in Z_A\}$ and $G_A = S_A \setminus S_A X^+$. A language $A$ is called *semi-singular* if $G_A$ is a maximal prefix code. The set of all semi-singular languages is denoted by $S(M)$. For example, let $X = \{a, b\}$ and $A = \{a, b, a^2, b^2\}$. From the definition of $S_A$, we can see $S_A = aX^* \cup bX^*$ and hence $G_A = \{a, b\}$ is a maximal prefix code. So $A \in S(M)$.

Let $X_x^+ = X^+ \setminus \{x\}$ and $L_A = \{x \in X^+ \mid vx \notin AX_x^+$ for some $v \in A\}$. A word $x \in X^+$ is called *A-inf-singular* if the following two conditions hold:

(i) $xX^* \subseteq L_A$;
(ii) for every $g \in G_A$ and $m \in X^*$, $g \neq xm$ and $x \neq gm$.

Let $I_A = \{x \in X^+ \mid x$ is $A$-inf-singular$\}$ and $H_A = I_A \setminus I_A X^+$. A language $A$ is called *inf-singular* if $H_A$ is a maximal prefix code. The set of all inf-singular languages is denoted by $I(M)$. In the following, we review some results which will be used in the rest of the paper.

LEMMA 2.1 [3] *A language A is a left cancellative language if and only if $G_A \cup H_A$ is a maximal prefix code.*

LEMMA 2.2 [3] $S(M) \subseteq D(M)$.

LEMMA 2.3 [3] $S_A \cap I_A = G_A \cap H_A = \emptyset$.

LEMMA 2.4 [3] $S_A = \{x \in X^+ | vx \notin AX_x^+$ *and* $vxX^* \cap A = \emptyset$ *for some* $v \in A\}$.

By the definition of $L_A = \{x \in X^+ | vx \notin AX_x^+$ for some $v \in A\}$, we have $S_A \subseteq L_A$.

LEMMA 2.5 [3] *If $A$ is a left singular language, then* $S_A = X^+$.

All left singular languages are in $S(M)$ because $G_A = X$ is a maximal prefix code by Lemma 2.5.

LEMMA 2.6 [3] *Let $A$ be a bounded language. Then $A \in D(M)$ if and only if $G_A$ is a maximal prefix code.*

All finite left cancellative languages are in $S(M)$ (see [11]), and if $A \in I(M)$ then $A$ is unbounded and infinite. The following lemma is an example of an inf-singular language.

LEMMA 2.7 [3] *Let* $X = \{a, b\}$ *and* $B = b^+a \cup (\bigcup_{i=1}^{\infty} b^i a X^i)$. *Then* $B \in I(M)$.

LEMMA 2.8 [3] *$S(M)$ is a sub-semi-group of $D(M)$.*

## 3. Properties of $S_A$, $L_A$ and $I_A$

The sets $S_A$, $L_A$ and $I_A$ are introduced by the definitions of semi-singular and inf-singular languages. Before we show some results on $S(M)$ and $I(M)$, we give some properties of these three sets which are often used in the later of the article. First, we cite a property of left cancellative languages which we need.

LEMMA 3.1 [3,10] *A language $A \in D(M)$ if and only if $AX^+ \neq AX_x^+$ (or $Z_A X^+ \neq AX_x^+$) for all $x \in X^+$.*

PROPOSITION 3.2 *A language $A$ is left cancellative if and only if $L_A = X^+$.*

*Proof* ($\Rightarrow$) Let $A$ be a left cancellative language. Then $AX^+ \neq AX_x^+$ for every $x \in X^+$ by Lemma 3.1. So for every $x \in X^+$, there exists $p \in A$ such that $px \notin AX_x^+$. Then $x \in L_A$ for every $x \in X^+$. So $X^+ \subseteq L_A$. On the other hand, $L_A \subseteq X^+$ by the definition of $L_A$. From above, we know $L_A = X^+$.

($\Leftarrow$) Let $L_A = X^+$. Then $x \in L_A$ for every $x \in X^+$. So for every $x \in X^+$ there exists $p \in A$ such that $px \notin AX_x^+$. Then $AX^+ \neq AX_x^+$ for every $x \in X^+$. So $A \in D(M)$ by Lemma 3.1. ■

We can now prove the following proposition.

PROPOSITION 3.3 *Let $A$ be a language. Then $A \in I(M)$ if and only if $I_A = X^+$.*

*Proof* ($\Rightarrow$) Let $A \in I(M)$. Then $H_A$ is a maximal prefix code. Suppose that $G_A \neq \emptyset$. Then there exists $x \in X^+$ such that $x \in G_A$. So $x \notin H_A$ by Lemma 2.3. Then $H_A \cup \{x\}$ is not a prefix code. Since $H_A \cup \{x\} \subseteq H_A \cup G_A$, $H_A \cup G_A$ is not a prefix code. This contradicts with Theorem 8 in [3]. So $G_A = \emptyset$. Then $H_A \cup G_A = H_A$ is a maximal prefix code. Then $A \in D(M)$. So $L_A = X^+$ by Proposition 3.2. For any $x \in X^+$, we have $xX^* \subseteq X^+ = L_A$. Since $G_A = \emptyset$, we have $x \in I_A$ for all $x \in X^+$. So $I_A = X^+$.

($\Leftarrow$) Let $I_A = X^+$. Then $x \in I_A$ for all $x \in X^+$. So $xX^* \subseteq L_A$ by the definition of $I_A$. Then $x \in L_A$ for all $x \in X^+$, which implies that $X^+ \subseteq L_A$. Then $A$ is a left cancellative language. By Lemma 2.1, we know that $G_A \cup H_A$ is a maximal prefix code. Since $I_A \cap S_A = \emptyset$ by Lemma 2.3 and $I_A = X^+$, then $S_A = \emptyset$. So $G_A = \emptyset$. Thus, $H_A$ is a maximal prefix code. So $A \in I(M)$. ■

Note:

(1) $I(M) \subseteq D(M)$.
(2) Let $A$ be any non-empty language. Then
 (i) $A \in S(M) \Leftrightarrow G_A$ is a maximal prefix code $\Leftrightarrow A \in D(M)$ and $H_A = \emptyset \Leftrightarrow A \in D(M)$ and $I_A = \emptyset$;
 (ii) $A \in I(M) \Leftrightarrow H_A$ is a maximal prefix code $\Leftrightarrow A \in D(M)$ and $G_A = \emptyset \Leftrightarrow A \in D(M)$ and $S_A = \emptyset$.

Let $A$ be a language and $l(A) = \{g \in A | gx \notin A$ for all $x \in X^+$ and $g = yz$ for some $z \in X^+$ implies $y \notin A\}$. If $l(A) \neq \emptyset$, then $A$ is called *a left singular language* [6,10,11]. By Lemma 2.5, we know if $A$ is a left singular language, then $S_A = X^+$. It is natural to ask whether or not $S_A = X^+$ for all $A \in S(M)$. The following example shows that there is a language $A \in S(M)$, but $S_A \neq X^+$.

*Example 3.4* Let $X = \{a, b\}$ and $A = \{a, b, a^3, b^3, aba, bab\}$. Then $Z_A = \{a, b\}$. For all $x \in X^+$, if $x \in XbX^*$, then $ax \notin AX_x^+$; if $x \notin XbX^*$, then $bx \notin AX_x^+$. Hence, $A$ is a left cancellative language by Lemma 3.1. Next, we will prove that $x \in S_A$ for all $x \in X^+ \setminus X$.

(1) If $x \in XbX^*$, then $axr = yz$ for some $y \in A,\ r, z \in X^*$ implies $y = a$.
(2) If $x \in XaX^*$, then $bxr = yz$ for some $y \in A,\ r, z \in X^*$ implies $y = b$.
(3) If $x \in X$, then $axa \in A$. So $x \notin S_A$ for all $x \in X$.

Therefore, $S_A = X^+ \setminus X = XX^+ \neq X^+$. So $G_A = X^2$ and $I_A = H_A = \emptyset$. Thus, $A \in S(M)$.

In the following, we discuss the relation between $S_{AB}$ and $S_B$, $I_{AB}$ and $I_B$, $L_{AB}$ and $L_B$ for two languages $A$ and $B$.

LEMMA 3.5 [3] *Let $A$ and $B$ be languages. Then $S_{AB} \subseteq S_B$.*

Before the relation between inf-singular words in $AB$ and $B$ is discussed, we propose an equivalent definition of inf-singular words once the following lemma is proved.

LEMMA 3.6 *Let $A$ be a language and $x \in X^+$. Then the following are equivalent*:

(1) *for every $g \in G_A$ and $m \in X^*$, $g \neq xm$ and $x \neq gm$;*
(2) *for every $g \in S_A$ and $m \in X^*$, $g \neq xm$ and $x \neq gm$.*

*Proof* ((2) $\Rightarrow$ (1)) It is obvious, since $G_A \subseteq S_A$.

((1) $\Rightarrow$ (2)) For every $g' \in G_A$ and $m \in X^*$, $g' \neq xm$ and $x \neq g'm$. Assume there exist $g \in S_A$ and $m \in X^*$ such that $g = xm$ or $x = gm$. Since $g \in S_A$ and $G_A = S_A \setminus S_A X^+$, then $g \in G_A$ or $g \in G_A X^+$. If $g \in G_A$, then $g \neq xm$ and $x \neq gm$ by Equation (1), which is a contradiction with the assumption that $g = xm$ or $x = gm$. If $g \in G_A X^+$, then $g = g_1 y$ for some $g_1 \in G_A$ and $y \in X^+$. When $g = xm$, we have $g_1 y = xm$. Now, we consider the following two cases. (a) If $lg(g_1) \geq lg(x)$, then $g_1 = xm_1$, where $m_1 \in X^*$ is a prefix of $m$. This is a contradiction for $g_1 \in G_A$. (b) If $lg(g_1) < lg(x)$, then $x = g_1 y_1$, where $y_1$ is a prefix of $y$. This is also a contradiction for $g_1 \in G_A$. When $x = gm$ and $g = g_1 y$, we have $x = (g_1 y)m = g_1(y_1 m)$. Then it is a contradiction for $g_1 \in G_A$. So Equation (1) implies Equation (2). ■

We can give now the alternative definition of inf-singular words through the form of a corollary.

Corollary 3.7 *Let A be a language. A word $x \in X^+$ is A-inf-singular if and only if the following two conditions hold*:

(i) $xX^* \subseteq L_A$;
(ii) *for every $g \in S_A$ and $m \in X^*$, $g \neq xm$ and $x \neq gm$.*

Proposition 3.8 *Let A and B be left cancellative languages. Then $I_B \subseteq I_{AB}$.*

*Proof* For any $u \in I_B$, we have $uX^* \subseteq L_B$. Since $A, B \in D(M)$, $AB \in D(M)$ (see [10]). Then $L_A = L_B = L_{AB} = X^+$ by Proposition 3.2. So $uX^* \subseteq L_{AB}$. From $u \in I_B$, we also have $s \neq um$ and $u \neq sm$ for all $s \in S_B$ and $m \in X^*$. Since $S_{AB} \subseteq S_B$, we get $s \neq um$ and $u \neq sm$ for all $s \in S_{AB}$ and $m \in X^*$. Then $u \in I_{AB}$ by Corollary 3.7. Thus $I_B \subseteq I_{AB}$. ■

Proposition 3.9 *Let A and B be languages. Then $L_{AB} \subseteq L_B$.*

*Proof* Let $x \notin L_B$. Then $qx \in BX_x^+$ for all $q \in B$. So there exist $y_q \in B$ and $z_q \in X_x^+$ such that $qx = y_q z_q$. For all $p \in A, q \in B$ and $r = pq \in AB$, we have $rx = pqx = py_q z_q$, where $py_q \in AB$ and $z_q \in X_x^+$. That is to say $rx \in ABX_x^+$. So $x \notin L_{AB}$. Thus, $L_{AB} \subseteq L_B$. ■

If $L_A = L_B$ and $S_A = S_B$, then $I_A = I_B$ by the definitions of $I_A$ and $I_B$. For $A, B \in D(M)$, we have $L_A = L_B = X^+$. Thus, if $S_A = S_B$ then $I_A = I_B$ for all $A, B \in D(M)$. But the converse is not true. That is, $I_A = I_B$ cannot imply $S_A = S_B$. For example, let $X = \{a, b\}$ and $A = \{a, b, a^3, b^3, aba, bab\}$. So $S_A = X^+ \setminus X = XX^+$ and $I_A = \emptyset$ by Example 3.4. Let $B$ be a left singular language. Then $S_B = X^+$ and $I_B = \emptyset$ by Lemmas 2.5 and 2.3. Thus, we have $I_A = I_B$, but $S_A \neq S_B$.

In general, $S_{AB} \subseteq S_B,\ I_B \subseteq I_{AB},\ L_{AB} \subseteq L_B$. When do the equations hold?

Theorem 3.10 *Let A and B be languages.*

(1) *If $L_A = X^+$, then $L_{AB} = L_B$.*
(2) *If $S_A = X^+$, then $S_{AB} = S_B$ and $I_{AB} = I_B$.*

*Proof* (1) By Proposition 3.9, we have $L_{AB} \subseteq L_B$ for all $A, B \in M$. In the following, we prove if $L_A = X^+$ then $L_B \subseteq L_{AB}$. We prove it by contradiction of the dual relation. Let $x \notin L_{AB}$. We will show that $x \notin L_B$. Since $L_A = X^+$, we have $qx \in X^+ = L_A$ for all $q \in B$. By $qx \in L_A$, there exists $p \in A$ such that $pqx \notin AX_{qx}^+$. Since $x \notin L_{AB}$, then $pqx \in ABX_x^+$ for the former $p \in A$ and $q \in B$. Then there exist $u \in A$, $v \in B$ and $w \in X_x^+$ such that $pqx = uvw$. By $pqx \notin AX_{qx}^+$, we have $p = u$ and $qx = vw$. Hence, $qx \in BX_x^+$ for all $q \in B$. So $x \notin L_B$ by definition. Thus, $L_B \subseteq L_{AB}$.

(2) By Lemma 3.5, we have $S_{AB} \subseteq S_B$ for all $A, B \in M$. Next we prove if $S_A = X^+$ then $S_B \subseteq S_{AB}$. Let $x \notin S_{AB}$. Since $S_A = X^+$, we have $qx \in X^+ = S_A$ for all $q \in B$. Then there exists $p \in A$ such that $pqx \notin AX_{qx}^+$ and $pqxX^* \cap A = \emptyset$. Since $x \notin S_{AB}$, we have $pqx \in ABX_x^+$ or $pqxX^* \cap AB \neq \emptyset$ by Lemma 2.4.

(i) If $pqx \in ABX_x^+$, then there exist $p_A \in A$, $q_B \in B$ and $u \in X_x^+$ such that $pqx = p_A q_B u$. If $p_A \neq p$, then $q_B u \neq qx$. So $pqx \in AX_{qx}^+$. This is a contradiction. So $p_A = p$. Then $qx = q_B u \in BX_x^+$. Thus, $x \notin S_B$.

(ii) If $pqxX^* \cap AB \neq \emptyset$, then there exist $r \in X^*$, $p'_A \in A$ and $q'_B \in B$ such that $pqxr = p'_A q'_B$. When $lg(pqx) \leq lg(p'_A)$, there exists $r_1 \in X^*$ such that $pqxr_1 = p'_A$. So $pqxX^* \cap A \neq \emptyset$. This is a contradiction. When $lg(pqx) > lg(p'_A)$, there exists $u \in X^+$ such that $pqx = p'_A u$. Since $pqx \notin$

$AX^+_{qx}$, we have $p = p'_A$. Then $qxr = q'_B$. So $qxrX^* \cap B \neq \emptyset$. Then $qxX^* \cap B \neq \emptyset$. Therefore, $x \notin S_B$.

From (i) and (ii) together, we can obtain if $x \notin S_{AB}$, then $x \notin S_B$. So $S_B \subseteq S_{AB}$ when $S_A = X^+$.

Since $S_A \subseteq L_A$ and $S_A = X^+$, we have $L_A = X^+$. By part (1) of this theorem, we get $L_{AB} = L_B$. Thus, $I_{AB} = I_B$ when $S_A = X^+$. ■

## 4. Semi-singular and inf-singular languages

By [3,10], we know $D(M)$ is a sub-semi-group of $M$ and $S(M)$ is a sub-semi-group of $D(M)$. We now show they are all strongly prefix sub-semi-groups. A sub-semi-group $T$ of a semi-group $S$ is called *a strongly prefix sub-semi-group* if for all $x, y \in S$, $xy \in T$ implies $y \in T$ (see [11]).

PROPOSITION 4.1 *$D(M)$ is a strongly prefix sub-semi-group of $M$.*

*Proof* Let $A$ and $B$ be languages and $AB$ be left cancellative language. Then $L_{AB} = X^+$ by Proposition 3.2. We know $L_{AB} \subseteq L_B$ by Proposition 3.9. So $L_B = X^+$. Then $B \in D(M)$. Thus, $D(M)$ is a strongly prefix sub-semi-group of $M$. ■

PROPOSITION 4.2 *$S(M)$ is a strongly prefix sub-semi-group of $D(M)$.*

*Proof* Let $A, B \in D(M)$ and $AB \in S(M)$. Then $I_{AB} = \emptyset$. By Proposition 3.8, we have $I_B \subseteq I_{AB}$. Then, $I_B = \emptyset$. So $B \in S(M)$. Thus, $S(M)$ is a strongly prefix sub-semi-group of $D(M)$. ■

The set of all semi-singular languages is a strongly prefix sub-semi-group of the monoid of left cancellative languages which is a generalization of Lemma 2.8. Then, how about the set of all inf-singular languages?

PROPOSITION 4.3 *$I(M)$ is a left ideal of $D(M)$.*

*Proof* For all $A \in D(M)$ and $B \in I(M)$, we have $S_B = \emptyset$. By Lemma 3.5, we have $S_{AB} \subseteq S_B$. Then $S_{AB} = \emptyset$. So $AB \in I(M)$. Thus, $I(M)$ is a left ideal of $D(M)$. ■

All left singular languages and all finite left cancellative languages are in $S(M)$. We will give a semi-singular language which is neither left singular nor finite.

*Example 4.4* Let $X = \{a, b\}$ and $A = X \cup a^2X^* \cup b^2X^*$. Then $Z_A = X$. For any $x \in X^+$, if $x \in aX^*$, then $bx \notin AX^+_x$ and $bxX^* \cap A = \emptyset$. So $bx$ is a $A$-semi-singular word. If $x \in bX^*$ then $ax \notin AX^+_x$ and $axX^* \cap A = \emptyset$. So $ax$ is a $A$-semi-singular word. Hence $S_A = X^+$. Then $A \in S(M)$. Clearly, $A$ is neither left singular nor finite.

Theorems 4.5 and 4.8 will give another two kinds of languages which are contained in $S(M)$. First, we cite some definitions from [11,13] which we need in the following. A language $A$ is said to be *right dense* if $wX^* \cap A \neq \emptyset$ for all $w \in X^*$. If $X^*wX^* \cap A \neq \emptyset$ for all $w \in X^*$, then $A$ is called *dense*. If $X^*wX^* \cap A = \emptyset$ for some $w \in X^+$, then $A$ is called *non-dense*. If $wX^* \cap A = \emptyset$ for some $w \in X^+$, then $A$ is called *non-right dense*.

THEOREM 4.5 *All non-dense left cancellative languages are in $S(M)$.*

*Proof* Let $A$ be a non-dense left cancellative language. Then there exists $w \in X^+$ such that $X^*wX^* \cap A = \emptyset$. Hence, $X^*uX^* \cap A = \emptyset$ for all $u \in X^*wX^*$. So for all $p \in A$, we have $puX^* \cap A = \emptyset$. Since $A \in D(M)$, then $AX^+ \neq AX_u^+$ for all $u \in X^+$ by Lemma 3.1. Then there exists $q \in A$ such that $qu \notin AX_u^+$. And $quX^* \cap A = \emptyset$. Hence, $u \in S_A$. So $X^*wX^* \subseteq S_A$. Then for all $x \in X^*$, we have $xw \in X^*wX^* \subseteq S_A$. Then $x \notin I_A$ by Corollary 3.7. Thus, $I_A = \emptyset$ and $A \in S(M)$. ■

The language $A$ in Example 3.4 is non-dense and $S_A \neq X^+$.

PROPOSITION 4.6 *Let $A$ be a language. If $S_A = X^+$, then $A$ is a non-right dense left cancellative language.*

*Proof* Let $S_A = X^+$. Then $x \in S_A$ for every $x \in X^+$. So for every $x \in X^+$ there exists $p \in A$ such that $px \notin AX_x^+$ and $pxX^* \cap A = \emptyset$ by Lemma 2.4. From $px \notin AX_x^+$, we know $AX^+ \neq AX_x^+$ for every $x \in X^+$. Then $A$ is a left cancellative language by Lemma 3.1. From $(px)X^* \cap A = \emptyset$, we know that $A$ is non-right dense. Thus, $A$ is a non-right dense left cancellative language. ■

By Lemma 2.5 and the proposition, we know that any left singular language is a non-right dense left cancellative language.

A left cancellative language $A$ is called *a maximal left cancellative language* if $A \cup \{x\}$ is not a left cancellative language for all $x \in X^+ \backslash A$ (see [12]). We will show that all maximal left cancellative languages are in $S(M)$. First, we cite a lemma which we need.

LEMMA 4.7 [12] *Let $A$ be a maximal left cancellative language. If $pw \in A$ for some $p \in Z_A$ and $w \in X^+$, then $pwX^* \subseteq A$.*

THEOREM 4.8 *All maximal left cancellative languages are in $S(M)$.*

*Proof* Let $A$ be a maximal left cancellative language. If for every $w \in X^+$ there exists $r \in X^*$ such that $wr \in S_A$, then $g = wr$ for some $g \in S_A$. So $w \notin I_A$ by Corollary 3.7. As $w$ was chosen arbitrarily, $I_A = \emptyset$. Thus $A \in S(M)$.

Next, we will prove for all $w \in X^+$ there exists $r \in X^*$ such that $wr \in S_A$.

First, we want to prove that for every $w \in X^+$, there exists $p_0 \in Z_A$ such that $p_0w \notin A$. Otherwise, assume that there exists $w' \in X^+$ such that $pw' \in A$ for all $p \in Z_A$. Then for all $q \in X^+$, we have $p(w'q) = (pw')q \in AX_{w'q}^+$. So $Z_AX^+ = AX_{w'q}^+$ for some $w'q \in X^+$. Then $A \notin D(M)$ by Lemma 3.1. This is a contradiction. So for every $w \in X^+$ there exists $p_0 \in Z_A$ such that $p_0w \notin A$. Then, we want to prove there exists $r \in X^*$ such that $p_0wrX^* \cap A = \emptyset$ and $p_0wr \notin AX_{wr}^+$.

(i) Assume that $p_0wrX^* \cap A \neq \emptyset$ for all $r \in X^*$. Let $B = A \cup \{p_0w\}$. First, since $p_0w \in AX^+ \subseteq BX^+$, then $p_0w \notin Z_B$. So $Z_A = Z_B$. Then, since $p_0w \notin A$ and $A$ is a maximal left cancellative language, $B \notin D(M)$. So there exists $u \in X^+$ such that $Z_BX^+ = BX_u^+$ by Lemma 3.1. Then $Z_AX^+ = Z_BX^+ = BX_u^+ = (A \cup \{p_0w\})X_u^+$. That is, $pu \in AX_u^+$ or $\{p_0w\}X_u^+$ for all $p \in Z_A$. Suppose that $p \in Z_A \setminus \{p_0\}$ and $pu \in \{p_0w\}X_u^+$. Then $pu = p_0wy_0$ for some $y_0 \in X_u^+$. Since $p, p_0 \in Z_A$, we have $p = p_0$. This is a contradiction. Thus,

$$pu \in AX_u^+ \text{ for all } p \in Z_A \setminus \{p_0\}.$$

For the word $p_0u$, it is in $AX_u^+$ or $\{p_0w\}X_u^+$. If $p_0u \in AX_u^+$, then $Z_AX^+ = AX_u^+$. Then $A \notin D(M)$. This is a contradiction. If $p_0u \in \{p_0w\}X_u^+$, then $p_0u = p_0wy_1$, where $y_1 \in X^+ \setminus \{u\}$. By assumption that $p_0wrX^* \cap A \neq \emptyset$ for all $r \in X^*$, we have $p_0wy_1X^* \cap A \neq \emptyset$ for the word $y_1$. Then there exists $s \in X^*$ such that $p_0wy_1s \in A$. Since $p_0u = p_0wy_1$, $p_0us = p_0wy_1s \in A$. For all

$t \in X^+$, we have $p_0(ust) = (p_0wy_1s)t \in AX^+_{ust}$. Since $pu \in AX^+_u$ for all $p \in Z_A \setminus \{p_0\}$, $pu = x_2y_2$ for some $x_2 \in A$ and $y_2 \neq u$. So $p(ust) = x_2(y_2st) \in AX^+_{ust}$ for all $p \in Z_A \setminus \{p_0\}$. Thus, there exists $ust \in X^+$ such that $Z_AX^+ = AX^+_{ust}$. Then $A \notin D(M)$. This is a contradiction. So there exists $r' \in X^*$ such that $p_0wr'X^* \cap A = \emptyset$.

(ii) Assume that $p_0wr' \in AX^+_{wr'}$. Then there exist $x_3 \in A$ and $y_3 \in X^+_{wr'}$ such that $p_0wr' = x_3y_3$ and $p_0 \neq x_3$. Since $p_0 \in Z_A$, $p_0w_1 = x_3 \in A$ or $p_0wr_1 = x_3 \in A$, where $w = w_1w_2,\ r' = r_1r_2$ and $w_1, r_1 \in X^+$, $w_2, r_2 \in X^*$. By Lemma 4.7, we have $p_0w \in A$ or $p_0wr' \in A$. If $p_0w \in A$, then we have a contradiction. If $p_0wr' \in A$, then $p_0wr'X^* \cap A \neq \emptyset$. This contradicts with (i). So $p_0wr' \notin AX^+_{wr'}$.

From (i) and (ii), we have $wr' \in S_A$. ■

In fact, the language $A = X \cup a^2X^* \cup b^2X^*$, where $X = \{a, b\}$ in Example 4.4 is a dense maximal left cancellative language. So it is a dense semi-singular language. First, $A$ is dense because for all $x \in X^+$, if $x \in aX^*$ then $ax \in A$; if $x \in bX^*$ then $bx \in A$. Second, we know that $A$ is left cancellative by Example 4.4. Finally, we will prove that $A$ is maximal. We can see $X^+ \setminus A = abX^* \cup baX^*$. Without loss of generality, for every $abx_1 \in abX^*$, we prove that $B = A \cup \{abx_1\}$ is not a left cancellative language. We can find a word $u = bx_1a$ such that $au = a(bx_1a) = (abx_1)a$; $bu = b(bx_1a) = (b^2x_1)a$; $(a^2x_2)u = (a^2x_2)(bx_1a) = (a^2x_2bx_1)a$; and $(b^2x_2)u = (b^2x_2)(bx_1a) = (b^2x_2bx_1)a$ for all $x_2 \in X^*$. Then $BX^+ = BX^+_u$. So $B$ is not left cancellative. Thus, $A$ is a maximal left cancellative language.

Theorem 4.5 tells us all non-dense left cancellative languages are semi-singular while the following theorem will show that all right dense left cancellative languages are inf-singular. Of course, not all dense left cancellative are inf-singular.

THEOREM 4.9 *All right dense left cancellative languages are in $I(M)$.*

*Proof* Let $A$ be a right dense left cancellative language. Then $xX^* \cap A \neq \emptyset$ for all $x \in X^+$. So $(px)X^* \cap A \neq \emptyset$ for all $p \in A$. Then $px$ is not an $A$-semi-singular word. So $x \notin S_A$ for all $x \in X^+$. Then $S_A = \emptyset$. Thus, $A \in I(M)$. ■

In the following example, we will construct a right dense left cancellative language. So it is an inf-singular language by the above theorem.

*Example 4.10* Let $X = \{a, b\}$ and $A = b^+a \cup (\bigcup_{i=0}^{\infty} b^iaX^iX^*)$. For any $x \in X^+$, assume $lg(x) = m$. Then $b^max \in AX^+ \setminus AX^+_x$. Hence, $A \in D(M)$. Next, we will prove that $A$ is right dense. For every $x \in X^+$ and $lg(x) = m$, if $x = b^m$ then $xa = b^ma \in A$. So $xX^* \cap A \neq \emptyset$. If $x = b^iax_1$, where $0 \leq i < m$ and $lg(x_1) = m - i - 1$, then we consider the following two cases.

(1) When $lg(x_1) = m - i - 1 \geq i$, we know that $x = b^iax_1 \in A$ by the construction of $A$. So $xX^* \cap A \neq \emptyset$.
(2) When $lg(x_1) = m - i - 1 < i$, for every $x_2 \in X^+$ and $lg(x_2) = 2i - m + 1$, we have $xx_2 = b^ia(x_1x_2) \in A$ by the construction of $A$. So $xX^* \cap A \neq \emptyset$.

From all above and $1 \cdot X^* \cap A = X^* \cap A = A \neq \emptyset$, we show that $A$ is a right dense left cancellative language. So $A \in I(M)$.

On the other hand, the following example will show that there is a left cancellative language in $I(M)$ which is not right dense but is dense.

*Example 4.11* Let $X = \{a, b\}$ and $B = b^+a \cup (\bigcup_{i=1}^{\infty} b^iaX^i)$. Then $B \in I(M)$ by Lemma 2.7. There is a word $a \in X^+$ such that $aX^* \cap B = \emptyset$. So $B$ is not right dense. Next, we will prove that

$B$ is dense. For every $x \in X^*$, we know that $x = 1$ or $x \in aX^*$ or $x \in bX^*$. When $x = 1$, we have $(ba)x = ba \in B$. So $X^*xX^* \cap B \neq \emptyset$. When $x \neq 1$, let $lg(x) = m$ where $m \geq 1$. If $x \in aX^*$, then there exists $x_1 \in X^*$ such that $x = ax_1$ where $lg(x_1) = m - 1 \geq 0$. So $b^m xb = b^m(ax_1)b = b^m a(x_1 b) \in b^m aX^m \subseteq B$. So $X^*xX^* \cap B \neq \emptyset$. If $x \in bX^*$ and $x = b^m$, then $xa = b^m a \in B$. So $X^*xX^* \cap B \neq \emptyset$. If $x = b^i ax_1$ where $1 \leq i < m$ and $lg(x_1) = m - i - 1$, then we consider the following three cases.

(1) When $lg(x_1) = m - i - 1 > i$, we know that $b^{m-2i-1}x = b^{m-2i-1}(b^i ax_1) = b^{m-i-1}ax_1 \in B$. So $X^*xX^* \cap B \neq \emptyset$.
(2) When $lg(x_1) = m - i - 1 = i$, we know that $x = b^i ax_1 \in B$. So $X^*xX^* \cap B \neq \emptyset$.
(3) When $lg(x_1) = m - i - 1 < i$, for every $x_2 \in X^+$ and $lg(x_2) = 2i - m + 1$, we have $xx_2 = b^i a(x_1 x_2) \in B$. So $X^*xX^* \cap B \neq \emptyset$. Thus, $B$ is dense.

In fact, the following theorem will tell us that all inf-singular languages are dense.

THEOREM 4.12 *All inf-singular languages are dense.*

*Proof* Let $A$ be an inf-singular language. Then $H_A$ is a maximal prefix code. Since $I(M) \subseteq D(M)$, then $A \in D(M)$. Then $G_A \cup H_A$ is a maximal prefix code. Thus, $G_A = \emptyset$ for $H_A$ is already a maximal prefix code. Then $S_A = \emptyset$. So for every $x \in X^+$, we have $x \notin S_A$. Then for all $p \in A$, one of the following holds: (i) $px \in AX_x^+$ or (ii) $pxX^* \cap A \neq \emptyset$. Since $A \in D(M)$, then $AX^+ \neq AX_x^+$ for all $x \in X^+$. Thus, there exists $q \in A$ such that $qx \notin AX_x^+$. Then we have $qxX^* \cap A \neq \emptyset$. As $x$ was chosen arbitrarily, $A$ is dense. ■

By Propositions 4.2 and 4.3, we know that $I(M)$ is a left ideal and $S(M)$ is a strongly prefix sub-semi-group of $D(M)$. Next we will show that: (1) $I(M)$ is not a strongly prefix sub-semi-group; (2) $S(M)$ is not a left ideal; (3) $S(M)$ and $I(M)$ are all not right ideals of $D(M)$. Let $A$ be a right dense left cancellative language over an alphabet $X$ and $B = \{a\}$, where $a \in X$. Then $A \in I(M)$ and $B \in S(M)$ for $B$ is a left singular language [6]. We can obtain that $AB$ is also a right dense left cancellative language. Then $AB \in I(M)$. Therefore, $S(M)$ is not a left ideal and $I(M)$ is not strongly prefix sub-semi-group. Since $I(M)$ is a left ideal, we have $D(M)I(M) \subseteq I(M)$. Since $S(M) \subseteq D(M)$ by Lemma 2.2, we know $S(M)I(M) \subseteq I(M)$. So for all $A \in S(M)$ and $B \in I(M)$, we have $AB \in I(M)$. Therefore, $AB \notin S(M)$ since $S(M) \cap I(M) = \emptyset$. Thus, $S(M)$ is not a right ideal. In order to explain that $I(M)$ is not a right ideal, we give the following example. Let $C = b^+a \cup (\bigcup_{i=1}^{\infty} b^i aX^i)$, $B = \{a\}$, where $X = \{a, b\}$. Then $C \in I(M)$ by Lemma 2.7 and $B \in S(M) \subseteq D(M)$. Then $CB = b^+a^2 \cup (\bigcup_{i=1}^{\infty} b^i aX^i a)$ is a left singular language because $(ba)^2 \in l(CB)$. Hence, $CB \in S(M)$. So $CB \notin I(M)$ by Lemma 2.3. Therefore, $I(M)$ is not a right ideal of $D(M)$.

Finally, we want to explain there is a left cancellative language which is neither semi-singular nor inf-singular. So $D(M)$ is the union of three disjoint sub-classes of $S(M)$, $I(M)$ and the rest.

THEOREM 4.13 $D(M) \setminus (S(M) \cup I(M))$ *is not empty.*

*Proof* Let $X = \{a, b\}$ and $A = b^+a \cup \{b^{i+1}a^2X^i | i \geq 0\}$. Then $Z_A = b^+a = \{b^i a | i \geq 1\}$. Next, we will calculate $G_A$ and $H_A$. For any $x \in X^+$, assume $lg(x) = n$. (1) If $x \in bX^*$, then $b^n ax \notin AX_x^+$ and $b^n axX^* \cap A = \emptyset$. So $bX^* \subseteq S_A$. (2) If $x \in aX^*$, then we want to prove $x \notin S_A$. For any $b^i a \in b^+a$, where $i \geq 1$, when $i < n$, we have $b^i ax \in AX_x^+$; when $i \geq n$, we have $b^i axX^* \cap A \neq \emptyset$. For any $b^{i+1}a^2w \in b^{i+1}a^2X^i$, where $i \geq 0$ and $w \in X^i$, we have $(b^{i+1}a^2w)x = (b^{i+1}a)(awx) \in AX_x^+$. So $x \notin S_A$ for all $x \in aX^*$. From Equations (1) and (2), we get $S_A = bX^*$ and $G_A = \{b\}$.

For all $x \in aX^*$, we want to prove that $xX^* \subseteq L_A$. Since $x \in aX^*$, we let $x = ax_1$, where $x_1 \in X^*$ and $lg(x_1) = n - 1$. For every $y \in X^*$, let $lg(y) = m$. We can find a word $v = b^{n+m}a \in A$ such that $v(xy) = (b^{n+m}a)(xy) = b^{n+m}a^2(x_1y) \in A$. So $v(xy) \notin AX^+_{xy}$, which implies that $xy \in L_A$. Then $xX^* \subseteq L_A$. And for all $x \in aX^*$ and $m \in X^*$, we have $x \neq bm$ and $b \neq xm$. Then $I_A = aX^*$ and $H_A = \{a\}$.

So $G_A \cup H_A = \{a, b\} = X$ is a maximal prefix code. Then $A \in D(M)$. But $G_A = \{b\}$ and $H_A = \{a\}$ are all not maximal prefix codes. So $A \notin S(M)$ and $A \notin I(M)$. Thus, $A \in D(M) \setminus (S(M) \cup I(M))$. ■

## Conclusion

In the monoid of left cancellative languages, all left singular languages, all non-dense left cancellative languages and all maximal left cancellative languages are semi-singular, while all right dense left cancellative languages are inf-singular. The dense language $A$ in Example 4.4 of the paper is semi-singular, while the dense language $B$ in Example 4.11 is inf-singular. To make a further study, it will be useful to determine which classes of dense cancellative languages are in $S(M)$ and which are in $I(M)$. From [12], we know that every maximal left cancellative language is left dense. So it will be an interesting thing to judge that left dense languages in $D(M)$ are semi-singular or inf-singular. Another kind of left cancellative languages which is rational can also be considered in the future. For the sub-class of $D(M) \setminus (S(M) \cup I(M))$, we have only found an example. Other properties about it can be investigated.

## Acknowledgements

The authors would like to thank the referees for their careful reading of the manuscript and useful suggestions. The research is supported by Natural Science Foundation of Yunnan Province of China #2010CD21.

## References

[1] J. Bestel and D. Perrin, *Theory of Codes*, Academic Press, Orlando, 1985.
[2] C. Chunhua, *The problem of freedom on certain subsemigroups of the submonoid of left singular languages*, Int. J. Comput. Math. 81 (2004), pp. 121–132.
[3] W.-Y. Hsieh and H.J. Shyr, *Left cancellative elements in the monoid of languages*, Soochow J. Math. 4 (1978), pp. 7–15.
[4] R. Johannesson, *Informationstheorie*, Grundlagen der Telekommunikation (in German), Addison-Wesley, 1992.
[5] J. Karhumaki, M. Latteux, and I. Petre, *Commutation with codes*, Theoret. Comput. Sci. 340 (2005), pp. 322–333.
[6] J.L. Lassez and H.J. Shyr, *Prefix properties and equations in the monoid of languages*, Tamkang J. Math. 9 (1978), pp. 5–14.
[7] M. Li and P.M.B. Vitányi, *An Introduction to Kolmogorov Complexity and its Applications*, Springer-Verlag, New York, 1993.
[8] S. Perkins and A.E. Escott, *Extended synchronizing code words for q-ary complete prefix codes*, Discrete Math. 231 (2001), pp. 391–401.
[9] M.P. Schutzenberger, *On an application of semigroup methods to some problems in coding*, IRE Trans. Inform. Theory IT-2 (1956), pp. 47–60.
[10] H.J. Shyr, *Left cancellative subsemigroup of a semigroup*, Soochow J. Math. Nat. Sci. 2 (1976), pp. 25–33.
[11] H.J. Shyr, *Free Monoids and Languages*, 3rd ed., Hon Min Book Company, Taichung, Taiwan, 2001.
[12] H.J. Shyr and S.S. Yu, *Some properties of left cancellative languages*, Proceedings 10th Symposium on Semigroups, Josei University, Japan, 1986, pp. 15–26.
[13] S.S. Yu, *Languages and Codes*, Tsang Hai Book Publishing Company, Taichung, Taiwan, 2005.